UNCERTAINTY AND INCONSISTENCY IN KNOWLEDGE REPRESENTATION

HABILITATIONSSCHRIFT

FACHBEREICH INFORMATIK
UNIVERSITÄT KOBLENZ-LANDAU

Dr. Matthias Thimm

Universität Koblenz-Landau — Fachbereich Informatik — Koblenz, Germany — 2016

ABSTRACT

This habilitation thesis collects works addressing several challenges on handling uncertainty and inconsistency in knowledge representation. In particular, this thesis contains works which introduce quantitative uncertainty based on probability theory into abstract argumentation frameworks. The formal semantics of this extension is investigated and its application for strategic argumentation in agent dialogues is discussed. Moreover, both the computational as well as the meaningfulness of approaches to analyze inconsistencies, both in classical logics as well as logics for uncertain reasoning is investigated. Finally, this thesis addresses the implementation challenges for various kinds of knowledge representation formalisms employing any notion of inconsistency tolerance or uncertainty.

PUBLICATIONS AND DISCLAIMER

This thesis contains six contributions published in different venues:

1. Matthias Thimm. A Probabilistic Semantics for Abstract Argumentation. In Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12). Montpellier, France, August 2012.

2. Tjitze Rienstra, Matthias Thimm, Nir Oren. Opponent Models with Uncertainty for Strategic Argumentation. In Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13). Beijing, China, August 2013.

3. Matthias Thimm. Coherence and Compatibility of Markov Logic Networks. In Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14). Prague, Czech Republic, August 2014.

4. Matthias Thimm. Stream-based Inconsistency Measurement. In International Journal of Approximate Reasoning, 68:68–87. January 2016.

5. Matthias Thimm. On the Expressivity of Inconsistency Measures. In Artificial Intelligence, 234:120–151. February 2016.

6. Matthias Thimm. Tweety - A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation. In Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14). Vienna, Austria, July 2014.

The publication *Opponent Models with Uncertainty for Strategic Argumentation* has been co-authored by Tjitze Rienstra (at the time of writing the publication he was a PhD student at the University of Luxembourg, he is now a Research Associate at the University of Luxembourg) and Nir Oren (Senior lecturer at the University of Aberdeen). The publication has been collaboratively written and my own contributions to the joint work consisted particularly of 1.) initializing the idea of the paper, 2.) drafting the general approaches, 3.) formalizing the third approach with the incorporation of virtual arguments, 4.) implementing all considered approaches, and 5.) performing the empirical evaluation.

All remaining contributions of this thesis have been developed and written by myself alone. But as scientific work is never the work of a single researcher, many people participated through discussions and collaboration on other papers in the material presented.

ACKNOWLEDGMENTS

I thank Steffen Staab for providing a work environment that allowed me to pursue ideas in many different directions, the opportunity to mature in teaching and project leadership, and his guidance through the obstacles of the life as a researcher.

I thank Gerhard Brewka and Ulrich Furbach for being part of my habilitation committee and encouragement.

I thank Tjitze Rienstra and Nir Oren for the joint work on opponent models for strategic argumentation. I thank Anthony Hunter for many discussions on incorporating probabilistic aspects into computational models of argumentation. I thank Gabriele Kern-Isberner for her continued support since writing my diploma thesis. I thank Kristian Kersting for giving me the idea to investigate the notion of coherence for Markov Logic Networks.

Although not reflected in the contributions of this thesis, I also thank the other co-authors of papers and co-editors of collections from the previous four years not already mentioned above (in alphabetical order): Christoph Beierle, Rita Berardi, Marco Casanova, Federico Cerutti, Renata Dividino, Marcelo Falappa, Marc Finthammer, Alejandro Garcia, Olaf Görlitz, Thomas Gottron, Gerd Gröner, Matthias Klusch, Patrick Krümpelmann, Jérôme Kunegis, Ralf Lämmel, Martin Leinberger, Marcin Paprzycki, Julia Perl, Nico Potyka, Cristina Sarasua, Stefan Scheglman, Marcelo Schiessl, Ansgar Scherp, Guillermo Simari, Hannes Strass, Luciano Tamargo, Ingo Timm, Mauro Vallati, Evelyne Viegas.

I thank all my colleagues at the Institute for Web Science at Technologies and the University of Koblenz for their friendship and help.

I thank Eva Thimm for her love and support.

CONTENTS

INTRODUCTION

Knowledge Representation and Reasoning (KR) (Brachman and Levesque, 2004) is the subfield of *Artificial Intelligence* (AI) (Russell and Norvig, 2003) that deals with the issues of logical formalizations of information and the modelling of rational reasoning behaviour. The methods developed within this field can be applied in all areas that benefit from automatic decision-support such as medicine (Shortliffe and Buchanan, 1975), accounting (Vasarhelyi *et al.*, 2005), chemistry (Judson, 2009), and law (Popple, 1996). A particularly important application area for knowledge representation lies in the *Semantic Web* (Antoniou and van Harmelen, 2004). Already today, there are many systems available that make use of semantically represented data like the *Google Knowledge Graph* (Singhal, 2012) or the *Wikidata project*[1] which allows for a structured access to the contents of *Wikipedia*[2]. Formal knowledge representation formalisms, that allow for a uniform method to exchange information, lie at the core of the semantic web. For those, research in the field of description logics (Baader *et al.*, 2003) and ontologies (Baader *et al.*, 2005) is applied in technologies like RDF (*Resource Description Framework*) and OWL (*Web Ontology Language*).

One of the main challenges in KR research is the handling of *uncertain* and *inconsistent* information which is essential for real-world applications. Unreliable sensor data, distorted communication channels, and other noisy data sources demand an uncertain treatment of information in order to produce reliable and robust results. The notion of *uncertainty* here refers to the graded or just unknown assessment of being "true" of some piece of information, from a subjective point of view of a decision-making agent such as a human being. Most of the information any agent possesses is not necessarily strictly true in the actual world and agents have to take into account both uncertainty of factual beliefs—such as "John was supposedly on vacation" and uncertainty on the applicability of rules when deriving new information—such as "When going on vacation, John usually takes his kids with him"—.

Furthermore, besides being uncertain, information may also be inconsistent. The notion of *inconsistency* refers (usually) to multiple pieces of information and represents a conflict between those, i. e., they cannot hold at the same time. The two statements "John is on vacation in California" and "John is at home in New York" represent inconsistent information and in order to draw meaningful conclusions from a knowledge base containing these statements, this conflict has to be consolidated somehow. Moreover, in real-world applications such as decision-support systems, a knowledge base is usually compiled by merging the formalized knowledge of many

1 http://www.wikidata.org
2 http://www.wikipedia.org

different experts. It is unavoidable that different experts contradict each other and that the merged knowledge base is inconsistent.

The notions of uncertainty and inconsistency are orthogonal to each other. For example, a piece of information may be uncertain such as "Tomorrow it will rain with probability 0.9". If this is the only belief an agent possesses, it is consistent (although vague). Multiple pieces of information may be inconsistent such as "The bird Tweety flies and the bird Tweety does not fly". These two pieces of information are contradictory but each one represents a certain statement. Moreover, multiple pieces of information can also be both uncertain and inconsistent such as "Tomorrow it will rain with probability 0.9 and tomorrow it will rain with probability 0.6". These beliefs are individually uncertain and taken together also inconsistent, cf. (Muiño, 2011). Finally, multiple pieces of information can, of course, also be both certain and consistent such as "The bird Tweety flies and the bird Opus does not fly".

The interplay of both uncertain and inconsistent beliefs is the central theme of this thesis. In particular, we investigate how approaches that allow for the treatment of inconsistent information can be extended by introducing (quantitative) uncertainty, and how approaches that allow for the treatment of uncertain information can be extended by introducing a treatment of inconsistencies.

The remainder of this introductory chapter is organized as follows. In Section 1 we give a short overview on formalisms for dealing with uncertainty and inconsistency, in particular on abstract argumentation and inconsistency measurement, which are the main approaches used in the contributions of this thesis. Afterwards we give an overview of these contributions and explain their relationships to each other in Section 2. We conclude this introductory chapter with some final remarks in Section 3.

1 APPROACHES IN KNOWLEDGE REPRESENTATION

In order to deal with the challenges imposed by uncertainty and inconsistency in knowledge representation, several formalisms have been developed that address one or both of these challenges.

For the challenge of representing and reasoning with uncertainty one usually distinguishes between formalisms for *qualitative uncertainty* and formalisms for *quantitative uncertainty*. The former comprise the large class of non-monotonic logics (Gabbay *et al.*, 1994), i. e., logics that do not generally satisfy the property of *monotonicity* of classical logics, which is that conclusions are preserved under the addition of new information. This property is responsible for the inadequacy of classical logic for reasoning under uncertainty as it basically demands that all derivations are strict and can never be given up. In order to allow for uncertain reasoning, these formalisms usually introduce—in one form or the other—some kind of rules that do not necessarily hold in all cases. The archetype of such a rule is the *default rule* from default logic (Reiter, 1980). A default rule represents basically a "rule of thumb" in commonsense reasoning, such as "birds usually fly" or "if the grass is wet then it likely rained" that are true under usual circumstances but can be ignored if more explicit information is available. Some further examples of formalisms following this line are, e. g., answer set programming (Gelfond and Leone, 2002), conditional logics (Nute and Cross, 2002), defeasible logics (Nute, 1994), and computational models of argumentation (see below). The outcome of these formalisms is usually a set of *plausible* inferences, i. e., formulas that can jointly be accepted. Some formalisms distinguish between *strict* inferences (conclusions that can be drawn from a classic-logical part of a knowledge base) and *defeasible* inferences (conclusions that are drawn by involving one or more non-strict rules), but other than that, no further distinction in terms of *quality of inference* is usually provided. This is different for approaches to *quantitative uncertainty* which allow a finer-grained representation of uncertainty. For example, in probabilistic logic (Paris, 1994; Kern-Isberner, 2001) classical (propositional) formulas in a knowledge base can be annotated with probabilities (or intervals of probabilities in e. g. (Lukasiewicz and Kern-Isberner, 1999)) and inferences can be drawn with annotated probabilities as well. This addition of expressivity also increases the computational complexity of these formalisms considerably and a lot of restrictions on probabilistic logics and alternative formalisms have been devised. Popular examples of formalisms that restrict the expressivity of probability theory are probabilistic networks such as Bayesian networks and Markov nets (Pearl, 1988). In these formalisms, certain (in-)dependence assumptions between probabilistic statements are made explicit in order to allow for performant reasoning capabilities. Alternative formalisms to probability theory such as fuzzy logics (Cintula *et al.*, 2011) and Dempster-Shafer theory (Shafer, 1976) also provide knowledge representation approaches that do not suffer from the computational complexity of probability theory.

For dealing with inconsistencies several approaches have been developed as well and some of the previously mentioned ones, such as computational models of argumentation, can also be regarded as instances of this class. Other examples include paraconsistent logics (Béziau *et al.*, 2007) which are formalisms based on classical logic that allow reasoning with inconsistent information by introducing new truth values for conflicting propositions. Approaches to inconsistency measurement (see below) can be used to analyse the severity of inconsistencies and to provide help in consolidating them. The fields of belief revision (Hansson, 2001) and belief merging (Cholvy and Hunter, 1997; Konieczny and Pérez, 1998) deal with the particular case of inconsistencies in dynamic settings. Usually, when new observations are made in a dynamic environment these observations can contradict with previously held beliefs and old beliefs have to be forgotten. While the fields of belief revision and belief merging are usually focusing on classical logic there are also approaches for non-monotonic formalisms such as answer set programming, cf. e. g. (Slota and Leite, 2012).

In the following, we give a brief overview on the two most relevant fields for this thesis: *Computational Models of Argumentation* and *Inconsistency Measurement*.

Computational Models of Argumentation

Computational models of argumentation (Besnard and Hunter, 2008; Rahwan and Simari, 2009) are approaches that deal with the representation and interaction of arguments and counterarguments. The seminal work of Dung (Dung, 1995) introduces an *abstract argumentation framework* as a directed graph $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ where the vertices Arg model the arguments and the edges $\rightarrow$ model a conflict relation between arguments. If for two arguments $\mathcal{A}, \mathcal{B}$ we have $(\mathcal{A}, \mathcal{B}) \in \rightarrow$—also denoted as $\mathcal{A} \rightarrow \mathcal{B}$—then we say that "$\mathcal{A}$ attacks $\mathcal{B}$". For example, the natural language argument "Tweety is a penguin and, thus, does not fly" is an argument that attacks the argument "Tweety is a bird and as birds typically fly, Tweety flies as well". A central notion in abstract argumentation frameworks is that of an *extension*, a subset of the arguments that can collectively be accepted, given the attacks between all arguments. In (Dung, 1995) the four most popular semantics for defining such extensions have been introduced, the grounded, complete, preferred, and stable semantics (for the formal definition see e. g. the contribution "*A Probabilistic Semantics for Abstract Argumentation*" in this thesis). Further semantics that address different points of view of natural argumentation have been defined in later works, see (Baroni *et al.*, 2011) for a survey. Abstract argumentation is a very general and abstract approach to knowledge representation and it could already been shown in (Dung, 1995) that it subsumes many other common KR approaches such as default logic and answer set programming. Further works that build on abstract argumentation frameworks address issues such as normative values of arguments (Bench-Capon, 2003), attacks on attacks (Modgil and Bench-Capon, 2011) and supporting arguments (Amgoud *et al.*, 2008), see also (Brewka *et al.*, 2014) for a recent survey.

Abstract argumentation frameworks are a very simple representation approach for argumentation and are not always suitable to model natural argumentation in a logically sound way (Amgoud and Besnard, 2013). Another area of research on computational models of argumentation addresses *structured argumentation* (Caminada and Amgoud, 2007; Besnard and Hunter, 2008). While abstract argumentation uses arguments as atomic entities, arguments have an inner structure in approaches to structured argumentation. For example, in the framework of *deductive argumentation* (Besnard and Hunter, 2008) classical logic—propositional and first-order logic—is used as the underlying knowledge representation formalism. Arguments are built from classical formulæ by identifying a set of classical formulæ as the *premise* and a single formula as the *conclusion* of an argument, such that the premise entails the conclusion. Therefore, arguments correspond to minimal proofs in the classic logical sense. If a classical logical knowledge base is inconsistent, arguments and counterarguments for different conclusions can be extracted from this knowledge base and put in relation to each other. While (Besnard and Hunter, 2008) bases its framework on classical logic other works such as ASPIC+ (Prakken, 2009) und *Defeasible Logic Programming (DeLP)* (Garcia and Simari, 2004) are based on non-classical formalisms that allow e. g. the use of default reasoning techniques for the construction of arguments. A hybrid approach between abstract and structured argumentation are *Abstract Dialectical Frameworks (ADFs)* (Brewka *et al.*, 2013) which are, like abstract argumentation frameworks, based on graph-theoretic notions but allow more complex conditions for accepting arguments. This framework is theoretically appealing as it has been shown that it subsumes many existing computational models of formal argumentation.

One of the most recent endeavors in research on argumentation is the integration of quantitative uncertainty (Li *et al.*, 2011; Rienstra, 2012; Grossi and van der Hoek, 2012; Dunne *et al.*, 2011; Thimm, 2012; Fazzinga *et al.*, 2013; Grossi and van der Hoek, 2013; Hunter, 2013; Hunter and Thimm, 2014d,a; Hunter, 2014; Gabbay, 2012; Verheij, 2014; Dondio, 2014; Baroni *et al.*, 2014; Polberg and Doder, 2014; Doder and Woltran, 2014) which is also the focus of the first two contributions of this thesis, "*A Probabilistic Semantics for Abstract Argumentation*" and "*Opponent Models with Uncertainty for Strategic Argumentation*".

Another focus of recent research is the computational complexity of argumentation (Dunne and Wooldridge, 2009) and the resulting algorithmic challenges for reasoning with argumentation systems (Bistarelli *et al.*, 2014). This is further exemplified by the *International Competition on Computational Models of Argumentation*[3] which is being conducted for the first time in 2015.

Further challenges in research on computational models of argumentation involve, among others, the dynamics of argumentation (Baumann, 2012; Coste-Marquis *et al.*, 2014; Baroni *et al.*, 2013) and the application of argumentation and negotiation in agent dialogues (Amgoud *et al.*, 2000; Karunatillake *et al.*, 2009; Black and Hunter, 2007; Rienstra *et al.*, 2013; Thimm, 2014b; Rahwan and Larson, 2009). The latter comprises of works on protocols and frameworks for dialogues (Amgoud *et al.*, 2000; Karunatillake *et al.*, 2009; Black and Hunter, 2007), as well as strategic aspects of argument selection (Rienstra *et al.*, 2013; Thimm, 2014b), and the relationships to game theory (Rahwan and Larson, 2009).

Inconsistency Measurement

The field of *Inconsistency Measurement* (Hunter and Konieczny, 2004; Grant and Hunter, 2006) is about quantitatively assessing the severity of inconsistency in knowledge bases. The main object of research are *inconsistency measures*, i. e., functions that assign a non-negative real value to a knowledge base with the informal meaning that larger values indicate a larger

3 http://argumentationcompetition.org

inconsistency. These kind of measures are useful for the tasks of analyzing knowledge bases in general (Thimm, 2014a), identifying the culprits of inconsistency (Hunter and Konieczny, 2010), as well as manual and automatic debugging of knowledge bases (Grant and Hunter, 2011; Potyka and Thimm, 2014) and inconsistent-tolerant reasoning (Potyka and Thimm, 2015). The traditional setting for inconsistency measurement is that of classical propositional logic and a lot of proposals of inconsistency measures have been made for this setting (Hunter and Konieczny, 2004, 2008, 2010; Ma *et al.*, 2009; Mu *et al.*, 2011a; Xiao and Ma, 2012; Grant and Hunter, 2011, 2013; McAreavey *et al.*, 2014; Jabbour *et al.*, 2014b; Thimm, 2014c). One simple example is the *MI-inconsistency measure* $\mathcal{I}_{\mathsf{MI}}$ where $\mathcal{I}_{\mathsf{MI}}(\mathcal{K})$ is defined as the number of minimal inconsistent subsets[4] of a knowledge base $\mathcal{K}$ of propositional formulæ. More elaborate inconsistency measures can be found in the mentioned references.

Besides measures for propositional logic, there have also been proposals for inconsistency measures for other logics, such as classical first-order logic (Grant and Hunter, 2008), description logics (Ma *et al.*, 2007; Zhou *et al.*, 2009), default logics (Doder *et al.*, 2010), and probabilistic and other weighted logics (Ma *et al.*, 2012; Thimm, 2013b, 2014a; Potyka, 2014; Potyka and Thimm, 2014, 2015; De Bona and Finger, 2015).

The development of inconsistency measures is based on the discussion of what can be regarded as more inconsistent than something else. In classical logic, inconsistency is defined as an absolute term which does not directly allow for an quantitative assessment, in contrast to *information measures* (Shannon, 1948; Lozinskii, 1994), i. e., measures that assess the amount of information in a knowledge base. Inconsistency measures can, however, be regarded as the logical counterpart to information measures, as "severe" inconsistency can be interpreted as "too much" information. Research in inconsistency measurement has therefore focused on formalizing "severity of inconsistency" and brought forward a series of *rationality postulates* aimed at addressing different aspects of inconsistency. One example of such a rationality postulate is *Monotony* which requires that the value of inconsistency cannot decrease when adding formulas to a knowledge base. The rationality of many of these postulates is heavily disputed in the community (Besnard, 2014) and, so far, no set of postulates has been generally acknowledged to be desirable.

As the work on theoretical foundations of inconsistency measurement lies still in the main focus of the community, the practical and algorithmic challenges of inconsistency measurement have only very recently been approached in a few works (Ma *et al.*, 2009; McAreavey *et al.*, 2014; Thimm, 2014c).

4 $M \subseteq \mathcal{K}$ is a minimal inconsistent subset if it is classically inconsistent and there is no $M' \subset M$ that is classically inconsistent as well.

2 OVERVIEW OF CONTRIBUTIONS

This thesis contains six contributions addressing various aspects of uncertainty and inconsistency in knowledge representation. The first two contributions, *A Probabilistic Semantics for Abstract Argumentation* and *Opponent Models with Uncertainty for Strategic Argumentation* deal with the introduction of probabilistic reasoning in computational models of argumentation. The following contribution *Coherence and Compatibility of Markov Logic Networks* introduces a notion of inconsistency in *Markov Logic Networks*, an approach for probabilistic reasoning with first-order logic. The contributions *Stream-based Inconsistency Measurement* and *On the Expressivity of Inconsistency Measures* deal with computational challenges and analyses of measuring inconsistency, respectively. The final contribution *Tweety: A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation* provides a practical perspective on all previous contributions by discussing implementation issues. In the following, we will provide a short summary of each contribution and describe their relationships.

A Probabilistic Semantics for Abstract Argumentation

The traditional semantics for abstract argumentation frameworks is essentially two-valued: an argument is either accepted, i. e., contained in an extension, or rejected, i. e., not contained in an extension. As an abstract argumentation framework may possess multiple extensions wrt. some semantics one usually takes either a skeptical perspective on reasoning with argumentation frameworks, where an argument is *skeptically accepted* if it is contained in *all* extensions and *skeptically rejected* when this is not the case, or a credulous perspective, where an argument is *credulously accepted* if it is contained in *some* extension and *credulously rejected* when this is not the case. However, the binary classification of the acceptance status of an argument is quite weak. The labeling-based semantics of Caminada (Caminada and Gabbay, 2009) identifies another value for the acceptance status of an argument on the level of an individual extension, the value *undecided*. This value is assigned to arguments that are neither contained in the extension nor attacked by it and is, in terms of *strength* of acceptance, located between accepted and rejected arguments. In (Wu and Caminada, 2010) this approach is extended by taking all extensions of a particular semantics into account, yielding a *justification status* of arguments ranging from *strong accept* (an argument is contained in all extensions) to *strong reject* (an argument is not contained in any extension). Between these two extreme statuses a lattice of intermediate statuses can be identified such as the status *weak accept* (for every extension the argument is either accepted or undecided).

The contribution *A Probabilistic Semantics for Abstract Argumentation* in this thesis continues the refinement of acceptance statuses of arguments outlined above by taking a quantitative perspective and introducing probabilistic assessments of the acceptance status of an argument. For that, the classical properties demanded from extensions are carried over to probabilistic interpretations. For example, the least disputable property of an extension is its conflict-freeness. For an abstract argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \to)$, a set $E \subseteq \mathsf{Arg}$ is *conflict-free*, if there are no arguments $\mathcal{A}, \mathcal{B} \in E$ such that $\mathcal{A} \to \mathcal{B}$. In other words, if an argument $\mathcal{A}$ is accepted, all arguments attacked by $\mathcal{A}$ must not be accepted. Carrying this notion over to a probabilistic setting and denoting by $P(\mathcal{A}), P(\mathcal{B})$ the probability of $\mathcal{A}$, $\mathcal{B}$, respectively (or the degree of belief one has on accepting $\mathcal{A}$, $\mathcal{B}$, respectively) we obtain the relationship $P(\mathcal{B}) \leq 1 - P(\mathcal{A})$ if $\mathcal{A} \to \mathcal{B}$. That is, the degree of belief in $\mathcal{B}$ is bounded by the inverse of the degree of belief in $\mathcal{A}$. In the extreme case, if $\mathcal{A}$ is completely acceptable, $P(\mathcal{A}) = 1$, we obtain $P(\mathcal{B}) = 0$ with the meaning that $\mathcal{B}$ is completely rejected. Other desirable properties of extensions can be phrased in a similar way, see the contribution and its follow-ups (Hunter and Thimm, 2014c,b,d) for details.

By formalizing desirable properties on the acceptance status of arguments in a probabilistic way, the classical notion of an extension becomes that of a probability function. This allows comparing abstract argumentation with the field of probabilistic reasoning (Pearl, 1988; Paris, 1994) on a technical level. One of the key insights gained in the contribution is that the grounded semantics of abstract argumentation is equivalent to probabilistic reasoning based on the principle of maximum entropy, the latter being a prominent method for reasoning with probabilistic logics (Paris, 1994). This result and others of the contribution bridge the gap between these two fields of study and the general approach of a *probabilistic semantics* for abstract argumentation shows the feasibility of incorporating quantitative uncertainty into computational models of argumentation.

Opponent Models with Uncertainty for Strategic Argumentation

Using probability theory as a means to provide uncertain assessments of the justification status of an argument allows a formal comparison of the fields of uncertain reasoning and abstract argumentation, as outlined above and presented in the contribution *A Probabilistic Semantics for Abstract Argumentation*. Besides utilizing these results for monological argumentation and, thus, decision-support systems, another application area are multiagent systems and negotiation processes. The contribution *Opponent Models with Uncertainty for Strategic Argumentation* investigates the use of probabilistic models of computational models of argumentation for the purpose of strategic move selection in dialogical argumentation, see (Thimm, 2014b) for a survey. Consider the following example, taken from (Thimm, 2014b), with two agents Anna and Bob discussing whether or not the moon-landing happened in 1969:

Anna: The pictures supposedly taken during the moon-landing cannot be authentic as several shadows are inconsistent. So the moon-landing did not happen in 1969.

Bob: Due to reflected light from the Earth, shadows may appear inconsistent but they are not.

Anna: But the American flag that was hissed by the astronauts, fluttered despite the lack of wind.

Bob: The flag did not flutter. Ripples on the flag originating from folding it made it seem to flutter on a picture.

The above dialogue exemplifies how an exchange of arguments can be used to reach a common consensus. These kinds of dialogues offer opportunities for strategic exploitation, in particular, when agents have knowledge about their opponents' skills and beliefs. For example, assume that Anna knows that Bob is not an expert on astronomical phenomena. Then she could bring forward the following argument:

Anna: The amount of Van Allen radiation the astronauts were exposed to during the trip would have been lethal.

In real-world settings for argumentation, there is usually no time to process all arguments to reach a consensus. In such a setting it would have a strategic advantage for Anna to put forward the above argument first, instead of the other ones. Then Bob may be convinced that Anna is right in claiming that the moon-landing did not happen.

The contribution *Opponent Models with Uncertainty for Strategic Argumentation* formalizes the scenarios of the form outlined above and presents different approaches to select the best argument to put forward in a dialogue based on an opponent model. In the simplest case, see also (Oren and Norman, 2009), an opponent model for an agent A consists of the arguments and attacks that A thinks another agent B is aware of. This model can also recursively be extended by also representing information A thinks that B thinks that A is aware of, etc. It can then be used by A to strategically select the next argument to put forward in order to persuade B of some opinion. It is clear that an opponent model can only be an uncertain approximation of what B actually knows and that it is subject to change when B puts forward arguments himself, see also (Hadjinikolis *et al.*, 2013). This can be modelled by more complex opponent models, based on probabilistic assessments of whether B knows certain arguments and attacks, and updating strategies of these models. The contribution extends the simple opponent model of (Oren and Norman, 2009) by incorporating these aspects and providing two novel approaches for strategic move selection. These approaches have also been empirically evaluated in the contribution, giving evidence that they outperform the naive approach in terms of successful persuasion.

Coherence and Compatibility of Markov Logic Networks

The knowledge representation formalism used in the contributions discussed so far was abstract argumentation, which has an explicit notion of *inconsistency* of information, i. e., a conflict between attacking arguments. Furthermore, the contributions extended this formalism by also introducing (quantitative) *uncertainty*. For another knowledge representation formalism this situation is reversed: Markov Logic Networks (Richardson and Domingos, 2006; Domingos and Lowd, 2009). A *Markov logic network* (MLN) L is set $L = \{(\phi_1, g_1), \ldots, (\phi_n, g_n)\}$ with first-order formulas ϕ_i and weights $g_i \in \mathbb{R}$, for $i = 1, \ldots, n$. An MLN L induces a probability function P (the log-linear model of L), based on the idea that formulas with larger weights receive a larger probability. Markov Logic Networks are thus an approach to probabilistic reasoning with first-order logic (Halpern, 1990; Jaeger, 1995), which is closely related to the field of *Statistical Relational Learning* (Getoor and Taskar, 2007; De Raedt, 2008). Interestingly, the approach of MLNs has no notion of inconsistency. Every MLN L has a unique and well-defined log-linear model[5]. However, it is not necessarily true that the log-linear model is *meaningful* and adequately represents the information in the network. For example, when representing weighted formulas such as $(\text{sunny}, 5)$ and $(\text{rain}, -20)$ one would probably expect that at least $P(\text{sunny}) > P(\text{rain})$ for the log-linear model P of the whole MLN. However, this is not guaranteed as other formulas may interfere in the computation of the final probabilities. Furthermore, consider the two weighted formulas $(\psi, 10)$ and $(\neg\psi, 10)$. The log-linear model P of only these two formulas is well-defined and has $P(\psi) = 0.5$. It is questionable whether these probabilities are appropriate and whether it would not be more appropriate to define this set of formulas as inconsistent. In particular, computing a log-linear model P' of an "empty" knowledge base yields $P'(\psi) = 0.5$ as well. Therefore, from the semantical point of view, the difference between inconsistency (contradicting formulas) and ignorance (no knowledge at all) cannot be recognized. This also makes it hard to detect modeling errors, particularly in large knowledge bases.

The contribution *Coherence and Compatibility of Markov Logic Networks* addresses the issue outlined above by developing a theory of coherence of MLNs. Informally speaking, an MLN is coherent if each weight of a formula in the MLN is adequately reflected in the log-linear model of the MLN. This informal notion is implemented in the contribution by defining a family of coherence measures that measure and aggregate the distances of each weighted formula to its value in the actual log-linear model. The main results of this contribution show that this family adequately represents a notion of inconsistency for MLNs and satisfies several desirable properties, such as monotonicity (adding formulas to the knowledge base can only decrease coherence). By utilizing the framework of coherence measurement

5 There is, however, a minor exception to this statement. If infinite weights are allowed in the representation of an MLN the log-linear model may be undefined.

for analyzing a given MLN the knowledge engineer is able to detect several design flaws, in particular when observing an unintuitive reasoning behavior and if future extensions of the MLN are envisaged:

1. If an MLN is coherent (i. e. has a comparatively large coherence value) but exhibits unintuitive inferences, then probably some weights have been chosen wrong (as there is only little interdependence between formulas).
2. If an MLN is coherent and exhibits no unintuitive inferences, then the MLN is a good representation of the given knowledge and it will probably be easier to extend it.
3. If an MLN is incoherent (i. e. has a comparatively low coherence value) and exhibits unintuitive inferences, then the knowledge engineer should have a look into the structure of the knowledge base as there may be unwanted interdependences amongst formulas.
4. If an MLN is incoherent but exhibits no unintuitive inferences, then the MLN may not be an adequate representation of the knowledge and further extensions might yield unintuitive results.

The coherence framework has also been applied to the problem of measuring the *compatibility* between different MLNs. The problem of knowledge aggregation (or information fusion) (Konieczny and Pérez, 1998; Cholvy and Hunter, 1997; Everaere *et al.*, 2015) is about combining knowledge bases from different sources into a joint knowledge base. For classical logics, this raises the issue of consistency-preservation if the joint knowledge base is inconsistent. As there is no notion of inconsistency for MLNs, the coherence framework can be used to judge the compatibility of different MLNs before merging them. This is done in the contribution by comparing the coherence of the individual MLNs with the coherence of the joint MLN and providing a quantitative measure for this judgement.

Stream-based Inconsistency Measurement

The computational challenges in determining the coherence value of an MLN are huge, see the contribution *Coherence and Compatibility of Markov Logic Networks* for details. This is true for most approaches to inconsistency measurement, included the coherence framework discussed above. Even for classical propositional logic the problem of determining consistency of a knowledge base is NP-complete (Papadimitriou, 1994). *Measuring* inconsistency can become even harder, in particular for large knowledge bases as they appear in, e. g., Semantic Web applications (Sacramento *et al.*, 2012).

The contribution *Stream-based Inconsistency Measurement* addresses the computational problem of measuring inconsistency by considering scenarios where the knowledge base can only be processed in a step-by-step fashion, i.e., in streams. The contribution develops a formal framework for inconsistency measurement in streams and defines an abstract notion for stream-based inconsistency measures. Classical approaches to inconsistency measurement in knowledge bases are adapted to the streaming scenario by providing window-based variants and their accuracy is formally investigated. The contribution also introduces a completely novel inconsistency measure based on *hitting sets* (see the contribution for technical details) and provides a stream-based approximation algorithm for it that is shown to outperform the other stream-based variants in terms of both accuracy and runtime performance.

The computational issues in measuring inconsistency are usually been ignored in the mostly theoretical field of inconsistency handling, as far as we know only the works (Ma *et al.*, 2009; McAreavey *et al.*, 2014) address similar issues as well. The contribution *Stream-based Inconsistency Measurement* therefore addresses this important need and shows that large-scale inconsistency measurement is feasible.

On the Expressivity of Inconsistency Measures

The contribution *Stream-based Inconsistency Measurement* discussed above addresses the computational challenges in measuring inconsistency. However, there is still the discussion on what a reasonable inconsistency measure exactly is, see (Besnard, 2014) for a recent discussion on this topic. The theory of inconsistency measurement is based on formal principles (postulates) that describe reasonable properties for inconsistency measurement, such as monotonicity when adding new formulas. Many of these postulates are disputed and there is still no consensus on how inconsistency should be quantified (Besnard, 2014). While the contribution *Stream-based Inconsistency Measurement* discusses computational complexity as another important dimension for deciding whether an inconsistency measure is reasonable, the contribution *On the Expressivity of Inconsistency Measures* introduces yet another important dimension, namely the *expressivity* of inconsistency measures. Informally speaking, an inconsistency measure $\mathcal{I}$ is more expressive than another inconsistency measure $\mathcal{I}'$ if $\mathcal{I}$ can distinguish between more inconsistent knowledge bases—i.e., assigns different values of inconsistency—than $\mathcal{I}'$. In the contribution this concept is formalized by introducing four different *expressivity characteristics*, i.e., measures that assign to any inconsistency measure the number of different values they can produce on some class of knowledge bases. The contribution surveys 15 different inconsistency measures from the recent literature and provides a thorough investigation of these measures wrt. the four proposed expressivity characteristics. As a result it could be shown that the measure $\mathcal{I}^{\Sigma}_{\text{dalal}}$ from (Grant and Hunter, 2013) and the measure $\mathcal{I}_{P_m}$ from (Jabbour and Raddaoui, 2013) are maximally expressive wrt. all four characteristics.

Besides introducing and investigating expressivity as a desirable property for inconsistency measures, the contribution *On the Expressivity of In-*

consistency Measures also uncovers some interesting relationships between the field of inconsistency measurement and several branches from mathematics. For example, there is a strong relationship between inconsistency measures based on the notion of minimal inconsistent sets and Sperner families[6] (Sperner, 1928) from set theory. Roughly speaking, the set of minimal inconsistent subsets of a knowledge base is also a Sperner family wrt. the knowledge base and, moreover, *every* Sperner family can be represented as the set of minimal inconsistent subsets of some knowledge base. From this strong relationship, existing results on Sperner families can be used to obtain interesting properties on measures based on minimal inconsistent subsets, such as, e. g., their values wrt. the introduced expressivity characteristics. Further interesting relationships discussed in *On the Expressivity of Inconsistency Measures* concern profiles of monotone Boolean functions and the Farey series[7], see the contribution for details.

Tweety: A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation

All the contributions discussed so far address the issues of uncertainty and inconsistency in knowledge representation in a mostly theoretical and analytical way, as it is common in the field of knowledge representation and reasoning. However, all technical approaches of the contributions have also been implemented and, if meaningful, the implementations have been practically evaluated as well. As the common framework for implementing these approaches, the *Tweety libraries for Knowledge Representation and Artificial Intelligence*[8] have been employed. The final contribution in this thesis, *Tweety: A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation*, presents this framework, which provides a general basis for implementing various approaches to knowledge representation and other formal approaches to artificial intelligence. As of now, it contains 33 different *Java* libraries that implement standard approaches such as propositional and first-order logic, a variety of approaches to formal argumentation such as abstract argumentation (Dung, 1995), deductive argumentation (Besnard and Hunter, 2001), *Defeasible Logic Programming* (Garcia and Simari, 2004), a general framework for belief revision (Hansson, 2001), answer set programming (Gelfond and Leone, 2002), and others. It also contains tools for working with mathematical expressions, linear programs, graphs, matrices, and provides bridges to third-party programs such as SAT-solvers, optimization solvers, and others.

Tweety is an open source project and can therefore be used and extended by everyone. In particular, instantiating the abstract Tweety classes for a

6 A Sperner family S wrt. some set X is a set of subsets of X where no subset is contained in another.

7 See `http://oeis.org/A005728`

8 `http://tweetyproject.org`

particular formalism is simple. Although Tweety is implemented in a object-oriented programming language it follows a strict declarative formal way to define concepts from theoretical knowledge representation research. The libraries of Tweety are under constant development with a current release cycle of six months.

3 CONCLUDING REMARKS

The contributions collected in this thesis deal with different aspects related to both uncertainty and inconsistency of information in knowledge representation. Of course, there are other works with similar aims and we refer the reader to the corresponding discussions on related works in the individual contributions for detailed analyses. Recent works of the author that address the same challenges but are not contained in this thesis can also be found in (Thimm, 2013b; Potyka and Thimm, 2014, 2015; Hunter and Thimm, 2014c,b,d,a; Thimm and Kern-Isberner, 2014).

A PROBABILISTIC SEMANTICS FOR ABSTRACT ARGUMENTATION

Matthias Thimm. A Probabilistic Semantics for Abstract Argumentation. In Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12). Montpellier, France, August 2012.

Abstract

Classical semantics for abstract argumentation frameworks are usually defined in terms of extensions or, more recently, labelings. That is, an argument is either regarded as accepted with respect to a labeling or not. In order to reason with a specific semantics one takes either a credulous or skeptical approach, i.e. an argument is ultimately accepted, if it is accepted in one or all labelings, respectively. In this paper, we propose a more general approach for a semantics that allows for a more fine-grained differentiation between those two extreme views on reasoning. In particular, we propose a probabilistic semantics for abstract argumentation that assigns probabilities or *degrees of belief* to individual arguments. We show that our semantics generalizes the classical notions of semantics and we point out interesting relationships between concepts from argumentation and probabilistic reasoning. We illustrate the usefulness of our semantics on an example from the medical domain.

1 INTRODUCTION

The field of computational models of argumentation (Rahwan and Simari, 2009) is concerned with non-monotonic reasoning mechanisms that focus on the role of arguments. An argument is an entity that represents some grounds to believe in a certain statement and that can be in conflict with arguments establishing contradictory claims. The most commonly used framework to talk about general issues of argumentation is that of abstract argumentation (Dung, 1995). In abstract argumentation, arguments are represented as atomic entities and the interrelationships between different arguments are modeled using an attack relation. Abstract argumentation has been thoroughly investigated in the past fifteen years and there is quite a lot of work on, e.g. extending abstract argumentation frameworks (Janssen *et al.*, 2008; Li *et al.*, 2011; Dunne *et al.*, 2011) and, in particular, semantical issues (Baroni *et al.*, 2005; Caminada, 2006; Baroni *et al.*, 2010; Wu and Caminada, 2010). Several different kinds of semantics for abstract argumentation frameworks have been proposed that highlight different aspects of argumentation. Usually, semantics are given to abstract argumentation frameworks in terms of extensions or, more recently, labelings. For a specific labeling an argument is either accepted, not accepted, or undecided. In a fixed semantical context, there is usually a set of labelings that is consistent with the semantical context. In order to reason with a semantics one has to take either a credulous or skeptical perspective. That is, an argument is ultimately accepted wrt. a semantics if the argument is accepted by at least one labeling consistent with that semantics (the credulous perspective) or if the argument is accepted by all labelings consistent with the semantics (the skeptical perspective). This extreme points of views may result in undesired results as in extreme cases the set of credulously accepted arguments may contain nearly the whole set of arguments and the set of skeptically accepted set of arguments may be nearly empty.

In this paper we propose a new way to assign semantics to abstract argumentation frameworks. More precisely, instead of using labelings we use probability functions on subsets of arguments as interpretations and define a probabilistic satisfaction relation that generalizes the notion of a complete labeling. In contrast to other works that combine abstract argumentation with quantitative uncertainty (Li *et al.*, 2011; Janssen *et al.*, 2008; Dung and Thang, 2010; Dunne *et al.*, 2011; Krause *et al.*, 1995; Kohlas, 2003; Alsinet *et al.*, 2008) we do not extend the underlying notion of an abstract argumentation framework but assess its inherent uncertainty using a more general semantics. In order to reason with this semantics we adopt notions from probabilistic reasoning for reasoning with sets of probability functions. We show that probabilistic semantics allow for a more fine-grained view on the relationships of arguments within an abstract argumentation framework.

On a more wider perspective, this paper also gives some first insights on the relationships between two of the most important sub-fields of artificial intelligence, namely argumentation and probabilistic reasoning (Pearl, 1998; Paris, 1994). In particular, we show that the grounded labeling in abstract argumentation corresponds to the maximum entropy model in probabilistic reasoning (wrt. our probabilistic semantics for abstract argumentation frameworks).

The rest of this paper is organized as follows. In Section 2 we give a brief overview on abstract argumentation and exemplify the problems raised above to motivate our approach. In Section 3 we introduce a probabilistic semantics for abstract argumentation frameworks and discuss its properties. We continue in Section 4 with a comparison of the probabilistic semantics and classical semantics and show interesting relationships between notions from argumentation and probabilistic reasoning. In Section 5 we illustrate the usefulness of the approach with a short example and discuss related work in Section 6. We conclude in Section 5 with a summary and discussion.

2 ABSTRACT ARGUMENTATION

Abstract argumentation frameworks (Dung, 1995) take a very simple view on argumentation as they do not presuppose any internal structure of an argument. Abstract argumentation frameworks only consider the interactions of arguments by means of an attack relation between arguments.

Definition 1 (Abstract Argumentation Framework)**.** An *abstract argumentation framework* AF is a tuple $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ where Arg is a set of arguments and $\rightarrow$ is a relation $\rightarrow \subseteq \mathsf{Arg} \times \mathsf{Arg}$.

For two arguments $\mathcal{A}, \mathcal{B} \in \mathsf{Arg}$ the relation $\mathcal{A} \rightarrow \mathcal{B}$ means that argument $\mathcal{A}$ attacks argument $\mathcal{B}$. Abstract argumentation frameworks can be concisely represented by directed graphs, where arguments are represented as nodes and edges model the attack relation.

Example 1. Consider the abstract argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ depicted in Fig. 1. Here it is $\mathsf{Arg} = \{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5\}$ and $\rightarrow = \{(\mathcal{A}_1, \mathcal{A}_2), (\mathcal{A}_2, \mathcal{A}_1), (\mathcal{A}_2, \mathcal{A}_3), (\mathcal{A}_3, \mathcal{A}_4), (\mathcal{A}_4, \mathcal{A}_5), (\mathcal{A}_5, \mathcal{A}_4), (\mathcal{A}_5, \mathcal{A}_3)\}$.

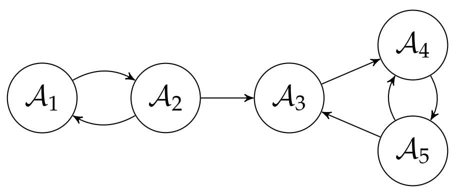


Figure 1: A simple argumentation framework

Semantics are usually given to abstract argumentation frameworks by means of extensions (Dung, 1995) or labelings (Wu and Caminada, 2010). An *extension* E of an argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ is a set of arguments $E \subseteq \mathsf{Arg}$ that gives some coherent view on the argumentation underlying AF. A labeling L is a function $L : \mathsf{Arg} \rightarrow \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ that assigns to each argument $\mathcal{A} \in \mathsf{Arg}$ either the value in, meaning that the argument is accepted, out, meaning that the argument is not accepted, or undec, meaning that the status of the argument is undecided. Let $\mathsf{in}(L) = \{\mathcal{A} \mid L(\mathcal{A}) = \mathsf{in}\}$ and $\mathsf{out}(L)$ resp. $\mathsf{undec}(L)$ be defined analogously. As extensions can be characterized by the arguments that labeled in in some labeling, we restrain our attention to labelings henceforth. In order to distinguish extension- and labeling-based semantics to the probabilistic semantics in the next section we denote the former *classical semantics*.

In the literature (Dung, 1995; Caminada, 2006) a wide variety of different types of classical semantics has been proposed. Arguably, the most important property of a semantics is its admissibility. A labeling L is called *admissible* if and only if for all arguments $\mathcal{A} \in \mathsf{Arg}$

1. if $L(\mathcal{A}) = \mathsf{out}$ then there is $\mathcal{B} \in \mathsf{Arg}$ with $L(\mathcal{B}) = \mathsf{in}$ and $\mathcal{B} \rightarrow \mathcal{A}$, and
2. if $L(\mathcal{A}) = \mathsf{in}$ then $L(\mathcal{B}) = \mathsf{out}$ for all $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \rightarrow \mathcal{A}$,

and it is called *complete* if, additionally, it satisfies

3. if $L(\mathcal{A}) = \mathsf{undec}$ then there is no $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \rightarrow \mathcal{A}$ and $L(\mathcal{B}) = \mathsf{in}$ and there is a $\mathcal{B}' \in \mathsf{Arg}$ with $\mathcal{B}' \rightarrow \mathcal{A}$ and $L(\mathcal{B}') \neq \mathsf{out}$.

The intuition behind admissibility is that an argument can only be accepted if there are no attackers that are accepted and if an argument is not accepted then there has to be some reasonable grounds. The idea behind the completeness property is that the status of argument is only undec if it cannot be classified as in or out. Different types of classical semantics can be phrased by imposing further constraints.

Definition 2. Let $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ be an abstract argumentation framework and $L : \mathsf{Arg} \rightarrow \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ a complete labeling.

- L is *grounded* if and only if $\mathsf{in}(L)$ is minimal.
- L is *preferred* if and only if $\mathsf{in}(L)$ is maximal.
- L is *stable* if and only if $\mathsf{undec}(L) = \emptyset$.
- L is *semi-stable* if and only if $\mathsf{undec}(L)$ is minimal.

All statements on minimality/maximality are meant to be with respect to set inclusion.

Note that a grounded labeling is uniquely determined and always exists (Dung, 1995). Besides the above mentioned types of classical semantics there are a lot of further proposals such as *CF2 semantics* (Baroni *et al.*, 2005). However, in this paper we focus on complete, grounded, preferred, stable, and semi-stable semantics.

Example 2. We continue Ex. 1. Consider the labeling L defined via

$$L(\mathcal{A}_1) = \mathsf{in} \qquad L(\mathcal{A}_2) = \mathsf{out} \qquad L(\mathcal{A}_3) = \mathsf{out}$$
$$L(\mathcal{A}_4) = \mathsf{out} \qquad L(\mathcal{A}_5) = \mathsf{in}$$

Clearly, L is an admissible labeling as it satisfies properties 1.) and 2.) from above. Additionally, it is complete and also preferred, stable, and semi-stable. Furthermore, consider the labeling L' defined via

$$L'(\mathcal{A}_1) = \mathsf{out} \qquad L'(\mathcal{A}_2) = \mathsf{in} \qquad L'(\mathcal{A}_3) = \mathsf{out}$$
$$L'(\mathcal{A}_4) = \mathsf{in} \qquad L'(\mathcal{A}_5) = \mathsf{out}$$

The labeling L' is also admissible, complete, preferred, stable, and semi-stable. Note, that the grounded labeling L_g is defined via $L_g(\mathcal{A}_1) = L_g(\mathcal{A}_2) = L_g(\mathcal{A}_3) = L_g(\mathcal{A}_4) = L_g(\mathcal{A}_5) = \mathsf{undec}$.

As one can see in the above example, most semantics are *multi-extension* semantics. That is, there is not always a unique labeling induced by the semantics. In order to reason with multi-extension semantics, usually, one takes either a credulous or skeptical perspective. That is, an argument $\mathcal{A}$ is *credulously inferred* with semantics $\mathcal{S} \in \{\text{complete}, \text{preferred}, \text{stable}, \text{semi-stable}\}$ if there is a $\mathcal{S}$-labeling L with $L(\mathcal{A}) = \mathsf{in}$. An argument $\mathcal{A}$ is *skeptically inferred* with semantics $\mathcal{S}$ if for all $\mathcal{S}$-labelings L it holds that $L(\mathcal{A}) = \mathsf{in}$. Taking either a credulous or skeptical perspective is a crucial choice as the set of inferred arguments might change drastically.

Example 3. We continue Ex. 2. Besides L and L' there is also another labeling L'' that is admissible, complete, preferred, stable, and semi-stable:

$$L''(\mathcal{A}_1) = \mathsf{out} \qquad L''(\mathcal{A}_2) = \mathsf{in} \qquad L''(\mathcal{A}_3) = \mathsf{out}$$
$$L''(\mathcal{A}_4) = \mathsf{out} \qquad L''(\mathcal{A}_5) = \mathsf{in}$$

With respect to complete, preferred, stable, and semi-stable semantics, it follows that no argument is skeptically inferred and all arguments but $\mathcal{A}_3$ are credulously inferred.

The example above shows that the difference of skeptical and credulous inference may be huge. Consequently, it is hard to assess the quality or strength of argument in an argumentation framework if only those types of inference are considered, cf. (Matt and Toni, 2008). Imagine that the arguments in Ex. 1 are interpreted within a decision-support system in the medical domain. That is, the arguments $\mathcal{A}_1, \ldots, \mathcal{A}_5$ represent different drugs for a specific disease and an attack means a negative "influence" of one drug to another. In this system, a decision comprises a set of drugs that are used for treatment and the question is how to select this set? With credulous semantics the recommendation is to administer almost all drugs and with skeptical semantics the recommendation is to administer no drug. None of these recommendations seem appropriate in the example. In particular, administering no drug at all may not be possible as some action may be required to be performed. Another possible way to select the set of drugs is to select the drugs from one specific labeling. But then the question arises which labeling to chose?

3 PROBABILISTIC SEMANTICS

In order to get a more fine-grained view on the status of arguments we propose a new semantics that generalizes classical semantics and is based on a probabilistic interpretation of arguments. For that, we need some further notation. Let $\mathfrak{P}(\mathcal{X})$ denote the power set of a set $\mathcal{X}$.

Definition 3. Let $\mathcal{X}$ be some finite set. A *probability function* P on $\mathcal{X}$ is a function $P : \mathfrak{P}(\mathcal{X}) \to [0,1]$ that satisfies

1. $P(\mathcal{X}) = 1$ and
2. $P(X_1 \cup X_2) = P(X_1) + P(X_2)$ for $X_1, X_2 \subseteq \mathcal{X}$, $X_1 \cap X_2 = \emptyset$.

For $x \in \mathcal{X}$ we write $P(x)$ instead of $P(\{x\})$. Here, a probability function is a function on the set of subsets of some (finite) set with two characteristic properties. First, the function must be *normalized*, i. e., the whole set must have probability one (property 1 above). Second, the probability of the union of two disjoint set is the sum of the probabilities of each set (property 2 above). These two properties are also called the *Kolmogorov properties of probability* (Jaynes, 2003). The following observation is easy to see and the proof can be found e. g. in (Paris, 1994).

Proposition 1. *For $X \subseteq \mathcal{X}$ and a probability function P on $\mathcal{X}$ it holds*

$$P(X) = \sum_{x \in X} P(x)$$

Due to the above proposition a probability function can be defined just by defining the probabilities for each $x \in \mathcal{X}$.

A probability function is usually used to model statistical events. Then $\mathcal{X}$ is the set of all possible atomic events and subset X of $\mathcal{X}$ represents the disjunction of the events in X. Given that $\mathcal{X}$ contains all possible events, property 1 above says that one event has to occur and property 2 states that atomic events are mutually exclusive.

In this paper, we use another interpretation for probability, that of *subjective probability* (Paris, 1994). There, a probability $P(X)$ for some $X \subseteq \mathcal{X}$ denotes the *degree of belief* we put into X. Then a probability function P can be seen as an epistemic state of some agent that has uncertain beliefs with respect to $\mathcal{X}$. In probabilistic reasoning (Pearl, 1998; Paris, 1994), this interpretation of probability is widely used to model uncertain knowledge representation and reasoning.

In the following, we consider probability functions on sets of arguments of an abstract argumentation frameworks. Let $\mathsf{AF} = (\mathsf{Arg}, \to)$ be some fixed abstract argumentation framework and let $\mathcal{E} = \mathfrak{P}(\mathsf{Arg})$ be the set of all sets of arguments. Let now $\mathcal{P}_{\mathsf{AF}}$ be the set of probability functions of the form $P : \mathfrak{P}(\mathcal{E}) \to [0,1]$. A probability function $P \in \mathcal{P}_{\mathsf{AF}}$ assigns to each set of possible extensions of AF a probability, i. e., $P(e)$ for $e \in \mathcal{E}$ is the probability that e is an extension and $P(E)$ for $E \subseteq \mathcal{E}$ is the probability that any of the sets in E is an extension. In particular, note the difference between e. g. $P(\{\mathcal{A}, \mathcal{B}\}) = P(\{\{\mathcal{A}, \mathcal{B}\}\})$ and $P(\{\{\mathcal{A}\}, \{\mathcal{B}\}\})$ for arguments $\mathcal{A}, \mathcal{B}$. While the former denotes the probability that $\{\mathcal{A}, \mathcal{B}\}$ is an extension the latter denotes the probability that $\{\mathcal{A}\}$ or $\{\mathcal{B}\}$ is an extension. In general, it holds $P(\{\mathcal{A}, \mathcal{B}\}) \neq P(\{\{\mathcal{A}\}, \{\mathcal{B}\}\})$.

For $P \in \mathcal{P}_{\mathsf{AF}}$ and $\mathcal{A} \in \mathsf{Arg}$ we abbreviate

$$P(\mathcal{A}) = \sum_{\mathcal{A} \in e \subseteq \mathsf{Arg}} P(e) \quad .$$

Given some probability function P, the probability $P(\mathcal{A})$ represents the degree of belief that $\mathcal{A}$ is in an extension (according to P), i. e., $P(\mathcal{A})$ is the sum of the probabilities of all possible extensions that contain $\mathcal{A}$. The set $\mathcal{P}_{\mathsf{AF}}$ contains all possible views one can take on the arguments of an abstract argumentation framework AF.

Example 4. We continue Ex. 1. Consider the function $P \in \mathcal{P}_{\mathsf{AF}}$ defined via $P(\{\mathcal{A}_1, \mathcal{A}_3, \mathcal{A}_5\}) = 0.3$, $P(\{\mathcal{A}_1, \mathcal{A}_4\}) = 0.45$, $P(\{\mathcal{A}_5, \mathcal{A}_2\}) = 0.1$, $P(\{\mathcal{A}_2, \mathcal{A}_4\}) = 0.15$, and $P(e) = 0$ for all remaining $e \in \mathcal{E}$. Due to Prop. 1 the function P is well-defined as e. g.

$$\begin{aligned} &P(\{\{\mathcal{A}_5, \mathcal{A}_2\}, \{\mathcal{A}_2, \mathcal{A}_4\}, \{\mathcal{A}_3\}\}) \\ &\quad = P(\{\mathcal{A}_5, \mathcal{A}_2\}) + P(\{\mathcal{A}_2, \mathcal{A}_4\}) + P(\{\mathcal{A}_3\}) \\ &\quad = 0.1 + 0.15 + 0 = 0.25 \end{aligned}$$

Therefore, P is a probability function according to Def. 3. According to P the probabilities of each argument of AF compute to $P(\mathcal{A}_1) = 0.75$, $P(\mathcal{A}_2) = 0.25$, $P(\mathcal{A}_3) = 0.3$, $P(\mathcal{A}_4) = 0.6$, and $P(\mathcal{A}_5) = 0.4$.

In the following, we are only interested in those probability functions of $\mathcal{P}_{\mathsf{AF}}$ that agree with our intuition on the interrelationships of arguments and attack. For example, if an argument $\mathcal{A}$ is not attacked we should completely believe in its validity if no further information is available. We propose the following notion of *justifiability* to describe this intuition.

Definition 4. A probability function $P \in \mathcal{P}_{\mathsf{AF}}$ is called *p-justfiable* wrt. AF, denoted by $P \models_J \mathsf{AF}$, if it satisfies for all $\mathcal{A} \in \mathsf{Arg}$

1. $P(\mathcal{A}) \leq 1 - P(\mathcal{B})$ for all $\mathcal{B}, \in \mathsf{Arg}$ with $\mathcal{B} \to \mathcal{A}$ and
2. $P(\mathcal{A}) \geq 1 - \sum_{\mathcal{B} \in \mathcal{F}} P(\mathcal{B})$ where $\mathcal{F} = \{\mathcal{B} \mid \mathcal{B} \to \mathcal{A}\}$.

Let $\mathcal{P}^J_{\mathsf{AF}}$ be the set of all p-justifiable probability functions wrt. AF.

The notion of p-justifiability generalizes the concept of complete semantics to the probabilistic setting. Property 1.) says that the degree of belief we assign to an argument $\mathcal{A}$ is bounded from above by the inverse degrees of belief we put into the attackers of $\mathcal{A}$. As a special case, note that if we completely believe in an attacker of $\mathcal{A}$, i. e. $P(\mathcal{B}) = 1$ for some $\mathcal{B}$ with $\mathcal{B} \to \mathcal{A}$, then it follows $P(\mathcal{A}) = 0$. This corresponds to property 1.) of a complete labeling, see Section 2. Property 2.) of Def. 4 says that the degree of belief we assign to an argument $\mathcal{A}$ is bounded from below by the inverse of the sum of the degrees of belief we we put into the attacks of $\mathcal{A}$. As a special case, note that if we completely disbelieve in all attackers of $\mathcal{A}$, i. e. $P(\mathcal{B}) = 0$ for all $\mathcal{B}$ with $\mathcal{B} \to \mathcal{A}$, then it follows $P(\mathcal{A}) = 1$. This corresponds to property 2.) of a complete labeling, see Section 2. The following proposition establishes the probabilistic analogue of the third property of a complete labeling.

Proposition 2. *Let P be p-justifiable and $\mathcal{A} \in \mathsf{Arg}$. If $P(\mathcal{A}) \in (0, 1)$ then*

1. *there is no $\mathcal{B} \in \mathsf{Arg}$ with $\mathcal{B} \to \mathcal{A}$ and $P(\mathcal{B}) = 1$ and*
2. *there is a $\mathcal{B}' \in \mathsf{Arg}$ with $\mathcal{B}' \to \mathcal{A}$ and $P(\mathcal{B}') > 0$.*

Before we investigate the relationships between our probabilistic semantics and classical argumentation semantics in more depth we analyze the properties of probabilistic semantics by itself.

Example 5. We continue Ex. 4. There, the probability function P is p-justifiable wrt. AF as e. g. $P(\mathcal{A}_1) \leq 1 - P(\mathcal{A}_2)$ and $P(\mathcal{A}_4) \geq 1 - P(\mathcal{A}_3) - P(\mathcal{A}_5)$.

The set of p-justifiable probability functions contains all probability functions that agree with our intuition of argumentation. This set has some nice properties as shown below.

Proposition 3. *The set $\mathcal{P}^J_{\mathsf{AF}}$ is non-empty and convex.*

The above proposition states that for every argumentation framework AF there is a p-justifiable probability function P wrt. AF. Furthermore, the set of p-justifiable probability functions is closed wrt. to convex combination. That is, given two p-justifiable probability function P_1, P_2 and some $\delta \in [0, 1]$ it follows that P_3 defined via $P_3(e) = \delta P_1(e) + (1 - \delta) P_2(e)$ for each $e \in \mathcal{E}$ is also p-justifiable.

In order to reason with a set of probability functions one can use *model-based inductive reasoning* techniques (Paris, 1994), i. e., instead of reasoning with the complete set one selects some appropriate representative and performs reasoning solely based on this representative. A very important approach for that is reasoning based on the *principle of maximum entropy* (Paris, 1994). For a probability function $P \in \mathcal{P}_{\mathsf{AF}}$ the *entropy* $H(P)$ of P is defined as $H(P) = -\sum_{e \in \mathcal{E}} P(e) \log P(e)$ with $0 \log 0 = 0$. The entropy measures the amount of indeterminateness of a probability function P. A probability function P_1 that describes absolute certain knowledge, i. e. $P_1(e) = 1$ for some $e \in \mathcal{E}$ and $P_1(e') = 0$ for every other $e' \in \mathcal{E}$, yields minimal entropy $H(P_1) = 0$. The uniform probability function P_0 with $P_0(e) = 1/|\mathcal{E}|$ for every $e \in \mathcal{E}$ yields maximal entropy $H(P_0) = -\log 1/|\mathcal{E}|$.

Definition 5. Let $\mathcal{P} \subseteq \mathcal{P}_{\mathsf{AF}}$ be a set of probability functions.

- $P^* \in \mathcal{P}$ is a *maximum entropy model* of $\mathcal{P}$ if $H(P^*)$ is maximal in $\{H(P) \mid P \in \mathcal{P}\}$. Let $\mathsf{MaxE}(\mathcal{P})$ be the set of all maximum entropy models of $\mathcal{P}$.
- $P^* \in \mathcal{P}$ is a *minimum entropy model* of $\mathcal{P}$ if $H(P^*)$ is minimal in $\{H(P) \mid P \in \mathcal{P}\}$. Let $\mathsf{MinE}(\mathcal{P})$ be the set of all minimum entropy models of $\mathcal{P}$.
- P_c is the *centroid* of $\mathcal{P}$ if

$$P_c(e) = \frac{\int_{\mathcal{P}} P(e)dP(e)}{\int_{\mathcal{P}} dP(e)} \quad \text{for all } e \in \mathcal{E}.$$

A maximum entropy model $P \in \mathsf{MaxE}(\mathcal{P})$ is as unbiased as possible among the probability functions in $\mathcal{P}$, i.e., it contains as less information as possible. Reasoning based on the principle of maximum entropy is a popular approach in probabilistic reasoning as it satisfies several nice properties (Paris, 1994). Here, we also consider minimum entropy models as they correspond to stable labelings (see below) and the centroid as further approaches for selecting specific models from a set of probability functions.

Proposition 4. *If $\mathcal{P}$ is a non-empty convex set of probability functions then $|\mathsf{MaxE}(\mathcal{P})| = 1$, i. e. a maximum entropy model exists and is uniquely determined.*

For the proof of the above proposition see e. g. (Paris, 1994). Taking together Propositions 3 and 4 we obtain the following nice observation as a simple corollary.

Corollary 1. *The maximum entropy model P^* of $\mathcal{P}^J_{\mathsf{AF}}$ exists and is uniquely determined.*

Note that the centroid P_c of $\mathcal{P}^J_{\mathsf{AF}}$ is, by definition, also uniquely determined[9] but this is, in general, not true for minimum entropy models.

Example 6. We continue Ex. 4. While both the maximum entropy model P^* and the centroid P_c of $\mathcal{P}^J_{\mathsf{AF}}$ are uniquely determined there are three minimum entropy models $P_1^{\min}, P_2^{\min}, P_3^{\min}$ of $\mathcal{P}^J_{\mathsf{AF}}$. The degrees of beliefs for the arguments of AF wrt. those models are given in Table 1, rounded to two decimal places. The maximum entropy model is as unbiased as possible, assigning a degree of belief of 0.5 to each argument, whereas the minimum entropy models have maximum information and take extreme values. The centroid P_c reflects the overall situation in AF. For example, argument $\mathcal{A}_3$ is attacked by two arguments and receives a small degree of belief. Furthermore, both $\mathcal{A}_2$ and $\mathcal{A}_5$ each attack two other arguments and also defend themselves against attacks, therefore getting a relatively high degree of belief of 0.57 and 0.64, respectively.

9 As $\mathcal{P}^J_{\mathsf{AF}}$ is convex it also holds that $P_c \in \mathcal{P}^J_{\mathsf{AF}}$.

	P^*	P_c	$P_1^{\min}$	$P_2^{\min}$	$P_3^{\min}$
$\mathcal{A}_1$	0.5	0.43	0	1	0
$\mathcal{A}_2$	0.5	0.57	1	0	1
$\mathcal{A}_3$	0.5	0.14	0	0	0
$\mathcal{A}_4$	0.5	0.36	1	0	0
$\mathcal{A}_5$	0.5	0.64	0	1	1

Table 1: Degrees of belief in Ex. 6

We take a closer look on the centroid of $\mathcal{P}^J_{\mathsf{AF}}$ in the next section.

4 COMPARISON WITH CLASSICAL SEMANTICS

In this section, we investigate the relationships between classical semantics and probabilistic semantics in more depth.

A probability function $P \in \mathcal{P}_{\mathsf{AF}}$ is a generalization of a labeling. For an argument $\mathcal{A} \in \mathsf{Arg}$, the probability $P(\mathcal{A}) = 1$ is equivalent to stating that the argument $\mathcal{A}$ is in and the probability $P(\mathcal{A}) = 0$ is equivalent to stating the $\mathcal{A}$ is out. A probability $P(\mathcal{A}) \in (0,1)$ generalizes the status undec while $P(\mathcal{A}) = 0.5$ is the most "unbiased undec". Labelings can be linked to probability functions as follows. For a labeling L the *characteristic probability function* P_L of L is defined via

1. if $\mathsf{undec}(L) = \emptyset$:

$$\begin{aligned} P_L(\mathsf{in}(L)) &= 1 \\ P_L(e') &= 0 \qquad \text{for all } e' \in \mathcal{E} \setminus \{\mathsf{in}(L)\} \end{aligned}$$

2. if $\mathsf{undec}(L) \neq \emptyset$:

$$\begin{aligned} P_L(\mathsf{in}(L)) &= P_L(\mathsf{in}(L) \cup \mathsf{undec}(L)) = 0.5 \\ P_L(e') &= 0 \qquad \text{for all } e' \in \mathcal{E} \setminus \{\mathsf{in}(L), \mathsf{in}(L) \cup \mathsf{undec}(L)\} \end{aligned}$$

Note that P_L is well-defined due to Prop. 1. It is easy to see, that $P_L(\mathcal{A}) = 1$ if and only if $L(\mathcal{A}) = \mathsf{in}$, $P_L(\mathcal{A}) = 0$ if and only if $L(\mathcal{A}) = \mathsf{out}$, and $P_L(\mathcal{A}) = 0.5$ if and only if $L(\mathcal{A}) = \mathsf{undec}$. Probability functions P_1, P_2 are *argument-equivalent*, denoted by $P_1 \equiv P_2$, if and only if $P_1(\mathcal{A}) = P_2(\mathcal{A})$ for all $\mathcal{A} \in \mathsf{Arg}$.

Theorem 1. *Let* AF *be some abstract argumentation framework and let* L *be some labeling.*

1. *If* L *is complete then* P_L *is p-justifiable.*
2. L *is grounded if and only if* $P_L \equiv P^*$ *for* $\{P^*\} = \mathsf{MaxE}(\mathcal{P}^J_{\mathsf{AF}})$.

3. *If stable labelings exist for* AF *then* L *is stable if and only if* $P_L \in MinE(\mathcal{P}^J_{\mathsf{AF}})$.

The above theorem establishes quite interesting relationships between our probabilistic semantics and classical semantics. First, the concept or p-justifiable generalizes complete semantics as every complete labeling induces a p-justifiable probability function. Second, the grounded labeling of an argumentation framework corresponds to the maximum entropy model of all p-justifiable probability functions (up to argument-equivalence). Third, the set of stable labeling corresponds to the set of minimum entropy models, provided that the former set is non-empty. The final two observations link the information-theoretic concept of entropy to classical argumentation semantics. The maximum entropy model of $\mathcal{P}^J_{\mathsf{AF}}$ is the probability function which is as unbiased as possible whereas the grounded labeling is the labeling which is as cautious as possible. Furthermore, a minimum entropy model of $\mathcal{P}^J_{\mathsf{AF}}$ is a probability function that maximizes information. Similarly, a stable labeling L has maximum information as it assigns to each argument either in or out.

Note that the converse of 1.) in Th. 1 does not hold in general.

Example 7. Consider the abstract argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ with $\mathsf{Arg} = \{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3\}$ and $\rightarrow$ as depicted in Fig. 2. Let P be a probability function defined as $P(\{\mathcal{A}_1\}) = P(\{\mathcal{A}_2\}) = 0.5$ and $P(e) = 0$ for all remaining $e \in \mathcal{E}$. Hence $P(\mathcal{A}_1) = P(\mathcal{A}_2) = 0.5$ and $P(\mathcal{A}_3) = 0$. Note that P is p-justifiable wrt. AF. However, there is no complete labeling L with $L(\mathcal{A}_1) = L(\mathcal{A}_2) = \mathsf{undec}$ and $L(\mathcal{A}_3) = \mathsf{out}$.

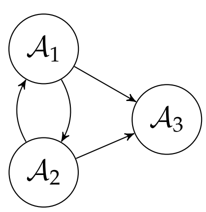


Figure 2: Argumentation framework from Ex. 7

Due to Th. 1 and Ex. 7 we have established that probabilistic semantics is a clear generalization of classical complete semantics. Therefore, we can integrate probabilistic semantics into the hierarchy of classical semantics as depicted in Fig. 3 (an arrow reads "is less general than").

The converse of 3.) in Th. 1 does not hold in general as well as a minimum entropy model even exists if AF has no stable labeling. However, the set $\mathsf{MinE}(\mathcal{P}^J_{\mathsf{AF}})$ can be characterized as follows.

$$\begin{tikzcd}
\text{stable semantics} \arrow[dr] & & & & \\
& \text{semi-stable semantics} \arrow[dr] & & & \\
& & \text{preferred semantics} \arrow[dr] & & \text{grounded semantics} \arrow[dl] \\
& & & \text{complete semantics} \arrow[d] & \\
& & & \text{probabilistic semantics} &
\end{tikzcd}$$

Figure 3: Relationships between semantics

Proposition 5. *Let* AF *be some abstract argumentation framework and let* L *be some labeling. Then* $P_L \in \mathit{MinE}(\mathcal{P}^J_{\mathsf{AF}})$ *if and only if* $\mathsf{undec}(L)$ *is minimal wrt. set cardinality.*

Note that the above proposition does not establish that $\mathsf{MinE}(\mathcal{P}^J_{\mathsf{AF}})$ is equivalent to the set of semi-stable labelings as a semi-stable labeling is characterized by having a minimal $\mathsf{undec}(L)$ wrt. set *inclusion*. However, it holds that L is a semi-stable labeling if $P_L \in \mathsf{MinE}(\mathcal{P}^J_{\mathsf{AF}})$.

In the previous section, the centroid P_c of $\mathcal{P}^J_{\mathsf{AF}}$ has proven to be a good candidate for representing the set $\mathcal{P}^J_{\mathsf{AF}}$ as a whole. We now turn to its relationship with classical semantics.

Theorem 2. *Let* $\{L_1, \ldots, L_m\}$ *be the set of complete labelings wrt.* AF. *Then the set* $\mathcal{P}^J_{\mathsf{AF}}$ *is a polytope where* $\{P_{L_1}, \ldots, P_{L_m}\}$ *is the set of its extremal points.*

The above theorem states that the set $\mathcal{P}^J_{\mathsf{AF}}$ is the convex hull of the characteristic probability functions of all complete labelings. It also leads to a very simple characterization of the centroid P_c of $\mathcal{P}^J_{\mathsf{AF}}$.

Corollary 2. *Let* $\{L_1, \ldots, L_m\}$ *be the set of complete labelings wrt.* $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ *and let* P_c *be the centroid of* $\mathcal{P}^J_{\mathsf{AF}}$. *Then*

$$P_c(\mathcal{A}) = \frac{\sum_{i=1}^{m} \delta(L_i, \mathcal{A})}{m} \quad \text{for all } \mathcal{A} \in \mathsf{Arg}$$

with $\delta(L, \mathcal{A}) = 1$ *if* $L(\mathcal{A}) = \mathsf{in}$, $\delta(L, \mathcal{A}) = 0$ *if* $L(\mathcal{A}) = \mathsf{out}$, *and* $\delta(L, \mathcal{A}) = 0.5$ *otherwise.*

In other words, the probability of an argument in the centroid of $\mathcal{P}^J_{\mathsf{AF}}$ is its average probability with respect to all complete labelings.

5 REASONING IN CRITICAL DOMAINS

In order to illustrate the usefulness of our non-classical semantics we elaborate on an example from the medical domain. This example is inspired by an example from (Modgil, 2009) and does not qualify for being medically accurate. Consider the argumentation framework $\mathsf{AF} = (\mathsf{Arg}, \rightarrow)$ with $\mathsf{Arg} = \{\mathcal{D}_1, \mathcal{D}_2, \mathcal{T}_1, \mathcal{T}_2, \mathcal{C}, \mathcal{P}_1, \mathcal{P}_2\}$ and $\rightarrow$ as depicted in Fig. 4. In AF, the arguments $\mathcal{D}_1$ and $\mathcal{D}_2$ are arguments for treating a patient suffering from blood clotting with aspirin and chlopidogrel, respectively. Both arguments attack each other as only one drug may be selected for treatment. The arguments $\mathcal{T}_1$ and $\mathcal{T}_2$ represent contradictory medical trials stating that aspirin is more effective than chlopidogrel ($\mathcal{T}_2$) and that chlopidogrel is more effective than aspirin ($\mathcal{T}_1$). Argument $\mathcal{C}$ states that chlopidogrel is too costly and should not be prescribed. Arguments $\mathcal{P}_1$ and $\mathcal{P}_2$ represent differing views on the importance of the features "health" and "low expenses": $\mathcal{P}_1$ states that the health of a patient if more important than expenses, therefore attacking argument $\mathcal{C}$. The argument $\mathcal{P}_2$ states that having low expenses is more important than a patient's health. It can easily be seen that the

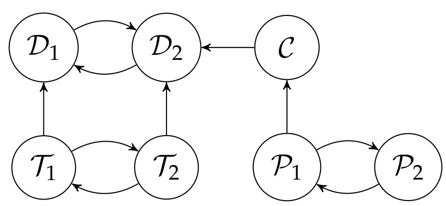


Figure 4: Argumentation framework for the medical domain

grounded labeling of AF declares each argument as undec. Therefore, for both credulous and skeptical inference no arguments can be established as ultimately accepted. For complete, preferred, semi-stable, stable semantics several labelings can be identified, some of the declaring $\mathcal{D}_1$ as in and some declaring $\mathcal{D}_2$ as in. However, for all those classical semantics each argument in AF can be credulously inferred and none can be skeptically inferred. The centroid P_c assigns to each argument except $\mathcal{D}_2$ a degree of belief of 0.5. The degree of belief of $\mathcal{D}_2$ is approximately 0.278. This assignment reflects the overall situation in AF as $\mathcal{D}_2$ is more controversial than $\mathcal{D}_1$ due to the former's cost. Although the degree of belief in $\mathcal{D}_1$ is not very high it is still higher than $\mathcal{D}_2$ which makes $\mathcal{D}_1$ a better recommendation. Furthermore, the centroid P_c also gives a concise overview on the uncertainty inherent in AF which supports the user in assessing his confidence when selecting a specific action.

6 RELATED WORK

The original definition of argumentation semantics by Dung (Dung, 1995) relies on the concept of an extension with a clear understanding of the status of an argument: an argument is either *in* an extension or not. Argument labelings (Wu and Caminada, 2010) generalize this view and make the (already implicitly existent) third status of an argument explicit by distinguishing between arguments that are *out* and arguments that are *undecided*. Our approach generalizes this idea even further by considering the whole interval $[0, 1]$ as the space for the status of an argument. As discussed in Section 4 the classical notions of in and out can be identified with the probabilities of 1 and 0, respectively, while the argument status undec corresponds to the whole open interval $(0, 1)$, with 0.5 being the most "unbiased" notion of undec.

To the best of our knowledge, the work reported here is the first that defines a probabilistic semantics for pure abstract argumentation frameworks. However, there are some works that extend abstract argumentation frameworks to incorporate some form of quantitative uncertainty, see e. g. (Janssen *et al.*, 2008; Dung and Thang, 2010; Li *et al.*, 2011; Dunne *et al.*, 2011). For example, the work (Li *et al.*, 2011) defines a probabilistic argumentation framework PAF via $\mathsf{PAF} = (\mathsf{Arg}, P_{\mathsf{Arg}}, \rightarrow, P_{\rightarrow})$ where $(\mathsf{Arg}, \rightarrow)$ is an abstract argumentation framework, P_{Arg} is a probability function on Arg, and $P_{\rightarrow}$ is a probability function on $\rightarrow$. A probabilistic argumentation framework PAF serves as a template for a set of abstract argumentation frameworks $\mathsf{AF}_1, \ldots, \mathsf{AF}_n$. Each AF_i ($i = 1, \ldots, n$) is a sub-framework of $(\mathsf{Arg}, \rightarrow)$ and has an associated probability $P(\mathsf{AF}_i)$ of its "occurrence" which is determined by the probabilities of arguments and attacks. By fixing a specific classical semantics, e. g. grounded semantics, in (Li *et al.*, 2011) a probabilistic interpretation $P(\mathcal{A})$ for an argument $\mathcal{A}$ is computed by summing up the probabilities of those AF_i in which $\mathcal{A}$ is in the grounded extension. Similarly, the work (Janssen *et al.*, 2008) extends abstract argumentation frameworks by allowing the attack relation $\rightarrow$ to be a *fuzzy relation*. Weighted argument systems (Dunne *et al.*, 2011) assign to each attack a positive real-value to represent its strength. Reasoning in weighted argument systems is performed by fixing some threshold β and focusing on those subsets of a system that neglects attacks with weights that sum up to at most β. The main difference between our approach and the approaches discussed so far is that they introduce additional uncertainty into the knowledge representation formalism while we assess the inherent uncertainty within abstract argumentation frameworks by a generalized semantics. A common ground of our approach and the approaches above is the focus on abstract argumentation frameworks and, therefore, the non-observance of uncertainty *within* the structure of arguments. There are also a few works that consider quantitative uncertainty within argument construction, see e. g. (Krause *et al.*, 1995; Kohlas, 2003; Alsinet *et al.*, 2008). In those works additional uncertainty is introduced by weighting formulas used for creating arguments.

Similarly to our approach, the work (Matt and Toni, 2008) also assigns degrees of strength to arguments of an abstract argumentation framework solely based on the framework's inherent uncertainty. In (Matt and Toni, 2008), an argumentation framework is interpreted within an argumentation dialogue and strengths indicate how defendable an argument is for a participant. In contrast to our work, the *degrees of acceptance* in (Matt and Toni, 2008) have no probabilistic interpretation and are computed in a propriety way as to reflect the situation of a competitive argumentation game. Furthermore, (Matt and Toni, 2008) is not concerned with semantical issues of abstract argumentation frameworks.

7 SUMMARY AND DISCUSSION

In this paper we proposed a new way for giving semantics to abstract argumentation frameworks. Instead of extensions or labelings we used probability functions to assign degrees of belief to arguments. We proposed a generalization of complete semantics and showed several interesting relationships between probabilistic and classical semantics on the one hand and abstract argumentation and probabilistic reasoning on the hand. In particular, we showed that the maximum entropy model of probabilistic reasoning corresponds to the grounded labeling in abstract argumentation. We also illustrated the usefulness of our approach in critical domains.

Probabilistic semantics generalizes the classical extension- and labeling-based semantics for abstract argumentation and allows for a more fine-grained differentiation of the status of arguments. For future work, we intend to investigate the relationship of probabilistic semantics with the notion of *accrual* (Prakken, 2005) which is concerned with effects of multiple arguments attacking another argument. Roughly, it is rational to assume that the more reasons there are against a single claim the less this claim is believed to be true. In our framework, accrual of arguments is already weakly adhered for by property 2.) of Def. 4 where, in particular, the number of attacks on an argument influences the lower bound for the degree of belief in that argument, see also Ex. 7. However, a deeper analysis of this issue is left for future work.

OPPONENT MODELS WITH UNCERTAINTY FOR STRATEGIC ARGUMENTATION

Tjitze Rienstra, Matthias Thimm, Nir Oren. Opponent Models with Uncertainty for Strategic Argumentation. In Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13). Beijing, China, August 2013.

Abstract

This paper deals with the issue of strategic argumentation in the setting of Dung-style abstract argumentation theory. Such reasoning takes place through the use of opponent models—recursive representations of an agent's knowledge and beliefs regarding the opponent's knowledge. Using such models, we present three approaches to reasoning. The first directly utilises the opponent model to identify the best move to advance in a dialogue. The second extends our basic approach through the use of quantitative uncertainty over the opponent's model. The final extension introduces *virtual arguments* into the opponent's reasoning process. Such arguments are unknown to the agent, but presumed to exist and interact with known arguments. They are therefore used to add a primitive notion of *risk* to the agent's reasoning. We have implemented our models and we have performed an empirical analysis that shows that this added expressivity improves the performance of an agent in a dialogue.

1 INTRODUCTION

Argumentation systems offer a natural, easily understood representation of non-monotonic reasoning, and have been applied to a variety of problem domains including planning and practical reasoning (Toniolo *et al.*, 2011) and legal reasoning (Grabmair and Ashley, 2010). Critically, many of these domains are adversarial, requiring an agent to identify and advance some set of arguments which are most likely to enable it to achieve its goals. In order to do so, the agent employs a strategy, typically in the form of a heuristic, which selects appropriate arguments given some contextual knowledge.

We describe one such strategy, and examine some of its properties. Our strategy assumes that an agent is not only aware of the arguments that it is permitted to advance, as well as what has already been stated, but that it also has a belief regarding its opponent's knowledge, and that this relationship is recursive unto some depth (i. e. an agent a believes some arguments, and believes that b knows some arguments, as well as believing that b believes that a knows some arguments, and so on). While (Oren and

Norman, 2009) have previously examined such a strategy, we extend and improve their work along several dimensions.

First, (Oren and Norman, 2009) assume that an agent holds only a single opponent model. However, uncertainty plays a crucial role in strategies, and we capture this uncertainty, associating different opponent models with different likelihoods. Second, agents are often unaware of all arguments in a domain, and we allow an agent to hold an opponent model containing arguments it itself is not aware of through the introduction of *virtual arguments*. Finally, we consider how an agent should update its knowledge and beliefs while taking part in a dialogue.

In (Prakken and Sartor, 2002) an influential four layered view of an argumentation system is described. The first two levels, consisting of the logical and dialectic layers, specify the content of an argument, and how arguments interact with each other. In our work, these layers are encapsulated within an abstract argumentation framework (Dung, 1995), which we summarise in Sec. 2. We encapsulate Prakken's *procedural layer*, which specifies how agents exchange arguments (via dialogue) via a general *discourse model*. This discourse model, described in Sec. 3, assumes only that agents take alternating turns in making moves, and further constrains the dialogue by limiting what moves can be made through a *legal moves* function. Section 4 then describes the *agent model*, associating a utility with specific arguments, and allowing for different types of belief states. This agent model captures Prakken's heuristic layer through the specification of agent strategy. We present three instances of the agent model, starting from the one described in (Oren and Norman, 2009), and repeatedly adding further layers of expressivity. Section 5 describes how an agent's beliefs should be updated as the dialogue progresses, following which we compare and empirically evaluate the different models in Sec. 6. Section 7 discusses related work, and we conclude in Sec. 8.

2 FORMAL PRELIMINARIES

In abstract argumentation theory, knowledge is represented by an abstract *argumentation framework* (or AF, in short), which is a set of arguments with an attack relation, cf. (Dung, 1995).

Definition 1. An *AF* is a pair (A, R) where A is the set of *arguments* and $R \subseteq A \times A$ is the *attack relation*.

The goal is to select sets of arguments, called *extensions*, that represent rational points of view on the acceptability of the arguments of the AF. The first condition for an extension to be rational is that it is *conflict-free*. Furthermore, if an argument is a member of an extension, it is assumed that it is *defended* by the extension. Formally:

Definition 2. Given an AF $F = (A, R)$, an *extension* is a set $E \subseteq A$. E is said to be *conflict-free* iff $\nexists x, y \in E, (x, y) \in R$. Given an argument $x \in A$, E is said to *defend* x iff $\forall y \in A$ s.t. $(y, x) \in R$, $\exists z \in E$ s.t. $(z, y) \in R$. We define $D_{(A,R)}(E)$ by $D_{(A,R)}(E) = \{x \in A \mid E \text{ defends } x\}$.

Using the notions of conflict-freeness and defense, we can define a number of argumentation semantics, each embodying a particular rationality criterion.

Definition 3. Let $F = (A, R)$ be and $E \subseteq A$ conflict-free extension. E qualifies as:

- **admissible** iff $E \subseteq D_{(A,R)}(E)$,
- **complete** iff $E = D_{(A,R)}(E)$,
- **grounded** iff E is minimal (w.r.t. set inclusion) among the set of complete extensions of F.

For the intuition behind the different semantics we refer the reader to (Dung, 1995).

3 THE DISCOURSE MODEL

The *discourse model* provides a way to specify the complete setting in which two agents (proponent $\mathcal{P}$ and opponent $\mathcal{O}$) engage in a certain type of dialogue. Recalling the four-layered model mentioned in the introduction, we first need a logical and dialectical layer. Here we use abstract argumentation theory, as presented in the previous section, and leave the logical content of arguments unspecified. We assume that there is a *universal AF* $(\mathcal{A}, \mathcal{R})$ which contains all arguments relevant to a particular discourse.

An agent $Ag \in \{\mathcal{P}, \mathcal{O}\}$ has limited knowledge and is aware only of some subset $B_{Ag} \subseteq \mathcal{A}$ of arguments which she can put forward in a dialogue. We assume that the attack relation is determined by the arguments, so that the knowledge of an agent can be identified with the set B_{Ag}, inducing an AF $(B_{Ag}, \mathcal{R} \cap (B_{Ag} \times B_{Ag}))$. In the remaining definitions, we assume $(\mathcal{A}, \mathcal{R})$ to be given.

Next, we need to fill in the procedural layer. The main object with which we are concerned here is a *dialogue trace*, which represents a dialogue between $\mathcal{P}$ and $\mathcal{O}$ by a sequence of moves (i.e. sets of arguments $M \subseteq \mathcal{A}$) made by $\mathcal{P}$ and $\mathcal{O}$ alternately, with $\mathcal{P}$ making the first move. Formally:

Definition 4 (Dialogue trace). A *dialogue trace* is a finite sequence $\pi = (M_1, \ldots, M_n)$ s.t. $M_1, \ldots, M_n \subseteq \mathcal{A}$. Every M_i is called a *move*. We define $A_\pi = M_1 \cup \ldots \cup M_n$ and $n_\pi = |M_1| + ... + |M_n|$. $\pi[n]$ denotes the dialogue trace consisting of the first n moves of π; $\pi[0]$ is the empty sequence. The set of all possible dialogue traces is denoted by S.

The rules of the dialogue are captured by the *legal move function* $\mathsf{legalmoves} : S \to 2^{2^{\mathcal{A}}}$ which returns valid follow-up moves for a particular dialogue trace. The heuristic, or strategic component is captured by an *agent model*, one for $\mathcal{P}$ and one for $\mathcal{O}$.

Definition 5 (Agent model)**.** An *abstract agent model* Δ is a triple $\Delta = (\mathcal{K}, \mathsf{move}, \mathsf{upd})$ where $\mathcal{K}$ is the *belief state*; move is a function mapping a dialogue trace and belief state to a set of moves, called the *move function*; and upd is a function mapping a belief state and a move to a new belief state, called the *update function*.

A belief state $\mathcal{K}$ captures the agent's knowledge, utility function and opponent model. The function move returns the set of moves for the agent, given her belief state and current dialogue trace and implements the agent's strategy. We assume that an agent's move function returns only legal moves. Note that an agent may be indifferent as to which move is best and can also decide to end the dialogue, i. e., move may return multiple or zero moves. Finally, the function upd takes a belief state $\mathcal{K}$ and the move made by the opponent and yields a new belief state $\mathcal{K}'$, the idea being that moves made by the opponent may change the agent's knowledge and beliefs.

Definition 6. A dialogue state is a pair $(\Delta_{\mathcal{P}}, \Delta_{\mathcal{O}})$ where $\Delta_{\mathcal{P}}, \Delta_{\mathcal{O}}$ are a proponent's and opponent's agent model. A dialogue trace $\pi = (M_1, \ldots, M_n)$ is called *valid* wrt. a legal move function $\mathsf{legalmoves}$ and a dialogue state $(\Delta_{\mathcal{P}}, \Delta_{\mathcal{O}})$ if and only if there exists a sequence of dialogue states $((\Delta^0_{\mathcal{P}}, \Delta^0_{\mathcal{O}}), \ldots, (\Delta^n_{\mathcal{P}}, \Delta^n_{\mathcal{O}}))$ with $\Delta^i_{Ag} = (\mathcal{K}^i_{Ag}, \mathsf{upd}_{Ag}, \mathsf{move}_{Ag})$ such that $\Delta^0_{\mathcal{P}} = \Delta_{\mathcal{P}}$, $\Delta^0_{\mathcal{O}} = \Delta_{\mathcal{O}}$ and, for $i = 1, \ldots, n$:

1. $M_i \in \mathsf{move}_{\mathcal{P}}(\pi[i-1], \mathcal{K}^{i-1}_{\mathcal{P}})$ if i is odd,
2. $M_i \in \mathsf{move}_{\mathcal{O}}(\pi[i-1], \mathcal{K}^{i-1}_{\mathcal{O}})$ if i is even,
3. $\mathcal{K}^i_{Ag} = \mathsf{upd}(\mathcal{K}^{i-1}_{Ag}, M_i)$ for $Ag \in \{\mathcal{P}, \mathcal{O}\}$.

No moves can be added to a dialogue trace if an agent decides to end the dialogue. A dialogue trace is then *complete*:

Definition 7. Let $\pi = (M_1, \ldots, M_n)$ be a valid dialogue trace with respect to a legal move function $\mathsf{legalmoves}$ and a dialogue state $(\Delta_{\mathcal{P}}, \Delta_{\mathcal{O}})$. We say π is *complete* if and only if there is no dialogue trace $\pi' = (M_1, \ldots, M_n, M_{n+1})$ which is valid with respect to $\mathsf{legalmoves}$ and $(\Delta_{\mathcal{P}}, \Delta_{\mathcal{O}})$.

Note that, because the move function may return more than one move, there may be more than one valid and complete dialogue trace for a given pair of initial agent models. Our discourse model is thus nondeterministic with respect to how a dialogue evolves.

In the following sections we present concrete instantiations of agent models. We demonstrate these models by fixing $(\mathcal{A}, \mathcal{R})$, specifying $\mathsf{legalmoves}$, and showing the resulting valid dialogue traces.

4 AGENT MODELS FOR STRATEGIC ARGUMENTATION

In this section we present three concrete instantiations of agent models. We focus here on the belief state $\mathcal{K}$ and move function. Each model extends the expressivity of the former. We show, by example, that these extensions are necessary to properly model strategically important beliefs of an agent. In each of the three agent models, the move function is based on a variant of the M* search algorithm (Carmel and Markovitch, 1996). We postpone the treatment of the third component of an agent model $(\mathcal{K}, \mathsf{move}, \mathsf{upd})$, namely the update function upd, returning to it in Section 5.

4.1 *The simple agent model*

The simple belief state of an agent (denoted by $\mathcal{K}_s$) consists, first, of a set $B \subseteq \mathcal{A}$ containing the arguments that the agent is aware of. The goals of the agent are encoded by the utility function u, that returns the utility (a real number) that the agent assigns to a particular dialogue trace $\pi \in S$ (cf. (Thimm and Garcia, 2010)). The agent's beliefs about the knowledge and goals of her opponent and about the beliefs about her opponent's belief about herself, etc., are modeled by simply nesting this structure, so that the third component of the simple belief state is again a simple belief state.

Definition 8. A *simple belief state* $\mathcal{K}_s$ is a tuple (B, u, E) where:

- $B \subseteq \mathcal{A}$ is the set of arguments the agent is aware of,
- $u : S \to \mathbb{R}$ is the utility function,
- $E = (B', u', E')$ is a simple belief state called the *opponent state*, such that $B' \subseteq B$.

The intuition behind $B' \subseteq B$ in the above stems from the common sense notion that an agent cannot have beliefs about whether or not her opponent is aware of an argument that she herself is not aware of. In other words, If an agent believes that her opponent knows argument a, then surely the agent herself also knows a. We refer to this requirement as the *awareness restriction*.

Except for this restriction, this model is the same as the one presented by Oren and Norman (Oren and Norman, 2009). They also present a variation of the maxmin algorithm (which is in turn a variation of the M* algorithm (Carmel and Markovitch, 1996)) that determines, given a belief state $\mathcal{K}_s$ and $\mathsf{legalmoves}$ function, the moves that yield the best expected outcome. We can use the same approach to define our move_s function. The algorithm that defines the $\texttt{move}_s$ function is shown in Algorithm 1. Note that the actual algorithm only needs to return the set of best moves. To simplify the algorithm, however, we define it to return both the set of best moves and the expected utility of these moves.

Algorithm 1 $\texttt{move}_s(\pi, (B, u, E))$

1: maxEU=$u(\pi)$
2: bestMoves=$\emptyset$
3: **for all** $M \in \textsf{legalmoves}(\pi)$ **do**
4:     $eu = 0$
5:     $(oUtil, oMoves) = \texttt{move}_s((\pi, M), E)$
6:     **for all** $M' \in oMoves$ **do**
7:         $(nUtil, nMoves) = \texttt{move}_s((\pi, M, M'), (B, u, E))$
8:         $eu = eu + nUtil * \frac{1}{|oMoves|}$
9:     **if** $eu > maxEU$ **then**
10:         $bestMoves = \emptyset$
11:     **if** $eu \geq maxEU$ **then**
12:         $bestMoves = bestMoves \cup \{M\}$
13:         $maxEU = eu$
14: **return** $(maxEU, bestMoves)$

The algorithm works as follows. Initially the *bestMoves* is empty and *maxEU*, acting as a lower bound on the expected utility to be improved upon, is set to the utility of the current trace. For every legal move M the set of best responses of the opponent is determined on line 5. Then, in lines 6 and 7 the expected utility of M is determined by again calling the move function. On line 8 we divide the expected utility of M by the number of opponent moves, taking into account all possible equally good moves the opponent can make without double counting. Lines 9–13 keep track of the moves, considered so far, that yield maximum expected utility. Note that, if *bestMoves* is empty at the end, not moving yields highest expected utility.

Note that we assume that the nesting of the belief state is sufficiently deep to run the algorithm. Alternatively, the algorithm can easily be adapted (as shown in (Oren and Norman, 2009)) to deal with belief states of insufficient depth, or to terminate at some fixed search depth.

In the rest of this text, we use a utility function to the effect that the dialogue is about grounded acceptance of an argument $x \in \mathcal{A}$. This has been called a *grounded game* in the literature (Modgil and Caminada, 2009). We subtract arbitrary small values ϵ, for each move in the dialogue, capturing the idea that shorter traces are preferred, effectively driving the agents to put forward only relevant moves. Formally:

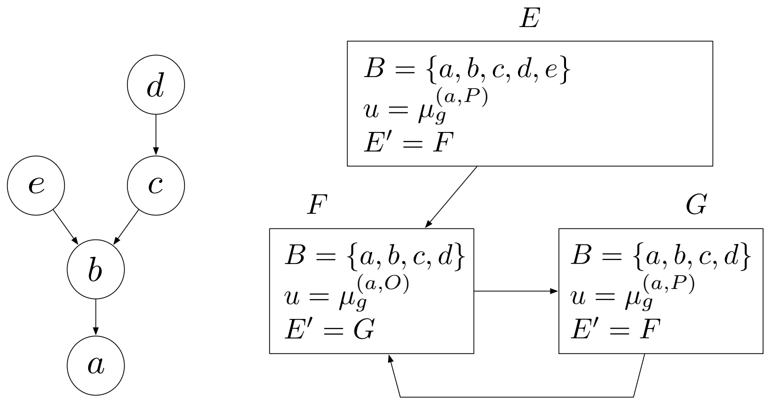


Figure 1: An AF and belief state structure.

Definition 9 (Grounded game utility function)**.** Let $x \in \mathcal{A}$ and $Ag \in \{\mathcal{P}, \mathcal{O}\}$. The *grounded game utility function*, denoted by $u_g^{(x,Ag)}$, evaluated over a dialogue trace π, is defined by:

$$u_g^{(x,Ag)}(\pi) = \begin{cases} v - n_\pi \epsilon & \text{if } x \text{ grounded in } F \\ -v - n_\pi \epsilon & \text{if } x \text{ attacked by gr. extension of } F \\ 0 & \text{otherwise} \end{cases}$$

where $F = (A_\pi, \mathcal{R} \cap (A_\pi \times A_\pi))$, and $v = 1$, if $Ag = \mathcal{P}$ and $v = -1$, if $Ag = \mathcal{O}$.

Following (Modgil and Caminada, 2009), we may refer to x as *in* if it is within the (uniquely determined) grounded extension, *out* if it is attacked by an element from this extension, and *undecided* otherwise.

For simplicity, we assume that utility functions are fixed in all models, i. e., both agents have correct beliefs about the opponent's utility function. Furthermore, the legalmoves function that we use simply forces moves to consist of a single argument, and is defined by $\mathsf{legalmoves}_1(\pi) = \{\{x\} \mid x \in \mathcal{A}\}$.

Example 1. Let $(\mathcal{A}, \mathcal{R})$, and E, F, G be the AF and belief states as shown in Figure 1. That is, $E = (B, u_g^{(a,P)}, F), F = (B', -u_g^{(a,O)}, G), G = (B', u_g^{(a,P)}, F)$ with $B = \{a, b, c, d, e\}$ and $B' = \{a, b, c, d\}$. We define the agent models $\Gamma = (\Delta_\mathcal{P}, \Delta_\mathcal{O})$ by $\Delta_\mathcal{P} = (E, \mathsf{moves}_s, \mathsf{upd})$ and $\Delta_\mathcal{O} = (F, \mathsf{moves}_s, \mathsf{upd})$, where upd is defined by $\mathsf{upd}(\mathcal{K}, M) = \mathcal{K}$. In words, $\mathcal{P}$ is aware of all arguments and (correctly) believes that $\mathcal{O}$ is aware of only a, b, c and d. There is a single valid dialogue trace w.r.t. Γ, namely $(\{a\}, \{b\}, \{e\})$. The first move is obvious, i. e., if $\mathcal{P}$ would not put forward the argument a under consideration, she loses because the opponent can end the dialogue. The second move b is made because $\mathcal{O}$ (not being aware of e) believes she can counter the only possible countermove after b (i. e., c) by d, resulting in a win. As it turns out, $\mathcal{P}$'s quickest way to win is by moving e and not c. This ends the dialogue, as $\mathcal{O}$ cannot increase utility by putting forward any remaining argument. The utilities of the trace are $1 - 3\epsilon$ for $\mathcal{P}$ and $-1 - 3\epsilon$ for $\mathcal{O}$.

4.2 *The uncertain agent model*

A limitation of the simple agent model is that it assumes certainty about the opponent model. In the *uncertain agent model*, denoted $\mathcal{K}_u$, we capture uncertainty by assigning (non-zero) probabilities to possible opponent modes:

Definition 10. An *uncertain belief state* $\mathcal{K}_u$ is a tuple $(B, u, \mathcal{E}, P)$ where:

- $B \subseteq \mathcal{A}$ is the set of arguments,
- $u : S \to \mathbb{R}$ is the utility function,
- $\mathcal{E}$ is a set of uncertain belief states (opponent belief states) such that $\forall (B', u', \mathcal{E}', P') \in \mathcal{E}$: $B' \subseteq B$,
- $P : \mathcal{E} \to (0, 1]$ is a probability function s.t. $\sum_{E \in \mathcal{E}} E = 1$.

The corresponding move_u function, defined by algorithm 2, is a straightforward generalization of move_s, taking into account probabilities of possible opponent models.

Algorithm 2 $\mathtt{move}_u(\pi, (B, u, \mathcal{E}, P))$

$maxEU = u(\pi)$
$bestMoves = \emptyset$
for all $M \in \mathsf{legalmoves}(\pi)$ **do**
  $eu = 0$
  for all $E \in \mathcal{E}$ **do**
    $(oUtil, oMoves) = (\mathtt{move}_u((\pi, M), E))$
    for all $M' \in oMoves$ **do**
      $(nUtil, nMoves) = \mathtt{move}_u((\pi, M, M'), (B, u, \mathcal{E}, P))$
      $eu = eu + nUtil * P(E') * \frac{1}{|oMoves|}$
  if $eu > maxEU$ **then**
    $bestMoves = \emptyset$
  if $eu \geq maxEU$ **then**
    $bestMoves = bestMoves \cup \{M\}$
    $maxEU = eu$
return $(maxEU, bestMoves)$

Example 2. Let $(\mathcal{A}, \mathcal{R}), E, F_1, F_2, G_1$ and G_2 be the AF and belief states as shown in Figure 2. We define the agent models $\Delta_{\mathcal{P}}$ by $\Delta_{\mathcal{P}} = (E, \mathsf{move}_u, \mathsf{upd})$

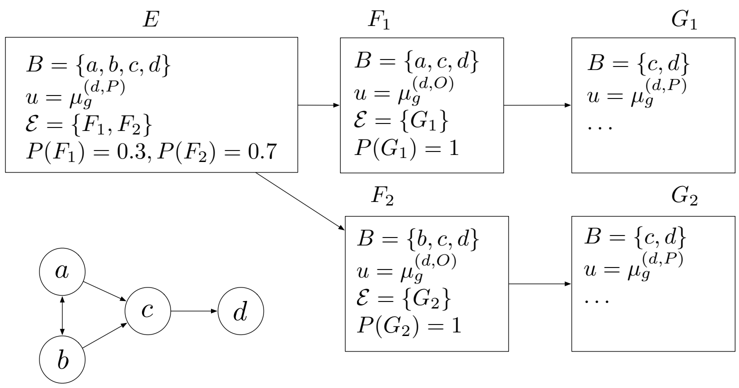


Figure 2: An AF (left) and belief state (right).

and $\Delta_{\mathcal{O}} = (F_1, \mathsf{move}_u, \mathsf{upd})$, where upd is defined by $\mathsf{upd}(\mathcal{K}, M) = \mathcal{K}$. In words, $\mathcal{P}$ is aware of all arguments and correctly believes $\mathcal{O}$ to be aware of c, d, but is uncertain about whether $\mathcal{O}$ knows a ($p = 0.3$) or b ($p = 0.7$). $\mathcal{O}$ in fact is aware of a and not of b. There is a single valid dialogue trace w.r.t. Γ, namely $(\{d\}, \{c\}, \{b\}, \{a\})$. Again, $\mathcal{P}$ first moves the argument d under consideration. $\mathcal{O}$, believing that $\mathcal{P}$ does not know a, replies with c, believing this will be successful. Now, $\mathcal{P}$ has to choose between putting forward a or b. Putting forward a while $\mathcal{O}$ knows b (which $\mathcal{O}$ believes to be more likely) will make d undecided. Thus $\mathcal{O}$ puts forward b. It turns out that $\mathcal{O}$ was in fact aware of a, and thus puts this argument forward, changing the status of d from out to undecided. The result is a 'tie-break', i. e., both $\mathcal{P}$ and $\mathcal{O}$ assign a utility of 0 +/- 4ϵ to the trace in which d is neither accepted nor rejected.

4.3 *The extended agent model*

The two models presented so far assume that an agent cannot have beliefs about whether or not her opponent is aware of an argument that she herself is not aware of. While this is a natural assumption, it limits the kind of situations we can model. An agent can still believe that her opponent knows *some* argument, even if she is not aware of this argument herself. We model such arguments as *virtual arguments*. These are believed to exist but cannot put it forward in a dialogue. For example, if one is engaged in a dialogue with a physicist about the speed of light one may assume that the physicist has an argument for the speed of light being larger than 50 kph. However, if one is not expert in physics, the exact nature of this argument might be unknown. Furthermore, we assume that if new argument is put forward, an agent knows whether or not this argument corresponds to a virtual argument she believed to exist. That is, she can *recognize* a new argument, and map it to a virtual argument. To model this, we add a set G of virtual arguments (such that $G \cap \mathcal{A} = \emptyset$), an attack relation R between

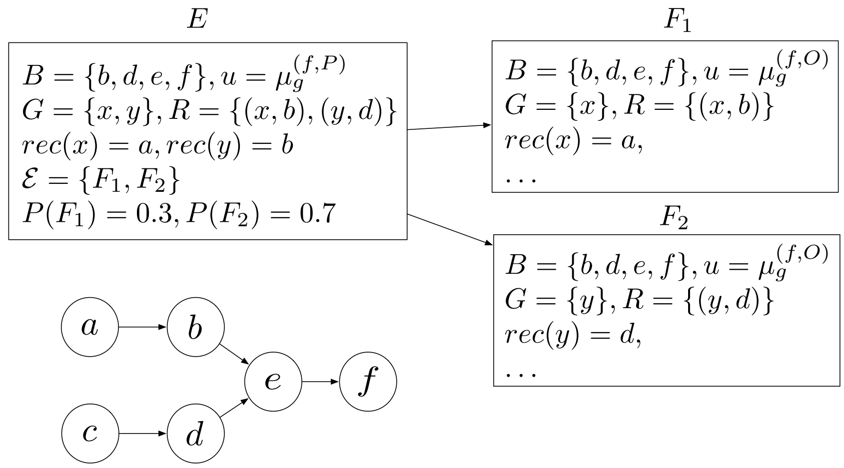


Figure 3: An AF and belief state structure.

virtual arguments and regular arguments, and a recognition function *rec* to the belief state. Formally:

Definition 11. Given a set of virtual arguments $\mathcal{G}$, an *extended belief state* $\mathcal{K}_e$ is a tuple $(B, u, G, R, rec, \mathcal{E}, P)$ where:

- $B \subseteq \mathcal{A}$ is the set of arguments,
- $u : S \to \mathbb{R}$ is the utility function,
- $G \subseteq \mathcal{G}$ is the set of *virtual arguments* believed to exist,
- $R \subseteq G \times A \cup A \times G \cup G \times G$ is the attack relation,
- $rec : \mathcal{A} \to 2^G$ is the *recognition function*,
- $\mathcal{E}$ is a set of extended belief states, called opponent belief states, s.t. $\forall (B', u', G', R', rec, \mathcal{E}', P') \in \mathcal{E}$: $B' \subseteq B$,
- $P : \mathcal{E} \to (0, 1]$ is a probability function s.t. $\sum_{E \in \mathcal{E}} = 1$.

Except for the added items, extended and uncertain belief states are similar and the $\texttt{move}_e$ function is, except for having differently typed parameters, is identical to $\texttt{move}_u$.

Example 3. Let $(\mathcal{A}, \mathcal{R}), E, F_1$ and F_2 be the AF and belief states shown in Figure 3. We define the agent models $\Delta_\mathcal{P}$ by $\Delta_\mathcal{P} = (E, \mathsf{move}_e, \mathsf{upd})$, where upd is defined by $\mathsf{upd}(\mathcal{K}, M) = \mathcal{K}$. In words, $\mathcal{P}$ is aware of b, d, e and f and believes $\mathcal{O}$ is also aware of these arguments. In addition, $\mathcal{P}$ believes $\mathcal{O}$ may have counterarguments to b and d, with probability 0.3 and 0.7 respectively. In the belief state of $\mathcal{P}$, not being aware of b and d, they are modeled as virtual arguments, i. e., x and y mapping to a and b respectively.

A possible dialogue trace is $(\{f\}, \{e\}, \{b\})$. Here, $\mathcal{P}$ puts forward f and $\mathcal{O}$ counters with e, believing this may be successful. For $\mathcal{P}$, the choice of whether to put forward b or d depends on her beliefs about (virtual) counterarguments. $\mathcal{P}$ believes it more likely (with $p = 0.7$) that $\mathcal{O}$ can counter d and therefore b is $\mathcal{P}$'s best move.

5 UPDATING OPPONENT MODELS

When an opponent puts forward a move, an opponent model needs to be updated to take into account the knowledge conveyed by this move. We propose a number of upd functions to model such updates. (Again, we assume that utility functions are fixed and therefore do not change.)

Definition 12. Let $\mathcal{K}_s = (B, u, E)$ be a simple belief state and $M \subseteq \mathcal{A}$ a move. The *simple update function* upd_s is defined by $\mathsf{upd}_s(\mathcal{K}_s, M) = (B \cup M, u, \mathsf{upd}_s(E, M))$.

Note that, if $\mathcal{K}_s$ satisfies the awareness restriction then $\mathsf{upd}_s(\mathcal{K}_s, M)$ does, too.

Definition 13. Let $\mathcal{K}_u = (B, u, \mathcal{E}, P)$ be an uncertain belief state and $M \subseteq \mathcal{A}$ a move. The *uncertain update function* upd_u is defined by $\mathsf{upd}_u(\mathcal{K}_u, M) = (B \cup M, u, \mathcal{E}', P')$ where

$$\mathcal{E}' = \bigcup_{E \in \mathcal{E}} \mathsf{upd}_u(E, M) \tag{2.1}$$

$$P'(E) = \sum_{E' \in \mathcal{E}, \mathsf{upd}_u(e,M)=E} P(E') \qquad \text{for } E \in \mathcal{E}' \tag{2.2}$$

In the above definition o we assume that arguments are observed independently of one another and, thus, probabilities stay robust in the light of observing unexpected moves. However, consider what occurs if the proponent believes that the opponent is aware of only of b with $p = 0.3$ and only of c with $p = 0.7$ how should those probabilities be adjusted when the opponent moves with argument a?

The problem is that this observation is inconsistent with the two opponent models considered possible. two ways exist to deal with this. First, we could switch to a uniform distribution, e. g. giving both states a probability of 0.5. Second, we could assume that the observations of arguments are probabilistically independent events. Taking the latter approach, we then add the observed move to every opponent model considered possible before, cf. Equation (2.1). Furthermore, some opponent models may collapse into one, so that we have to sum up probabilities for such states, cf. Equation (2.2).

Definition 14. Let $\mathcal{K}_e = (B, u, G, R, rec, \mathcal{E}, P)$ be an extended belief state and $M \subseteq \mathcal{A}$ a move. The *extended update function* upd_e is defined via $\mathsf{upd}_e(\mathcal{K}_e, M) = (B \cup M, u, G', R', rec, \mathcal{E}', P')$ where

$$G' = G \setminus \bigcup_{a \in M} rec(a) \tag{2.3}$$

$$R' = R \cap (G' \times B \cup B \times G' \cup G' \times G') \tag{2.4}$$

$$\mathcal{E}' = \bigcup_{E \in \mathcal{E}} \mathsf{upd}_3(E, M) \tag{2.5}$$

$$P'(E) = \sum_{E' \in \mathcal{E}, \mathsf{upd}_3(E', M) = E} P(E') \qquad \text{for all } E \in \mathcal{E}' \tag{2.6}$$

The following proposition establishes a strict hierarchy of our three models w.r.t. expressivity. For example, our approaches for strategic argument selection and update coincide when restricting to less expressive models. For that, we say that a simple belief state $E = (A, u, \hat{E})$ and an uncertain belief state $E' = (A', u', \mathcal{E}, P)$ are *equivalent*, denoted $E \sim E'$, if $A = A'$, $u = u'$, $\mathcal{E} = \{\hat{E}'\}$, $P(\hat{E}') = 1$, and $\hat{E}' \sim \hat{E}$ recursively. In other words, $E \sim E'$ if E' does not provide any information beyond E. Similarly, we define equivalence to an extended belief state E'' if E'' adds no virtual arguments.

Proposition 1. *1. If $E = (A, u, E)$ is a simple belief state then*

a) $E' = (A, u, \mathcal{E}, P)$ with $\mathcal{E} = \{E\}$ and $P(E) = 1$ is an uncertain belief state and

$$\mathtt{move}_s(\pi, E) = \mathtt{move}_u(\pi, E') \qquad \text{for every } \pi$$
$$\mathsf{upd}_s(E, M) \sim \mathsf{upd}_u(E', M) \qquad \text{for every } M$$

b) $E' = (A, u, G, R, rec, \mathcal{E}, P)$ with $G = R = \emptyset$, $rec(a) = \emptyset$ for every $a \in A$, $\mathcal{E} = \{E\}$ and $P(E) = 1$ is an extended belief state and

$$\mathtt{move}_s(\pi, E) = \mathtt{move}_e(\pi, E') \qquad \text{for every } \pi$$
$$\mathsf{upd}_s(E, M) \sim \mathsf{upd}_e(E', M) \qquad \text{for every } M$$

2. If $E = (A, u, \mathcal{E}, P)$ is an uncertain belief state then $E' = (A, u, G, R, rec, \mathcal{E}, P)$ with $G = R = \emptyset$ and $rec(a) = \emptyset$ for every $a \in A$ is an extended belief state and

$$\mathtt{move}_u(\pi, E) = \mathtt{move}_e(\pi, E') \qquad \text{for every } \pi$$
$$\mathsf{upd}_u(E, M) \sim \mathsf{upd}_e(E', M) \qquad \text{for every } M$$

Proofs are omitted due to space restrictions.

6 IMPLEMENTATION AND EVALUATION

We implemented the three different opponent models using Java in the *Tweety library for artificial intelligence*[1]. Our AF allows for the automatic generation of random abstract argumentation theories and simulates a dialogue between multiple agents. We used this AF to conduct experiments with our models and to evaluate their effectiveness in practice.

For evaluating performance we generated a random abstract argumentation theory with 10 arguments, ensuring that the argument under consideration is in its grounded extension, i. e. under perfect information the proponent should win the dialogue. However, from these 10 arguments only 50 % are known by the proponent but 90 % by the opponent. We used a proponent without opponent model and generated an extended belief state for the opponent (with maximum recursion depth 3). From this extended belief state we derived an uncertain belief state by simply removing the virtual arguments. From this uncertain belief state we derived a simple belief state by sampling a nested opponent model from the probability function in the uncertain belief state. For each belief state we simulated a dialogue against the same opponent and counted the number of wins. We repeated the experiment 5000 times, with Figure 1 showing our results. As seen, increasing the complexity of the belief state yields better overall performance. In particular, note that the difference between the performances of the simple and uncertain belief states is larger than between uncertain and extended belief states. However, this observation is highly depended on the actual number of virtual arguments used (which was around 30 % of all arguments in this experiment) and is different for larger values (due to space restrictions we do not report on the results of those experiments).

7 RELATED WORK

Recently, interest has arisen in combining probability with argumentation. (Hunter, 2012) describes two systems which concern themselves with the likelihood that an agent knows a specific argument, and we can view the possible argument AFs that can be induced from these likelihoods as possible models of agent knowledge. (Thimm, 2012) investigates probabilistic interpretations of abstract argumentation and relationships to approaches for probabilistic reasoning. Furthermore, (Oren *et al.*, 2012) investigated strategies in such a probabilistic setting but concerned themselves with monologues rather than dialogues.

Our work concerns itself with identifying the arguments an agent should advance at any point in a dialogue. Other work in this vein includes (Oren *et al.*, 2006), which aims to minimise the cost of moves, with no concern to the opponent's knowledge, and without looking more than one step ahead when reasoning. Such a strategy can easily be encoded by our approach. By

1 http://tinyurl.com/tweety-opp

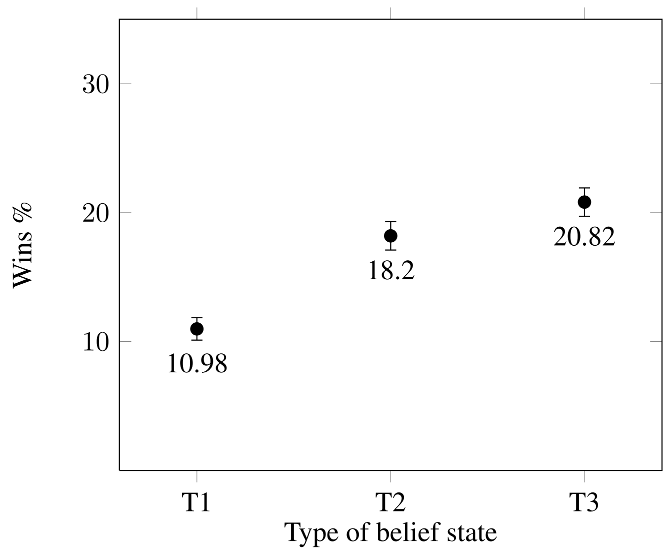


Figure 4: Performance of the simple (T1), uncertain (T2), and extended (T3) belief states in comparison (with Binomial proportion confidence intervals)

assigning probabilities to arguments, (Roth *et al.*, 2007) constructed a game tree allowing dialogue participants to maximise the likelihood of some argument being accepted or rejected. The probabilities in that system arose from *a priori* knowledge, and no consideration was given to the possibility of an opponent model.

(Rahwan and Larson, 2008; Rahwan *et al.*, 2009) consider a very different aspect of strategy, attempting to identify situations which are *strategy-proof*, that is, when full revelation of arguments is the best course of action to follow. Similarly, (Thimm and Garcia, 2010) extends that work to structured AFs and also proposes some simple dominant strategies for other specific situations. This can be contrasted with our work, where e. g. withholding information can result in a better outcome for the agent than revealing all its arguments.

8 CONCLUSIONS AND FUTURE WORK

We proposed three structures for modeling an opponents belief in strategic argumentation. Our simple model uses a recursive structure to hold the beliefs an agent has on the other agent's beliefs. We extended this model to incorporate quantitative uncertainty on the actual opponent model and qualitative uncertainty on the set of believed arguments. All our models have been implemented and we tested their performance in a series of experiments. As expected, increasing the complexity of the opponent modelling structure resulted in improved outcomes for the agent.

We consider several avenues of future work. First, agents using our strategies attempt to maximise their outcome, with no consideration for *risk*. We seek to extend our work to cater for this notion by introducing second order probabilities into our system. We also intend to investigate whether virtual arguments are equivalent to a simpler system wherein no attacks between virtual arguments can exist. Furthermore, while it is difficult to obtain large scale argument graphs obtained from real world domains, we hope to validate our approach over such corpora. Finally, while our results (for clarity of presentation) focus on abstract argument, (Hadjinikolis *et al.*, 2012) has highlighted the need for strategies when structured argumentation is used. Since the work presented here can easily be extended to this domain, we are in the process of adapting our algorithms to deal with dialogues built on top of structured argumentation.

COHERENCE AND COMPATIBILITY OF MARKOV LOGIC NETWORKS

Matthias Thimm. Coherence and Compatibility of Markov Logic Networks. In Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14). Prague, Czech Republic, August 2014.

Abstract

Markov logic is a robust approach for probabilistic relational knowledge representation that uses a log-linear model of weighted first-order formulas for probabilistic reasoning. This log-linear model always exists but may not represent the knowledge engineer's intentions adequately. In this paper, we develop a general framework for measuring this *coherence* of Markov logic networks by comparing the resulting probabilities in the model with the weights given to the formulas. Our measure takes the interdependence of different formulas into account and analyzes the degree of impact they have on the probabilities of other formulas. This approach can be used by the knowledge engineer in constructing a well-formed Markov logic network if data for learning is not available. We also apply our approach to the problem of assessing the *compatibility* of multiple Markov Logic networks, i. e., to measure to what extent the merging of these networks results in a change of probabilities.

1 INTRODUCTION

Statistical relational learning (De Raedt and Kersting, 2008) is a research area that deals with knowledge representation and learning in probabilistic first-order logics. Therein, a particularly popular approach is *Markov Logic* (Richardson and Domingos, 2006). A *Markov logic network* (MLN) is a set of weighted first-order formulas where a larger weight means that the formula is more likely to be true. The semantics of an MLN is given via a log-linear model that takes the weights of formulas into account in order to determine probabilities for classical first-order interpretations. Markov logic networks have been used for e. g. diagnosis of bronchial carcinoma on ion mobility spectrometry data (Finthammer *et al.*, 2010) or social network analysis (Domingos and Lowd, 2009).

In knowledge representation and reasoning *consistency* is a crucial issue and in order to cope with inconsistency different formalisms use different techniques. For example, most belief revision approaches (Hansson, 2001) have to maintain consistency by altering the represented information, and default logics and the like (Reiter, 1980; Gelfond and Leone, 2002) use a non-monotonic inference procedure that bypasses classical inconsistency. Still, even a default theory can be inconsistent in a non-classical sense if there are two complementary defaults present in the theory. In Markov logic, inconsistency is not an issue as every MLN has a well-defined log-linear model (ignoring MLNs that contain infinite weights on two contradictory formulas). Therefore, every MLN is consistent by definition. However, it is not necessarily true that the log-linear model is *meaningful* and adequately represents the information in the network. For example, when representing weighted formulas such as $(\text{sunny}, 5)$ and $(\text{rain}, -20)$ one would probably expect that at least $P(\text{sunny}) > P(\text{rain})$ for the log-linear model P of the whole MLN. However, this is not guaranteed as other formulas may interfere in the computation of the final probabilities. Furthermore, consider the two weighted formulas $(\psi, 10)$ and $(\neg\psi, 10)$. The log-linear model P of only these two formulas is well-defined and has $P(\psi) = 0.5$. It is questionable whether these probabilities are appropriate and whether it would not be more appropriate to define this set of formulas as inconsistent. In particular, computing a log-linear model P' of an "empty" knowledge base yields $P'(\psi) = 0.5$ as well. Therefore, from the semantical point of view, the difference between inconsistency (contradicting formulas) and ignorance (no knowledge at all) cannot be recognized. This also makes it hard to detect modeling errors, particularly in large knowledge bases.

In this paper, we introduce the notion of *coherence* for MLNs. Informally, an MLN is coherent if it is "adequately" represented by its log-linear model. We develop a general framework for coherence measurement that bases on a notion of distance between the log-linear model and the weights of the formulas of the MLN. This measure is able to identify the amount of interference between different formulas of the MLN and thus gives an estimation of whether inference based on the log-linear model might result in counterintuitive results. We discuss one particular application of our framework for merging multiple MLNs into a single one. This is a typical scenario when multiple (domain) experts have to share their knowledge in order to solve a more general task. When merging multiple MLNs, the formulas of one MLN might influence the probabilities previously determined by another MLN which might give unintuitive results. By comparing the coherence of the merged MLN with the coherence of the individual MLNs we define a notion of compatibility for the merging scenario. In summary, the contributions of this paper are as follows:

1. We introduce the notion of *coherence* as a measure for assessing the adequateness of the log-linear model of an MLN (Section 3).

2. We show that our measure satisfies several desirable properties such as monotonicity and independence of irrelevant information. We also present a methodology for using the notion of coherence for knowledge engineering (Section 4).

3. We apply the notion of coherence to the problem of merging multiple MLNs and show that our measure is able to identify incompatibilities (Section 5).

4. We briefly describe our implementation of the coherence measurement framework (Section 6).

Proofs of technical results can be found in the appendix.

2 MARKOV LOGIC NETWORKS

Markov logic (Richardson and Domingos, 2006) is a statistical relational framework which combines Markov networks (Pearl, 1988) with aspects of first-order logic. The Markov logic syntax complies with first-order logic without functions where each formula is quantified by an additional weight. Let Pred be a finite set of predicates, C a finite set of constants, V a set of variables, and $\mathcal{L}_C$ be the functor-free first-order language on Pred, C, and V. For what remains we assume Pred and V to be fixed.

Definition 1. A *Markov logic network* (MLN) L on $\mathcal{L}_C$ is a finite ordered set of tuples $L = \langle(\phi_1, g_1), \ldots, (\phi_n, g_n)\rangle$ with $\phi_1, \ldots, \phi_n \in \mathcal{L}_C$ and $g_1, \ldots, g_n \in \mathbb{R}$.

In contrast to the original literature on MLNs (Richardson and Domingos, 2006) we define an MLN to be an *ordered* set of tuples (ϕ_i, g_i) $(i = 1, \ldots, n)$. This order can be arbitrary and has no special meaning other than to enumerate the elements of an MLN in an unambiguous manner. Any set operation on an MLN is defined in the same way as without an explicit order.

Note, that the weights of an MLN L have no obvious probabilistic interpretation (Fisseler, 2008) and are interpreted relative to each other when defining the joint probability function for L (see below).

Example 1. We adopt the standard example (Domingos and Lowd, 2009) to illustrate the intuition behind MLNs. Define $L_{sm} = \langle(\phi_1, 0.7), (\phi_2, 2.3), (\phi_3, 1.5), (\phi_4, 1.1), (\phi_5, \infty)\rangle$ via

$$\phi_1 = \mathsf{friends}(X, Y) \wedge \mathsf{friends}(Y, Z) \Rightarrow \mathsf{friends}(X, Z)$$
$$\phi_2 = \neg(\exists Y : \mathsf{friends}(X, Y)) \Rightarrow \mathsf{smokes}(X)$$
$$\phi_3 = \mathsf{smokes}(X) \Rightarrow \mathsf{cancer}(X)$$
$$\phi_4 = \mathsf{friends}(X, Y) \Rightarrow (\mathsf{smokes}(X) \Leftrightarrow \mathsf{smokes}(Y))$$
$$\phi_5 = \mathsf{friends}(X, Y) \Leftrightarrow \mathsf{friends}(Y, X)$$

The above MLN models uncertain relationships of smoking habits and people. Formula ϕ_1 means that being friends is a transitive relation, ϕ_2 means that people without friends usually smoke, ϕ_3 that smoking causes cancer, ϕ_4 that friends have similar smoking habits, and ϕ_5 that being friends is a symmetric relation. The formula ϕ_5 has an infinite weight which results in ϕ_5 being a *hard* constraint that must be satisfied.

Semantics are given to an MLN L by grounding L appropriately in order to build a Markov net and its corresponding log-linear model. Let $\Omega(C)$ be the set of (Herbrand) interpretations for Pred and C. For $\phi \in \mathcal{L}_C$ let $\mathsf{gnd}_C(\phi)$ denote the set of ground instances of ϕ wrt. C. Let $\omega \in \Omega(C)$ and define $n_\phi^C(\omega) = |\{\phi' \in \mathsf{gnd}_C(\phi) \mid \omega \models \phi' \}|$. The term $n_\phi^C(\omega)$ denotes the number of instances of ϕ that are satisfied in ω. Then a probability function $P_{L,C} : \Omega(C) \to [0,1]$ can be defined as

$$P_{L,C}(\omega) = \frac{1}{Z_C} \exp\left(\sum_{(\phi,g)\in L} n_\phi^C(\omega) g\right) \tag{3.7}$$

with

$$Z_C = \sum_{\omega\in\Omega(C)} \exp\left(\sum_{(\phi,g)\in L} n_\phi^C(\omega) g\right)$$

being a normalization constant and $\exp(x) = e^x$ is the exponential function with base e. By defining $P_{L,C}$ in this way, worlds that violate fewer instances of formulas are more probable than worlds that violate more instances (depending on the weights of the different formulas). Hence, the fundamental idea for MLNs is that first-order formulas are not handled as hard constraints. Instead, each formula is more or less softened depending on its weight. Hence, a possible world *may* violate a formula without necessarily receiving a zero probability. A formula's weight specifies how strong the formula is, i. e., how much the formula influences the probability of a satisfying world versus a violating world. This way, the weights of all formulas influence the determination of a possible world's probability in a complex manner. One clear advantage of this approach is that MLNs can directly handle contradictions in a knowledge base, since the (contradictory) formulas are weighted against each other.

The probability function $P_{L,C}$ can be extended to sentences (ground formulas) of $\mathcal{L}_C$ via

$$P_{L,C}(\phi) = \sum_{\omega\models\phi} P_{L,C}(\omega) \tag{3.8}$$

for ground $\phi \in \mathcal{L}_C$.

Determining the probability of a sentence ϕ using Equations (3.7) and (3.8) is merely manageable for very small sets of constants, but intractable for domains of a more realistic size. While $P_{L,C}(\phi)$ can be approximated using Markov chain Monte-Carlo methods (MCMC methods) performance might still be too slow in practice (Richardson and Domingos, 2006). There are more sophisticated and efficient methods to perform approximated inference if ϕ is a conjunction of ground literals, cf. (Richardson and Domingos, 2006). Also, approaches for lifted inference exploit symmetries in the graph models which can speed up performance quite impressively, see e.g. (Kersting, 2012) for an overview.

3 MEASURING COHERENCE

Representing knowledge using Markov Logic requires defining the weights for the qualitative parts of the knowledge. In (Richardson and Domingos, 2006) it is suggested that weights of formulas have to be learned from data. Nonetheless, in (Domingos and Lowd, 2009) and (Fisseler, 2008) a heuristic is discussed that determines weights of formulas from probabilities. There, an interpretation of the weight g of a formula ϕ is provided as the log-odd between a world where ϕ is true and a world where ϕ is false, other things being equal, i.e., given some probability $p \in [0,1]$ and a formula ϕ the corresponding *Markov weight* $g_{p,\phi}$ of p is defined by

$$g_{p,\phi} = \ln \frac{p}{1-p} r_\phi \tag{3.9}$$

where $\ln x$ is the natural logarithm of x and r_ϕ is the ratio of the number of worlds not satisfying and the number of worlds satisfying some ground instance of ϕ[1], see also (Fisseler, 2008) for a discussion. The justification for this heuristic comes from the general observation that for a ground formula ϕ and an MLN $L = \langle(\phi, g_{p,\phi})\rangle$, one exactly obtains $P_{L,C}(\phi) = p$. Arguably, it is easier for an expert to express uncertainty in the truth of a formula in form of a probability instead of a weight on a logarithmic scale. When defining an MLN L in this way one has to be aware of the fact that the probabilistic model $P_{L,C}$ of L and a set of constants C may not completely reproduce those intended probabilities.

Example 2. Consider the MLN $L = \langle(A(X), 2), (A(c_1), -5)\rangle$ and $C = \{c_1, c_2, c_3\}$. Assume that the weights of the formulas of L have been defined using the schema of Markov weights, i.e., the probability of $A(X)$ is intended to be approximately 0.881 ($g_{0.881,A(X)} \approx 2$) and of $A(c_1)$ it is approximately 0.0067 ($g_{0.0067,A(X)} \approx -5$). However, we obtain $P_{L,C}(A(c_1)) = 0.047$ which matches neither probability.

In contrast to other probabilistic logics such as classical probabilistic logic (Paris, 1994) or Bayes nets (Pearl, 1988), weights in Markov Logic are not handled as constraints but as factors that influence the determination of probabilities. By accepting this behavior the observation made in Example 2 is understandable. However, due to this behavior it is hard to verify whether some formalization is adequate for a representation problem and whether it is *robust* with respect to extensions:

1 For example, it is $r_\phi = 1$ for a ground atom ϕ and $r_\phi = 1/3$, $r_\phi = 3$ for a disjunction resp. conjunction of ground atoms.

Example 3. Assume we want to model an MLN L such that its model gives a probability 0.5 for each instance $A(c_1), A(c_2), A(c_3)$. This can be achieved by modeling $L = \langle (A(X), -10), (A(X), 10) \rangle$ and $C = \{c_1, c_2, c_3\}$. Assume now we want to incorporate a new piece of information such that $P_{L,C}(A(c_1)) = 0.9$ but still $P_{L,C}(A(c_2)) = P_{L,C}(A(c_3)) = 0.5$. In order to realize this one has to add a new weighted formula $(A(c_1), g)$ to L with some weight g. Due to the interference with the other formulas g cannot easily be determined. This results from the inadequate modeling of the initial knowledge via the MLN L. In this case, the empty MLN would have been a better fit to represent the intended uniform probability distribution. Also, the extended MLN $L' = \langle (A(c_1), 2.2) \rangle$ ($2.2 \approx \ln({}^{0.9}/_{1-0.9})$) yields $P_{L,C}(A(c_1) \approx 0.9$ and $P_{L,C}(A(c_2) = P_{L,C}(A(c_3) = 0.5$.

In the rest of this section, we investigate the issue of assessing how well the probabilistic model $P_{L,C}$ of an MLN L and a set of constants C reflects the probabilities used for defining L. For that we employ the Markov weights as a comparison criterion, i. e., we compare the probability of every formula of L in the probabilistic model $P_{L,C}$ with the probability this formula would have in the probabilistic model $P_{L',C}$ of the MLN L' that only consists of this formula. Note that our approach could also be formulated using any other (surjective) function g'_p that assigns weights to probabilities.

Similarly as consistency is defined for classical logics we also define a strict version of coherence. In particular, we say that L is *perfectly coherent* wrt. C if $P_{L,C}$ assigns to each formula the same probability as prescribed by the Markov weights. More formally:

Definition 2. Let $L = \langle (\phi_1, g_1), \ldots, (\phi_n, g_n) \rangle$ be an MLN. We say that L is *perfectly coherent* if and only if for all $i = 1, \ldots, n$ and $\phi' \in \mathsf{gnd}_C(\phi_i)$ it holds $P_{L,C}(\phi') = p$ and $g_i = g_{p,\phi_i}$.

If $g = g_{p,\phi}$ is a Markov weight observe that

$$p = p_{g,\phi} = \frac{\exp(g)}{r_\phi + \exp(g)}$$

with $p_{g,\phi} = 1$ if $g = \infty$ and $p_{g,\phi} = 0$ if $g = -\infty$. We also call $p_{g,\phi}$ a *Markov probability*. Following the spirit of inconsistency measures for probabilistic logics (Thimm, 2013b) we take a more graded approach to coherence analysis and, consequently, in the following we will consider the problem of defining coherence values.

Before formalizing our coherence measurement framework we need some further notation. Let C be a set of constants and $\phi \in \mathcal{L}_C$. The *ground vector* of ϕ with respect to C is defined via $\mathsf{gnd}^{\rightarrow}_C(\phi) = \langle \phi_1, \ldots, \phi_n \rangle$ where $\mathsf{gnd}_C(\phi) = \{\phi_1, \ldots, \phi_n\}$ and $\phi_1, \ldots, \phi_n$ is some arbitrary but fixed canonical ordering of $\mathsf{gnd}_C(\phi)$. If $\langle \phi_1, \ldots, \phi_n \rangle \in \mathcal{L}_C^n$ is a vector of formulas and P a probability function then we write

$$P(\langle \phi_1, \ldots, \phi_n \rangle) = \langle P(\phi_1), \ldots, P(\phi_n) \rangle$$

As a central tool for measuring coherence we use (weak) distance measures.

Definition 3. Let $n \in \mathbb{N}^+$. A function $d : [0,1]^n \times [0,1]^n \to [0, \infty)$ is called a weak *distance measure* if it satisfies 1.) $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$ (*reflexivity*) and 2.) $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$ (*symmetry*), for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Note that weak distance measures differ from standard distance measures by not requiring the *triangle inequality* to hold. In the following, we will refer to these measures as *distance measures* anyway for reasons of brevity. In this work we consider the following measures (let $\vec{x} = \langle x_1, \ldots, x_n \rangle, \vec{y} = \langle y_1, \ldots, y_n \rangle \in [0,1]^n$, $p \in \mathbb{N}^+$): 1.) $d_p(\vec{x}, \vec{y}) = \sqrt[p]{|x_1 - y_1|^p + \ldots + |x_n - y_n|^p}$ (p-norm distance), 2.) $d_{p,0}(\vec{x}, \vec{y}) = \sqrt[p]{|x_1 - y_1|^p + \ldots + |x_n - y_n|^p} / \sqrt[p]{n}$ (normalized p-norm distance), 3.) $d_{\max}(\vec{x}, \vec{y}) = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\}$ (max-distance), 4.) $d_{\min}(\vec{x}, \vec{y}) = \min\{|x_1 - y_1|, \ldots, |x_n - y_n|\}$ (min-distance), and 5.) $d_{\text{avg}}(\vec{x}, \vec{y}) = (|x_1 - y_1| + \ldots + |x_n - y_n|)/n$ (average distance).

In the following, we will use distance measures to measure the differences between vectors of probabilities that arise for each formula of an MLN upon grounding and the corresponding expected probabilities. In order to aggregate the distances of each formula we use aggregation functions.

Definition 4. A function $\theta : [0,1]^n \to [0,1]$ is called an *aggregation function*.

We consider the following aggregation functions (let $\vec{x} = \langle x_1, \ldots, x_n \rangle \in [0,1]^n$): 1.) $\theta_{\max}(\vec{x}) = \max\{x_1, \ldots, x_n\}$ (maximum), 2.) $\theta_{\min}(\vec{x}) = \min\{x_1, \ldots, x_n\}$ (minimum), and 3.) $\theta_{\text{avg}}(\vec{x}) = (x_1 + \ldots + x_n)/n$ (average).

Using distance measures and aggregation functions we define the *coherence* of an MLN L as how well L reflects the probabilities that are intended to be modeled by weights.

Definition 5. Let d be a distance measure, θ an aggregation function, $L = \langle (\phi_1, g_1), \ldots, (\phi_n, g_n) \rangle$ an MLN, and C a set of constants. Then the *coherence* $\mathsf{coh}_C^{d,\theta}(L)$ of L wrt. C and given d, θ is defined via

$$\mathsf{coh}_C^{d,\theta}(L) = 1 - \theta\left(\left\langle d\left(P_{L,C}(\mathsf{gnd}_C^{\rightarrow}(\phi_i)), \Pi_{\phi_i,g_i}^{|\mathsf{gnd}_C^{\rightarrow}(\phi_i)|}\right)\right\rangle_{i=1,\ldots,n}\right) \tag{3.10}$$

with

$$\Pi_{\phi,g}^n = \langle \underbrace{p_{g,\phi}, \ldots, p_{g,\phi}}_{\text{n times}} \rangle$$

The intuition behind the above definition is as follows. Assume that $(\phi(X), g) \in L$ and that $\{\phi(c_1), \ldots, \phi(c_n)\}$ are the groundings of $\phi(X)$. Then $P_{L,C}$ assigns to each of this ground formulas a (potentially different) probability $P_{L,C}(\phi(c_i))$ ($i = 1, \ldots, n$). First, we compute the distance of the vector $\langle P_{L,C}(\phi(c_1), \ldots, P_{L,C}(\phi(c_n)\rangle$ to the vector $\langle p_{g,\phi(X)}, \ldots, p_{g,\phi(X)}\rangle$ (the uniform vector of the probability corresponding to the weight g). Finally, we aggregate the distances of all these vectors for all formulas in L. Therefore, $\mathsf{coh}_C^{d,\theta}(L)$ provides an aggregated assessment of how close the actual probabilities match the weights.

As we are in a probabilistic framework, one might wonder why we use ordinary distance measures and aggregation functions for defining a measure of coherence. A seemingly better alternative should be e. g. the Kullback-Leibler divergence (Kullback and Leibler, 1951) which has a well-defined meaning when measuring the difference between two probability functions. However, in our setting we compare a probability function $P_{L,C}$ with a set of probabilities derived from the weights of the MLN L. In particular, the latter is usually contradictory (unless L is perfectly coherent), so the meaning of the Kullback-Leibler divergence in this context is not clear. We leave this issue for future work and consider now the distance measures defined so far.

4 ANALYSIS

To further illustrate the meaning of the Definition 5 let us consider the coherence measure $\mathsf{coh}_C^{d_{\max},\theta_{\max}}$ and an MLN $L = \langle(\phi_1, g_1), \ldots(\phi_n, g_n)\rangle$. Then $\mathsf{coh}_C^{d_{\max},\theta_{\max}}(L)$ is one minus the maximum deviation of the probability of some ground instance ϕ_i of L in $P_{L,C}$ to the probability p_i estimated by its weight g_i, assumed that g_i has been determined by setting $g_i = \ln \frac{p_i}{1-p_i} r_{\phi_i}$.

Example 4. Consider the MLN $L = \langle(A(X), 2)\rangle$ and $C = \{c_1, c_2, c_3\}$. Note that the probability p intended to be modeled by the weight 2 is $p = p_{2,A(X)} = \frac{\exp(2)}{1+exp(2)} \approx 0.881$ (note that $r_{A(X)} = 1$). As there is only one formula in L it also follows directly that $P_{L,C}(A(c_1)) = P_{L,C}(A(c_2)) = P_{L,C}(A(c_3)) \approx 0.881$ as well. It follows that $\mathsf{coh}_C^{d_{\max},\theta_{\max}}(L) = 1 - 0 = 1$.

Example 5. We continue Example 2 and consider the MLN $L = \langle(A(X), 2), (A(c_1), -5)\rangle$ and $C = \{c_1, c_2, c_3\}$. Note that the probability p_1 intended to be modeled by the weight 2 is $p_1 = p_{2,A(X)} = \frac{\exp(2)}{1+\exp(2)} \approx 0.881$ and for the weight -5 it is $p_2 = p_{-1,A(c_1)} = \frac{\exp(-5)}{1+\exp(-5)} \approx 0.0067$. For $P_{L,C}$ we obtain

$P_{L,C}(A(c_1)) \approx 0.041$ and $P_{L,C}(A(c_2)) = P_{L,C}(A(c_3)) \approx 0.881$. Then $\mathsf{coh}_C^{d_{\max},\theta_{\max}}(L)$ computes to

$$\begin{aligned}\mathsf{coh}_C^{d_{\max},\theta_{\max}}(L) &= 1 - \max\{|P_{L,C}(A(c_1)) - p_1|,\\ &\quad |P_{L,C}(A(c_2)) - p_1|, |P_{L,C}(A(c_3)) - p_1|,\\ &\quad |P_{L,C}(A(c_1)) - p_2|\}\\ &\approx 0.16\end{aligned}$$

In the introduction we gave an example illustrating that MLNs are not always capable of differentiating between (logical) inconsistency and ignorance. However, using our notion of coherence we are able to detect this difference.

Example 6. Consider the MLN $L = \langle (A, -10), (A, 10) \rangle$ with a proposition (a predicate without parameters) A and $C = \{c_1, c_2, c_3\}$. The probabilities p_1, p_2 intended to modeled by the weights -10, and 10 are (respectively) $p_1 = p_{-10,A} \approx 0$ and $p_2 = p_{10,A} \approx 1$ and for $P_{L,C}$ we obtain $P_{L,C}(A) = 0.5$. Then we have

$$\begin{aligned}\mathsf{coh}_C^{d_{\max},\theta_{\max}}(L) &= 1 - \max\{|P_{L,C}(A) - p_1|, |P_{L,C}(A) - p_2|\}\\ &= 0.5\end{aligned}$$

Furthermore, for the empty MLN $L' = \langle\rangle$ and an arbitrary C we always have $\mathsf{coh}_C^{d,\theta}(L) = 1$ for any $d \in \{d_p, d_{p,0}, d_{\max}, d_{\min}, d_{\text{avg}}\}$ and $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{\text{avg}}\}$.

We now turn to the formal properties of $\mathsf{coh}_C^{d,\theta}$.

Proposition 1. *For $d \in \{d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$ and $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{avg}\}$ we have $\mathsf{coh}_C^{d,\theta}(L) \in [0,1]$ for every L and C.*

The above proposition shows that many coherence measures are normalized on $[0,1]$ and, therefore, different MLNs can be compared and categorized by their coherence values. Note that the Proposition 1 does not hold in general for d_p.

Proposition 2. *If d satisfies reflexivity and θ satisfies $\theta(x_1, \ldots, x_n) = 0$ iff $x_1 = \ldots = x_n = 0$ then $\mathsf{coh}_C^{d,\theta}(L) = 1$ iff L is perfectly coherent wrt. C.*

The above proposition states that our framework satisfies the basic property of detecting whether an MLN is perfectly coherent, given some minimal requirements of both distance measure and aggregation function.

Corollary 1. *If $d \in \{d_p, d_{p,0}, d_{\max}, d_{avg}\}$ ($p \in \mathbb{N}^+$) and $\theta \in \{\theta_{\max}, \theta_{avg}\}$ then $\mathsf{coh}_C^{d,\theta}(L) = 1$ iff L is perfectly coherent wrt. C.*

Note that the above statement does not hold for $d_{\min}$ and $\theta_{\min}$.

Next we look into the behavior of $\mathsf{coh}_C^{d,\theta}$ under changes of L and C.

Proposition 3. *For any d it holds $\mathsf{coh}_C^{d,\theta_{\max}}(L)$ is monotonically decreasing in L, i.e. $\mathsf{coh}_C^{d,\theta_{\max}}(L) \geq \mathsf{coh}_C^{d,\theta_{\max}}(L \cup \{(\phi, g)\})$.*

This property states that $\mathsf{coh}_C^{d,\theta_{\max}}(L)$ cannot get more coherent under addition of formulas. This corresponds to the classical concept of inconsistency insofar that an inconsistent knowledge base of classical logical formulas cannot get consistent when adding new information. Note that the above property does not hold in general for $\theta_{\min}$ and θ_{avg}. For a special case of a new formula we make the following observation.

Proposition 4. *For any d, if a consistent ϕ shares no predicate with L then $\mathsf{coh}_C^{d,\theta_{\max}}(L) = \mathsf{coh}_C^{d,\theta_{\max}}(L \cup \{(\phi, g)\})$ for every $g \in \mathbb{R}$.*

In other words, if we add totally unrelated (but consistent) information to an MLN this does not change its coherence.

Proposition 5. *For $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{avg}\}$ it holds that $\mathsf{coh}_C^{d_{\min},\theta}(L)$ is monotonically increasing in C, i.e. $\mathsf{coh}_C^{d_{\min},\theta}(L) \leq \mathsf{coh}_{C \cup \{c\}}^{d_{\min},\theta}(L)$.*

This result states that considering more individuals increases the coherence of the MLN wrt. $d_{\min}$. The rationality of satisfying this property is evident as by taking more individuals into account exceptions to formulas become negligible. Consider the MLN L of Example 5 which specifies a general rule ($A(X)$ has to hold in general) and an exception (c_1 does not satisfy $A(X)$). However, the general rule dominates the coherence value the more individuals actually satisfy it.

Example 7. We continue Example 5 but consider varying sizes of the domain. So let $L = \langle (A(X), 2), (A(c_1), -5) \rangle$ and $C_i = \{c_1, \ldots, c_i\}$ for $i \in \mathbb{N}$. Figure 1 shows the behavior of four different coherence measures when the domain increases in size.

The framework proposed so far can be utilized by a knowledge engineer when debugging MLNs. In particular, a coherence measure can be used to evaluate whether the semantics of an MLN adequately represents its intended meaning if no data for learning is available. Note that this tool can be applied even if the heuristic for defining weights from probabilities may not seem adequate as the tool uses these only for assessing the influence one formula has on another.

Example 7 showed that, in particular, distance measures based on the p-norm may give a more fine-grained view on the evolution of coherence values (however, note that these distance measures do not satisfy monotonicity wrt. the domain in general). Independently of the actually chosen

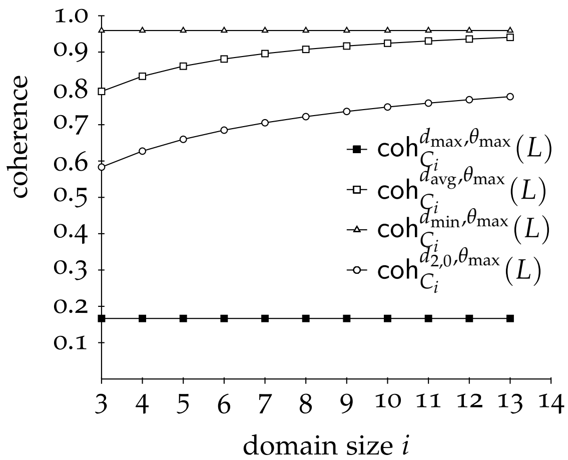


Figure 1: Coherence values for Example 7

combination of distance measure and aggregation function, by utilizing the framework of coherence measurement for analyzing a given MLN the knowledge engineer is already able to detect several design flaws:

1. If an MLN is coherent (i. e. has a comparatively large coherence value) but exhibits unintuitive inferences, then probably some weights have been chosen wrong (as there is only little interdependence between formulas).

2. If an MLN is coherent and exhibits no unintuitive inferences, then the MLN is a good representation of the given knowledge and it will probably be easier to extend it.

3. If an MLN is incoherent (i. e. has a comparatively low coherence value) and exhibits unintuitive inferences, then the knowledge engineer should have a look into the structure of the knowledge base as there may be unwanted interdependences amongst formulas.

4. If an MLN is incoherent but exhibits no unintuitive inferences, then the MLN may not be an adequate representation of the knowledge and further extensions might yield unintuitive results.

As a final remark, observe that our notion of coherence is also compatible with the usual notion of probabilistic consistency. In particular, starting from a consistent probabilistic view in form of a probability function, we can always find a perfectly coherent MLN representing this probability function.

Proposition 6. *Let $P : \Omega(C) \to [0,1]$ be any probability function. Then there is a perfectly coherent MLN L on $\mathcal{L}_C$ with $P_{L,C} = P$. In particular, it holds $\mathsf{coh}^{d,\theta}_C(L) = 1$ for any $d \in \{d_p, d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$ and $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{avg}\}$.*

As for every MLN L the probability function $P_{L,C}$ is always well-defined the above observation can also be used to transform an incoherent MLN L into a coherent MLN L' that weighs formulas more adequately. However, note that the formulas in L' need not necessarily to be the same as in L.

5 APPLICATION: COMPATIBILITY OF MLNS

A particular use case for applying our framework arises when considering a knowledge merging scenario. Consider the case of multiple experts merging their knowledge in order to obtain a broader picture on some problem domain. Then, the individual pieces of information of each expert contribute to the overall probabilities obtained from the log-linear model of the merged MLN. Given that the experts have contradictory views on some parts of the modeled knowledge the merged MLN might not adequately reflect the joined knowledge. In order to analyze whether the merging of MLNs gives rise to a potentially meaningless joint MLN we employ our framework of coherence measurement as follows.

Definition 6. Let d be a distance measure, θ an aggregation function, $L_1, \ldots, L_m$ MLNs, and $C_1, \ldots, C_m$ sets of constants. The *compatibility* $\mathsf{comp}^{d,\theta}_{C_1,\ldots,C_m}(L_1, \ldots, L_m)$ of $L_1, \ldots, L_m$ wrt. $C_1, \ldots, C_m$ given d, θ is defined via

$$\begin{aligned}
&\mathsf{comp}^{d,\theta}_{C_1,\ldots,C_m}(L_1, \ldots, L_m) \\
&= \frac{1}{2}\left(1 + \mathsf{coh}^{d,\theta}_{C_1\cup\ldots\cup C_m}(L_1 \cup \ldots \cup L_m) - \frac{1}{m}\sum_{i=1}^{m} \mathsf{coh}^{d,\theta}_{C_i}(L_i)\right)
\end{aligned}$$

The value $\mathsf{comp}^{d,\theta}_{C_1,\ldots,C_m}(L_1, \ldots, L_m)$ describes how well the MLNs $L_1, \ldots, L_n$ can be merged. In essence, it measures how much the coherence of the joint MLN differs from the average coherence of all input MLNs. Intuitively, the larger the value of $\mathsf{comp}^{d,\theta}_{C_1,\ldots,C_m}(L_1, \ldots, L_m)$ the more *compatible* the MLNs should be. The exact form of the compatibility measure has been chosen like this to satisfy the *normalization property*, see Proposition 8 below.

Example 8. Consider the three MLNs $L_1 = \langle(\phi_1, 1.85), (\phi_2, 1.85)\rangle, L_2 = \langle(\phi_3, \infty)\rangle$, $L_3 = \langle(\phi_4, 1.1), (\phi_5, \infty)\rangle$ defined via

$$\phi_1 = \mathsf{quaker}(X) \Rightarrow \mathsf{pacifist}(X)$$
$$\phi_2 = \mathsf{republican}(X) \Rightarrow \neg\mathsf{pacifist}(X)$$
$$\phi_3 = \mathsf{republican}(\mathit{nixon}) \wedge \mathsf{quaker}(\mathit{nixon}) \wedge \mathsf{president}(\mathit{nixon})$$
$$\phi_4 = \mathsf{president}(X) \Rightarrow \neg\mathsf{actor}(X)$$
$$\phi_5 = \mathsf{president}(\mathit{reagan}) \wedge \mathsf{actor}(\mathit{reagan})$$

which model an extended version of the Nixon diamond. Using $\mathsf{coh}_{C}^{d_{\max},\theta_{\max}}$ we obtain

$$\mathsf{coh}_{\{d\}}^{d_{\max},\theta_{\max}}(L_1) \approx 0.982$$
$$\mathsf{coh}_{\{nixon\}}^{d_{\max},\theta_{\max}}(L_2) = 1$$
$$\mathsf{coh}_{\{reagan\}}^{d_{\max},\theta_{\max}}(L_3) = 0.9$$

and for the merged MLN $L = L_1 \cup L_2 \cup L_3$ we obtain

$$\mathsf{coh}_{\{d,nixon,reagan\}}^{d_{\max},\theta_{\max}}(L) \approx 0.55$$

This leads to

$$\mathsf{comp}_{\{d\},\{nixon\},\{reagan\}}^{d_{\max},\theta_{\max}}(L_1, L_2, L_3) \approx 0.295$$

Furthermore, note that $\mathsf{coh}_{\{d,nixon\}}^{d_{\max},\theta_{\max}}(L_1 \cup L_2) = 0.55$ and $\mathsf{coh}_{\{nixon,reagan\}}^{d_{\max},\theta_{\max}}(L_2 \cup L_3) = 0.85$ and, therefore, L_2 and L_3 are more compatible than L_1 and L_2:

$$\mathsf{comp}_{\{d\},\{nixon\}}^{d_{\max},\theta_{\max}}(L_1, L_2) \approx 0.2795$$
$$\mathsf{comp}_{\{nixon\},\{reagan\}}^{d_{\max},\theta_{\max}}(L_2, L_3) = 0.45$$

Our compatibility measure gives meaningful results in the above example. We now investigate how it behaves in the general case.

Proposition 7. *It holds* $\mathsf{comp}_{C_1,\ldots,C_m}^{d,\theta_{\max}}(L_1, \ldots, L_m) \in [0, 1]$ *for every* $d \in \{d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$*,*

The statement above says that the compatibility measure is normalized and therefore comparable.

Proposition 8. *For every* $d \in \{d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$ *it is* $\mathsf{comp}_{C_1,\ldots,C_m}^{d,\theta_{\max}}(L_1, \ldots, L_m) = 0$ *if and only if* $\mathsf{coh}_{C_1\cup\ldots\cup C_m}^{d,\theta_{\max}}(L_1 \cup \ldots \cup L_m) = 0$ *and* $\mathsf{coh}_{C_i}^{d,\theta_{\max}}(L_i) = 1$ *for all* $i = 1, \ldots, m$*.*

The above proposition states that a set of MLNs is completely incompatible if and only if each individual MLN is perfectly coherent and the merged MLN is completely incoherent.

6 IMPLEMENTATION

The framework for measuring coherence of MLNs has been implemented in the *Tweety library for artificial intelligence*[2]. The framework contains imple-

2 http://tinyurl.com/MLNCoherence2

mentations for all distance measures and aggregation functions discussed above and we provided both a naive and complete MLN reasoner and a wrapper for using the *Alchemy*[3] MLN reasoner. While the naive MLN reasoner implements Equations (3.7) and (3.8) in a straightforward way by simply computing the probabilities $P_{L,C}(\omega)$ for all $\omega \in \Omega(C)$, the *Alchemy* MLN reasoner supports different approximate methods such as Markov chain Monte-Carlo. Computing the coherence value $\mathsf{coh}_C^{d,\theta}(L)$ is computationally quite expensive as it involves calls to the MLN reasoner for every ground instance of a formula in L. Therefore, using the naive MLN reasoner is only feasible for small examples. However, in its current version the *Alchemy* MLN reasoner does not support querying the probabilities of arbitrary ground formulas but only for ground atoms. In order to obtain the probability of an arbitrary ground formula ϕ using *Alchemy* it has to be incorporated into the MLN via adding a strict formula $\phi \Leftrightarrow a$ with some new ground atom a. Then the probability of a can be queried which is, in theory, the same as the probability of ϕ. However, during our experiments we discovered that internal optimization mechanisms of *Alchemy* might change the probabilities of other formulas when adding the strict formula $\phi \Leftrightarrow a$. This observation also raises the need for the development of an MLN reasoner that supports querying for arbitrary ground formulas. Recent developments such as (Niepert, 2013) are gaining to close this gap.

7 DISCUSSION AND CONCLUSION

We introduced coherence as an indicator of how the weighted formulas of an MLN interact with each other. We used distance measures and aggregation functions to measure coherence by comparing the observed probabilities with the ones stemming from a naive probabilistic interpretation of the weights. By doing so, we came up with a meaningful assessment tool that satisfies several desirable properties. As an application for our framework we investigated the issue of merging and developed an indicator for quantifying the compatibility of different MLNs.

The approach presented in this paper can be used by a knowledge engineer to determine suitable weights for formulas, thus complementing the work of Pápai et al. (Pápai *et al.*, 2012) where MLNs are constructed by taking subjective probabilities of an expert into account. In particular, (Pápai *et al.*, 2012) already discusses the issue of consistent and inconsistent input probabilities and that in the latter case, parameters for the probability distribution have to be averaged, thus also resulting in an incoherent MLN in the sense of our work. By assessing the representation quality of MLNs using our approach experts can be guided to carefully choose correct weights/probabilities or re-structure the knowledge base.

To the best of our knowledge this work is the first that deals with quantifying the representation quality of an MLN by investigating the interrela-

3 http://alchemy.cs.washington.edu

tionships of its formulas. The work presented is inspired by works on measuring the inconsistency in probabilistic conditional logic (Thimm, 2013b). The work (Thimm, 2013b) defines an inconsistency measure by measuring the distance of an inconsistent knowledge base to the next consistent one. In this aspect, our framework uses similar methods. But as the concept of consistency is not applicable for MLNs we used a probabilistic interpretation of weights as a reference for assessing the coherence of an MLN. The term *coherence* has also been used before to describe "appropriateness" of a knowledge base or a model in other contexts. For example, in (Meijs, 2005) a set of propositional formulas is said to be coherent with respect to a probability function if the probability of each single formula increases when conditioning on the other formulas (there are also other similar notions considered).

Although MLNs are quite a mature framework for dealing with first-order probabilistic information, the lack of powerful and flexible MLN reasoner became evident in our experiments. Besides *Alchemy* we also looked at other available reasoning systems for MLNs such as thebeast[4] and Tuffy[5] but all lacked the crucial feature of computing the probabilities of arbitrary ground formulas. For future work, we consider to approach this problem and develop an MLN reasoner that can specifically be used for measuring coherence. Another direction for future work is the problem of deciding whether a coherent MLN can be learned from data and how to do this.

APPENDIX: PROOFS OF TECHNICAL RESULTS

Proposition 1. *For $d \in \{d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$ and $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{avg}\}$ we have $\mathsf{coh}_C^{d,\theta}(L) \in [0,1]$ for every L and C.*

Proof Sketch. Consider Equation (4) and observe that both $P_{L,C}(\mathsf{gnd}_C^{\vec{\to}}(\phi_i))$ and $\Pi_{\phi_i,g_i}^{|\mathsf{gnd}_C^{\vec{\to}}(\phi_i)|}$ are vectors in $[0,1]^{|\mathsf{gnd}_C^{\vec{\to}}(\phi_i)|}$, for all $i = 1,\ldots,n$. For $\vec{x},\vec{y} \in [0,1]^n$ observe that $d(\vec{x},\vec{y}) \in [0,1]$ as well for $d \in \{d_{p,0}, d_{\max}, d_{\min}, d_{\text{avg}}\}$. Finally, if $\vec{x} \in [0,1]^m$ we also have that $\theta(\vec{x}) \in [0,1]$ for $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{\text{avg}}\}$. It follows that $\mathsf{coh}_C^{d,\theta}(L) \in [0,1]$. □

Proposition 2. *If d satisfies reflexivity and θ satisfies $\theta(x_1,\ldots,x_n) = 0$ iff $x_1 = \ldots = x_n = 0$ then $\mathsf{coh}_C^{d,\theta}(L) = 1$ iff L is perfectly coherent wrt. C.*

Proof Sketch. Let d satisfy reflexivity, let θ satisfy $\theta(x_1,\ldots,x_n) = 0$ iff $x_1 = \ldots = x_n = 0$, and consider Equation (4). Let now $\mathsf{coh}_C^{d,\theta}(L) = 1$ which is equivalent to

$$\theta\left(\left\langle d\left(P_{L,C}(\mathsf{gnd}_C^{\vec{\to}}(\phi_i)), \Pi_{\phi_i,g_i}^{|\mathsf{gnd}_C^{\vec{\to}}(\phi_i)|}\right)\right\rangle_{i=1,\ldots,n}\right) = 0$$

4 http://code.google.com/p/thebeast/
5 http://hazy.cs.wisc.edu/hazy/tuffy/

Due to $\theta(x_1, \ldots, x_n) = 0$ iff $x_1 = \ldots = x_n = 0$ the above is only valid iff

$$d\left(P_{L,C}(\mathsf{gnd}_C^{\rightarrow}(\phi_i)), \Pi_{\phi_i,g_i}^{|\mathsf{gnd}_C^{\rightarrow}(\phi_i)|}\right) = 0$$

for $i = 1, \ldots, n$. As d satisfies reflexivity the above is only valid iff

$$P_{L,C}(\mathsf{gnd}_C^{\rightarrow}(\phi_i)) = \Pi_{\phi_i,g_i}^{|\mathsf{gnd}_C^{\rightarrow}(\phi_i)|}$$

which is equivalent to stating that L is perfectly coherent wrt. C. □

Proposition 3. *For any d it holds $\mathsf{coh}_C^{d,\theta_{\max}}(L)$ is monotonically decreasing in L, i. e. $\mathsf{coh}_C^{d,\theta_{\max}}(L) \geq \mathsf{coh}_C^{d,\theta_{\max}}(L \cup \{(\phi, g)\})$.*

Proof Sketch. This follows from the fact that $\max\{a_1, \ldots, a_n\} \leq \max\{a_1, \ldots, a_{n+1}\}$ for any $a_1, \ldots, a_{n+1} \in \mathbb{R}$. □

Proposition 4. *For any d, if a consistent ϕ shares no predicate with L then $\mathsf{coh}_C^{d,\theta_{\max}}(L) = \mathsf{coh}_C^{d,\theta_{\max}}(L \cup \{(\phi, g)\})$ for every $g \in \mathbb{R}$.*

Proof Sketch. Let Pred be the set of predicates and $\mathsf{Pred}_L, \mathsf{Pred}_\phi \subseteq \mathsf{Pred}$ with $\mathsf{Pred}_L \cap \mathsf{Pred}_\phi$ and w.l.o.g. $\mathsf{Pred}_L \cup \mathsf{Pred}_\phi = \mathsf{Pred}$ the sets of predicates appearing in L and ϕ, respectively. For $\omega \in \Omega(C)$ let ω_L, ω_ϕ denote the projection of ω on Pred_L and Pred_ϕ, respectively. Let $\Omega(C)^L$ and $\Omega(C)^\phi$ denote the corresponding sets of projected interpretations. If ϕ is consistent and shares no predicates with L, it can be shown that $P_{L\cup\{(\phi,g)\},C}$ (defined as a probability function on $\Omega(C)$) factorizes into the two probability functions $P_{L,C}, P_{\{(\phi,g)\},C}$ (defined on $\Omega(C)^L$ and $\Omega(C)^\phi$, respectively) such that

$$P_{L\cup\{(\phi,g)\},C}(\omega) = P_{L,C}(\omega_L) P_{\{(\phi,g)\},C}(\omega_\phi)$$

As ϕ is consistent we have that $\{(\phi, g)\}$ is perfectly coherent wrt. C (there is nothing to contradict ϕ). It follows that the coherence of $P_{L\cup\{(\phi,g)\},C}$ depends only on $P_{L,C}(\omega_L)$. Note that the probability of every formula ψ, which does not contain predicates from ϕ, can be determined by $P_{L,C}(\omega_L)$ alone:

$$\begin{aligned}
P_{L\cup\{(\phi,g)\},C}(\psi) &= \sum_{\omega\in\Omega(C),\omega\models\psi} P_{L\cup\{(\phi,g)\},C}(\omega) \\
&= \sum_{\omega\in\Omega(C),\omega\models\psi} P_{L,C}(\omega_L) P_{\{(\phi,g)\},C}(\omega_\phi) \\
&= \sum_{\omega_L\in\Omega(C)^L,\omega_\phi\in\Omega(C)^\phi,\omega_L\models\psi} P_{L,C}(\omega_L) P_{\{(\phi,g)\},C}(\omega_\phi) \\
&= \sum_{\omega_L\in\Omega(C)^L,\omega_L\models\psi} P_{L,C}(\omega_L) \sum_{\omega_\phi\in\Omega(C)^\phi} P_{\{(\phi,g)\},C}(\omega_\phi) \\
&= \sum_{\omega_L\in\Omega(C)^L,\omega_L\models\psi} P_{L,C}(\omega_L) \\
&= P_{L,C}(\psi)
\end{aligned}$$

Therefore, we obtain $\mathsf{coh}_C^{d,\theta_{\max}}(L) = \mathsf{coh}_C^{d,\theta_{\max}}(L \cup \{(\phi, g)\})$. $\square$

Proposition 5. *For $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{avg}\}$ it holds that $\mathsf{coh}_C^{d_{\min},\theta}(L)$ is monotonically increasing in C, i.e. $\mathsf{coh}_C^{d_{\min},\theta}(L) \leq \mathsf{coh}_{C\cup\{c\}}^{d_{\min},\theta}(L)$.*

Proof Sketch. First note that $\mathsf{gnd}_C(\phi) \subseteq \mathsf{gnd}_{C\cup\{c\}}(\phi)$ for every formula ϕ. Therefore, the vectors in Equation (4) given to the distance function $d_{\min}$ also increase in size. Furthermore, note that

$$\begin{aligned}
&d_{\min}(\langle x_1, \dots, x_n\rangle, \langle y_1, \dots, y_n\rangle) \geq \\
&\qquad d_{\min}(\langle x_1, \dots, x_{n+1}\rangle, \langle y_1, \dots, y_{n+1}\rangle)
\end{aligned}$$

Finally, note that for $z_1, \dots, z_m, z'_1, \dots, z'_m$ with $z_i \geq z'_i$ (for $i = 1, \dots, m$) we have that

$$\theta(z_1, \dots, z_m) \geq \theta(z'_1, \dots, z'_m)$$

for $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{\text{avg}}\}$. It follows $\mathsf{coh}_C^{d_{\min},\theta}(L) \leq \mathsf{coh}_{C\cup\{c\}}^{d_{\min},\theta}(L)$. $\square$

Proposition 6. *Let $P : \Omega(C) \to [0, 1]$ be any probability function. Then there is a perfectly coherent MLN L on $\mathcal{L}_C$ with $P_{L,C} = P$. In particular, it holds $\mathsf{coh}_C^{d,\theta}(L) = 1$ for any $d \in \{d_p, d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$ and $\theta \in \{\theta_{\max}, \theta_{\min}, \theta_{avg}\}$.*

Proof Sketch. Let $P : \Omega(C) \to [0, 1]$ and for each $\omega \in \Omega(C)$ let ϕ_ω be a formula that characterizes ω, i.e., $\omega \models \phi_\omega$ and there is not $\omega' \in \Omega(C)$ with $\omega' \neq \omega$ and $\omega' \models \phi_\omega$ (if ω is an Herbrand interpretation ω_ϕ can be the conjunction of all literals appearing in ω and the negation of all literals not appearing in ω). Consider the MLN L given by

$$L = \left\{ \left(\phi_\omega, \ln\left(\frac{P(\omega)}{1 - P(\omega)} r_{\phi_\omega} \right) \right) \mid \omega \in \Omega(C) \right\}$$

In other words, L is constructed by assigning each world the weight corresponding to its probability. By construction we obtain $P = P_{L,C}$ and L is perfectly coherent wrt. C. $\square$

Proposition 7. *It holds $\mathsf{comp}_{C_1,\dots,C_m}^{d,\theta_{\max}}(L_1, \dots, L_m) \in [0, 1]$ for every $d \in \{d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$,*

Proof. By Proposition 1 we obtain

$$\begin{aligned}
&\mathsf{comp}_{C_1,\dots,C_m}^{d,\theta}(L_1, \dots, L_m) \\
&= 1/2 \left(1 + \mathsf{coh}_{C_1\cup\dots\cup C_m}^{d,\theta}(L_1 \cup \dots \cup L_m) - \frac{1}{m} \sum_{i=1}^{m} \mathsf{coh}_{C_i}^{d,\theta}(L_i) \right) \\
&\geq 1/2 \left(1 + 0 - \frac{1}{m} \sum_{i=1}^{m} 1 \right) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
&\mathsf{comp}^{d,\theta}_{C_1,\ldots,C_m}(L_1,\ldots,L_m) \\
&= 1/2\left(1+\mathsf{coh}^{d,\theta}_{C_1\cup\ldots\cup C_m}(L_1\cup\ldots\cup L_m)-\frac{1}{m}\sum_{i=1}^{m}\mathsf{coh}^{d,\theta}_{C_i}(L_i)\right) \\
&\leq 1/2\left(1+1-\frac{1}{m}\sum_{i=1}^{m}0\right) \\
&= 1
\end{aligned}$$

□

Proposition 8. *For every* $d \in \{d_{p,0}, d_{\max}, d_{\min}, d_{avg}\}$ *it is* $\mathsf{comp}^{d,\theta_{\max}}_{C_1,\ldots,C_m}(L_1,\ldots,L_m) = 0$ *if and only if* $\mathsf{coh}^{d,\theta_{\max}}_{C_1\cup\ldots\cup C_m}(L_1\cup\ldots\cup L_m) = 0$ *and* $\mathsf{coh}^{d,\theta_{\max}}_{C_i}(L_i) = 1$ *for all* $i = 1,\ldots,m$.

Proof. This is a straightforward corollary from the proof of Proposition 7. □

STREAM-BASED INCONSISTENCY MEASUREMENT

Matthias Thimm. Stream-based Inconsistency Measurement. In International Journal of Approximate Reasoning, 68:68–87. January 2016..

Abstract

Inconsistency measures have been proposed to assess the severity of inconsistencies in knowledge bases of classical logic in a quantitative way. In general, computing the value of inconsistency is a computationally hard task as it is based on the satisfiability problem which is itself NP-complete. In this work, we address the problem of measuring inconsistency in knowledge bases that are accessed in a *stream* of propositional formulæ. That is, the formulæ of a knowledge base cannot be accessed directly but only once through processing of the stream. This work is a first step towards practicable inconsistency measurement for applications such as Linked Open Data, where huge amounts of information is distributed across the web and a direct assessment of the quality or inconsistency of this information is infeasible due to its size. Here we discuss the problem of stream-based inconsistency measurement on classical logic, in order to make use of existing measures for classical logic. However, it turns out that inconsistency measures defined on the notion of minimal inconsistent subsets are usually not apt to be used in the streaming scenario. In order to address this issue, we adapt measures defined on paraconsistent logics and also present a novel inconsistency measure based on the notion of a *hitting set*. We conduct an extensive empirical analysis on the behavior of these different inconsistency measures in the streaming scenario, in terms of runtime, accuracy, and scalability. We conclude that for two of these measures, the stream-based variant of the new inconsistency measure and the stream-based variant of the *contension* inconsistency measure, large-scale inconsistency measurement in streaming scenarios is feasible.

1 INTRODUCTION

Inconsistency measurement is a subfield of Knowledge Representation and Reasoning (KR) that is concerned with the quantitative assessment of the severity of inconsistencies in knowledge bases. Consider the following two knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ formalized in propositional logic:

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\} \qquad \mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

Both knowledge bases are classically inconsistent as for $\mathcal{K}_1$ we have $\{a, \neg a \wedge \neg b\} \models \perp$ and for $\mathcal{K}_2$ we have, e. g., $\{a, \neg a\} \models \perp$. These inconsistencies render the knowledge bases useless for reasoning if one wants to use classical reasoning techniques. In order to make the knowledge bases useful again, one can either use non-monotonic/paraconsistent reasoning techniques (Makinson, 2005; Priest, 1979) or one revises the knowledge bases appropriately to make them consistent (Hansson, 2001). Looking at the knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ one can observe that the *severity* of their inconsistency is different. In $\mathcal{K}_1$, only two out of four formulæ (a and $\neg a \wedge \neg b$) are *participating* in making $\mathcal{K}_1$ inconsistent while for $\mathcal{K}_2$ all formulæ contribute to its inconsistency. Furthermore, for $\mathcal{K}_1$ only two propositions (a and b) are conflicting and using e. g. paraconsistent reasoning one could still infer meaningful statements about c and d. For $\mathcal{K}_2$ no such statement can be made. This leads to the assessment that $\mathcal{K}_2$ should be regarded *more* inconsistent than $\mathcal{K}_1$. Inconsistency measures can be used to quantitatively assess the inconsistency of knowledge bases and to provide a guide for how to repair them. Moreover, they can be used as an analytical tool to assess the quality of knowledge representation. For example, one simple inconsistency measure, see e. g. (Grant and Hunter, 2011), is to take the number of *minimal inconsistent subsets* (MIs) as an indicator for the inconsistency: the more MIs a knowledge base contains, the more inconsistent it is. For $\mathcal{K}_1$ we have then 1 as its inconsistency value and for $\mathcal{K}_2$ we have 2. A lot of different approaches of inconsistency measures and postulates for inconsistency measures have been proposed, mostly for classical propositional logic (Knight, 2001; Hunter, 2002; Hunter and Konieczny, 2004, 2008, 2010; Ma *et al.*, 2009; Mu *et al.*, 2011a,b; Xiao and Ma, 2012; Grant and Hunter, 2011, 2013; Besnard, 2014; McAreavey *et al.*, 2014; Jabbour *et al.*, 2014b), but also for classical first-order logic (Grant and Hunter, 2006, 2008), description logics (Ma *et al.*, 2007; Deng *et al.*, 2007; Qi and Hunter, 2007; Zhou *et al.*, 2009), default logics (Doder *et al.*, 2010), and probabilistic and other weighted logics (Daniel, 2009; Muiño, 2011; Ma *et al.*, 2012; Thimm, 2013b, 2014a; Potyka, 2014; Mu *et al.*, 2014).

Inconsistencies arise easily when many experts share their knowledge in order to construct a joint knowledge base, particularly for large knowledge bases as they appear in, e. g., Semantic Web applications (Sacramento *et al.*, 2012). So far, the field of inconsistency measurement is focused on the problem on what a *reasonable* inconsistency measure is and what properties it should satisfy. In this paper, we consider the *computational* problems of inconsistency measurement, particularly with respect scenarios where the knowledge base can only be processed in a step-by-step fashion, i. e., in *streams*. More precisely, we consider a scenario where, instead of a knowledge base $\mathcal{K}$ we are faced with a stream $\mathcal{S}$ that for any point in time $i \in \mathbb{N}$ gives us a propositional formula $\phi = \mathcal{S}(i)$. The measures we are interested in update for every time step i the currently computed inconsistency value and therefore approximate the actual inconsistency value of $\bigcup_{j=1}^{i}\{\mathcal{S}(j)\}$ with the limiting case $i \to \infty$.

To address the issue of stream-based inconsistency measurement, we present a novel inconsistency measure $\mathcal{I}_{hs}$ that is inspired by the η-inconsistency measure of (Knight, 2002) and is particularly apt to be applied to the streaming scenario. This measure bases on the notion of a *hitting set* which (in our context) is a minimal set of classical interpretations such that every formula of a knowledge base is satisfied by at least one element of the set. We then formalize the problem of stream-based inconsistency measurement, describe desirable properties of stream-based inconsistency measures by relating the problem to the classical setting of inconsistency measurement, and propose specific instantiations for stream-based inconsistency measures. We investigate the properties and the behavior of our new measures both analytically and empirically. For the latter, we conduct an extensive empirical evaluation on artificial data. Our findings show that the stream-based variant of our novel measure, as well as a measure based on paraconsistent logics are suitable in terms of runtime, accuracy, and scalability for the stream-based scenario. In summary, the contributions of this paper are as follows:

1. We present a novel inconsistency measure $\mathcal{I}_{hs}$ based on hitting sets and show how this measure relates to other measures (Section 3).

2. We formalize a theory of inconsistency measurement in streams and relate it to the classical setting of inconsistency measurement (Section 4).

3. We provide a window-based approach for applying classical inconsistency measures to the streaming case and develop specific approaches for some concrete classical measures (Section 5).

4. We conduct an extensive empirical study on the behavior of those inconsistency measures in terms of runtime, accuracy, and scalability. In particular, we show that the stream variants of $\mathcal{I}_{hs}$ and of the *contension* measure $\mathcal{I}_c$ are effective and accurate for measuring inconsistency in the streaming scenario (Section 6).

Additionally, we give necessary preliminaries for propositional logic in Section 2, provide some review of related work in Section 7 and conclude the paper in Section 5. Proofs of technical results can be found in the appendix. This paper extends and revises the previously published paper (Thimm, 2014c) by correcting and extending technical results, providing proofs, and adding further discussion.

2 PRELIMINARIES

Let At be a propositional signature, i.e., a (finite) set of propositions (also called atoms), and let $\mathcal{L}(\mathsf{At})$ the corresponding propositional language constructed using the usual connectives $\wedge$ (*and*), $\vee$ (*or*), and $\neg$ (*negation*).

Definition 1. A knowledge base $\mathcal{K}$ is a finite set of formulæ $\mathcal{K} \subseteq \mathcal{L}(\mathsf{At})$. Let $\mathbb{K}(\mathsf{At})$ be the set of all knowledge bases.

We write $\mathbb{K}$ instead of $\mathbb{K}(\mathsf{At})$ when there is no ambiguity regarding the signature. If X is a formula or a set of formulaæ we write $\mathsf{At}(X)$ to denote the set of propositions appearing in X. Semantics to a propositional language $\mathcal{L}(\mathsf{At})$ is given by *interpretations* and an *interpretation* ω on Atis a function $\omega : \mathsf{At} \rightarrow \{\mathsf{true}, \mathsf{false}\}$. Let $\mathsf{Int}(\mathsf{At})$ denote the set of all interpretations for At. An interpretation ω *satisfies* (or is a *model* of) an atom $a \in \mathsf{At}$, denoted by $\omega \models a$, if and only if $\omega(a) = \mathsf{true}$. For $\omega \in \mathsf{Int}(\mathsf{At})$ and $\phi, \phi' \in \mathcal{L}(\mathsf{At})$ we define

- $\omega \models \neg\phi$ if and only if $\omega \not\models \phi$
- $\omega \models \phi \wedge \phi'$ if and only if $\omega \models \phi$ and $\omega \models \phi'$
- $\omega \models \phi \vee \phi'$ if and only if $\omega \models \phi$ or $\omega \models \phi'$

As an abbreviation we sometimes identify an interpretation ω with its *complete conjunction*, i. e., if $a_1, \ldots, a_n \in \mathsf{At}$ are those propositions that are assigned true by ω and $a_{n+1}, \ldots, a_m \in \mathsf{At}$ are those propositions that are assigned false by ω we identify ω by $a_1 \ldots a_n \overline{a_{n+1}} \ldots \overline{a_m}$ (or any permutation of this). For example, the interpretation ω_1 on $\{a, b, c\}$ with $\omega(a) = \omega(c) =$ true and $\omega(b) =$ false is abbreviated by $a\bar{b}c$.

For $\Phi \subseteq \mathcal{L}(\mathsf{At})$ we also define $\omega \models \Phi$ if and only if $\omega \models \phi$ for every $\phi \in \Phi$. Define furthermore the set of models $\mathsf{Mod}(X) = \{\omega \in \mathsf{Int}(\mathsf{At}) \mid \omega \models X\}$ for every formula or set of formulæ X. Two formulæ or sets of formulæ X and Y are *equivalent*, denoted by $X \equiv Y$, if and only if $\mathsf{Mod}(X) = \mathsf{Mod}(Y)$. Furthermore, two knowledge bases $\mathcal{K}$, $\mathcal{K}'$ are *semi-extensionally equivalent* ($\mathcal{K} \equiv^{\sigma} \mathcal{K}'$) if there is a bijection $\sigma : \mathcal{K} \rightarrow \mathcal{K}'$ such that for all $\alpha \in \mathcal{K}$ we have $\alpha \equiv \sigma(\alpha)$ (Thimm, 2013b). If $\mathsf{Mod}(X) = \emptyset$ we also write $X \models \perp$ and say that X is *inconsistent*. Note that checking $X \not\models \perp$ is an NP-complete problem as it is equivalent to the satisfiability problem SAT (Cook, 1971).

Let $\mathbb{R}_0^+$ be the set of non-negative real numbers. Inconsistency measures are functions $\mathcal{I} : \mathbb{K} \rightarrow \mathbb{R}_0^+ \cup \{\infty\}$ that aim at assessing the severity of the inconsistency in a knowledge base $\mathcal{K}$, cf. (Grant and Hunter, 2011). The basic idea is that the larger the inconsistency in $\mathcal{K}$ the larger the value $\mathcal{I}(\mathcal{K})$. However, inconsistency is a concept that is not easily quantified and there have been a couple of proposals for inconsistency measures so far, in particular for classical propositional logic, see e. g. (Knight, 2001; Hunter, 2002; Hunter and Konieczny, 2004, 2008, 2010; Ma *et al.*, 2009; Mu *et al.*, 2011a,b; Xiao and Ma, 2012; Grant and Hunter, 2011, 2013; Besnard, 2014; McAreavey *et al.*, 2014; Jabbour *et al.*, 2014b). Below we recall some popular measures but we first introduce some necessary notations. Let $\mathcal{K} \in \mathbb{K}$ be some knowledge base.

Definition 2. A set $M \subseteq \mathcal{K}$ is called *minimal inconsistent subset* (MI) of $\mathcal{K}$ if $M \models \perp$ and there is no $M' \subset M$ with $M' \models \perp$. Let $\mathsf{MI}(\mathcal{K})$ be the set of all MIs of $\mathcal{K}$.

Definition 3. A formula $\alpha \in \mathcal{K}$ is called *free formula* of $\mathcal{K}$ if there is no $M \in \mathsf{MI}(\mathcal{K})$ with $\alpha \in M$. Let $\mathsf{Free}(\mathcal{K})$ denote the set of all free formulæ of $\mathcal{K}$.

We adopt the following definition of a (basic) inconsistency measure from (Grant and Hunter, 2011).

Definition 4. A *basic inconsistency measure* is a function $\mathcal{I} : \mathbb{K} \to \mathbb{R}_0^+ \cup \{\infty\}$ that satisfies the following three conditions:

1. $\mathcal{I}(\mathcal{K}) = 0$ if and only if $\mathcal{K}$ is consistent,
2. if $\mathcal{K} \subseteq \mathcal{K}'$ then $\mathcal{I}(\mathcal{K}) \leq \mathcal{I}(\mathcal{K}')$, and
3. for all $\alpha \in \mathsf{Free}(\mathcal{K})$ we have $\mathcal{I}(\mathcal{K}) = \mathcal{I}(\mathcal{K} \setminus \{\alpha\})$.

The first property (also called *consistency*) of a basic inconsistency measure ensures that all consistent knowledge bases receive a minimal inconsistency value and every inconsistent knowledge base receives a positive inconsistency value. The second property (also called *monotony*) states that the value of inconsistency cannot decrease when adding new information. The third property (also called *free formula independence*) states that removing harmless formulæ from a knowledge base—i. e., formulæ that do not contribute to the inconsistency—does not change the value of inconsistency. If $\mathcal{I}$ is a basic inconsistency measure and $\mathcal{K} \in \mathbb{K}$ is a knowledge base we say that $\mathcal{I}(\mathcal{K})$ is the *inconsistency value* of $\mathcal{K}$ wrt. $\mathcal{I}$. In the following we will drop the "basic" and refer to measures satisfying the above three conditions simply as inconsistency measures. For the remainder of this paper we consider the following selection of inconsistency measures: the MI measure $\mathcal{I}_{\mathsf{MI}}$, the MI^c measure $\mathcal{I}_{\mathsf{MI}^c}$, the contension measure $\mathcal{I}_c$, and the η-measure $\mathcal{I}_\eta$, which will be defined below, cf. (Grant and Hunter, 2011; Knight, 2002).

In order to define the contension measure $\mathcal{I}_c$ we need to consider three-valued interpretations for propositional logic (Priest, 1979). A three-valued interpretation v on At is a function $v : \mathsf{At} \to \{T, F, B\}$ where the values T and F correspond to the classical truth values true and false, respectively. The additional truth value B stands for *both* and is meant to represent a conflicting truth value for a proposition. The function v is extended to arbitrary formulæ as shown in Table 2. Then, an interpretation v satisfies a formula α, denoted by $v \models^3 \alpha$ if either $v(\alpha) = T$ or $v(\alpha) = B$.

For defining the η-inconsistency measure (Knight, 2002) we need to consider probability functions P of the form $P : \mathsf{Int}(\mathsf{At}) \to [0, 1]$ with $\sum_{\omega \in \mathsf{Int}(\mathsf{At})} P(\omega) = 1$. Let $\mathcal{P}(\mathsf{At})$ be the set of all those probability functions and for a given probability function $P \in \mathcal{P}(\mathsf{At})$ define the probability of an arbitrary formula α via $P(\alpha) = \sum_{\omega \models \alpha} P(\omega)$.

α	β	$v(\alpha \wedge \beta)$	$v(\alpha \vee \beta)$
T	T	T	T
T	B	B	T
T	F	F	T
B	T	B	T
B	B	B	B
B	F	F	B
F	T	F	T
F	B	F	B
F	F	F	F

α	$v(\neg\alpha)$
T	F
B	B
F	T

Table 1: Truth tables for propositional three-valued logic (Priest, 1979).

Definition 5. Let $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^c}$, $\mathcal{I}_c$, and $\mathcal{I}_\eta$ be defined via

$$\begin{aligned}
\mathcal{I}_{\mathsf{MI}}(\mathcal{K}) &= |\mathsf{MI}(\mathcal{K})| \\
\mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}) &= \sum_{M \in \mathsf{MI}(\mathcal{K})} \frac{1}{|M|} \\
\mathcal{I}_c(\mathcal{K}) &= \min\{|v^{-1}(B)| \mid v \models^3 \mathcal{K}\} \\
\mathcal{I}_\eta(\mathcal{K}) &= 1 - \max\{\xi \mid \exists P \in \mathcal{P}(\mathsf{At}) : \forall \alpha \in \mathcal{K} : P(\alpha) \geq \xi\}
\end{aligned}$$

The measure $\mathcal{I}_{\mathsf{MI}}$ takes the number of minimal inconsistent subsets of a knowledge base as an indicator for the amount of inconsistency: the more minimal inconsistent subsets the more severe the inconsistency. The measure $\mathcal{I}_{\mathsf{MI}^c}$ refines this idea by also taking the size of the minimal inconsistent subsets into account. Here the idea is that larger minimal inconsistent subsets are less severe than smaller minimal inconsistent subsets (the less formulæ are needed to produce an inconsistency the more "obvious" the inconsistency). The measure $\mathcal{I}_c$ considers the set of three-valued models of a knowledge base (which is always non-empty) and uses the minimal number of propositions with conflicting truth value as an indicator for inconsistency. Finally, the measure $\mathcal{I}_\eta$ (which always assigns an inconsistency value between 0 and 1) looks for the maximal probability one can assign to every formula of a knowledge base[1]. All these measures are basic inconsistency measures as defined in Definition 4.

Example 1. For the knowledge bases $\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$ and $\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$ from the introduction we obtain the following inconsistency values.

The knowledge base $\mathcal{K}_1$ contains one minimal inconsistent subset $\{a, \neg a \wedge \neg b\}$, i.e. $\mathsf{MI}(\mathcal{K}_1) = \{\{a, \neg a \wedge \neg b\}\}$, while $\mathcal{K}_2$ contains two minimal inconsistent subsets $\{a, \neg a\}$ and $\{b, \neg b\}$, i.e. $\mathsf{MI}(\mathcal{K}_2) = \{\{a, \neg a\}, \{b, \neg b\}$. This results in $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_1) = 1$ and $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_2) = 2$.

1 Note that we modified the definition of $\mathcal{I}_\eta$ slightly compared to the original definition in order to fit our framework.

As the size of the only minimal inconsistent subset of $\mathcal{K}_1$ is 2 we have $\mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_1) = 1/2$. For $\mathcal{K}_2$ we have two minimal inconsistent subsets each of size 2, resulting in $\mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_2) = 1/2 + 1/2 = 1$.

For the propositional signature $\mathsf{At}_1 = \{a, b, c, d\}$ consider the three-valued interpretation v_1 defined via

$$v_1(a) = B \qquad v_1(b) = F \qquad v_1(c) = T \qquad v_1(d) = T$$

and observe $v_1 \models^3 \mathcal{K}_1$. Note that $|v^{-1}(B)| = 1$ and that there cannot be another v which assigns to fewer propositions the value B. So we have $\mathcal{I}_c(\mathcal{K}_1) = 1$. For the propositional signature $\mathsf{At}_2 = \{a, b\}$ consider the three-valued interpretation v_2 defined via

$$v_2(a) = B \qquad\qquad v_2(b) = B$$

and observe $v_2 \models^3 \mathcal{K}_2$. Note that $|v^{-1}(B)| = 2$ and that there cannot be another v which assigns to fewer propositions the value B. So we have $\mathcal{I}_c(\mathcal{K}_2) = 2$.

For At_1 consider the probability function $P_1 : \mathsf{Int}(\mathsf{At}_1) \to [0,1]$ defined via $P_1(abcd) = 1/2$, $P_1(\overline{a}\overline{b}cd) = 1/2$ and $P_1(\omega) = 0$ for all remaining $\omega \in \mathsf{Int}(\mathsf{At}_1)$. Then we have

$$\begin{aligned}
P_1(a) &= P_1(abcd) = 0.5\\
P_1(b \vee c) &= P_1(abcd) + P_1(\overline{a}\overline{b}cd) = 0.5 + 0.5 = 1\\
P_1(\neg a \wedge \neg b) &= P_1(\overline{a}\overline{b}cd) = 0.5\\
P_1(d) &= P_1(abcd) + P_1(\overline{a}\overline{b}cd) = 0.5 + 0.5 = 1
\end{aligned}$$

and therefore for all $\alpha \in \mathcal{K}_1$ it is $P_1(\alpha) \geq 1/2$. Observe that there cannot be another P which assigns larger probability to all formulas, so we have $\mathcal{I}_\eta(\mathcal{K}_1) = 1 - 1/2 = 1/2$. For At_2 consider the probability function $P_2 : \mathsf{Int}(\mathsf{At}_2) \to [0,1]$ defined via $P_2(ab) = 1/2$, $P_2(\overline{a}\overline{b}) = 1/2$ and $P_2(\omega) = 0$ for all remaining $\omega \in \mathsf{Int}(\mathsf{At}_2)$. Then we have $P_2(a) = P_2(b)$ and also $\mathcal{I}_\eta(\mathcal{K}_2) = 1 - 1/2 = 1/2$.

In summary, these are the inconsistency values obtained for the discussed inconsistent measures:

$$\begin{array}{llll}
\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_1) = 1 & \mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_1) = 1/2 & \mathcal{I}_c(\mathcal{K}_1) = 1 & \mathcal{I}_\eta(\mathcal{K}_1) = 1/2\\
\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_2) = 2 & \mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_2) = 1 & \mathcal{I}_c(\mathcal{K}_2) = 2 & \mathcal{I}_\eta(\mathcal{K}_2) = 1/2
\end{array}$$

Example 2. In the previous example, all considered inconsistency measures agreed that $\mathcal{K}_1$ is not more inconsistent than $\mathcal{K}_2$. While e. g. $\mathcal{I}_\eta$ is indifferent about $\mathcal{K}_1$ and $\mathcal{K}_2$ the measure $\mathcal{I}_{\mathsf{MI}^c}$ evaluates $\mathcal{K}_1$ to be less inconsistent than

$\mathcal{K}_2$. It can also be the case that inconsistency measures behave completely incomparable. Consider the knowledge bases $\mathcal{K}_3$ and $\mathcal{K}_4$ defined via

$$\mathcal{K}_3 = \{a, b, c, d, \neg(a \vee b \vee c \vee d), e, f, g, h, \neg(e \vee f \vee g \vee h)\}$$
$$\mathcal{K}_4 = \{a, \neg a\}$$

Observe that $\mathsf{MI}(\mathcal{K}_3) = \{m_1, m_2\}$ with $m_1 = \{a, b, c, d, \neg(a \vee b \vee c \vee d)\}$ and $m_2 = \{e, f, g, h, \neg(e \vee f \vee g \vee h)\}$, and $\mathsf{MI}(\mathcal{K}_4) = \{m_3\}$ with $m_3 = \{a, \neg a\}$. Then we have

$$\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_3) = |\mathsf{MI}(\mathcal{K}_3)| = 2$$
$$\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_4) = |\mathsf{MI}(\mathcal{K}_4)| = 1$$

but

$$\mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_3) = \frac{1}{|m_1|} + \frac{1}{|m_2|} = \frac{2}{5}$$
$$\mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_4) = \frac{1}{|m_3|} = \frac{1}{2}$$

So $\mathcal{I}_{\mathsf{MI}}$ and $\mathcal{I}_{\mathsf{MI}^c}$ completely disagree on the order of $\mathcal{K}_3$ and $\mathcal{K}_4$.

For a more detailed introduction to inconsistency measures see e.g. (Grant and Hunter, 2006) and for some recent developments see e.g. (Besnard, 2014; Jabbour *et al.*, 2014a; Mu *et al.*, 2014; McAreavey *et al.*, 2014; Jabbour *et al.*, 2014b).

3 AN INCONSISTENCY MEASURE BASED ON HITTING SETS

The basic idea of our novel inconsistency measure $\mathcal{I}_{hs}$ is inspired by the measure $\mathcal{I}_\eta$ which seeks a probability function that maximizes the probability of all formulæ of a knowledge base. Basically, the measure $\mathcal{I}_\eta$ looks for a minimal number of models of parts of the knowledge base and maximizes their probability in order to maximize the probability of the formulæ. By just considering this basic idea we arrive at the notion of a *hitting set* for inconsistent knowledge bases.

Definition 6. A subset $H \subseteq \mathsf{Int}(\mathsf{At})$ is called a *hitting set* of $\mathcal{K}$ if for every $\alpha \in \mathcal{K}$ there is $\omega \in H$ with $\omega \models \alpha$.

Some observations on hitting sets are as follows.

Proposition 1. *Let $\mathcal{K}$ be a knowledge base. The following two statements are equivalent:*

1. *there is no $\phi \in \mathcal{K}$ with $\phi \models \perp$*
2. *there exists a hitting set H of $\mathcal{K}$*

Proposition 2. *Let $\mathcal{K}$ be a knowledge base.*

1. *If H is a hitting set of $\mathcal{K}$ then every H' with $H \subseteq H'$ is a hitting set of $\mathcal{K}$.*
2. *$H = \emptyset$ is a hitting set of $\mathcal{K}$ if and only if $\mathcal{K} = \emptyset$.*
3. *$\mathcal{K}$ is consistent if and only if there is a hitting set H of $\mathcal{K}$ with $|H| = 1$.*
4. *If H is a hitting set of $\mathcal{K}$ then H is a hitting set of every $\mathcal{K}'$ with $\mathcal{K}' \subseteq \mathcal{K}$.*

We then define the measure $\mathcal{I}_{hs}$ as the minimal cardinality of a hitting set of the knowledge base minus one.

Definition 7. The function $\mathcal{I}_{hs} : \mathbb{K} \to \mathbb{R}_0^+ \cup \{\infty\}$ is defined via

$$\mathcal{I}_{hs}(\mathcal{K}) = \min\{|H| \mid H \text{ is a hitting set of } \mathcal{K}\} - 1$$

with $\min \emptyset = \infty$ for every $\mathcal{K} \in \mathbb{K} \setminus \{\emptyset\}$ and $\mathcal{I}_{hs}(\emptyset) = 0$.

Note that, following Proposition 1, we have $\mathcal{I}_{hs}(\mathcal{K}) = \infty$ if and only if $\mathcal{K}$ contains a contradictory formula (e. g. $a \wedge \neg a$). Observe also that we need a case differentiation for $\mathcal{K} = \emptyset$ as $\emptyset$ has also the hitting set $\emptyset$, see Proposition 2.

Example 3. We continue Example 1 and consider $\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$ and $\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$. Let $H_1 \subseteq \mathsf{Int}(\mathsf{At})$ be defined via $H_1 = \{abcd, \overline{a}\overline{b}cd\}$. Observe that for both $\mathcal{K}_1$ and $\mathcal{K}_2$ we have that H_1 is a hitting set, i. e., every formula of the knowledge base is satisfied by at least one interpretation of H_1. Furthermore, H_1 is also a minimal hitting set (with respect to set cardinality) as, e. g., for $\mathcal{K}_2$ the two formulas a and $\neg a$ require at least two different interpretations to be satisfied. Therefore we have $\mathcal{I}_{hs}(\mathcal{K}_1) = \mathcal{I}_{hs}(\mathcal{K}_2) = 1$.

Example 4. Consider the knowledge base $\mathcal{K}_5$ defined via

$$\mathcal{K}_5 = \{a \vee d, a \wedge b \wedge c, b, \neg b \vee \neg a, a \wedge b \wedge \neg c, a \wedge \neg b \wedge c\}$$

Then $H_2 = \{abcd, a\overline{b}cd, ab\overline{c}d\} \subseteq \mathsf{Int}(\mathsf{At})$ is a minimal hitting set for $\mathcal{K}_5$ and therefore $\mathcal{I}_{hs}(\mathcal{K}_5) = 2$.

As the following result shows, $\mathcal{I}_{hs}$ is indeed a suitable inconsistency measure.

Proposition 3. *The function $\mathcal{I}_{hs}$ is a (basic) inconsistency measure.*

The result below shows that $\mathcal{I}_{hs}$ also behaves well with some more properties mentioned in the literature (Hunter and Konieczny, 2010; Thimm, 2013b).

Proposition 4. *The measure $\mathcal{I}_{hs}$ satisfies the following properties:*

- *If* $\alpha \in \mathcal{K}$ *is such that* $\mathsf{At}(\alpha) \cap \mathsf{At}(\mathcal{K} \setminus \{\alpha\}) = \emptyset$ *then* $\mathcal{I}_{hs}(\mathcal{K}) = \mathcal{I}_{hs}(\mathcal{K} \setminus \{\alpha\})$ *(*safe formula independence*).*
- *If* $\mathcal{K} \equiv^{\sigma} \mathcal{K}'$ *then* $\mathcal{I}_{hs}(\mathcal{K}) = \mathcal{I}_{hs}(\mathcal{K}')$ *(*irrelevance of syntax*).*
- *If* $\alpha \models \beta$ *and* $\alpha \not\models \perp$ *then* $\mathcal{I}_{hs}(\mathcal{K} \cup \{\alpha\}) \geq \mathcal{I}_{hs}(\mathcal{K} \cup \{\beta\})$ *(*dominance*).*

However, $\mathcal{I}_{hs}$ is incompatible with some other properties such as *super-additivity* and *MinInc-separability* (Hunter and Konieczny, 2010).

Example 5. A measure $\mathcal{I}$ satisfies *super-additivity* if for $\mathcal{K} \cap \mathcal{K}' = \emptyset$ we have $\mathcal{I}(\mathcal{K} \cup \mathcal{K}') \geq \mathcal{I}(\mathcal{K}) + \mathcal{I}(\mathcal{K}')$. A measure $\mathcal{I}$ satisfies *MinInc-separability* if $\mathsf{MI}(\mathcal{K} \cup \mathcal{K}') = \mathsf{MI}(\mathcal{K}) \cup \mathsf{MI}(\mathcal{K}')$ and $\mathsf{MI}(\mathcal{K}) \cap \mathsf{MI}(\mathcal{K}') = \emptyset$ implies $\mathcal{I}(\mathcal{K} \cup \mathcal{K}') = \mathcal{I}(\mathcal{K}) + \mathcal{I}(\mathcal{K}')$. Consider

$$\mathcal{K}_4 = \{a, \neg a\}$$
$$\mathcal{K}_6 = \{b, \neg b\}$$

Then we have $\mathcal{I}_{hs}(\mathcal{K}_4) = \mathcal{I}_{hs}(\mathcal{K}_6) = 1$. However, we also have $\mathcal{I}_{hs}(\mathcal{K}_4 \cup \mathcal{K}_6) = 1$ as $\{ab, \overline{a}\overline{b}\}$ is a hitting set of $\mathcal{K}_4 \cup \mathcal{K}_6$. It follows that $\mathcal{I}_{hs}$ violates *super-additivity*. Furthermore, observe that $\mathsf{MI}(\mathcal{K}_4) = \{\mathcal{K}_4\}$ and $\mathsf{MI}(\mathcal{K}_6) = \{\mathcal{K}_6\}$ and therefore $\mathsf{MI}(\mathcal{K}_4 \cup \mathcal{K}_6) = \mathsf{MI}(\mathcal{K}_4) \cup \mathsf{MI}(\mathcal{K}_6)$ and $\mathsf{MI}(\mathcal{K}_4) \cap \mathsf{MI}(\mathcal{K}_6) = \emptyset$. It follows that $\mathcal{I}_{hs}$ also violates *MinInc-separability*.

The measure $\mathcal{I}_{hs}$ can also be nicely characterized by a consistent *partitioning* of a knowledge base.

Definition 8. A set $\Phi = \{\Phi_1, \ldots, \Phi_n\}$ with $\Phi_1 \cup \ldots \cup \Phi_n = \mathcal{K}$ and $\Phi_i \cap \Phi_j = \emptyset$ for $i, j = 1, \ldots, n$, $i \neq j$, is called a *partitioning* of $\mathcal{K}$. A partitioning $\Phi = \{\Phi_1, \ldots, \Phi_n\}$ is consistent if $\Phi_i \not\models \perp$ for $i = 1, \ldots, n$.

Proposition 5. *For every knowledge base* $\mathcal{K}$

$$\mathcal{I}_{hs}(\mathcal{K}) = \min\{|\Phi| \mid \Phi \textit{ is a consistent partitioning of } \mathcal{K}\} - 1$$

with $\min \emptyset = \infty$ *for every* $\mathcal{K} \in \mathbb{K} \setminus \{\emptyset\}$ *and* $\mathcal{I}_{hs}(\emptyset) = 0$*.*

As $\mathcal{I}_{hs}$ is inspired by $\mathcal{I}_{\eta}$ we go on by comparing these two measures.

Proposition 6. *Let* $\mathcal{K}$ *be a knowledge base. If* $\infty > \mathcal{I}_{hs}(\mathcal{K}) > 0$ *then*

$$\mathcal{I}_{\eta}(\mathcal{K}) \leq 1 - \frac{1}{\mathcal{I}_{hs}(\mathcal{K}) + 1}$$

Note that for $\mathcal{I}_{hs}(\mathcal{K}) = 0$ we always have $\mathcal{I}_{\eta}(\mathcal{K}) = 0$ as well, as both are basic inconsistency measures. Furthermore, $\mathcal{I}_{hs}(\mathcal{K}) = \infty$ is equivalent to the existence of a $\phi \in \mathcal{K}$ with $\phi \models \perp$, cf. Proposition 1, which is equivalent to $\mathcal{I}_{\eta}(\mathcal{K}) = 1$ (Knight, 2002). Although Proposition 6 describes a loose relationship between $\mathcal{I}_\eta$ and $\mathcal{I}_{hs}$ both measures are in general different as we will see below.

We say that an inconsistency measure $\mathcal{I}_1$ *is subsumed by* an inconsistency measure $\mathcal{I}_2$, denoted by $\mathcal{I}_1 \sqsubseteq \mathcal{I}_2$, if the order on knowledge bases imposed by $\mathcal{I}_1$ is a subset of the order imposed by $\mathcal{I}_2$. More formally, $\mathcal{I}_1 \sqsubseteq \mathcal{I}_2$ if and only if $\mathcal{I}_1(\mathcal{K}) < \mathcal{I}_1(\mathcal{K}')$ implies $\mathcal{I}_2(\mathcal{K}) < \mathcal{I}_2(\mathcal{K}')$ for all $\mathcal{K}, \mathcal{K}' \in \mathbb{K}$. Two inconsistency measures $\mathcal{I}_1$ and $\mathcal{I}_2$ are *equivalent*, denoted by $\mathcal{I}_1 \simeq \mathcal{I}_2$, if and only if $\mathcal{I}_1 \sqsubseteq \mathcal{I}_2$ and $\mathcal{I}_2 \sqsubseteq \mathcal{I}_1$.

It turns out that $\mathcal{I}_{hs}$ is neither equivalent nor is subsumed by any of the previously discussed inconsistency measures[2].

Proposition 7. *There is no subsumption relation between $\mathcal{I}_{hs}$ and any $\mathcal{I} \in \{\mathcal{I}_{\mathsf{MI}}, \mathcal{I}_{\mathsf{MI}^c}, \mathcal{I}_c, \mathcal{I}_\eta\}$.*

Corollary 1. *$\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_{\mathsf{MI}^c}$, $\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_c$, and $\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_\eta$.*

Example 6. Consider the knowledge bases $\mathcal{K}_7$ and $\mathcal{K}_8$ given as

$$\mathcal{K}_7 = \{a \wedge b \wedge c, \neg a \wedge \neg b \wedge \neg c\}$$
$$\mathcal{K}_8 = \{a \wedge b, \neg a \wedge b, a \wedge \neg b\}$$

Then we have e.g. $\mathcal{I}_{hs}(\mathcal{K}_7) = 2 < 3 = \mathcal{I}_{hs}(\mathcal{K}_8)$ but $\mathcal{I}_c(\mathcal{K}_7) = 3 > 2 = \mathcal{I}_c(\mathcal{K}_8)$.

4 INCONSISTENCY MEASUREMENT IN STREAMS

In the following, we introduce and formalize the problem of inconsistency measurement in streams of propositional formulaæ. The goal of this formalization is to obtain stream-based inconsistency measures that approximate given inconsistency measures when the latter would have been applied to the knowledge base as a whole. We first formalize this setting and, afterwards, provide concrete approaches for some inconsistency measures.

We use a very simple formalization of a stream that is sufficient for our needs.

Definition 9. A *propositional stream* $\mathcal{S}$ is a function $\mathcal{S} : \mathbb{N} \rightarrow \mathcal{L}(\mathsf{At})$. Let $\mathbb{S}$ be the set of all propositional streams.

A propositional stream models a sequence of propositional formulæ. On a wider scope, a propositional stream can also be interpreted as a very general abstraction of the output of a linked open data crawler (such as LDSpider (Isele *et al.*, 2010)) that crawls knowledge formalized as RDF (*Resource Description Framework*) from the web, possibly enriched with OWL semantics to have a well-defined notion of consistency. For notational

2 Note that this result corrects Corollary 1 from (Thimm, 2014c) where $\mathcal{I}_{hs} \sqsubseteq \mathcal{I}_\eta$ was claimed.

convenience, we write a propositional stream $\mathcal{S}$ with $\mathcal{S}(0) = \phi_0$, $\mathcal{S}(1) = \phi_1$, $\mathcal{S}(2) = \phi_2, \ldots$ also as a tuple $\mathcal{S} = \langle \phi_0, \phi_1, \phi_2, \ldots \rangle$

Using the abstraction of a propositional stream, we can also model large knowledge bases by propositional streams that indefinitely repeat the formulæ of the knowledge base. For that, we assume for a knowledge base $\mathcal{K} = \{\phi_1, \ldots, \phi_n\}$ the existence of a *canonical enumeration* $\mathcal{K}^c = \langle \phi_1, \ldots, \phi_n \rangle$ of the elements of $\mathcal{K}$. This enumeration can be arbitrary and has no specific meaning other than to enumerate the elements in an unambiguous way.

Definition 10. Let $\mathcal{K}$ be a knowledge base and $\mathcal{K}^c = \langle \phi_1, \ldots, \phi_n \rangle$ its canonical enumeration. The *$\mathcal{K}$-stream* $\mathcal{S}_\mathcal{K}$ is defined as $\mathcal{S}_\mathcal{K}(i) = \phi_{(i \bmod n)+1}$ for all $i \in \mathbb{N}$.

Using $\mathcal{K}$-streams we can formalize the desired behavior of stream-based inconsistency measures as follows. Given a $\mathcal{K}$-stream $\mathcal{S}_\mathcal{K}$ and an inconsistency measure $\mathcal{I}$ we aim at defining a measure $\mathcal{J}_\mathcal{I}$ that processes the elements of $\mathcal{S}_\mathcal{K}$ one by one and approximates (or converges to) $\mathcal{I}(\mathcal{K})$.

Definition 11. A *stream-based inconsistency measure* $\mathcal{J}$ is a function $\mathcal{J} : \mathbb{S} \times \mathbb{N} \to \mathbb{R}_0^+ \cup \{\infty\}$.

Definition 12. Let $\mathcal{I}$ be an inconsistency measure and $\mathcal{J}$ a stream-based inconsistency measure. Then $\mathcal{J}$ *approximates* (or *is an approximation of*) $\mathcal{I}$ if for all $\mathcal{K} \in \mathbb{K}$ we have $\lim_{i\to\infty} \mathcal{J}(\mathcal{S}_\mathcal{K}, i) = \mathcal{I}(\mathcal{K})$.

A stream-based inconsistency measure $\mathcal{J}$ is supposed to maintain some state information (which is hidden in the formal definition) that is updated when processing the i-th element of a propositional stream. For $i \in \mathbb{N}$ we say that $\mathcal{J}(\mathcal{S}, i)$ is the *inconsistency value* of $\mathcal{S}$ wrt. $\mathcal{J}$ at time point i. We also require that $\mathcal{J}$ is not able to process formulas from the future, i. e., the value $\mathcal{J}(\mathcal{S}, i)$ is independent of every value $\mathcal{S}(j)$ for $j > i$. More formally:

Definition 13. A stream-based inconsistency measure $\mathcal{J}$ is *future-ignorant* if and only if for all $\mathcal{S}, \mathcal{S}' \in \mathbb{S}$, if $\mathcal{S}(i) = \mathcal{S}'(i)$ for all $i = 0, \ldots, n$ then $\mathcal{J}(\mathcal{S}, n) = \mathcal{J}(\mathcal{S}', n)$.

In the following, we only consider future-ignorant stream-based inconsistency measures.

5 STREAM-BASED INCONSISTENCY MEASURES

In this section we develop concrete approaches for adopting classical inconsistency measures, including the $\mathcal{I}_{hs}$ measure developed above, to the streaming scenario. First, we present an approach based on considering a window on the stream at any time point $i \in \mathbb{N}$. Second, we provide approximation algorithms for both $\mathcal{I}_{hs}$ and $\mathcal{I}_c$ that use concepts of the programming paradigms of *simulated annealing* and *evolutionary algorithms*.

5.1 *A Window-based Approach for Stream-based Inconsistency Measures*

The simplest form of implementing a stream-based variant of any algorithm or function is to use a window-based approach, i. e., to consider at any time point a specific excerpt from the stream and apply the original algorithm or function on this excerpt, cf. (Beck *et al.*, 2015). This approach gives us for each time point $i \in \mathbb{N}$ the inconsistency value of the considered excerpt. In order to not dismiss the inconsistency value determined at time point i in time point $i+1$, we aggregate the newly determined inconsistency value at time point $i+1$ with the one from the previous step using an aggregation function.

Definition 14. An *aggregation function* g is a function $g : (\mathbb{R}_0^+ \cup \{\infty\}) \times (\mathbb{R}_0^+ \cup \{\infty\}) \to \mathbb{R}_0^+ \cup \{\infty\}$ with

1. $g(x,y) \in [\min\{x,y\}, \max\{x,y\}]$ for all $x, y \in \mathbb{R}_0^+$,
2. $g(x,\infty) \geq x$ for all $x \in \mathbb{R}_0^+$,
3. $g(\infty,y) \geq y$ for all $y \in \mathbb{R}_0^+$, and
4. $g(\infty,\infty) = \infty$.

Possible aggregation functions are, e. g., the maximum function max or a smoothing function $g_\alpha(x,y) = \alpha x + (1-\alpha)y$ for some $\alpha \in [0,1]$ (for every $x, y \in \mathbb{R}_0^+ \cup \{\infty\}$).[3]

For any propositional stream $\mathcal{S}$ let $\mathcal{S}^{i,j}$ (for $i \leq j$) be the knowledge base obtained by taking the formulæ from $\mathcal{S}$ between positions i and j, i. e., $\mathcal{S}^{i,j} = \{\mathcal{S}(i), \ldots, \mathcal{S}(j)\}$.

Definition 15. Let $\mathcal{I}$ be an inconsistency measure, $w \in \mathbb{N} \cup \{\infty\}$, and g an aggregation function. We define the *window-based inconsistency measure* $\mathcal{J}_{\mathcal{I}}^{w,g} : \mathsf{S} \times \mathbb{N} \to \mathbb{R}_0^+ \cup \{\infty\}$ via

$$\mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S},i) = \begin{cases} \mathcal{I}(\{\mathcal{S}(0)\}) & \text{if } i = 0 \\ g(\mathcal{I}(\mathcal{S}^{\max\{0,i-w+1\},i}), \mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S},i-1)) & \text{otherwise} \end{cases}$$

for every $\mathcal{S}$ and $i \in \mathbb{N}$.[4]

The intuition behind the window-based inconsistency measure $\mathcal{J}_{\mathcal{I}}^{w,g}$ is as follows. At a specific time point i the current inconsistency value $\mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S},i)$ is determined by, first, determining the inconsistency value of the knowledge base obtained from joining the previously encountered $w \in \mathbb{N}$ formulas, and then aggregating this value with the previously determined value $\mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S},i-1)$. If $w = \infty$, the value $\mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S},i)$ is determined by aggregating

3 With $\max(x,\infty) = \max(\infty,x) = \max(\infty,\infty) = \infty$, $g_\alpha(x,\infty) = g_\alpha(\infty,x) = g_\alpha(\infty,\infty) = \infty$ for $x \in \mathbb{R}_0^+$.

4 For $w = \infty$ we define $\max\{0, -\infty\} = 0$

the inconsistency value of the union of all but the last encountered formula with the inconsistency value of the union of all encountered formulas. Observe that $\mathcal{J}_{\mathcal{I}}^{w,g}$ is indeed a future-ignorant stream-based inconsistency measure.

Example 7. Consider the propositional stream $\mathcal{S}_1$ given via

$$\mathcal{S}_1 = \langle a \wedge b, \neg a, \neg b, a \vee b, \neg b \wedge \neg a, \ldots \rangle$$

All further elements of $\mathcal{S}_1$ are unimportant for this example. Consider further the inconsistency measure $\mathcal{I}_{\mathsf{MI}}$, the aggregation function $g_{0.7}$ (smoothing function for $\alpha = 0.7$), and the window size 3. We consider the first four timepoints in the evaluation of $\mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}$.

- At timepoint $i = 0$ we obtain

$$\mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}(\mathcal{S}_1, 0) = \mathcal{I}_{\mathsf{MI}}(\{\mathcal{S}_1(0)\}) = \mathcal{I}_{\mathsf{MI}}(\{a \wedge b\}) = 0$$

- For $i = 1$ note that $\mathcal{S}^{\max\{0,i-w+1\},i} = \mathcal{S}^{\max\{0,-1\},1} = \mathcal{S}^{0,1} = \{a \wedge b, \neg a\}$ and we have

$$\begin{aligned}\mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}(\mathcal{S}_1, 1) &= g_{0.7}(\mathcal{I}_{\mathsf{MI}}(\mathcal{S}_1^{\max\{0,1-w\},1}), \mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}(\mathcal{S}_1, 0))\\ &= g_{0.7}(\mathcal{I}_{\mathsf{MI}}(\{a \wedge b, \neg a\}), 0)\\ &= 0.7 \cdot 1 + (1 - 0.7) \cdot 0 = 0.7\end{aligned}$$

- For $i = 2$ we have

$$\begin{aligned}\mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}(\mathcal{S}_1, 2) &= g_{0.7}(\mathcal{I}_{\mathsf{MI}}(\mathcal{S}_1^{\max\{0,2-w\},2}), \mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}(\mathcal{S}_1, 1))\\ &= g_{0.7}(\mathcal{I}_{\mathsf{MI}}(\{a \wedge b, \neg a, \neg b\}), 0.7)\\ &= 0.7 \cdot 2 + (1 - 0.7) \cdot 0.7 = 1.61\end{aligned}$$

- For $i = 3$ we have

$$\begin{aligned}\mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}(\mathcal{S}_1, 3) &= g_{0.7}(\mathcal{I}_{\mathsf{MI}}(\mathcal{S}_1^{\max\{0,3-w+1\},3}), \mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{3,g_{0.7}}(\mathcal{S}_1, 2))\\ &= g_{0.7}(\mathcal{I}_{\mathsf{MI}}(\{\neg a, \neg b, a \vee b\}), 1.61)\\ &= 0.7 \cdot 1 + (1 - 0.7) \cdot 1.61 = 1.183\end{aligned}$$

Some observations on the properties of $\mathcal{J}_{\mathcal{I}}^{w,g}$ are as follows.

Proposition 8. *Let $\mathcal{I}$ be an inconsistency measure, $w \in \mathbb{N} \cup \{\infty\}$, and g an aggregation function.*

1. *If w is finite then $\mathcal{J}_{\mathcal{I}}^{w,g}$ is* not *an approximation of $\mathcal{I}$.*
2. *If $w = \infty$ and $g(x, y) \geq (x + y)/2$ then $\mathcal{J}_{\mathcal{I}}^{w,g}$ is an approximation of $\mathcal{I}$.*

3. $\mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S}_{\mathcal{K}}, i) \leq \mathcal{I}(\mathcal{K})$ *for every* $\mathcal{K} \in \mathbb{K}$ *and* $i \in \mathbb{N}$.

As can be seen from Example 7 and item 1. of Proposition 8, the main issue with the window-based approach to measuring inconsistency in streams is that only local information (wrt. the current window) can be used to determine the inconsistency value. If, for example, there is a minimal inconsistent subset not covered by any window—as $\{\mathcal{S}_1(0), \mathcal{S}_1(4)\} = \{a \wedge b, \neg b \wedge \neg a\}$ in Example 7—the inconsistency value obtained by the window-based approach will always be an underestimation of the actual inconsistency value, cf. item 2 of Proposition 8.

5.2 *Stream-based Approximation Algorithms for $\mathcal{I}_{hs}$ and $\mathcal{I}_c$*

The approximation algorithms for $\mathcal{I}_{hs}$ and $\mathcal{I}_c$ that are presented in this subsection are using concepts of the programming paradigms of *simulated annealing* and *evolutionary algorithms*, which are both approaches to solve non-convex optimization problems (Lawrence, 1987). Let $f : X \to \mathbb{R}$ be some function that has to be maximized. The basic idea of evolutionary algorithms is to maintain a population of domain elements $X_p \subseteq X$. In each iteration step a subset of X_p with maximal values wrt. f is selected and the rest discarded. From the selected set, new domain elements are generated by *crossover* (combing two or more of the selected domain elements) and *mutation* (random alteration of a selected domain element). This process is repeated until some convergence criterium is satisfied. In the best case the population converges to the global maximum of f. Simulated annealing works roughly as follows. In the beginning, a single domain element $x \in X$ is selected at random. In each iteration a random choice is made whether to locally improve x (select an $x' \in X$ in the vicinity that has larger value wrt. f than x) or to "jump" to a different part in X. The probability of jumping decreases with the number of iterations (this feature is also called *cooling*) and the algorithm stops in some local maximum, which is, in the best case, also the global maximum.

The basic idea for the stream-based approximation of $\mathcal{I}_{hs}$ is as follows. At any processing step we maintain a candidate set $C \in 2^{\mathsf{Int(At)}}$ (initialized with the empty set) that approximates a hitting set of the underlying knowledge base. At the beginning of a processing step we make a random choice (with decreasing probability the more formulæ we already encountered) whether to remove some element of C. This action ensures that C does not contain superfluous elements and mirrors the cooling step in simulated annealing. Afterwards we check whether there is still an interpretation in C that satisfies the currently encountered formula. If this is not the case we add some random model of the formula to C (as in the mutation step of evolutionary algorithms). Finally, we update the previously computed inconsistency value with $|C| - 1$, taking also some aggregation function g (as for the window-based approach) into account. In order to increase the probability of successfully finding a minimal hitting set we do not maintain

Algorithm 1 $\texttt{hs-stream}^{m,g,f}(\mathcal{S}, i)$

$currentValue = 0$
$Cand = \{\emptyset, \ldots, \emptyset\}$
$N = 0$
for all $j = 0, \ldots, i$ **do**
    $currentValue = \texttt{update}_{hs}^{m,g,f}(\mathcal{S}(j))$
return $currentValue$

Algorithm 2 $\texttt{update}_{hs}^{m,g,f}(form)$

$N = N + 1$
if $form \equiv \perp$ **then**
    $currentValue = \infty$
if $currentValue = \infty$ **then**
    return $currentValue$
$newValue = \infty$
for all $C \in Cand$ **do**
    $rand \in [0, 1]$
    if $rand < f(N)$ **then**
        Remove some random ω from C
    if $\neg\exists \omega \in C : \omega \models form$ **then**
        Add random $\omega \in \mathsf{Mod}(form)$ to C
    $newValue = min(newValue, (|C| - 1))$
$currentValue = g(newValue, currentValue)$
return $currentValue$

a single candidate set C but a (multi-)set $Cand = \{C_1, \ldots, C_m\}$ (as in evolutionary algorithms) for some previously specified parameter $m \in \mathbb{N}$ and use the minimum size of these candidate hitting sets.

We call a function $f : \mathbb{N} \to [0, 1]$ *null-bound* if there is $k > 0$ such that f is strictly decreasing on $\{0, \ldots, k\}$ and $f(i) = 0$ for all $i > k$.

Definition 16. Let $m \in \mathbb{N}$, g an aggregation function, and $f : \mathbb{N} \to [0, 1]$ null-bound. We define the approximation algorithm $\mathcal{J}_{hs}^{m,g,f}$ via[5]

$$\mathcal{J}_{hs}^{m,g,f}(\mathcal{S}, i) = \texttt{hs-stream}^{m,g,f}(\mathcal{S}, i)$$

for every $\mathcal{S}$ and $i \in \mathbb{N}$. The algorithm $\texttt{hs-stream}^{m,g,f}(\mathcal{S}, i)$ is given in Algorithm 1 and its subroutine $\texttt{update}_{hs}^{m,g,f}$ is depicted in Algorithm 2.

At the first call of the algorithm $\texttt{hs-stream}^{m,g,f}$ the value of $currentValue$ (which contains the currently estimated inconsistency value) is initialized to 0, the (multi-)set $Cand \subseteq 2^{\mathsf{Int(At)}}$ (which contains a population of candidate

5 Note that $\mathcal{J}_{hs}^{m,g,f}$ is not strictly a stream-based inconsistency measure (as a mathematical function) according to Definition 11 as it is a randomized algorithm.

hitting sets) is initialized with m empty sets, and N is initialized with 0. The function f can be any null-bound function and ensures that every candidate C reaches some stable result eventually. The parameter m increases the probability that at least one of the candidate hitting sets attains the global optimum of a minimal hitting set. The algorithm $\texttt{hs-stream}^{m,g,f}$ then repeatedly calls the subroutine $\texttt{update}_{hs}^{m,g,f}$ that updates the inconsistency value with each formula in the stream, up to the i-th formula.

In order to address the special case of contradictory formulas, lines 2–5 of $\texttt{update}_{hs}^{m,g,f}$ ensure that *currentValue* and thus the inconsistency value is set to ∞.

Example 8. We continue Example 7 and consider again the propositional stream $\mathcal{S}_1$ given via

$$\mathcal{S}_1 = \langle a \wedge b, \neg a, \neg b, a \vee b, \neg b \wedge \neg a, \ldots \rangle$$

We consider $m = 1$ (only one candidate hitting set is maintained) and use the maximum function max as our aggregation function ($g = \max$). Let $f = f_0$ be defined as $f_0(n) = 1/(n+1)$ for $n \in \{0, \ldots, 10\}$ and $f_0(n) = 0$ for $n > 10$. We consider the first four timepoints in the evaluation of $\mathcal{J}_{hs}^{1,\max,f_1}$.

- For $i = 0$, we first initialize $currentValue = 0$, $Cand = \{\emptyset\}$, and $N = 0$ in Algorithm 1, and then set $N = 1$ in line 1 of Algorithm 2. Lines 2–5 in Algorithm 2 are skipped as we do not have any contradictory formulas in $\mathcal{S}_1$. In line 6 we set $newValue = \infty$. In line 7 we select $C_1 = \emptyset$. Suppose in line 8 we determine $rand = 0.7$. As $f_0(1) = 1/2$ we do not execute line 10. As there is no $\omega \in C_1$ that satisfies $form = \mathcal{S}_1(0) = a \wedge b$ we add some model, e. g., ab, to C_1 in line 12. We set $newValue = 0$ in line 13 and $currentValue = \max(0, 0) = 0$ in line 14.

- For $i = 1$, suppose in line 8 of Algorithm 2 we determine $rand = 0.4$. As $f_0(2) = 1/3$ we do not execute line 10. As there is no $\omega \in C_1 = \{ab\}$ that satisfies $form = \mathcal{S}_1(1) = \neg a$ we add some model, e. g., $\bar{a}\bar{b}$, to C_1 in line 12 (now we have $C_1 = \{ab, \bar{a}\bar{b}\}$). We set $newValue = 1$ in line 13 and $currentValue = \max(1, 0) = 1$ in line 14.

- For $i = 2$, suppose in line 8 of Algorithm 2 we determine $rand = 0.5$. As $f_0(3) = 1/4$ we do not execute line 10. Note that there is $\omega \in C_1$ with $\omega \models form = \mathcal{S}_1(2) = \neg b$ ($\omega = \bar{a}\bar{b}$). Therefore we skip line 12 and set $newValue = 1$ in line 13 and $currentValue = \max(1, 1) = 1$ in line 14.

- For $i = 3$, suppose in line 8 of Algorithm 2 we determine $rand = 0.1$. As $f_0(3) = 1/5 > rand$ we execute line 10 and remove ab from C_1. As there is no $\omega \in C_1$ that satisfies $form = \mathcal{S}_1(3) = a \vee b$ we add some model, e. g., $a\bar{b}$, to C_1 in line 12. We set $newValue = 1$ in line 13 and $currentValue = \max(1, 1) = 1$ in line 14.

Algorithm 3 $\texttt{c-stream}^{m,g,f}(\mathcal{S}, i)$

$currentValue = 0$
$Cand = \{v_1, \ldots, v_m\}$
$N = 0$
for all $j = 0, \ldots, i$ **do**
    $currentValue = \texttt{update}_c^{m,g,f}(\mathcal{S}(j))$
return $currentValue$

As $\mathcal{J}_{hs}^{m,g,f}$ is a random process we cannot show that $\mathcal{J}_{hs}^{m,g,f}$ is an approximation of $\mathcal{I}_{hs}$ in the general case. However, we can give the following result.

Proposition 9. *For every $p \in [0, 1)$, g some aggregation function with $g(x, y) \geq (x + y)/2$, $f : \mathbb{N} \to [0, 1]$ a null-bound function, and $\mathcal{K} \in \mathbb{K}$ there is $m \in \mathbb{N}$ such that with probability greater or equal p it is the case that $\lim_{i \to \infty} \mathcal{J}_{hs}^{m,g,f}(\mathcal{S}_{\mathcal{K}}, i) = \mathcal{I}_{hs}(\mathcal{K})$.*

This result states that $\mathcal{J}_{hs}^{m,g,f}$ indeed approximates $\mathcal{I}_{hs}$ if we choose the number of populations large enough. In the next section we will provide some empirical evidence that even for small values of m results are satisfactory. As for the runtime, note that in lines 2 and 11 of Algorithm 2 an FNP-complete problem is solved (determining some model of a propositional formula). However, under the reasonable assumption that formulas are usually quite small compared to the size of the whole knowledge base the impact of this step is negligible.

Both Definition 16 and Algorithms 1 and 2 can be modified slightly in order to approximate $\mathcal{I}_c$ instead of $\mathcal{I}_{hs}$, yielding a new measure $\mathcal{J}_c^{m,g,f}$.

Definition 17. Let $m \in \mathbb{N}$, g an aggregation function, and $f : \mathbb{N} \to [0, 1]$ some null-bound function. We define the approximation algorithm $\mathcal{J}_c^{m,g,f}$ via[6]

$$\mathcal{J}_c^{m,g,f}(\mathcal{S}, i) = \texttt{c-stream}^{m,g,f}(\mathcal{S}, i)$$

for every $\mathcal{S}$ and $i \in \mathbb{N}$. The algorithm $\texttt{c-stream}^{m,g,f}(\mathcal{S}, i)$ is given in Algorithm 3 and its subroutine $\texttt{update}_c^{m,g,f}$ is depicted in Algorithm 4.

In $\texttt{c-stream}^{m,g,f}(\mathcal{S}, i)$ and $\texttt{update}_c^{m,g,f}$, the set of candidates $Cand$ contains three-valued interpretations instead of sets of classical interpretations (initialized with randomly chosen interpretations $v_1, \ldots, v_m$ with $v_i^{-1}(B) = \emptyset$ for $i = 1, \ldots, m$). In line 6 of $\texttt{update}_c^{m,g,f}$, we flip some arbitrary proposition from B to T or F. Similarly, in lines 8–13 of $\texttt{update}_c^{m,g,f}$ we flip some

6 Note that $\mathcal{J}_c^{m,g,f}$ is not strictly a stream-based inconsistency measure (as a mathematical function) according to Definition 11 as it is a randomized algorithm.

Algorithm 4 `update`$_c^{m,g,f}(form)$

$N = N + 1$
$newValue = \infty$
for all $v \in Cand$ **do**
    rand $\in [0,1]$
    if $rand < f(N)$ and $v^{-1}(B) \neq \emptyset$ **then**
        Set random proposition in v from B to T or F
    if $v \not\models^3 form$ **then**
        Select random $\omega \in \mathsf{Mod}(form)$
        for all $p \in \mathsf{At}$ **do**
            if $\omega \models p$ and $v(p) = F$ **then**
                $v(p) = B$
            if $\omega \not\models p$ and $v(p) = T$ **then**
                $v(p) = B$
    $newValue = min(newValue, |v^{-1}(B)|)$
$currentValue = g(newValue, currentValue)$
return $currentValue$

propositions to B in order to satisfy the new formula. Finally, the inconsistency value is determined by taking the number of B-valued propositions (the minimum of all candidates in $Cand$).

With respect to the accuracy of $\mathcal{J}_c^{m,g,f}$, we can make a similar statement as for $\mathcal{J}_{hs}^{m,g,f}$.

Proposition 10. *For every* $p \in [0,1)$*,* g *some aggregation function with* $g(x,y) \geq (x+y)/2$*,* $f : \mathbb{N} \to [0,1]$ *a null-bound function, and* $\mathcal{K} \in \mathbb{K}$ *there is* $m \in \mathbb{N}$ *such that with probability greater or equal* p *it is the case that* $\lim_{i\to\infty} \mathcal{J}_c^{m,g,f}(\mathcal{S}_\mathcal{K}, i) = \mathcal{I}_c(\mathcal{K})$*.*

In order to evaluate the accuracy and performance of these stream-based inconsistency measures in more detail, we perform some empirical experiments in the following section.

6 EMPIRICAL EVALUATION

In this section we describe our empirical experiments on runtime, accuracy, and scalability of the discussed stream-based inconsistency measures. Our Java implementations[7] have been added to the *Tweety Libraries for Knowledge Representation* (Thimm, 2014d).

7 $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^c}$, $\mathcal{I}_\eta$, $\mathcal{J}_\mathcal{I}^{w,g}$: `http://mthimm.de/r?r=tweety-inc-commons`
$\mathcal{I}_c$, $\mathcal{I}_{hs}$: `http://mthimm.de/r?r=tweety-inc-pl`
$\mathcal{J}_{hs}^{m,g,f}$: `http://mthimm.de/r?r=tweety-stream-hs`
$\mathcal{J}_c^{m,g,f}$: `http://mthimm.de/r?r=tweety-stream-c`
Evaluation framework: `http://mthimm.de/r?r=tweety-stream-eval`

6.1 *Evaluated Approaches*

For our evaluation, we considered the inconsistency measures $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^c}$, $\mathcal{I}_\eta$, $\mathcal{I}_c$, and $\mathcal{I}_{hs}$. We used the SAT solver *lingeling*[8] for the sub-problems of determining consistency and to compute a model of a formula.

For enumerating the set of MIs of a knowledge base (as required by $\mathcal{I}_{\mathsf{MI}}$ and $\mathcal{I}_{\mathsf{MI}^c}$) we used MARCO[9], a tool for computing all minimal unsatisfiable sets of clauses from a knowledge base given in conjunctive normal form (CNF). In order to apply MARCO to our general non-CNF knowledge bases, we used the following approach. First, a general knowledge base $\mathcal{K}$ is converted to CNF, i. e., each formula of $\mathcal{K}$ is converted to a set of clauses. In doing so, we retain a mapping from each original formula to its set of clauses (note that clauses may appear multiple times in the resulting knowledge base, if they originate from different formulas). On the knowledge base in CNF we apply MARCO, which returns the set of all minimal sets of unsatisfiable clauses. Using the mapping to the original formulas, from each minimal set of unsatisfiable clauses a set of formulas is derived. By construction, the resulting set of formulas is inconsistent, but not necessarily minimally inconsistent. Therefore, after all these sets have been computed, a final minimality check is performed and all non-minimal sets are filtered out. This approach is similar to the one employed by MIMUS (McAreavey *et al.*, 2014), a tool which also determines MIs from a general knowledge base and is based on CAMUS[10]. We decided to use MARCO with the above preprocessing step instead of MIMUS directly, as initial experiments suggested that the former one is usually faster if the knowledge base contains at least one minimal inconsistent subset (which is the standard case in our evaluation). This observation is consistent with the observations made by (Liffiton and Malik, 2013), where CAMUS is criticized to be slower than MARCO for determining many minimal unsatisfiable sets. While CAMUS is a multi-purpose tool that also computes *minimal correction sets*, MARCO is optimized for computing minimal unsatisfiable sets of clauses fast.

The measure $\mathcal{I}_\eta$ was implemented using the linear optimization solver *lp_solve*[11].

The measures $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^c}$, and $\mathcal{I}_\eta$ were used to define three different versions of the window-based measure $\mathcal{J}_{\mathcal{I}}^{w,g}$ (with $w = 500, 1000, 2000$ and $g = \max$). For the measures $\mathcal{I}_c$ and $\mathcal{I}_{hs}$ we tested each three versions of their streaming variants $\mathcal{J}_c^{m,g_{0.75},f_1}$ and $\mathcal{J}_{hs}^{m,g_{0.75},f_1}$ (with $m = 10, 100, 500$) with $f_1 : \mathbb{N} \to [0,1]$ defined via $f_1(i) = 1/(i+1)$ for all $i \in \mathbb{N}$ with $i \leq 2^{32}$ and $f_1(i) = 0$ otherwise. Furthermore, $g_{0.75}$ is the smoothing function for $\alpha = 0.75$ as defined in the previous section.

8 http://fmv.jku.at/lingeling/
9 http://sun.iwu.edu/~mliffito/marco/
10 http://sun.iwu.edu/~mliffito/camus/
11 http://lpsolve.sourceforge.net

6.2 *Experiment Setup*

For measuring the runtime of the different approaches we generated 100 random knowledge bases in CNF with each 5000 formulæ and 30 propositions.[12] A knowledge base was generated by randomly determining the propositions appearing in a clause (uniformly distributed and up to a maximum of 4) and randomly negating some of these propositions (uniformly for each proposition). For each generated knowledge base $\mathcal{K}$ we considered its $\mathcal{K}$-stream and processing of the stream was aborted after 40000 iterations. We fed the $\mathcal{K}$-stream to each of the evaluated stream-based inconsistency measures and measured the average runtime per iteration and the total runtime. For each iteration, we set a time-out of 2 minutes and aborted processing of the stream completely if a time-out occurred.

In order to measure accuracy, for each of the considered approaches we generated another 100 random knowledge bases (not necessarily in CNF) with specifically set inconsistency values, used otherwise the same settings as above, and measured the returned inconsistency values.

To evaluate the scalability of our stream-based approach of $\mathcal{I}_{hs}$ we conducted a third experiment[13] where we fixed the number of propositions (60) and the specifically set inconsistency value (200) and varied the size of the knowledge bases from 5000 to 50000 (with steps of 5000 formulæ). We measured the total runtime up to the point when the inconsistency value was within a tolerance of ± 1 of the expected inconsistency value.

The experiments were conducted on a server with two Intel Xeon X5550 QuadCore (2.67 GHz) processors with 8 GB RAM running SUSE Linux 2.6.

6.3 *Results*

Our first observation concerns the inconsistency measure $\mathcal{I}_\eta$ which proved to be not suitable to work on large knowledge bases. Computing the value $\mathcal{I}_\eta(\mathcal{K})$ for some knowledge base $\mathcal{K}$ includes solving a linear optimization problem over a number of variables which is (in the worst-case) exponential in the number of propositions of the signature. In our setting with $|\mathsf{At}| = 30$ the generated optimization problem contained therefore $2^{30} = 1073741824$ variables. Hence, even the optimization problem itself could not be constructed within the timeout of 2 minutes for every step. In the following, we will therefore not report on further results for $\mathcal{I}_\eta$.

As for the runtime of the window-based approaches of $\mathcal{I}_{\mathsf{MI}}$ and $\mathcal{I}_{\mathsf{MI}^c}$ and our stream-based approaches for $\mathcal{I}_c$ and $\mathcal{I}_{hs}$ see Table 2. There one can see that $\mathcal{J}^{w,g}_{\mathcal{I}_{\mathsf{MI}}}$ and $\mathcal{J}^{w,g}_{\mathcal{I}_{\mathsf{MI}^c}}$ on the one hand, and $\mathcal{J}^{m,g,f}_{c}$ and $\mathcal{J}^{m,g,f}_{hs}$ on the other hand, have comparable runtimes, respectively. The former two have almost identical runtimes, which is obvious as the determination of the MIs is the

12 All sampling algorithms can be found at `http://mthimm.de/r?r=tweety-sampler`

13 We did the same experiment with our stream-based approach of $\mathcal{I}_c$ but do not report the results due to the similarity to $\mathcal{I}_{hs}$.

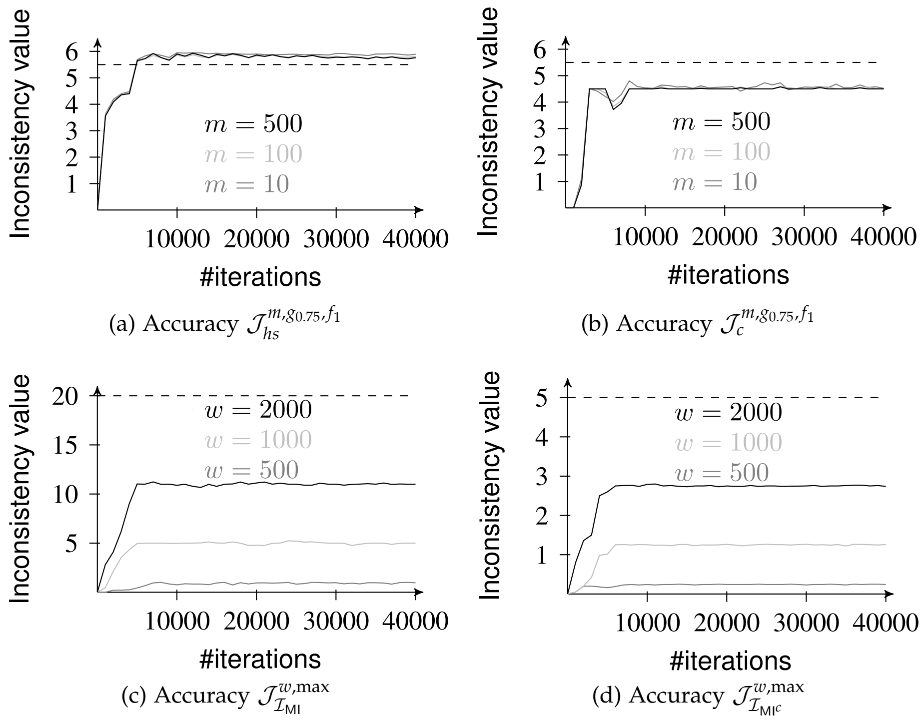
Measure	RT (iteration)	RT (total)	Measure	RT (iteration)	RT (total)
$\mathcal{J}^{500,\max}_{\mathcal{I}_{\mathsf{MI}}}$	198ms	133m	$\mathcal{J}^{10,g_{0.75},f_1}_{c}$	0.16ms	6.406s
$\mathcal{J}^{1000,\max}_{\mathcal{I}_{\mathsf{MI}}}$	359ms	240m	$\mathcal{J}^{100,g_{0.75},f_1}_{c}$	1.1ms	43.632s
$\mathcal{J}^{2000,\max}_{\mathcal{I}_{\mathsf{MI}}}$	14703ms	9812m	$\mathcal{J}^{500,g_{0.75},f_1}_{c}$	5.21ms	208.422s
$\mathcal{J}^{500,\max}_{\mathcal{I}_{\mathsf{MI}^c}}$	198ms	134m	$\mathcal{J}^{10,g_{0.75},f_1}_{hs}$	0.07ms	2.788s
$\mathcal{J}^{1000,\max}_{\mathcal{I}_{\mathsf{MI}^c}}$	361ms	241m	$\mathcal{J}^{100,g_{0.75},f_1}_{hs}$	0.24ms	9.679s
$\mathcal{J}^{2000,\max}_{\mathcal{I}_{\mathsf{MI}^c}}$	14812ms	9874m	$\mathcal{J}^{500,g_{0.75},f_1}_{hs}$	1.02ms	40.614s

Table 2: Runtimes for the evaluated measures; each value is averaged over 100 random knowledge bases of 5000 formulæ; the total runtime is after 40000 iterations

main problem in both their computations. Clearly, $\mathcal{J}^{m,g,f}_{c}$ and $\mathcal{J}^{m,g,f}_{hs}$ are significantly faster per iteration (and in total) than $\mathcal{J}^{w,g}_{\mathcal{I}_{\mathsf{MI}}}$ and $\mathcal{J}^{w,g}_{\mathcal{I}_{\mathsf{MI}^c}}$, only very few milliseconds for the latter and several hundreds and thousands of milliseconds for the former (for all variants of m and w). The impact of increasing m for $\mathcal{J}^{m,g,f}_{c}$ and $\mathcal{J}^{m,g,f}_{hs}$ is expectedly linear while the impact of increasing the window size w for $\mathcal{J}^{w,g}_{\mathcal{I}_{\mathsf{MI}}}$ and $\mathcal{J}^{w,g}_{\mathcal{I}_{\mathsf{MI}^c}}$ is exponential (this is also clear as both solve an FNP-hard problem).

As for the accuracy of the different approaches see Figure 1. There one can see that both $\mathcal{J}^{m,g,f}_{hs}$ and $\mathcal{J}^{m,g,f}_{c}$ (Figures 1a and 1b) converge quite quickly (almost right after the knowledge base has been processed once) into a $[-1,1]$ interval of the actual inconsistency value, where $\mathcal{J}^{m,g,f}_{hs}$ is even closer to it. The window-based approaches (Figures 1c and 1d) have a comparable bad performance (this is clear as those approaches cannot *see* all MIs at any iteration due to the limited window size). Surprisingly, the impact of larger values of m for $\mathcal{J}^{m,g,f}_{hs}$ and $\mathcal{J}^{m,g,f}_{c}$ is rather small in terms of accuracy which suggests that the random process of our algorithm is quite robust. Even for $m = 10$ the results are quite satisfactory.

As for the scalability of $\mathcal{J}^{m,g_{0.75},f_1}_{hs}$ see Figure 2. There one can observe a linear increase in the runtime of all variants wrt. the size of the knowledge base. Furthermore, the difference between the variants is also linear in the parameter m (which is also clear as each population is an independent random process). It is noteworthy, that the average runtime for $\mathcal{J}^{10,g_{0.75},f_1}_{hs}$ is about 66.1 seconds for knowledge bases with 50000 formulæ. As the significance of the parameter m for the accuracy is also only marginal, the measure $\mathcal{J}^{10,g_{0.75},f_1}_{hs}$ is clearly an effective and accurate stream-based inconsistency measure.



(a) Accuracy $\mathcal{J}_{hs}^{m,g_{0.75},f_1}$

(b) Accuracy $\mathcal{J}_{c}^{m,g_{0.75},f_1}$

(c) Accuracy $\mathcal{J}_{\mathcal{I}_{\mathsf{MI}}}^{w,\max}$

(d) Accuracy $\mathcal{J}_{\mathcal{I}_{\mathsf{MI}^c}}^{w,\max}$

Figure 1: Accuracy performance for the evaluated measures (dashed line is actual inconsistency value); each value is averaged over 100 random knowledge bases of 5000 formulæ (30 propositions) with varying inconsistency values

7 RELATED WORK

This work is the first to address inconsistency measurement in streaming scenarios. The closest family of related works are approaches for efficient inconsistency measure computation for the classical setting (where a knowledge base is given as a whole).

In (Ma *et al.*, 2009, 2010; Xiao *et al.*, 2010) Ma and colleagues present an anytime algorithm that approximates an inconsistency measure based on a 4-valued paraconsistent logic (similar to the contension inconsistency measure). The algorithm provides lower and upper bounds for this measure and can be stopped at any point in time with some guaranteed quality. The main difference between our framework and the algorithm of (Ma *et al.*, 2009, 2010) is that the latter needs to process the whole knowledge base in each atomic step and is therefore not directly applicable for the streaming scenario. The empirical evaluation in (Ma *et al.*, 2009, 2010) also suggests that our streaming variant of $\mathcal{I}_{hs}$ is much more performant as Ma et al. report an average runtime of their algorithm of about 240 seconds on a knowledge base with 120 formulæ and 20 propositions (no evaluation on

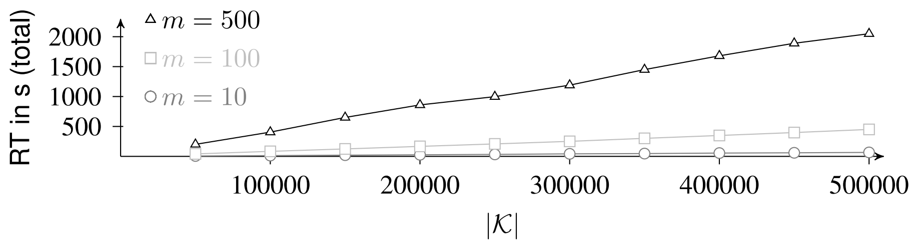


Figure 2: Evaluation of the scalability of $\mathcal{J}_{hs}^{m,g_{0.75},f_1}$; each value is averaged over 10 random knowledge bases of the given size

larger knowledge bases is given) while our measure has a runtime of only a few seconds for knowledge bases with 5000 formulæ with comparable accuracy[14].

In (McAreavey *et al.*, 2014) an approach is developed for computing measures based on minimal inconsistent subsets (such as $\mathcal{I}_{\mathsf{MI}}$ and $\mathcal{I}_{\mathsf{MI}^c}$) more efficiently. Core to the computation of these measures is the determination of $\mathsf{MI}(\mathcal{K})$ for an arbitrary knowledge base $\mathcal{K}$. While the computational challenges of determining the set of minimal inconsistent subsets for knowledge bases in CNF has been studied for some time in the SAT community—see e. g. (Liffiton and Sakallah, 2005; Büning and Kullmann, 2009; Liffiton and Malik, 2013; Previti and Marques-Silva, 2013)—additional issues arise when considering knowledge bases that are not in CNF. These issues are addressed by (McAreavey *et al.*, 2014) where an approach for efficiently computing minimal inconsistent sets for arbitrary knowledge is presented. The approach has been implemented in the tool MIMUS which is based on the tool CAMUS[15] for computing minimal inconsistent subsets of knowledge bases in CNF. In (McAreavey *et al.*, 2014) this tool has been empirically evaluated, also in the context of measuring inconsistency with measures based on minimal inconsistent subsets. Compared to (McAreavey *et al.*, 2014) we consider arbitrary inconsistency measures and not just those based on minimal inconsistent subsets. Still, the work of McAreavey et al. is relevant for applying those inconsistency measures to our streaming scenario. We slightly adapted the general approach of (McAreavey *et al.*, 2014) and used MARCO[16], instead of the predecessor CAMUS, for our empirical evaluation (see Section 6).

8 SUMMARY AND CONCLUSION

In this paper we introduced and discussed the problem of stream-based inconsistency measurement. For that, we developed a novel inconsistency

14 Although hardware specifications for these experiments are different this huge difference is quite relevant.
15 `http://sun.iwu.edu/~mliffito/camus/`
16 `http://sun.iwu.edu/~mliffito/marco/`

measure $\mathcal{I}_{hs}$ that is based on the notion of a *hitting set* and analyzed its properties. We presented a general framework for applying classical inconsistency measures to the streaming scenario and developed specific approximation algorithms for $\mathcal{I}_{hs}$ and the contension measure $\mathcal{I}_c$. Our empirical evaluation showed that the latter two approaches outperform the baseline window-based approaches to measure inconsistency and streams and provide general evidence of the feasibility of stream-based inconsistency measurement.

All discussed inconsistency measures (classical and stream-based ones), as well as the evaluation framework have been implemented in Java and added to the open source project *Tweety*[17] (Thimm, 2014d). Current work is about the application of our work on linked open data sets (Isele *et al.*, 2010) enriched with OWL semantics.

APPENDIX: PROOFS OF TECHNICAL RESULTS

Proposition 1. *Let $\mathcal{K}$ be a knowledge base. The following two statements are equivalent:*

1. *there is no $\phi \in \mathcal{K}$ with $\phi \models \perp$*
2. *there exists a hitting set H of $\mathcal{K}$*

Proof. Let $\mathcal{K} = \{\phi_1, \ldots, \phi_n\}$. First, assume that there is no $\phi_i \in \mathcal{K}$ with $\phi_i \models \perp$ for $i = 1, \ldots, n$. It follows $\mathsf{Mod}(\phi_i) \neq \emptyset$ for every $i = 1, \ldots, n$. Let $\omega_i \in \mathsf{Mod}(\phi_i)$, then by definition $\{\omega_1, \ldots, \omega_n\}$ is a hitting set of $\mathcal{K}$. Let now $H = \{\omega_1, \ldots, \omega_m\}$ be a hitting set of $\mathcal{K}$. Then for every $\phi \in \mathcal{K}$ there is $\omega \in H$ with $\omega \models \phi$. Therefore there can be no $\phi \in \mathcal{K}$ with $\mathsf{Mod}(\phi) \neq \emptyset$. □

Proposition 2. *Let $\mathcal{K}$ be a knowledge base.*

1. *If H is a hitting set of $\mathcal{K}$ then every H' with $H \subseteq H'$ is a hitting set of $\mathcal{K}$.*
2. *$H = \emptyset$ is a hitting set of $\mathcal{K}$ if and only if $\mathcal{K} = \emptyset$.*
3. *$\mathcal{K}$ is consistent if and only if there is a hitting set H of $\mathcal{K}$ with $|H| = 1$.*
4. *If H is a hitting set of $\mathcal{K}$ then H is a hitting set of every $\mathcal{K}'$ with $\mathcal{K}' \subseteq \mathcal{K}$.*

Proof. Let $\mathcal{K} = \{\phi_1, \ldots, \phi_n\}$.

1. Let H be a hitting set of $\mathcal{K}$ and let H' be such that $H \subseteq H'$. Then for every $\phi \in \mathcal{K}$ we have $\omega \in H \subseteq H'$ such that $\omega \models \phi$. Therefore H' is a hitting set of $\mathcal{K}$.
2. Let $\mathcal{K} = \emptyset$. Then $H = \emptyset$ is a trivial hitting set of $\mathcal{K}$ by definition of universal quantification. Note also that for any $\mathcal{K}$ with $\mathcal{K} \neq \emptyset$ the set $H = \emptyset$ cannot be a hitting set.

17 http://tweetyproject.org

3. Let $\mathcal{K} \neq \emptyset$ be consistent. Then there is $\omega \in \mathsf{Int}(\mathsf{At})$ with $\omega \models \mathcal{K}$, i. e., $\omega \models \phi$ for every $\phi \in \mathcal{K}$. Therefore, $\{\omega\}$ is a hitting set of $\mathcal{K}$ with $|\{\omega\}| = 1$. Let H be any hitting set of $\mathcal{K}$ with $|H| = 1$, i. e., $H = \{\omega\}$. Then $\omega \models \phi$ for all $\phi \in \mathcal{K}$ and, hence, $\phi \models \mathcal{K}$. Therefore, $\mathcal{K}$ is consistent. For the case $\mathcal{K} = \emptyset$ note that every subset of $\mathsf{Int}(\mathsf{At})$ is a hitting set of $\mathcal{K}$.

4. Let H be a hitting set of $\mathcal{K}$ and let $\mathcal{K}' \subseteq \mathcal{K}$. Then for every $\phi \in \mathcal{K}'$ there is $\omega \in H$ with $\omega \models \phi$ as $\mathcal{K}' \subseteq \mathcal{K}$. Hence, H is a hitting set of $\mathcal{K}$. □

Proposition 3. *The function $\mathcal{I}_{hs}$ is a (basic) inconsistency measure.*

Proof. We have to show that properties 1.), 2.), and 3.) of Definition 4 are satisfied.

1. This follows directly from items 2.) and 3.) of Proposition 2.

2. This follows directly from item 4.) of Proposition 2.

3. Let $\alpha \in \mathsf{Free}(\mathcal{K})$ and define $\mathcal{K}' = \mathcal{K} \setminus \{\alpha\}$. Let H be a hitting set of $\mathcal{K}'$ with $|H|$ being minimal and let $\omega \in H$. Furthermore, let $\mathcal{K}'' \subseteq \mathcal{K}'$ be the set of all formulæ β such that $\omega \models \beta$. It follows that $\mathcal{K}''$ is consistent. As α is a free formula it follows that $\mathcal{K}'' \cup \{\alpha\}$ is also consistent (otherwise there would be a minimal inconsistent subset of $\mathcal{K}''$ containing α). Let ω' be a model of $\mathcal{K}'' \cup \{\alpha\}$. Then $H' = (H \setminus \{\omega\}) \cup \{\omega'\}$ is a hitting set of $\mathcal{K}$ and due to 2.) also of minimal cardinality. Hence, we have $\mathcal{I}_{hs}(\mathcal{K}') = \mathcal{I}_{hs}(\mathcal{K})$. □

Proposition 4. *The measure $\mathcal{I}_{hs}$ satisfies the following properties:*

- *If $\alpha \in \mathcal{K}$ is such that $\mathsf{At}(\alpha) \cap \mathsf{At}(\mathcal{K} \setminus \{\alpha\}) = \emptyset$ then $\mathcal{I}_{hs}(\mathcal{K}) = \mathcal{I}_{hs}(\mathcal{K} \setminus \{\alpha\})$ (*safe formula independence*).*

- *If $\mathcal{K} \equiv^{\sigma} \mathcal{K}'$ then $\mathcal{I}_{hs}(\mathcal{K}) = \mathcal{I}_{hs}(\mathcal{K}')$ (*irrelevance of syntax*).*

- *If $\alpha \models \beta$ and $\alpha \not\models \perp$ then $\mathcal{I}_{hs}(\mathcal{K} \cup \{\alpha\}) \geq \mathcal{I}_{hs}(\mathcal{K} \cup \{\beta\})$ (*dominance*).*

Proof.

- This is satisfied as safe formula independence follows from free formula independence, cf. (Hunter and Konieczny, 2010; Thimm, 2013b).

- Let H be a hitting set of $\mathcal{K}$ with minimal cardinality. So, for every $\alpha \in \mathcal{K}$ we have $\omega \in H$ with $\omega \models \alpha$. Due to $\alpha \equiv \sigma(\alpha)$ we also have $\omega \models \sigma(\alpha)$ and, thus for very $\beta \in \mathcal{K}'$ we have $\omega \in H$ with $\omega \models \beta$. So H is also a hitting set of $\mathcal{K}'$. Minimality follows from the fact that σ is a bijection.

- Let H be a minimal hitting set of $\mathcal{K}_1 = \mathcal{K} \cup \{\alpha\}$ with minimal cardinality and let $\omega \in H$ be such that $\omega \models \alpha$. Then we also have that $\omega \models \beta$ and H is also a hitting set of $\mathcal{K}_2 = \mathcal{K} \cup \{\beta\}$. Hence, $\mathcal{I}_{hs}(\mathcal{K}_1) \geq \mathcal{I}_{hs}(\mathcal{K}_2)$. □

Proposition 5. *For every knowledge base $\mathcal{K}$*

$$\mathcal{I}_{hs}(\mathcal{K}) = \min\{|\Phi| \mid \Phi \textit{ is a consistent partitioning of } \mathcal{K}\} - 1$$

with $\min \emptyset = \infty$ *for every* $\mathcal{K} \in \mathbb{K} \setminus \{\emptyset\}$ *and* $\mathcal{I}_{hs}(\emptyset) = 0$.

Proof. For $\mathcal{K} = \emptyset$ and the case that $\mathcal{K}$ contains ϕ with $\phi \models \perp$, the statement is trivially satisfied so assume $\mathcal{K} \neq \emptyset$ and that $\mathcal{K}$ does not contain an inconsistent formula. Let $\Phi = \{\Phi_1, \ldots, \Phi_n\}$ be a consistent partitioning with $|\Phi|$ being minimal and let $\omega_i \in \mathsf{Int}(\mathsf{At})$ be such that $\omega_i \models \Phi_i$ (for $i = 1, \ldots, n$). Observe that $\omega_i \neq \omega_j$ for all $i \neq j$, otherwise $\Phi_i \cup \Phi_j$ would have a model $\omega_i = \omega_j$ and $\Phi' = \Phi \setminus \{\Phi_i, \Phi_j\} \cup \{\Phi_i \cup \Phi_j\}$ would be a consistent partitioning with $|\Phi'| < |\Phi|$. Then $H = \{\omega_1, \ldots, \omega_n\}$ is a hitting set of $\mathcal{K}$ and we have $\mathcal{I}_{hs}(\mathcal{K}) \leq |\Phi| - 1$. Let now $H = \{\omega_1, \ldots, \omega_n\}$ be a hitting set of $\mathcal{K}$ with $|H|$ being minimal. Let $\Phi = \{\Phi_1, \ldots, \Phi_n\}$ be a set such that $\phi \in \Phi_i$ implies $\omega_i \models \phi$ for every $\phi \in \mathcal{K}$ (note that there may be multiple partitionings satisfying this property but they all have the same cardinality). Note that Φ is a partitioning of $\mathcal{K}$ and that Φ_i is consistent for every $i = 1, \ldots, n$. It follows $\mathcal{I}_{hs}(\mathcal{K}) \geq |\Phi| - 1$ and therefore the claim. □

Proposition 6. *Let $\mathcal{K}$ be a knowledge base. If $\infty > \mathcal{I}_{hs}(\mathcal{K}) > 0$ then*

$$\mathcal{I}_\eta(\mathcal{K}) \leq 1 - \frac{1}{\mathcal{I}_{hs}(\mathcal{K}) + 1}$$

Proof. Let H be a minimal hitting set of $\mathcal{K}$ with minimal cardinality, i. e., we have $\mathcal{I}_{hs}(\mathcal{K}) = |H| - 1$. Define a probability function $P : \mathsf{Int}(\mathsf{At}) \to [0, 1]$ via $P(\omega) = 1/|H|$ for every $\omega \in H$ and $P(\omega') = 0$ for every $\omega' \in \mathsf{Int}(\mathsf{At}) \setminus H$ (note that P is indeed a probability function). As H is a hitting set of $\mathcal{K}$ we have that $P(\phi) \geq 1/|H|$ for every $\phi \in \mathcal{K}$ as at least one model of ϕ gets probability $1/|H|$ in P. So we have $\mathcal{I}_\eta(\mathcal{K}) \leq 1 - 1/|H| = 1 - 1/(\mathcal{I}_{hs}(\mathcal{K}) + 1)$. □

Proposition 7. *There is no subsumption relation between $\mathcal{I}_{hs}$ and any $\mathcal{I} \in \{\mathcal{I}_{\mathsf{MI}}, \mathcal{I}_{\mathsf{MI}^c}, \mathcal{I}_c, \mathcal{I}_\eta\}$.*

Proof.

1. $\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_{\mathsf{MI}}$: Consider the knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ given as

$$\mathcal{K}_8 = \{a \wedge b, \neg a \wedge b, a \wedge \neg b\}$$
$$\mathcal{K}_9 = \{a, b, c, \neg a \wedge \neg b \wedge \neg c\}$$

   Then we have $\mathcal{I}_{hs}(\mathcal{K}_9) = 1 < 2 = \mathcal{I}_{hs}(\mathcal{K}_8)$ but $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_9) = 3 = \mathcal{I}_{\mathsf{MI}}(\mathcal{K}_8)$.

2. $\mathcal{I}_{\mathsf{MI}} \not\sqsubseteq \mathcal{I}_{hs}$: Consider the knowledge bases $\mathcal{K}_4$ and $\mathcal{K}_{10}$ given as

$$\begin{aligned}\mathcal{K}_4 &= \{a, \neg a\}\\ \mathcal{K}_{10} &= \{a, \neg a \wedge \neg b, b\}\end{aligned}$$

Then we have $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_4) = 1 < 2 = \mathcal{I}_{\mathsf{MI}}(\mathcal{K}_{10})$ but $\mathcal{I}_{hs}(\mathcal{K}_4) = 1 = \mathcal{I}_{hs}(\mathcal{K}_{10})$.

3. $\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_{\mathsf{MI}^c}$: For the knowledge bases from item 1.) we also have $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_8) = 1.5 = \mathcal{I}_{\mathsf{MI}}(\mathcal{K}_9)$.

4. $\mathcal{I}_{\mathsf{MI}^c} \not\sqsubseteq \mathcal{I}_{hs}$: For the knowledge bases from item 2.) we also have $\mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_4) = 1/2 < 1 = \mathcal{I}_{\mathsf{MI}^c}(\mathcal{K}_{10})$.

5. $\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_c$: see Example 6.

6. $\mathcal{I}_c \not\sqsubseteq \mathcal{I}_{hs}$: see Example 6.

7. $\mathcal{I}_{hs} \not\sqsubseteq \mathcal{I}_\eta$: Consider the knowledge bases $\mathcal{K}_4$ and $\mathcal{K}_{11}$ given as

$$\begin{aligned}\mathcal{K}_4 = \{&a, \neg a\}\\ \mathcal{K}_{11} = \{&(a \wedge b \wedge c) \vee (a \wedge b \wedge \neg c) \vee (a \wedge \neg b \wedge c),\\ &(a \wedge b \wedge c) \vee (a \wedge b \wedge \neg c) \vee (a \wedge \neg b \wedge \neg c),\\ &(a \wedge b \wedge c) \vee (a \wedge b \wedge \neg c) \vee (\neg a \wedge b \wedge c),\\ &(a \wedge b \wedge c) \vee (a \wedge \neg b \wedge c) \vee (a \wedge \neg b \wedge \neg c),\\ &(a \wedge b \wedge c) \vee (a \wedge \neg b \wedge c) \vee (\neg a \wedge b \wedge c),\\ &(a \wedge b \wedge c) \vee (a \wedge \neg b \wedge \neg c) \vee (\neg a \wedge b \wedge c),\\ &(a \wedge b \wedge \neg c) \vee (a \wedge \neg b \wedge c) \vee (a \wedge \neg b \wedge \neg c),\\ &(a \wedge b \wedge \neg c) \vee (a \wedge \neg b \wedge c) \vee (\neg a \wedge b \wedge c),\\ &(a \wedge b \wedge \neg c) \vee (a \wedge \neg b \wedge \neg c) \vee (\neg a \wedge b \wedge c),\\ &(a \wedge \neg b \wedge c) \vee (a \wedge \neg b \wedge \neg c) \vee (\neg a \wedge b \wedge c)\}\end{aligned}$$

Then we have $\mathcal{I}_{hs}(\mathcal{K}_4) = 1 < 2 = \mathcal{I}_{hs}(\mathcal{K}_{11})$ but $\mathcal{I}_\eta(\mathcal{K}_4) = 1/2 > 2/5 = \mathcal{I}_\eta(\mathcal{K}_{11})$. Let us discuss $\mathcal{K}_{11}$ a bit more. Consider the five interpretations $\omega_1, \ldots, \omega_5$ of the propositional signature $\mathsf{At} = \{a, b, c\}$ defined via

$$\begin{aligned}\omega_1 &= abc\\ \omega_2 &= ab\overline{c}\\ \omega_3 &= a\overline{b}c\\ \omega_4 &= a\overline{b}\overline{c}\\ \omega_5 &= \overline{a}bc\end{aligned}$$

Then $\mathcal{K}_{11}$ comprises of formulas ϕ that are satisfied by exactly three out of these five interpretations (and for each 3-element subset of

$\{\omega_1, \ldots, \omega_5\}$ there is exactly one corresponding formula). It follows that a probability function P assigning probability $1/5$ to each of these five interpretations (and zero to the remaining interpretations) yields $P(\phi) = 3/5$ for each $\phi \in \mathcal{K}_2$ (and this is maximal), thus yielding $\mathcal{I}_\eta(\mathcal{K}_{11}) = 1 - 3/5 = 2/5$. Further, any 3-element subset of $\{\omega_1, \ldots, \omega_5\}$ is also a hitting set of $\mathcal{K}_{11}$: as every $\phi \in \mathcal{K}_2$ is satisfied by exactly three interpretations, one can remove any two of them and still maintain the hitting set property. So $H = \{\omega_1, \omega_2, \omega_3\}$ is a hitting set and one can easily see that there is no smaller one, yielding $\mathcal{I}_{hs}(\mathcal{K}_{11}) = |H| - 1 = 2$.

8. $\mathcal{I}_\eta \not\sqsubseteq \mathcal{I}_{hs}$: Consider the knowledge bases $\mathcal{K}_4$ and $\mathcal{K}_{12}$ given as

$$\mathcal{K}_4 = \{a, \neg a\}$$
$$\mathcal{K}_{12} = \{a, b, \neg a \vee \neg b\}$$

Then we have $\mathcal{I}_\eta(\mathcal{K}_{12}) = 1/3 < 1/2 = \mathcal{I}_\eta(\mathcal{K}_4)$ but $\mathcal{I}_{hs}(\mathcal{K}_{12}) = 1 = \mathcal{I}_{hs}(\mathcal{K}_4)$. □

Corollary 1. *$\mathcal{I}_{hs} \not\approx \mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{hs} \not\approx \mathcal{I}_{\mathsf{MI}^c}$, $\mathcal{I}_{hs} \not\approx \mathcal{I}_c$, and $\mathcal{I}_{hs} \not\approx \mathcal{I}_\eta$.*

Proof. This follows directly from Proposition 7 and the definition of equivalence. □

Proposition 8. *Let $\mathcal{I}$ be an inconsistency measure, $w \in \mathbb{N} \cup \{\infty\}$, and g an aggregation function.*

1. *If w is finite then $\mathcal{J}_{\mathcal{I}}^{w,g}$ is* not *an approximation of $\mathcal{I}$.*
2. *$\mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S}_{\mathcal{K}}, i) \leq \mathcal{I}(\mathcal{K})$ for every $\mathcal{K} \in \mathbb{K}$ and $i \in \mathbb{N}$.*
3. *If $w = \infty$ and $g(x, y) \geq (x + y)/2$ then $\mathcal{J}_{\mathcal{I}}^{w,g}$ is an approximation of $\mathcal{I}$.*

Proof.

1. Assume $\mathcal{K}$ is a minimal inconsistent set with $|\mathcal{K}| > w$. Then $\mathcal{I}(\mathcal{S}^{\max\{0,i-w\},i}) = 0$ for all $i > 0$ (as every subset of $\mathcal{K}$ is consistent) and $\mathcal{J}_{\mathcal{I}}^{w,g}(\mathcal{S}, i) = 0$ for all $i > 0$ as well. As $\mathcal{I}$ is an inconsistency measure $\mathcal{I}(\mathcal{K}) > 0$ and, hence, $\mathcal{J}_{\mathcal{I}}^{w,g}$ does not approximate $\mathcal{I}$.
2. This follows from the fact that $\mathcal{I}$ is a basic inconsistency measure and therefore satisfies $\mathcal{I}(\mathcal{K}) \leq \mathcal{I}(\mathcal{K}')$ for $\mathcal{K} \subseteq \mathcal{K}'$.
3. If $w = \infty$ there is $i_0 \in \mathbb{N}$ such that $\mathcal{I}(\mathcal{S}^{\max\{0,i-w\},i}) = \mathcal{I}(\mathcal{K})$ for all $i > i_0$. Due to item 2 above (all previous values estimated the inconsistency value from below) and as $g(x, y) > (x + y)/2$ (in each step the new value is the average of the previous value and the actual inconsistency value) the value $\mathcal{I}(\mathcal{K})$ will be approximated by $\mathcal{J}_{\mathcal{I}}^{w,g}$ eventually. □

Proposition 9. *For every $p \in [0,1)$, g some aggregation function with $g(x,y) \geq (x+y)/2$, $f : \mathbb{N} \rightarrow [0,1]$ a null-bound function, and $\mathcal{K} \in \mathbb{K}$ there is $m \in \mathbb{N}$ such that with probability greater or equal p it is the case that $\lim_{i\rightarrow\infty} \mathcal{J}_{hs}^{m,g,f}(\mathcal{S}_{\mathcal{K}}, i) = \mathcal{I}_{hs}(\mathcal{K})$.*

Proof. Let $p \in [0,1)$, g some aggregation function with $g(x,y) \geq (x+y)/2$, $f : \mathbb{N} \rightarrow [0,1]$ a null-bound function, and $\mathcal{K} \in \mathbb{K}$. Let $H = \{\omega_1, \ldots, \omega_h\} \subseteq \mathsf{Int}(\mathsf{At})$ be a hitting set of $\mathcal{K}$ such that $\mathcal{I}_{hs}(\mathcal{K}) = |H| - 1$. Consider the evolution of a single candidate set during the iterated execution of $\texttt{update}_{hs}^{m,g,f}(form)$. If $\mathcal{K}$ contains a contradictory formula then lines 2–5 ensure that the return value of Algorithm 2 is always ∞ and thus the claim holds trivially. We now assume that $\mathcal{K}$ contains no contradictory formula.

Let $C_0 = \emptyset$ be the initial candidate set and let C_i for $i \in \mathbb{N}$ denote the candidate set after iteration i. In the first iteration, C_0 does not contain any interpretations yet, so lines 9 and 10 of Algorithm 2 are vacuous. As the condition in line 11 evaluates to true, we add some interpretation to C_0. As H is a hitting set of $\mathcal{K}$ there is $\omega \in H$ with $\omega \in \mathsf{Mod}(form)$. The probability of choosing ω in line 12, so the probability of C_0 evolving to $C_1 = \{\omega\}$, is $p_0 = 1/|\mathsf{Mod}(form)| > 0$. In the second iteration, the probability q_1 that line 10 is *not* executed is greater than zero (as f is bounded by 1 and the condition involves a strictly less comparison). Assume that $\omega \not\models form$ (otherwise simply continue with the next formula in the next iteration). Then, again, there is $\omega' \in H$ with $\omega' \models form$ and the probability of choosing ω' in line 12 is $1/|\mathsf{Mod}(form)| > 0$. Therefore, the probability of C_0 evolving through C_1 to $C_2 = \{\omega, \omega'\}$ is $p_1 = p_0 q_1/|\mathsf{Mod}(form)| > 0$. It follows that the probability of C_0 evolving to $C_h = H$ in its h-th iteration is strictly greater than zero. Note that beginning in the $h+1$-th iteration the condition in line 11 is not satisfied anymore and that $C_i = C_{i+1}$ for every $i > h$ with positive probability as well as there is a positive probability that line 10 will not be executed (with increasing probability over the iterations as f is a null-bound function). So for every candidate set $C \in Cand$ there is a positive probability $\hat{p}$ that C evolves to H and does not change anymore thereafter.

In general, observe that every candidate set $C \in Cand$ evolves into a hitting set of $\mathcal{K}$ as the probability of executing line 10 becomes zero eventually and lines 11 and 12 ensure that every formula in $\mathcal{K}$ has a model in C. Furthermore, each evolution of a candidate set $C \in Cand$ is an independent random process. So for $|Cand| = m$ the probability that at least one element of $Cand$ evolves to H in the above described manner is $1 - (1 - \hat{p})^m$ (here $(1 - \hat{p})^m$ is the probability that none of the candidate sets evolve to H). Observe that $1 - (1 - \hat{p})^m$, due to $\hat{p} > 0$, is monotonously increasing in m with $\lim_{m\rightarrow\infty} 1 - (1 - \hat{p})^m = 1$. Therefore we can choose $\hat{m}$ such that $1 - (1 - \hat{p})^{\hat{m}} \geq p \in [0,1)$ and with probability at least p the set of candidate sets $Cand$ with $|Cand| = \hat{m}$ contains at least one candidate set that evolves to H. Let $C_0, C_1, \ldots, C_h, C_{h+1}, \ldots$ be the evolution of this candidate set with

$C_0 \subseteq C_1 \ldots \subseteq C_h = C_{h+1} = \ldots$ with $C_h = H$. It follows that the variable *newValue* has never a value larger than $|H| - 1$ whenever line 14 is executed. As *currentValue* is initialized with 0 it also follows that *currentValue* has never a larger value than $|H| - 1$ in line 15, as g is an aggregation function.

Consider now an iteration step i where all candidate sets in *Cand* have stabilized and do not change thereafter. Then *newValue* always has the value $|H| - 1$ in line 14. If at iteration step i *currentValue* has already the value $|H| - 1$ then the value of *currentValue* is not changed (as g is an aggregation function). Then we have that the return value of Algorithm 2 in line 15 is always $|H| - 1 = \mathcal{I}_{hs}(\mathcal{K})$ and therefore $\lim_{i\to\infty} \mathcal{J}_{hs}^{m,g,f}(\mathcal{S}_{\mathcal{K}}, i) = \mathcal{I}_{hs}(\mathcal{K})$. If at iteration step i *currentValue* has a value $\alpha_0 < |H| - 1$ then observe that *currentValue* is updated to some value $\alpha_1 \geq (\alpha_0 + (|H| - 1))/2$ with $\alpha_1 \leq |H| - 1$ (as g is an aggregation function). In subsequent iterations this value is updated while satisfying $\alpha_j \geq (\alpha_{j+1} + (|H| - 1))/2$ with $\alpha_j \leq |H| - 1$ which converges to $|H| - 1$ and thus proves the claim. □

Proposition 10. *For every $p \in [0, 1)$, g some aggregation function with $g(x, y) \geq (x + y)/2$, $f : \mathbb{N} \to [0, 1]$ a null-bound function, and $\mathcal{K} \in \mathbb{K}$ there is $m \in \mathbb{N}$ such that with probability greater or equal p it is the case that $\lim_{i\to\infty} \mathcal{J}_c^{m,g,f}(\mathcal{S}_{\mathcal{K}}, i) = \mathcal{I}_c(\mathcal{K})$.*

Proof. Let $p \in [0, 1)$, g some aggregation function with $g(x, y) \geq (x + y)/2$, $f : \mathbb{N} \to [0, 1]$ a null-bound function, and $\mathcal{K} \in \mathbb{K}$. Let $\hat{v} : \mathsf{At} \to \{T, F, B\}$ be a three-valued interpretation of the atoms appearing in $\mathcal{K}$ such that $\hat{v} \models^3 \mathcal{K}$ and $\mathcal{I}_c(\mathcal{K}) = |\hat{v}^{-1}(B)| = t$. Consider the evolution of a single three-valued interpretation $v \in \mathit{Cand}$ during the iterated execution of $\mathtt{update}_c^{m,g,f}(\mathit{form})$.

Let v_0 be the initial interpretation with $v^{-1}(B) = \emptyset$ set and let v_i for $i \in \mathbb{N}$ denote the interpretation after iteration i. As f is null-bound there is an iteration $k > 0$ from which on line 6 is not executed anymore. Furthermore, observe that lines 7–13 ensure that v is changed in such a way that it satisfies the formula form. Note that once we are in an iteration $k' \geq k$ and v satisfies all formulas in $\mathcal{K}$ lines 7–13 will also not be executed anymore. Consequently, the evolution of v always converges at some iteration $l > 0$ and v_l satisfies all formulas in $\mathcal{K}$. Similarly to the analysis in the proof of Proposition 9 the probability of v evolving to $\hat{v}$, i. e., $v_l = \hat{v}$, is strictly greater than zero (albeit potentially quite small). In particular, the probability $\hat{p}$ of the evolution $v_0, v_1, \ldots, v_t, \ldots$ with $|v_0^{-1}(B)| = 0$,$|v_1^{-1}(B)| = 1$,…, $|v_t^{-1}(B)| = t$, and $v_{j+1} = v_j = \hat{v}$ for all $j \geq t$ is strictly greater zero. So for $|\mathit{Cand}| = m$ the probability that at least one element of *Cand* evolves to $\hat{v}$ in the above described manner is $1 - (1 - \hat{p})^m$ (here $(1 - \hat{p})^m$ is the probability that none of the interpretations evolve to $\hat{v}$). Therefore we can choose $\hat{m}$ such that $1 - (1 - \hat{p})^{\hat{m}} \geq p \in [0, 1)$ and with probability at least p the set *Cand* with $|\mathit{Cand}| = \hat{m}$ contains at least one interpretation that evolves to $\hat{v}$. It follows that the variable *newValue* has never a value larger than t whenever line 15 is executed. As *currentValue* is initialized with 0 it also follows that *currentValue* has never a larger value than t in line 16, as g is an aggregation function.

Consider now an iteration step i where all interpretations in *Cand* have stabilized and do not change thereafter. Then *newValue* always has the value t in line 15. If at iteration step i *currentValue* has already the value t then the value of *currentValue* is not changed (as g is an aggregation function). Then we have that the return value of Algorithm 4 in line 16 is always $t = \mathcal{I}_c(\mathcal{K})$ and therefore $\lim_{i\to\infty} \mathcal{J}_c^{m,g,f}(\mathcal{S}_{\mathcal{K}}, i) = \mathcal{I}_c(\mathcal{K})$. If at iteration step i *currentValue* has a value $\alpha_0 < t$ then observe that *currentValue* is updated to some value $\alpha_1 \geq (\alpha_0 + t)/2$ with $\alpha_1 \leq t$ (as g is an aggregation function). In subsequent iterations this value is updated while satisfying $\alpha_j \geq (\alpha_{j+1} + t))/2$ with $\alpha_j \leq t$ which converges to t and thus proves the claim. $\square$

ON THE EXPRESSIVITY OF INCONSISTENCY MEASURES

Matthias Thimm. In Artificial Intelligence, 234:120–151. February 2016.

Abstract

We survey recent approaches to inconsistency measurement in propositional logic and provide a comparative analysis in terms of their *expressivity*. For that, we introduce four different expressivity characteristics that quantitatively assess the number of different knowledge bases that a measure can distinguish. Our approach aims at complementing ongoing discussions on rationality postulates for inconsistency measures by considering expressivity as a desirable property. We evaluate 16 different measures on the proposed characteristics and conclude that the distance-based measure $\mathcal{I}^{\Sigma}_{\text{dalal}}$ from (Grant and Hunter, 2013) and the proof-based measure $\mathcal{I}_{P_m}$ from (Jabbour and Raddaoui, 2013) have maximal expressivity along all considered characteristics. In our study, we discovered several interesting relationships of inconsistency measurement to e. g. set theory and Boolean functions and we also report on these findings.

1 INTRODUCTION

Inconsistency measurement is about the quantitative assessment of the severity of inconsistencies in knowledge bases. Consider the following two knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ formalised in propositional logic:

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\} \qquad \mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

Both knowledge bases are classically inconsistent as for $\mathcal{K}_1$ we have $\{a, \neg a \wedge \neg b\} \models \perp$ and for $\mathcal{K}_2$ we have, e. g., $\{a, \neg a\} \models \perp$. These inconsistencies render the whole knowledge bases useless for reasoning if one wants to use classical reasoning techniques. In order to make the knowledge bases useful again, one can either rely on non-monotonic/paraconsistent reasoning techniques (Makinson, 2005; Priest, 1979) or one revises the knowledge bases appropriately to make them consistent (Hansson, 2001). Looking at the knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ one can observe that the *severity* of their inconsistency is different. In $\mathcal{K}_1$, only two out of four formulas (a and $\neg a \wedge \neg b$) are "participating" in making $\mathcal{K}_1$ inconsistent while for $\mathcal{K}_2$ all formulas contribute to its inconsistency. Furthermore, for $\mathcal{K}_1$ only two propositions (a and b) are conflicting and using e. g. paraconsistent reasoning one could still infer meaningful statements about c and d. For $\mathcal{K}_2$ no such statement can be made. This leads to the assessment that $\mathcal{K}_2$ should be regarded *more* inconsistent than $\mathcal{K}_1$.

Inconsistency measures can be used to analyse inconsistencies and to provide insights on how to repair them. An inconsistency measure $\mathcal{I}$ is a function on knowledge bases, such that the larger the value $\mathcal{I}(\mathcal{K})$ the more severe the inconsistency in $\mathcal{K}$. A lot of different approaches of inconsistency measures have been proposed, mostly for classical propositional logic (Hunter and Konieczny, 2004, 2008, 2010; Ma *et al.*, 2009; Mu *et al.*, 2011a; Xiao and Ma, 2012; Grant and Hunter, 2011, 2013; McAreavey *et al.*, 2014; Jabbour *et al.*, 2014b), but also for classical first-order logic (Grant and Hunter, 2008), description logics (Ma *et al.*, 2007; Zhou *et al.*, 2009), default logics (Doder *et al.*, 2010), and probabilistic and other weighted logics (Ma *et al.*, 2012; Thimm, 2013b; Potyka, 2014). Due to this plethora of inconsistency measures it is hard to determine which measure to use for an application and which measure is meaningful. Rationality postulates have been proposed that address the issue of assessing the quality of a measure—see e. g. (Hunter and Konieczny, 2006; Mu *et al.*, 2011a)—but many of these properties have been criticised to address only a specific point of view, see (Besnard, 2014) for a recent discussion on this topic.

In this paper, we take a different perspective on the evaluation of inconsistency measures by considering a *quantitative* analysis of their *expressivity*, that is, we study how many different (inconsistent) knowledge bases can be distinguished by a given inconsistency measure. By the term *expressivity* we here refer to the property of a semantical concept—here, an inconsistency measure—and its capability to distinguish syntactical constructs—here, knowledge bases—, similarly as it has been done for the analysis of expressivity of semantics for other logical languages, see e. g. skepticism relations for formal argumentation (Baroni and Giacomin, 2008). Our analysis is meant to complement the study on rationality postulates and is, of course, not meaningful on its own as the compliance of measures with the basic intuitions behind inconsistency measures can only be assessed by rationality postulates. However, we introduce expressivity of inconsistency measures as an *additional* method to evaluate their quality. In particular, we propose four different *expressivity characteristics* that quantify the relation between the number of different values of an inconsistency measure wrt. different notions of the size of the knowledge base, such as number of formulas or number of propositions. We conduct a thorough comparative analysis of 16 different inconsistency measures from the literature (Hunter and Konieczny, 2008, 2010; Grant and Hunter, 2011; Knight, 2002; Thimm, 2016; Grant and Hunter, 2013; Mu *et al.*, 2011a; Jabbour and Raddaoui, 2013; Xiao and Ma, 2012; Doder *et al.*, 2010) and classify these measures in a hierarchy of expressivity. In our study, we made several interesting observations, such as the relation between the measure $\mathcal{I}_{\mathsf{MI}}$ (Grant and Hunter, 2011) and Sperner families (Sperner, 1928) and of the measure $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$ (Grant and Hunter, 2011) with profiles of Boolean functions. One of our results is that the distance-based measure $\mathcal{I}^{\Sigma}_{\mathrm{dalal}}$ from (Grant and Hunter, 2013) and the proof-based measure $\mathcal{I}_{P_m}$ from (Jabbour and Raddaoui, 2013) have maximal expressivity along all considered characteristics.

In summary, the contributions of this paper are as follows:

1. We conduct a focused survey of 16 inconsistency measures from the recent literature (Section 3).
2. We propose four different expressivity characteristics, evaluate the considered inconsistency measures wrt. these characteristics, and study our findings (Section 4).
3. We classify the evaluated measures into hierarchies of expressivity and thus provide a means to quantitatively compare different measures (Section 5).

We give necessary preliminaries in Section 2 and provide a summary in Section 5. 6 contains proofs of technical results and 6 lists all example knowledge bases and families of knowledge bases used in the paper. All inconsistency measures discussed in this paper have been implemented and an online interface to try out these measures is available[1].

2 PRELIMINARIES

Let At be some fixed propositional signature, i. e., a (possibly infinite) set of propositions, and let $\mathcal{L}(\mathsf{At})$ be the corresponding propositional language constructed using the usual connectives $\wedge$ (*and*), $\vee$ (*or*), and $\neg$ (*negation*).

Definition 1. A knowledge base $\mathcal{K}$ is a finite set of formulas $\mathcal{K} \subseteq \mathcal{L}(\mathsf{At})$. Let $\mathbb{K}$ be the set of all knowledge bases.

If X is a formula or a set of formulas we write $\mathsf{At}(X)$ to denote the set of propositions appearing in X. Semantics to a propositional language is given by *interpretations* and an *interpretation* ω on At is a function $\omega : \mathsf{At} \rightarrow \{\mathsf{true}, \mathsf{false}\}$. Let $\mathsf{Int}(\mathsf{At})$ denote the set of all interpretations for At. An interpretation ω *satisfies* (or is a *model* of) a proposition $a \in \mathsf{At}$, denoted by $\omega \models a$, if and only if $\omega(a) = \mathsf{true}$. The satisfaction relation $\models$ is extended to formulas in the usual way.

As an abbreviation we sometimes identify an interpretation ω with its *complete conjunction*, i. e., if $a_1, \ldots, a_n \in \mathsf{At}$ are those propositions that are assigned true by ω and $a_{n+1}, \ldots, a_m \in \mathsf{At}$ are those propositions that are assigned false by ω we identify ω by $a_1 \ldots a_n \overline{a_{n+1}} \ldots \overline{a_m}$ (or any permutation of this). For example, the interpretation ω_1 on $\{a, b, c\}$ with $\omega(a) = \omega(c) = \mathsf{true}$ and $\omega(b) = \mathsf{false}$ is abbreviated by $a\bar{b}c$.

For $\Phi \subseteq \mathcal{L}(\mathsf{At})$ we also define $\omega \models \Phi$ if and only if $\omega \models \phi$ for every $\phi \in \Phi$. Define furthermore the set of models $\mathsf{Mod}(X) = \{\omega \in \mathsf{Int}(\mathsf{At}) \mid \omega \models X\}$ for every formula or set of formulas X. If $\mathsf{Mod}(X) = \emptyset$ we also write $X \models \perp$ and say that X is *inconsistent*.

1 http://tweetyproject.org/w/incmes/

3 INCONSISTENCY MEASURES

Let $\mathbb{R}^{\infty}_{\geq 0}$ be the set of non-negative real values including ∞. Inconsistency measures are functions $\mathcal{I} : \mathbb{K} \rightarrow \mathbb{R}^{\infty}_{\geq 0}$ that aim at assessing the severity of the inconsistency in a knowledge base $\mathcal{K}$, cf. (Grant and Hunter, 2011). The basic idea is that the larger the inconsistency in $\mathcal{K}$ the larger the value $\mathcal{I}(\mathcal{K})$ and $\mathcal{I}(\mathcal{K}) = 0$ if and only if $\mathcal{K}$ is consistent. However, inconsistency is a concept that is not easily quantified and there have been a couple of proposals for inconsistency measures so far, in particular for classical propositional logic, see e. g. (Besnard, 2014; McAreavey *et al.*, 2014; Jabbour *et al.*, 2014b; Hunter *et al.*, 2014) for some recent works. We selected 16 inconsistency measures from the literature in order to conduct our analysis on expressivity, taken from (Hunter and Konieczny, 2008, 2010; Grant and Hunter, 2011; Knight, 2002; Thimm, 2016; Grant and Hunter, 2013; Mu *et al.*, 2011a; Jabbour and Raddaoui, 2013; Xiao and Ma, 2012; Doder *et al.*, 2010). We briefly introduce these measures in this section for the sake of completeness, but we refer for a detailed explanation to the corresponding original papers.

To illustrate the different inconsistency measures we will use the knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ from the introduction as running examples.

Example 1. Let $\mathsf{At} = \{a, b, c, d\}$ and define knowledge bases $\mathcal{K}_1$ and $\mathcal{K}_2$ as follows:

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

A summary of the formal definitions of the considered inconsistency measures can be found in Table 1. We will discuss these measures in more detail below.

3.1 *The drastic inconsistency measure*

The basic motivation for measuring inconsistency is to provide a *graded* assessment of inconsistency and not only a "consistent"/"inconsistent" assessment. However, in order to evaluate more sophisticated measures and the usefulness of rationality postulates (cf. beginning of Section 4), the *drastic inconsistency measure* $\mathcal{I}_d$—that can only distinguish between consistent and inconsistent knowledge bases—is usually used as a baseline approach, cf. e. g. (Hunter and Konieczny, 2008).

Definition 2. The *drastic inconsistency measure* $\mathcal{I}_d : \mathbb{K} \rightarrow \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_d(\mathcal{K}) = \begin{cases} 1 & \text{if } \mathcal{K} \models \perp \\ 0 & \text{otherwise} \end{cases}$$

$$\mathcal{I}_d(\mathcal{K}) = \begin{cases} 1 & \text{if } \mathcal{K} \models \perp \\ 0 & \text{otherwise} \end{cases}$$

$$\mathcal{I}_{\mathsf{MI}}(\mathcal{K}) = |\mathsf{MI}(\mathcal{K})|$$

$$\mathcal{I}_{\mathsf{MI}^C}(\mathcal{K}) = \sum_{M \in \mathsf{MI}(\mathcal{K})} \frac{1}{|M|}$$

$$\mathcal{I}_\eta(\mathcal{K}) = 1 - \max\{\xi \mid \exists P \in \mathcal{P}(\mathsf{At}) : \forall \alpha \in \mathcal{K} : P(\alpha) \geq \xi\}$$

$$\mathcal{I}_c(\mathcal{K}) = \min\{|v^{-1}(B)| \mid v \models^3 \mathcal{K}\}$$

$$\mathcal{I}_{LP_m}(\mathcal{K}) = \mathcal{I}_c(\mathcal{K}) / |\mathsf{At}(\mathcal{K})|$$

$$\mathcal{I}_{mc}(\mathcal{K}) = |\mathtt{MC}(\mathcal{K})| + |\mathsf{SC}(\mathcal{K})| - 1$$

$$\mathcal{I}_p(\mathcal{K}) = |\bigcup_{M \in \mathsf{MI}(\mathcal{K})} M|$$

$$\mathcal{I}_{hs}(\mathcal{K}) = \min\{|H| \mid H \text{ is a hitting set of } \mathcal{K}\} - 1$$

$$\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}) = \min\{\sum_{\alpha \in \mathcal{K}} d_{\text{d}}(\mathsf{Mod}(\alpha), \omega) \mid \omega \in \mathsf{Int}(\mathsf{At})\}$$

$$\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}) = \min\{\max_{\alpha \in \mathcal{K}} d_{\text{d}}(\mathsf{Mod}(\alpha), \omega) \mid \omega \in \mathsf{Int}(\mathsf{At})\}$$

$$\mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K}) = \min\{|\{\alpha \in \mathcal{K} \mid d_{\text{d}}(\mathsf{Mod}(\alpha), \omega) > 0\}| \mid \omega \in \mathsf{Int}(\mathsf{At})\}$$

$$\mathcal{I}_{D_f}(\mathcal{K}) = 1 - \Pi_{i=1}^{|\mathcal{K}|}(1 - R_i(\mathcal{K})/i)$$

$$\mathcal{I}_{P_m}(\mathcal{K}) = \sum_{a \in \mathsf{At}} |P_m^{\mathcal{K}}(a)| \cdot |P_m^{\mathcal{K}}(\neg a)|$$

$$\mathcal{I}_{mv}(\mathcal{K}) = \frac{|\bigcup_{M \in \mathsf{MI}(\mathcal{K})} \mathsf{At}(M)|}{|\mathsf{At}(\mathcal{K})|}$$

$$\mathcal{I}_{nc}(\mathcal{K}) = |\mathcal{K}| - \max\{n \mid \forall \mathcal{K}' \subseteq \mathcal{K} : |\mathcal{K}'| = n \Rightarrow \mathcal{K}' \not\models \perp\}$$

Table 1: Definitions of the considered inconsistency measures

for $\mathcal{K} \in \mathbb{K}$.

In other words, $\mathcal{I}_d(\mathcal{K}) = 1$ if and only if $\mathcal{K}$ is inconsistent (and 0 otherwise).

Example 2. We continue Example 1 and consider

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

As both $\mathcal{K}_1$ and $\mathcal{K}_2$ are inconsistent we obtain $\mathcal{I}_d(\mathcal{K}_1) = \mathcal{I}_d(\mathcal{K}_2) = 1$.

3.2 *Inconsistency Measures based on Minimal Inconsistencies*

One approach to assess the severity of inconsistency in a knowledge base is to focus on its set of *minimal inconsistent subsets*. A set $M \subseteq \mathcal{K}$ is called *minimal inconsistent subset* (MI) of $\mathcal{K}$ if $M \models\perp$ and there is no $M' \subset M$ with $M' \models\perp$. Let $\mathsf{MI}(\mathcal{K})$ be the set of all MIs of $\mathcal{K}$. Informally speaking, minimal inconsistent subsets contain the *essence* of the inconsistency in a knowledge base. Every formula participating in creating an inconsistency is part of at least one minimal inconsistent subset.

A straightforward approach to use minimal inconsistent subsets for measuring inconsistency is to take their number as an indicator (Hunter and Konieczny, 2008).

Definition 3. The MI-*inconsistency measure* $\mathcal{I}_{\mathsf{MI}} : \mathbb{K} \rightarrow \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_{\mathsf{MI}}(\mathcal{K}) = |\mathsf{MI}(\mathcal{K})|$$

for $\mathcal{K} \in \mathbb{K}$.

One drawback of $\mathcal{I}_{\mathsf{MI}}$ is that it treats every inconsistent subset of $\mathcal{K}$ equally. A knowledge base with one minimal inconsistent subset of size 2 has the same inconsistency value as another knowledge base with one minimal inconsistent subset of size 10. It is usually acknowledged that a smaller minimal inconsistent subset is more severe than a larger one, cf. (Hunter and Konieczny, 2008). The MI^c *inconsistency measure* takes this into account and is defined as follows.

Definition 4. The MI^c-*inconsistency measure* $\mathcal{I}_{\mathsf{MI}^C} : \mathbb{K} \rightarrow \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_{\mathsf{MI}^C}(\mathcal{K}) = \sum_{M \in \mathsf{MI}(\mathcal{K})} \frac{1}{|M|}$$

for $\mathcal{K} \in \mathbb{K}$.

In other words, for every minimal inconsistent subset M of $\mathcal{K}$, $1/|M|$ is added up in order to obtain $\mathcal{I}_{\mathsf{MI}^C}(\mathcal{K})$. In this way, larger minimal inconsistent subset contribute less to the overall inconsistency value than smaller ones.

Example 3. We continue Example 1 and consider

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

Here we have

$$\mathsf{MI}(\mathcal{K}_1) = \{\{a, \neg a \wedge \neg b\}\}$$
$$\mathsf{MI}(\mathcal{K}_2) = \{\{a, \neg a\}, \{b, \neg b\}\}$$

Therefore we obtain $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_1) = 1$ and $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_2) = 2$ and

$$\mathcal{I}_{\mathsf{MI}^{\mathsf{C}}}(\mathcal{K}_1) = \frac{1}{|\{a, \neg a \wedge \neg b\}|} = \frac{1}{2}$$
$$\mathcal{I}_{\mathsf{MI}^{\mathsf{C}}}(\mathcal{K}_2) = \frac{1}{|\{a, \neg a\}|} + \frac{1}{|\{b, \neg b\}|} = 1$$

The measure $\mathcal{I}_{\mathsf{MI}^{\mathsf{C}}}$ has been further extended in (Mu *et al.*, 2011a) where not only the sizes of the different minimal inconsistent subsets but also their *distribution* in a knowledge base has been considered. For every knowledge base $\mathcal{K}$ and $i = 1, \ldots, |\mathcal{K}|$, define

$$\mathsf{MI}^{(i)}(\mathcal{K}) = \{M \in \mathsf{MI}(\mathcal{K}) \mid |M| = i\}$$
$$\mathsf{CN}^{(i)}(\mathcal{K}) = \{C \subseteq \mathcal{K} \mid |C| = i \wedge C \not\models \perp\}$$

That is, $\mathsf{MI}^{i}(\mathcal{K})$ is the set of minimal inconsistent subsets of $\mathcal{K}$ of size i and $\mathsf{CN}^{(i)}(\mathcal{K})$ is the set of consistent subsets of $\mathcal{K}$ of size i. Furthermore define

$$R_i(\mathcal{K}) = \begin{cases} 0 & \text{if } |\mathsf{MI}^{(i)}(\mathcal{K})| + |\mathsf{CN}^{(i)}(\mathcal{K})| = 0 \\ |\mathsf{MI}^{(i)}(\mathcal{K})| / (|\mathsf{MI}^{(i)}(\mathcal{K})| + |\mathsf{CN}^{(i)}(\mathcal{K})|) & \text{otherwise} \end{cases}$$

for $i = 1, \ldots, |\mathcal{K}|$. The value $R_i(\mathcal{K})$ thus gives the ratio of minimal inconsistent sets of size i to the number of minimal inconsistent and consistent subsets of size i. Obviously, if for two knowledge bases $\mathcal{K}$ and $\mathcal{K}'$, a value $R_i(\mathcal{K})$ is larger than $R_i(\mathcal{K}')$ (for some i and everything else being equal) than $\mathcal{K}$ should be regarded more inconsistent than $\mathcal{K}'$. The idea of the approach of (Mu *et al.*, 2011a) is to weigh these values also wrt. the sizes, i. e., a large $R_i(\mathcal{K})$ has more impact than a large $R_j(\mathcal{K})$, with $j > i$.

Definition 5. The *D_f-inconsistency measure* $\mathcal{I}_{D_f} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_{D_f}(\mathcal{K}) = 1 - \Pi_{i=1}^{|\mathcal{K}|}(1 - R_i(\mathcal{K})/i)$$

for $\mathcal{K} \in \mathbb{K}$.

Note that the above definition of $\mathcal{I}_{D_f}$ represents only a single instance of the family introduced in (Mu *et al.*, 2011a). Other variants can be obtained by other ways of aggregating the values $R_1(\mathcal{K}), \ldots, R_{|\mathcal{K}|}(\mathcal{K})$.

Example 4. We continue Example 3 and recall

$$\mathsf{MI}(\mathcal{K}_1) = \{\{a, \neg a \wedge \neg b\}\}$$
$$\mathsf{MI}(\mathcal{K}_2) = \{\{a, \neg a\}, \{b, \neg b\}\}$$

Then we have

$$\mathsf{MI}^{(2)}(\mathcal{K}_1) = \{\{a, \neg a \wedge \neg b\}\} \qquad \text{and } \mathsf{MI}^{(i)}(\mathcal{K}_1) = \emptyset \quad \text{for } i \neq 2$$
$$\mathsf{MI}^{(2)}(\mathcal{K}_2) = \{\{a, \neg a\}, \{b, \neg b\}\} \qquad \text{and } \mathsf{MI}^{(i)}(\mathcal{K}_2) = \emptyset \quad \text{for } i \neq 2$$

Furthermore, we have

$$\mathsf{CN}^{(1)}(\mathcal{K}_1) = \{\{a\}, \{b \vee c\}, \{\neg a \wedge \neg b\}, \{d\}\}$$
$$\mathsf{CN}^{(2)}(\mathcal{K}_1) = \{\{a, b \vee c\}, \{a, d\}, \{b \vee c, \neg a \wedge \neg b\}, \{b \vee c, d\}, \{\neg a \wedge b, d\}\}$$
$$\mathsf{CN}^{(3)}(\mathcal{K}_1) = \{\{a, b \vee c, d\}, \{b \vee c, \neg a \wedge \neg b, d\}\}$$
$$\mathsf{CN}^{(4)}(\mathcal{K}_1) = \emptyset$$

and

$$\mathsf{CN}^{(1)}(\mathcal{K}_2) = \{\{a\}, \{\neg a\}, \{b\}, \{\neg b\}\}$$
$$\mathsf{CN}^{(2)}(\mathcal{K}_2) = \{\{a, b\}, \{a, \neg b\}, \{\neg a, b\}, \{\neg a, \neg b\}\}$$
$$\mathsf{CN}^{(3)}(\mathcal{K}_2) = \emptyset$$
$$\mathsf{CN}^{(4)}(\mathcal{K}_2) = \emptyset$$

yielding

$$R_1(\mathcal{K}_1) = 0 \qquad R_2(\mathcal{K}_1) = 1/6 \qquad R_3(\mathcal{K}_1) = 0 \qquad R_4(\mathcal{K}_1) = 0$$

and

$$R_1(\mathcal{K}_2) = 0 \qquad R_2(\mathcal{K}_2) = 2/6 \qquad R_3(\mathcal{K}_2) = 0 \qquad R_4(\mathcal{K}_2) = 0$$

Finally, we obtain

$$\begin{aligned}\mathcal{I}_{D_f}(\mathcal{K}_1) &= 1 - (1 - R_1(\mathcal{K}_1))(1 - R_2(\mathcal{K}_1)/2)(1 - R_3(\mathcal{K}_1)/3)(1 - R_4(\mathcal{K}_1)/4) \\ &= \frac{1}{6 \cdot 2} = \frac{1}{12} \\ \mathcal{I}_{D_f}(\mathcal{K}_2) &= 1 - (1 - R_1(\mathcal{K}_2))(1 - R_2(\mathcal{K}_2)/2)(1 - R_3(\mathcal{K}_2)/3)(1 - R_4(\mathcal{K}_2)/4) \\ &= \frac{2}{6 \cdot 2} = \frac{1}{6}\end{aligned}$$

Another simple approach for utilizing minimal inconsistent subsets is to use the number of formulas occurring in some minimal inconsistent subsets—that is, the number of *problematic* formulas—as the inconsistency value, cf. e. g. (Grant and Hunter, 2011).

Definition 6. The *problematic inconsistency measure* $\mathcal{I}_p : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_p(\mathcal{K}) = |\bigcup_{M \in \mathsf{MI}(\mathcal{K})} M|$$

for $\mathcal{K} \in \mathbb{K}$.

Example 5. We continue Example 3 and recall

$$\begin{aligned}\mathsf{MI}(\mathcal{K}_1) &= \{\{a, \neg a \wedge \neg b\}\}\\ \mathsf{MI}(\mathcal{K}_2) &= \{\{a, \neg a\}, \{b, \neg b\}\}\end{aligned}$$

Then $\mathcal{I}_p(\mathcal{K}_1) = 2$ and $\mathcal{I}_p(\mathcal{K}_2) = 4$.

3.3 *Inconsistency Measures based on Maximal Consistency*

Another family closely related to the family of measures based on minimal inconsistent subsets is the one based on maximal consistent subsets. Let $\mathsf{MC}(\mathcal{K})$ be the set of maximal consistent subsets of $\mathcal{K}$, i. e.

$$\mathsf{MC}(\mathcal{K}) = \{\mathcal{K}' \subseteq \mathcal{K} \mid \mathcal{K}' \not\models \perp \wedge \forall \mathcal{K}'' \supsetneq \mathcal{K}' : \mathcal{K}'' \models \perp\}$$

Note, that a maximal consistent subset can be obtained by removing one formula from each minimal inconsistent subset from the knowledge base. Thus, the number of maximal consistent subsets and the number of minimal inconsistent sets correlate.

Furthermore, let $\mathsf{SC}(\mathcal{K})$ be the set of self-contradictory formulas of $\mathcal{K}$, i. e.

$$\mathsf{SC}(\mathcal{K}) = \{\phi \in \mathcal{K} \mid \phi \models \perp\}$$

An inconsistency measure that takes both maximal consistent subsets and self-contradictory formulas into account can be defined as follows, cf. e. g. (Grant and Hunter, 2011).

Definition 7. The $\mathtt{MC}$-*inconsistency measure* $\mathcal{I}_{mc} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_{mc}(\mathcal{K}) = |\mathtt{MC}(\mathcal{K})| + |\mathsf{SC}(\mathcal{K})| - 1$$

for $\mathcal{K} \in \mathbb{K}$.

Note that the subtraction of 1 in the definition of $\mathcal{I}_{mc}$ is to ensure that a consistent knowledge base has inconsistency value 0 (a consistent knowledge base has one maximal consistent subset, itself, and no self-contradictory formulas).

Another approach utilizing the idea of maximum consistency is the approach of (Doder *et al.*, 2010).

Definition 8. The *nc-inconsistency measure* $\mathcal{I}_{nc} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_{nc}(\mathcal{K}) = |\mathcal{K}| - \max\{n \mid \forall \mathcal{K}' \subseteq \mathcal{K} : |\mathcal{K}'| = n \Rightarrow \mathcal{K}' \not\models \perp\}$$

for $\mathcal{K} \in \mathbb{K}$.

In other words, the inconsistency of $\mathcal{K}$ is assessed by seeking a maximal value $n \in \{1, \ldots, |\mathcal{K}|\}$ such that all subsets of size n of $\mathcal{K}$ are consistent. The larger this value n, the smaller the inconsistency. Note that the above definition of $\mathcal{I}_{nc}$ differs from the original definition in (Doder *et al.*, 2010) (where only the max-term was considered) in order to ensure that consistent knowledge bases receive a value of zero and the inconsistency value increases with increasing inconsistency.

Example 6. We continue Example 1 and consider

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

Here we have

$$\mathtt{MC}(\mathcal{K}_1) = \{\{a, b \vee c, d\}, \{b \vee c, \neg a \wedge \neg b, d\}\}$$
$$\mathtt{MC}(\mathcal{K}_2) = \{\{a, b\}, \{a, \neg b\}, \{\neg a, b\}, \{\neg a, \neg b\}\}$$

and $\mathsf{SC}(\mathcal{K}_1) = \mathsf{SC}(\mathcal{K}_2) = \emptyset$. Therefore we obtain

$$\mathcal{I}_{mc}(\mathcal{K}_1) = 1$$
$$\mathcal{I}_{mc}(\mathcal{K}_2) = 3$$

Furthermore, note that for both $\mathcal{K}_1$ and $\mathcal{K}_2$ we can find subsets of size 2 that are inconsistent: $\{a, \neg a \wedge \neg b\}$ for $\mathcal{K}_1$ and $\{a, \neg a\}$ for $\mathcal{K}_2$. Furthermore, all one-element subsets of $\mathcal{K}_1$ and $\mathcal{K}_2$ are consistent, respectively. Therefore, we obtain

$$\mathcal{I}_{nc}(\mathcal{K}_1) = 3$$
$$\mathcal{I}_{nc}(\mathcal{K}_2) = 3$$

3.4 *Probabilistic Inconsistency Measures*

One of the first approaches to measuring inconsistency is Knight's measure $\mathcal{I}_{\eta}$, which is based on probability functions over the underlying propositional language (Knight, 2002). Recall that $\mathsf{Int}(\mathsf{At})$ is the set of interpretations of the propositional language $\mathcal{L}_(\mathsf{At})$. A *probability function P* on $\mathcal{L}_(\mathsf{At})$

is a function $P : \mathsf{Int}(\mathsf{At}) \to [0,1]$ with $\sum_{\omega \in \mathsf{Int}(\mathsf{At})} P(\omega) = 1$. We extend P to assign a probability to any formula $\phi \in \mathcal{L}_(\mathsf{At})$ by defining

$$P(\phi) = \sum_{\omega \models \phi} P(\omega)$$

Let $\mathcal{P}(\mathsf{At})$ be the set of all those probability functions. The idea of (Knight, 2002) is to seek a probability function that maximizes the probability of each formula of a knowledge base $\mathcal{K}$. If we can find a probability function that assigns probability 1 to each formula this means that the knowledge base is consistent. If the knowledge base is inconsistent, then the probability mass must be distributed (recall that an inconsistent set of formulas cannot be satisfied by a single interpretation ω; thus the probability $P(\omega)$ can only be associated with a subset of this set). Therefore, the smaller the maximal probability that can be assigned to all formulas the more inconsistent the knowledge base.

Definition 9. The *η-inconsistency measure* $\mathcal{I}_\eta : \mathbb{K} \to \mathbb{R}^\infty_{\geq 0}$ is defined as

$$\mathcal{I}_\eta(\mathcal{K}) = 1 - \max\{\xi \mid \exists P \in \mathcal{P}(\mathsf{At}) : \forall \alpha \in \mathcal{K} : P(\alpha) \geq \xi\}$$

for $\mathcal{K} \in \mathbb{K}$.

Note that we modified the definition of $\mathcal{I}_\eta$ slightly compared to the original definition in order to ensure that consistent knowledge bases receive an inconsistency value of zero.

Instead of seeking a probability function to maximize the probabilities of the formulas in $\mathcal{K}$, one can simplify this idea and seek only a minimal set of interpretations that need to receive a positive probability in order to ensure that every formula has a positive probability (Thimm, 2016). In other words, a subset $H \subseteq \mathsf{Int}(\mathsf{At})$ is called a *hitting set* of $\mathcal{K}$ if for every $\phi \in \mathcal{K}$ there is $\omega \in H$ with $\omega \models \phi$. Focusing only on minimizing the number of interpretations needed to form a hitting set we can define another measure as follows.

Definition 10. The *hitting-set inconsistency measure* $\mathcal{I}_{hs} : \mathbb{K} \to \mathbb{R}^\infty_{\geq 0}$ is defined as

$$\mathcal{I}_{hs}(\mathcal{K}) = \min\{|H| \mid H \text{ is a hitting set of } \mathcal{K}\} - 1$$

for $\mathcal{K} \in \mathbb{K}$ with $\min \emptyset = \infty$.

Note that $\mathcal{I}_{hs}(\mathcal{K}) = \infty$ if and only if $\mathcal{K}$ contains a self-contradictory formula, i. e., $\alpha \in \mathcal{K}$ with $\alpha \models \perp$. In this case, no hitting set of $\mathcal{K}$ exists.

Example 7. We continue Example 1 and consider

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

Consider the probability function $P_1 \in \mathcal{P}(\{a, b, c, d\})$ defined via

$$P_1(abcd) = P_1(\overline{a}\overline{b}cd) = 0.5$$
$$P_1(\omega) = 0 \quad \text{for } \omega \in \mathsf{Int}(\{a, b, c, d\}) \setminus \{abcd, \overline{a}\overline{b}cd\}$$

Then we obtain

$$P_1(a) = P_1(\neg a \wedge \neg b) = 0.5$$
$$P_1(b \vee c) = P_1(d) = 1$$

and thus $P_1(\phi) \geq 0.5$ for all $\phi \in \mathcal{K}_1$. Furthermore, there can be no other P' that assigns larger probability to all $\phi \in \mathcal{K}_1$. Hence, we have $\mathcal{I}_\eta(\mathcal{K}_1) = 1 - 0.5 = 0.5$. The function P_1 can also be used to determine $\mathcal{I}_\eta(\mathcal{K}_2) = 0.5$.

The set $H_1 = \{abcd, \overline{a}\overline{b}cd\}$ is also a hitting set of both $\mathcal{K}_1$ and $\mathcal{K}_2$ and there is no smaller set that is a hitting set. Therefore we obtain $\mathcal{I}_{hs}(\mathcal{K}_1) = \mathcal{I}_{hs}(\mathcal{K}_2) = 1$.

3.5 *Variable-based Inconsistency Measures*

Another approach to assess the severity of inconsistency is to take the number of propositions from At that participate in the inconsistency. The approach of (Xiao and Ma, 2012) is to take the ratio of the propositions appearing in a minimal inconsistent subset wrt. the total number of propositions as the inconsistency value.

Definition 11. The *mv inconsistency measure* $\mathcal{I}_{mv} : \mathbb{K} \to \mathbb{R}^\infty_{\geq 0}$ is defined as

$$\mathcal{I}_{mv}(\mathcal{K}) = \frac{|\bigcup_{M \in \mathsf{MI}(\mathcal{K})} \mathsf{At}(M)|}{|\mathsf{At}(\mathcal{K})|}$$

for $\mathcal{K} \in \mathbb{K}$.

In other words, $\mathcal{I}_{mv}(\mathcal{K})$ is the ratio of the signature involved in minimal inconsistent subsets.

Instead of utilizing minimal inconsistent subsets one can also using paraconsistent semantics to identify the part of the signature involved in inconsistency. In this paper, we will only consider the *contension measure* $\mathcal{I}_c$—cf. e. g. (Grant and Hunter, 2011)—and its normalized variant $\mathcal{I}_{LP_m}$ from (Hunter and Konieczny, 2010) as representatives of this family of measures.

α	β	$v(\alpha \wedge \beta)$	$v(\alpha \vee \beta)$
T	T	T	T
T	B	B	T
T	F	F	T
B	T	B	T
B	B	B	B
B	F	F	B
F	T	F	T
F	B	F	B
F	F	F	F

α	$v(\neg\alpha)$
T	F
B	B
F	T

Table 2: Truth tables for propositional three-valued logic (Priest, 1979).

Similar approaches relying on the same ideas can be found in e. g. (Ma *et al.*, 2007, 2011).

The contension measure $\mathcal{I}_c$ utilizes three-valued interpretations for propositional logic (Priest, 1979). A three-valued interpretation v on At is a function $v : \mathsf{At} \to \{T, F, B\}$ where the values T and F correspond to the classical true and false, respectively. The additional truth value B stands for *both* and is meant to represent a conflicting truth value for a proposition. The function v is extended to arbitrary formulas as shown in Table 2. Then, an interpretation v satisfies a formula α, denoted by $v \models^3 \alpha$ if either $v(\alpha) = T$ or $v(\alpha) = B$. Then inconsistency can be measured by seeking an interpretation v that assigns B to a minimal number of propositions.

Definition 12. The *contension inconsistency measure* $\mathcal{I}_c : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_c(\mathcal{K}) = \min\{|v^{-1}(B)| \mid v \models^3 \mathcal{K}\}$$

for $\mathcal{K} \in \mathbb{K}$.

In (Hunter and Konieczny, 2010) a variant $\mathcal{I}_{LP_m}$ of this measure was defined that further normalizes the inconsistency value by the number of propositions appearing in $\mathcal{K}$.

Definition 13. The *normalized contension inconsistency measure* $\mathcal{I}_{LP_m} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_{LP_m}(\mathcal{K}) = \frac{\mathcal{I}_c(\mathcal{K})}{|\mathsf{At}(\mathcal{K})|}$$

for $\mathcal{K} \in \mathbb{K}$.

Example 8. We continue Example 1 and consider

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

and recall

$$\mathsf{MI}(\mathcal{K}_1) = \{\{a, \neg a \wedge \neg b\}\}$$
$$\mathsf{MI}(\mathcal{K}_2) = \{\{a, \neg a\}, \{b, \neg b\}\}$$

Then we have

$$\mathcal{I}_{mv}(\mathcal{K}_1) = \frac{|\{a,b\}|}{|\{a,b,c,d\}|} = 1/2$$
$$\mathcal{I}_{mv}(\mathcal{K}_2) = \frac{|\{a,b\}|}{|\{a,b\}|} = 1$$

Furthermore, consider $v_1 : \{a,b,c,d\} \to \{T,F,B\}$ defined via

$$v_1(a) = B \qquad v_1(b) = F \qquad v_1(c) = v_1(d) = T$$

Then $v_1 \models^3 \phi$ for all $\phi \in \mathcal{K}_1$ and there is no other v' that assigns B to fewer propositions, yielding $\mathcal{I}_c(\mathcal{K}_1) = 1$ and $\mathcal{I}_{LP_m}(\mathcal{K}_1) = 1/4$. For $v_2 : \{a,b\} \to \{T,F,B\}$ defined via

$$v_2(a) = v_2(b) = B$$

we have $v_2 \models^3 \phi$ for all $\phi \in \mathcal{K}_2$ and there is no other v' that assigns B to fewer propositions, yielding $\mathcal{I}_c(\mathcal{K}_2) = 2$ and $\mathcal{I}_{LP_m}(\mathcal{K}_2) = 2/2 = 1$.

3.6 *Distance-based Inconsistency Measures*

In (Grant and Hunter, 2013) three families of inconsistency measures are defined that are based on a notion of distance to consistency. More precisely, an *interpretation distance* d is a function $d : \mathsf{Int}(\mathsf{At}) \times \mathsf{Int}(\mathsf{At}) \to [0, \infty)$ that satisfies (let $\omega, \omega', \omega'' \in \mathsf{Int}(\mathsf{At})$)

1. $d(\omega, \omega') = 0$ if and only if $\omega = \omega'$ (*reflexivity*),
2. $d(\omega, \omega') = d(\omega', \omega)$ (*symmetry*), and
3. $d(\omega, \omega'') \leq d(\omega, \omega') + d(\omega', \omega'')$ (*triangle inequality*).

One prominent example of such a distance is the *Dalal distance* d_{d} defined via

$$d_{\mathrm{d}}(\omega, \omega') = |\{a \in \mathsf{At} \mid \omega(a) \neq \omega'(a)\}|$$

for all $\omega, \omega' \in \mathsf{Int}(\mathsf{At})$. In other words, $d_d(\omega, \omega')$ is the number of propositions where ω and ω' assign different truth values. If $X \subseteq \mathsf{Int}(\mathsf{At})$ is a set of interpretations we define $d_\mathrm{d}(X, \omega) = \min_{\omega' \in X} d_\mathrm{d}(\omega', \omega)$ (if $X = \emptyset$ we define $d_\mathrm{d}(X, \omega) = \infty$). While (Grant and Hunter, 2013) consider arbitrary distances, we will focus here on the Dalal distance for reasons of simplicity.

The basic idea of the approaches in (Grant and Hunter, 2013) is to measure and aggregate the distances of the models of the formulas in a knowledge base $\mathcal{K}$. For example, if a knowledge base $\mathcal{K}$ has two formulas and their models have a large distance to each other, then $\mathcal{K}$ should be regarded as more inconsistent compared to a knowledge base $\mathcal{K}'$ with two formulas, where this is not the case. More precisely, the first approach from (Grant and Hunter, 2013) considered here seeks an interpretation ω such that the sum of all distances of the sets of models to ω is minimal.

Definition 14. The Σ-*distance inconsistency measure* $\mathcal{I}^{\Sigma}_{\mathrm{dalal}} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}^{\Sigma}_{\mathrm{dalal}}(\mathcal{K}) = \min \left\{ \sum_{\alpha \in \mathcal{K}} d_\mathrm{d}(\mathsf{Mod}(\alpha), \omega) \mid \omega \in \mathsf{Int}(\mathsf{At}) \right\}$$

for $\mathcal{K} \in \mathbb{K}$.

Another approach of (Grant and Hunter, 2013) is to seek an interpretation ω such that the maximum distance of the sets of models is minimal.

Definition 15. The max-*distance inconsistency measure* $\mathcal{I}^{\max}_{\mathrm{dalal}} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}^{\max}_{\mathrm{dalal}}(\mathcal{K}) = \min \left\{ \max_{\alpha \in \mathcal{K}} d_\mathrm{d}(\mathsf{Mod}(\alpha), \omega) \mid \omega \in \mathsf{Int}(\mathsf{At}) \right\}$$

for $\mathcal{K} \in \mathbb{K}$.

The final approach of (Grant and Hunter, 2013) is to minimize the number of formulas, where their corresponding sets of models have a positive distance.

Definition 16. The *hit-distance inconsistency measure* $\mathcal{I}^{\mathrm{hit}}_{\mathrm{dalal}} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}^{\mathrm{hit}}_{\mathrm{dalal}}(\mathcal{K}) = \min \left\{ |\{\alpha \in \mathcal{K} \mid d_\mathrm{d}(\mathsf{Mod}(\alpha), \omega) > 0\}| \mid \omega \in \mathsf{Int}(\mathsf{At}) \right\}$$

for $\mathcal{K} \in \mathbb{K}$.

Example 9. We continue Example 1 and consider

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

Observe that for the interpretation $\omega_1 = a\bar{b}cd \in \mathsf{Int}(\{a,b,c,d\})$ we have

$$\begin{aligned} d_{\mathrm{d}}(\mathsf{Mod}(a), \omega_1) &= 0 \\ d_{\mathrm{d}}(\mathsf{Mod}(b \vee c), \omega_1) &= 0 \\ d_{\mathrm{d}}(\mathsf{Mod}(\neg a \wedge \neg b), \omega_1) &= 1 \\ d_{\mathrm{d}}(\mathsf{Mod}(d), \omega_1) &= 0 \end{aligned}$$

and therefore $\sum_{\alpha \in \mathcal{K}_1} d_{\mathrm{d}}(\mathsf{Mod}(\alpha), \omega_1) = 1$. There is no other interpretation ω' with a smaller total distance, so we have $\mathcal{I}^{\Sigma}_{\mathrm{dalal}}(\mathcal{K}_1) = 1$. Furthermore, we have $\max_{\alpha \in \mathcal{K}_1} d_{\mathrm{d}}(\mathsf{Mod}(\alpha), \omega_1) = 1$ and there is also no other interpretation ω' with a smaller maximum distance. Hence, we have $\mathcal{I}^{\max}_{\mathrm{dalal}}(\mathcal{K}_1) = 1$ and similarly $\mathcal{I}^{\mathrm{hit}}_{\mathrm{dalal}}(\mathcal{K}_1) = 1$. For $\mathcal{K}_2$ we obtain

$$\begin{aligned} &\mathcal{I}^{\Sigma}_{\mathrm{dalal}}(\mathcal{K}_2) = 2 \\ &\mathcal{I}^{\max}_{\mathrm{dalal}}(\mathcal{K}_2) = 1 \end{aligned}$$

with a similar argumentation as above. For $\mathcal{I}^{\mathrm{hit}}_{\mathrm{dalal}}(\mathcal{K}_2)$ observe that every interpretation ω must always falsify exactly one formula in $\{a, \neg a\}$ and exactly one formula in $\{b, \neg b\}$. Therefore we obtain $\mathcal{I}^{\mathrm{hit}}_{\mathrm{dalal}}(\mathcal{K}_2) = 2$.

3.7 *Proof-based Inconsistency Measures*

The final measure we consider in this paper is the proof-based measure from (Jabbour and Raddaoui, 2013). The basic idea is to count, for all propositions $x \in \mathsf{At}$, both the number of *minimal proofs* for x and its negation $\neg x$. If there are many proofs for both, this indicates a large inconsistency in the knowledge base. In the context of (Jabbour and Raddaoui, 2013) a *minimal proof* for $\alpha \in \{x, \neg x \mid x \in \mathsf{At}\}$ in $\mathcal{K}$ is a set $\pi \subseteq \mathcal{K}$ such that

1. α appears as a literal in π
2. $\pi \models \alpha$, and
3. π is minimal wrt. set inclusion.

Note that this definition does not require that π is consistent. In particular, the set $\{a \wedge \neg a\}$ is a minimal proof for both a and $\neg a$. Note furthermore, that item 1.) requires that α appears in the exact same form in π, e.g. a appears in $a \wedge b$ but not in $\neg a \wedge b$ (this is a syntactic criterion).

Let $P^{\mathcal{K}}_m(x)$ be the set of all minimal proofs of x in $\mathcal{K}$. The proof-based measure of (Jabbour and Raddaoui, 2013) can then be defined by summing up the products of the number of minimal proofs for complementary literals.

Definition 17. The *proof-based inconsistency measure* $\mathcal{I}_{P_m} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ is defined as

$$\mathcal{I}_{P_m}(\mathcal{K}) = \sum_{a \in \mathsf{At}} |P^{\mathcal{K}}_m(a)| \cdot |P^{\mathcal{K}}_m(\neg a)|$$

for $\mathcal{K} \in \mathbb{K}$.

Note that the definition of $\mathcal{I}_{P_m}$ is not the original definition but a characterisation also provided in (Jabbour and Raddaoui, 2013).

Example 10. We continue Example 1 and consider

$$\mathcal{K}_1 = \{a, b \vee c, \neg a \wedge \neg b, d\}$$
$$\mathcal{K}_2 = \{a, \neg a, b, \neg b\}$$

Observe that

$$\begin{aligned}
P_m^{\mathcal{K}_1}(a) &= \{\{a\}\} \\
P_m^{\mathcal{K}_1}(\neg a) &= \{\{\neg a \wedge \neg b\}\} \\
P_m^{\mathcal{K}_1}(b) &= \{\{a, b \vee c, \neg a \wedge \neg b\}\} \\
P_m^{\mathcal{K}_1}(\neg b) &= \{\{\neg a \wedge \neg b\}\} \\
P_m^{\mathcal{K}_1}(c) &= \{\{a, b \vee c, \neg a \wedge \neg b\}\} \\
P_m^{\mathcal{K}_1}(\neg c) &= \emptyset \\
P_m^{\mathcal{K}_1}(d) &= \{\{d\}\} \\
P_m^{\mathcal{K}_1}(\neg d) &= \emptyset
\end{aligned}$$

and

$$\begin{aligned}
P_m^{\mathcal{K}_2}(a) &= \{\{a\}\} \\
P_m^{\mathcal{K}_2}(\neg a) &= \{\{\neg a\}\} \\
P_m^{\mathcal{K}_2}(b) &= \{\{b\}\} \\
P_m^{\mathcal{K}_2}(\neg b) &= \{\{\neg b\}\}
\end{aligned}$$

It follows that

$$\mathcal{I}_{P_m}(\mathcal{K}_1) = 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 0 + 1 \cdot 0 = 2$$
$$\mathcal{I}_{P_m}(\mathcal{K}_2) = 1 \cdot 1 + 1 \cdot 1 = 2$$

4 EXPRESSIVITY CHARACTERISTICS

In the literature, inconsistency measures are usually analytically evaluated on a set of *rationality postulates*.[2] Some basic example postulates given in (Hunter and Konieczny, 2006) are the following (let $\mathcal{I}$ be any inconsistency measure)

2 Some few works also consider empirical evaluation on computational performance and accuracy of algorithms approximating existing inconsistency measures, see e. g. (Ma *et al.*, 2009; McAreavey *et al.*, 2014; Thimm, 2016)

Consistency $\mathcal{I}(\mathcal{K}) = 0$ if and only if $\mathcal{K}$ is consistent

Monotony if $\mathcal{K} \subseteq \mathcal{K}'$ then $\mathcal{I}(\mathcal{K}) \leq \mathcal{I}(\mathcal{K}')$

Independence for all $\alpha \in \mathcal{K}$, if $\alpha \notin M$ for every $M \in \mathsf{MI}(\mathcal{K})$ then $\mathcal{I}(\mathcal{K}) = \mathcal{I}(\mathcal{K} \setminus \{\alpha\})$

Satisfaction of the property *consistency* ensures that all consistent knowledge bases receive a minimal inconsistency value and every inconsistent knowledge base receives a positive inconsistency value (we already implicitly required satisfaction of this postulate in the definition of an inconsistency measure). The postulate *monotony* states that the value of inconsistency can only increase when adding new information. *Independence* states that removing "harmless" formulas from a knowledge base does not change the value of inconsistency. Besides these three postulates a series of other postulates have been proposed in the literature, see e. g. (Hunter and Konieczny, 2006; Mu *et al.*, 2011a; Besnard, 2014). However, some of these postulates are disputed as each of them usually covers only a single aspect of inconsistency, such as *independence*, which focuses on the role of minimal inconsistent subsets. An excellent discussion on the rationality of various postulates for inconsistency measures can be found in (Besnard, 2014). Besides Besnard, several other authors have also criticised the rationality of individual postulates—discussions can be found in almost all papers cited before—and so there is some disagreement on which postulates are meaningful and which are not. One the one hand this calls for more work on rationality postulates and, on the other hand, it also suggests to investigate additional means for comparison. In the following, we propose a novel quantitative approach to evaluate and compare inconsistency measures that aims at complementing the existing approach of rationality postulates.

The drastic inconsistency measure $\mathcal{I}_d$ (see Table 1) is usually considered as a very naive baseline approach for inconsistency measurement. Surprisingly, this measure already satisfies many rationality postulates such as *consistency*, *monotony*, and *independence* (the proofs are straightforward). What sets it apart from other more sophisticated inconsistency measures is that it cannot differentiate between different inconsistent knowledge bases. However, this demand is exactly what inconsistency measures are supposed to satisfy. While the *qualitative* behaviour of inconsistency measures is being discussed quite deeply using rationality postulates, their *quantitative* properties in terms of *expressivity* have been almost neglected so far.[3] With expressivity of inconsistency measures we here mean the number of different values an inconsistency measure can attain. We investigate the expressivity of inconsistency measures along four different dimensions of subclasses of knowledge bases.

3 Some few rationality postulates such as *Attenuation* (Mu *et al.*, 2011a) are addressing this issue only in some very limited form and from a particular point of view.

Definition 18. Let ϕ be a formula. The *length* $\mathsf{l}(\phi)$ of ϕ is recursively defined as

$$\mathsf{l}(\phi) = \begin{cases} 1 & \text{if } \phi \in \mathsf{At} \\ 1 + \mathsf{l}(\phi') & \text{if } \phi = \neg\phi' \\ 1 + \mathsf{l}(\phi_1) + \mathsf{l}(\phi_2) & \text{if } \phi = \phi_1 \wedge \phi_2 \\ 1 + \mathsf{l}(\phi_1) + \mathsf{l}(\phi_2) & \text{if } \phi = \phi_1 \vee \phi_2 \end{cases}$$

Definition 19. Define the following subclasses of the set of all knowledge bases $\mathbb{K}$:

$$\begin{aligned} \mathbb{K}^v(n) &= \{\mathcal{K} \in \mathbb{K} \mid |\mathsf{At}(\mathcal{K})| \leq n\} \\ \mathbb{K}^f(n) &= \{\mathcal{K} \in \mathbb{K} \mid |\mathcal{K}| \leq n\} \\ \mathbb{K}^l(n) &= \{\mathcal{K} \in \mathbb{K} \mid \forall\phi \in \mathcal{K} : \mathsf{l}(\phi) \leq n\} \\ \mathbb{K}^p(n) &= \{\mathcal{K} \in \mathbb{K} \mid \forall\phi \in \mathcal{K} : |\mathsf{At}(\phi)| \leq n\} \end{aligned}$$

In other words, $\mathbb{K}^v(n)$ is the set of all knowledge bases that mention at most n different propositions, $\mathbb{K}^f(n)$ is the set of all knowledge bases that contain at most n formulas, $\mathbb{K}^l(n)$ is the set of all knowledge bases that contain only formulas with maximal length n, and $\mathbb{K}^p(n)$ is the set of all knowledge bases that contain only formulas that mention at most n different propositions each. The motivation for considering these particular subclasses of knowledge bases is that each of them considers a different aspect of the size of a knowledge base. As a syntactical object, a knowledge base is a set of formulas, and both the number of formulas (considered by the class $\mathbb{K}^f(n)$) and the length of each formula ($\mathbb{K}^l(n)$) are the essential parameters that define its size. From a semantical point of view, the number of propositions appearing in each formula ($\mathbb{K}^p(n)$) and in the complete knowledge base ($\mathbb{K}^v(n)$) define the scope of the knowledge. Larger numbers for both of them also indicate larger scope and thus greater size. Inconsistency measures should adhere to the size of the knowledge base in terms of their expressivity. For example, the number of possible inconsistency values of a particular measure should not decrease when moving from a set $\mathbb{K}^v(n)$ to set $\mathbb{K}^v(n')$ with $n' > n$, as knowledge bases with n' formulas should provide a larger variety in terms of inconsistency as knowledge bases of size n. Indeed, this property is true for all considered measures as $\mathbb{K}^v(n) \subseteq \mathbb{K}^v(n')$ (the same holds for all classes above).

The aim of our expressivity analysis is to investigate the number of different values that a specific inconsistency measure can attain on different subclasses of knowledge bases. We formalise this idea using *expressivity characteristics* as follows.

Definition 20. Let $\mathcal{I}$ be an inconsistency measure and $n > 0$. Let $\alpha \in \{v, f, l, p\}$. The *$\alpha$-characteristic* $\mathcal{C}^\alpha(\mathcal{I}, n)$ of $\mathcal{I}$ wrt. n is defined as

$$\mathcal{C}^\alpha(\mathcal{I}, n) = |\{\mathcal{I}(\mathcal{K}) \mid \mathcal{K} \in \mathbb{K}^\alpha(n)\}|$$

In other words, $\mathcal{C}^\alpha(\mathcal{I}, n)$ is the number of different inconsistency values $\mathcal{I}$ assigns to knowledge bases from $\mathbb{K}^\alpha(n)$. Note that these characteristics are not always the same as the *maximal* value of an inconsistency measure on a specific set of knowledge bases, even if the codomain of the measure is the natural numbers. Indeed, it can be the case that intermediate values cannot be attained.

Example 11. Consider $\mathcal{I}_\eta$ which has the codomain $[0,1]$ as each value $\mathcal{I}_\eta(\mathcal{K})$ can be associated with a probability value, cf. Table 1. In (Knight, 2002) it has already been shown that $\mathcal{I}_\eta(\mathcal{K})$ is always a rational number, so $\sqrt{2}/2 \notin \mathsf{Im}\ \mathcal{I}_\eta$.[4] Furthermore, the possible values of $\mathcal{I}_\eta$ are further constrained when considering specific subclasses from above. For example, for every arbitrary knowledge base $\mathcal{K}$ which contains at most 2 formulas, the only possible values of $\mathcal{I}_\eta(\mathcal{K})$ are $0, 1/2, 1$, so we have $\mathcal{C}^f(\mathcal{I}_\eta, 2) = 3$.

We now come to the main contribution of this paper, which is a thorough study of the 16 considered inconsistency measures in terms of our four proposed expressivity characteristics.

Theorem 1. *The α-characteristics $\mathcal{C}^\alpha(\mathcal{I}, n)$ ($\alpha \in \{f, v, l, p\}$) for the inconsistency measures $\mathcal{I}_d$, $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$, $\mathcal{I}_\eta$, $\mathcal{I}_c$, $\mathcal{I}_{LP_m}$, $\mathcal{I}_{mc}$, $\mathcal{I}_p$, $\mathcal{I}_{hs}$, $\mathcal{I}^{\Sigma}_{dalal}$, $\mathcal{I}^{\max}_{dalal}$, $\mathcal{I}^{hit}_{dalal}$, $\mathcal{I}_{D_f}$, $\mathcal{I}_{P_m}$, $\mathcal{I}_{mv}$, and $\mathcal{I}_{nc}$ are as shown in Table 3.*

The complete proof of the above theorem can be found in the appendix. However, some of these proofs provide some interesting insights into the behaviour of particular inconsistency measures and provide relations to other specific mathematical branches. Therefore, we will discuss these insights in more detail in the following subsections before we continue with the actual discussion on comparing expressivity in Section 5.

4.1 *Sperner Families and Minimal Inconsistent Sets*

Many of the inconsistency measures discussed above use the notion of minimal inconsistent subset as a central tool for assessing the inconsistency of a knowledge base. In its simplest implementation, the inconsistency measure $\mathcal{I}_{\mathsf{MI}}$ is defined to be exactly the number of the minimal inconsistent subsets of a knowledge base. Accordingly, in order to determine the number $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n)$ it is necessary to investigate how many different minimal inconsistent subsets a knowledge base with n formulas may possess. This question has already been investigated from a more abstract perspective in set theory under the notion of *Sperner families* (also called *independent systems*).

4 $\mathsf{Im}\ f$ is the image of a function f.

	$\mathcal{C}^v(\mathcal{I}$,n)	$\mathcal{C}^f(\mathcal{I}$,n)	$\mathcal{C}^l(\mathcal{I}$,n)	$\mathcal{C}^p(\mathcal{I}$,n)
$\mathcal{I}_d$	2	2	2^*	2
$\mathcal{I}_{\mathsf{MI}}$	∞	$\binom{n}{\lfloor n/2 \rfloor}+1$	∞^*	∞
$\mathcal{I}_{\mathsf{MI}^C}$	∞	$\leq \Psi(n)^\ddagger$	∞^*	∞
$\mathcal{I}_\eta$	$\Phi(2^n)^\dagger$	$\leq \Phi(\binom{n}{\lfloor n/2 \rfloor})^\dagger$	∞^{**}	∞^*
$\mathcal{I}_c$	$n+1$	∞	∞^*	∞
$\mathcal{I}_{LP_m}$	$\Phi(n)$	∞	∞^*	∞
$\mathcal{I}_{mc}$	∞	$\binom{n}{\lfloor n/2 \rfloor}^{**}$	∞^*	∞
$\mathcal{I}_p$	∞	$n+1$	∞^*	∞
$\mathcal{I}_{hs}$	2^n+1	$n+1$	∞^{**}	∞^*
$\mathcal{I}^{\Sigma}_{\text{dalal}}$	∞	∞^*	∞^*	∞
$\mathcal{I}^{\max}_{\text{dalal}}$	$n+2$	∞^*	$\lfloor (n+7)/3 \rfloor^{**}$	$n+2$
$\mathcal{I}^{\text{hit}}_{\text{dalal}}$	∞	$n+1$	∞^*	∞
$\mathcal{I}_{D_f}$	∞	$\leq \Psi(n)^\ddagger$	∞^*	∞
$\mathcal{I}_{P_m}$	∞	∞	∞^*	∞
$\mathcal{I}_{mv}$	$n+1$	∞^*	∞^*	∞
$\mathcal{I}_{nc}$	∞	$n+1$	∞^*	∞

Table 3: Characteristics of inconsistency measures ($n \geq 1$)
*only for $n > 1$
**only for $n > 3$
$^\dagger\Phi(x)$ is the number of fractions in the Farey series of order x and can be defined as $\Phi(x) = |\{k/l \mid l = 1, \ldots, x, k = 0, \ldots, l\}|$, see e. g. `http://oeis.org/A005728`
$^\ddagger\Psi(n)$ is the number of profiles of monotone Boolean functions of n variables, see e. g. `http://oeis.org/A220880`

Definition 21. Let $\mathcal{S} = \{S_1, \ldots, S_n\}$ with $n \geq 1$ be a family of sets over a set $X \neq \emptyset$, i. e., $S_i \subseteq X$ for all $i = 1, \ldots, n$. The family S is called a *Sperner family* over X if for all $S, S' \in \mathcal{S}$ with $S \neq S'$, $S \not\subseteq S'$.

In other words, $\mathcal{S}$ is a Sperner family if none of its elements is contained in another. It can easily be seen that for every *inconsistent* knowledge base $\mathcal{K}$ the set $\mathsf{MI}(\mathcal{K})$ is also a Sperner family over $\mathcal{K}$: for $M, M' \in \mathsf{MI}(\mathcal{K})$ with $M \neq M'$ it cannot hold $M \subseteq M'$, otherwise M' would not be a minimal inconsistent set. Note that a consistent knowledge base yields $\mathsf{MI}(\mathcal{K}) = \emptyset$ which is not covered by the above definition.

As $\mathsf{MI}(\mathcal{K})$ is a Sperner family over $\mathcal{K}$ its maximal cardinality is bounded by the maximal cardinality of any Sperner family over $\mathcal{K}$. For the latter we have the following result.

Theorem 2 ((Sperner, 1928)). *Let X be a set with $n = |X|$.*

1. *There is a Sperner family $\mathcal{S}_{\max}$ over X with $|\mathcal{S}_{\max}| = \binom{n}{\lfloor n/2 \rfloor}$.*[5]

2. *For every Sperner family $\mathcal{S}'$ over X, $|\mathcal{S}'| \leq |\mathcal{S}_{\max}|$.*

5 $\mathcal{S}_{\max}$ can be constructed by taking the set of all subsets $S \subseteq X$ with $|S| = \lfloor n/2 \rfloor$.

A corollary of the above theorem is that $|\mathsf{MI}(\mathcal{K})| \leq \binom{|\mathcal{K}|}{\lfloor |\mathcal{K}|/2 \rfloor}$ for every knowledge base $\mathcal{K}$. Also taking into account that $|\mathsf{MI}(\mathcal{K})|$ is always a non-negative integer we can directly entail $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n) \leq \binom{n}{\lfloor n/2 \rfloor} + 1$ (the addition of 1 is due to the fact that a consistent knowledge base $\mathcal{K}$ yields $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}) = 0$, which is not covered by the above theorem).

Interestingly, the set $\mathsf{MI}(\mathcal{K})$ is not only a Sperner family, but every Sperner family can be represented as $\mathsf{MI}(\mathcal{K})$ of some knowledge base $\mathcal{K}$. Let $X = \{\alpha_1, \ldots, \alpha_n\}$ be any set and let $\mathcal{S}$ be a Sperner family over X. Consider a propositional signature $\mathsf{At} = \{a_1, \ldots, a_n\}$ where each proposition $a_i \in \mathsf{At}$ corresponds to the element $\alpha_i \in X$. Consider now a knowledge base $\mathcal{K}_n^{\mathcal{S}} = \{\phi_1, \ldots, \phi_n\}$ defined via

$$\phi_i = a_i \wedge \bigwedge_{M \in \mathcal{S}, \alpha_i \in M} \ \bigvee_{\alpha_j \in M \setminus \{\alpha_i\}} \neg a_j$$

for $i = 1, \ldots, n$. Every ϕ_i states that a_i is accepted and for each set M in $\mathcal{S}$ which contains α_i at least one of the other elements must not be accepted. Thus, every ϕ_i lists the conditions under which any set containing α_i does *not* contain an element of the Sperner family.

Example 12. Let $X = \{\alpha, \beta, \gamma, \delta\}$ be a set and consider the Sperner family $\mathcal{S} = \{S_1, S_2, S_3\}$ over X defined via

$$S_1 = \{\beta, \gamma, \delta\} \qquad S_2 = \{\alpha, \beta\} \qquad S_3 = \{\alpha, \gamma\}$$

Consider now the signature $\mathsf{At} = \{a, b, c, d\}$ where the proposition a corresponds to α, b to β, c to γ, and d to δ. Then $\mathcal{K}_4^{\mathcal{S}} = \{\phi_1, \phi_2, \phi_3, \phi_4\}$ is defined via

$$\begin{aligned}
\phi_1 &= a \wedge \neg b \wedge \neg c \\
\phi_2 &= b \wedge (\neg c \vee \neg d) \wedge \neg a \\
\phi_3 &= c \wedge (\neg b \vee \neg d) \wedge \neg a \\
\phi_4 &= d \wedge (\neg b \vee \neg c)
\end{aligned}$$

For example, ϕ_3 states that if some set S contains γ (c), either not β or not δ ($\neg b \vee \neg d$), and not α ($\neg a$), then S does not contain any element of $\mathcal{S}$.

By construction, it follows that $M = \{\alpha_{k_1}, \ldots, \alpha_{k_m}\}$ (for some $k_1, \ldots, k_m \in \{1, \ldots, n\}$) is an element of the Sperner family $\mathcal{S}$ if and only if the set $\{\phi_{k_1}, \ldots, \phi_{k_m}\}$ is a minimal inconsistent subset of $\mathcal{K}_n^{\mathcal{S}} = \{\phi_1, \ldots, \phi_n\}$.

Example 13. We continue Example 12 and consider $S_1 = \{\beta, \gamma, \delta\} \in \mathcal{S}$. The set S_1 corresponds to the set $\{\phi_2, \phi_3, \phi_4\} \subseteq \mathcal{K}_4^{\mathcal{S}}$ with

$$\phi_2 = b \wedge (\neg c \vee \neg d) \wedge \neg a$$
$$\phi_3 = c \wedge (\neg b \vee \neg d) \wedge \neg a$$
$$\phi_4 = d \wedge (\neg b \vee \neg c)$$

As one can see $\{\phi_2, \phi_3, \phi_4\}$ is also a minimal inconsistent subset of $\mathcal{K}_4^{\mathcal{S}}$. Further, the set of minimal inconsistent subsets of $\mathcal{K}_4^{\mathcal{S}}$ is indeed

$$\mathsf{MI}(\mathcal{K}_4^{\mathcal{S}}) = \{\{\phi_2, \phi_3, \phi_4\}, \{\phi_1, \phi_2\}, \{\phi_1, \phi_3\}\}$$

which is in direct correspondence to $\mathcal{S}$.

From these observations it follows that $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_n^{\mathcal{S}}) = |\mathsf{MI}(\mathcal{K}_n^{\mathcal{S}})| = |\mathcal{S}|$. Observe that if $\mathcal{S}$ is a Sperner family over X and $S \in \mathcal{S}$ then $\mathcal{S} \setminus \{S\}$ is also a Sperner family over X (provided that it is non-empty). Then for a set X with $|X| = n$ and for every $i = 1, \ldots, \binom{n}{\lfloor n/2 \rfloor}$ we can define a Sperner family $\mathcal{S}_i$ over X with $|\mathcal{S}_i| = i$. This means in our setting, for every $i = 1, \ldots, \binom{n}{\lfloor n/2 \rfloor}$ we can define a knowledge base $\mathcal{K}$ with $|\mathcal{K}| = n$ such that $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}) = i$. Together with the fact that 0 is also a possible value of $\mathcal{I}_{\mathsf{MI}}$ we obtain $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n) = \binom{n}{\lfloor n/2 \rfloor} + 1$.

Furthermore, also the set of maximal consistent subsets $\mathtt{MC}(\mathcal{K})$ of a knowledge base $\mathcal{K}$ form a Sperner family over $\mathcal{K}$, which is the reason why the value $\mathcal{C}^f(\mathcal{I}_{mc}, n) = \binom{n}{\lfloor n/2 \rfloor}$ is almost the same as $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n)$ (please see the appendix for the detailed proof of $\mathcal{C}^f(\mathcal{I}_{mc}, n) = \binom{n}{\lfloor n/2 \rfloor}$).

4.2 *Monotone Boolean Functions and Minimal Inconsistent Sets*

Another interesting observation is the relation of the measures $\mathcal{I}_{\mathsf{MI}^{\mathsf{C}}}$ and $\mathcal{I}_{D_f}$ with the number $\Psi(n)$ of profiles of monotone Boolean functions of n variables (Stephen and Yusun, 2012). Let us recall some basics on Boolean functions.

Definition 22. A *Boolean function* f of n variables ($n \in \mathbb{N}_0$) is a function $f : \{0,1\}^n \to \{0,1\}$.

Let $B(n)$ be the set of all Boolean functions of n variables. It can easily be seen that $|B(n)| = 2^{2^n}$.

Definition 23. A Boolean function f of n variables ($n \in \mathbb{N}_0$) is *monotone* if for every $i = 1, \ldots, n$ we have

$$f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) \leq f(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n)$$

Let $MB(n)$ be the set of all monotone Boolean functions of n variables. Obviously $|MB(n)| \leq |B(n)|$. However, no closed form for $m(n) = |MB(n)|$—which is also called *Dedekind number*[6] (Dedekind, 1897)—is known and only values up to $n = 8$ could be determined yet, cf. (Stephen and Yusun, 2012). The Dedekind number $m(n)$ has also a meaning in the context of Sperner families (see previous subsection). In fact, $m(n)$ is also the number of *different* Sperner families over a set X with $|X| = n$.

For the issue of analyzing inconsistency measures, the Dedekind number $m(n)$ itself is not directly applicable as it is somewhat *syntax-sensitive*. More precisely, consider the set $X = \{\alpha, \beta, \gamma\}$ and the two Sperner families $\{\{\alpha, \beta\}\}$ and $\{\{\beta, \gamma\}\}$. Each of these families count one in the Dedekind number $m(3)$. From the perspective of inconsistency measurement, the sets of minimal inconsistent subsets $\mathsf{MI}(\mathcal{K}_1) = \{\{\alpha, \beta\}\}$ and $\mathsf{MI}(\mathcal{K}_2) = \{\{\alpha, \gamma\}\}$ for some knowledge bases $\mathcal{K}_1, \mathcal{K}_2 \in \mathbb{K}^f(3)$ are *indistinguishable* for all inconsistency measures solely based on utilizing minimal inconsistent sets. While for $\mathcal{I}_{\mathsf{MI}}$ only the *number* of minimal inconsistent sets is important, even for more elaborate measures such as $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$ and $\mathcal{I}_{D_f}$ these sets are equivalent as they coincide in both the number of minimal inconsistent sets and the cardinalities of each of those. Recall that in order to define the measure $\mathcal{I}_{D_f}$ we defined for a knowledge base $\mathcal{K}$ the sets

$$\mathsf{MI}^{(i)}(\mathcal{K}) = \{M \in \mathsf{MI}(\mathcal{K}) \mid |M| = i\}$$

for $i = 1, \ldots, |\mathcal{K}|$. Given these values, we can also redefine the inconsistency measure $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$ via

$$\mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}) = \sum_{i=1}^{|\mathcal{K}|} \frac{|\mathsf{MI}^{(i)}(\mathcal{K})|}{i}$$

Let us call $\mathsf{profile}(\mathcal{K}) = (|\mathsf{MI}^{(1)}(\mathcal{K})|, \ldots, |\mathsf{MI}^{(|\mathcal{K}|)}(\mathcal{K})|) \in \mathbb{N}_0^{|\mathcal{K}|}$ the MI-*profile* of $\mathcal{K}$. One can see that $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$ depends only on $\mathsf{profile}(\mathcal{K})$ and not the actual structure of the minimal inconsistent subsets.[7] However, this property of indifference has a corresponding property in the context of Boolean functions.

Definition 24. Two Boolean functions f_1, f_2 of n variables ($n \in \mathbb{N}_0$) are *equivalent*, denoted by $f_1 \sim f_2$, if there is a permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that

$$f_1(x_1, \ldots, x_n) = f_2(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

for all $x_1, \ldots, x_n \in \{0, 1\}$.

6 See also `http://oeis.org/A000372`

7 The same holds for $\mathcal{I}_{D_f}$ as an MI-profile $(\mathsf{MI}^0(\mathcal{K}), \ldots, \mathsf{MI}^n(\mathcal{K}))$ also uniquely determines the corresponding CN-profile $(\mathsf{CN}^0(\mathcal{K}), \ldots, \mathsf{CN}^n(\mathcal{K}))$

Let $MB(n)/_\sim$ be the quotient set of $MB(n)$ wrt. $\sim$, i. e.

$$MB(n)/_\sim = \{[f]_\sim | f \in MB(n)\}$$

where $[f]_\sim$ is the *equivalence class* (also called *profile*) of f, i. e., $[f]_\sim = \{f' \in MB(n) \mid f' \sim f\}$. Then $\Psi(n)$ is defined as $\Psi(n) = |MB(n)/_\sim|$, i. e., $\Psi(n)$ is the number of profiles of monotone Boolean functions of n variables. As for the Dedekind number $m(n)$, no closed form for $\Psi(n)$ is known (Stephen and Yusun, 2012).[8] However, we can observe an intriguing relationship of this number to inconsistency measures via

$$\Psi(n) = |\{\mathsf{profile}(\mathcal{K}) \mid \mathcal{K} \in \mathbb{K}^f(n)\}|$$

In other words, the number of different MI-profiles of knowledge bases of size n is the same as the number of profiles of monotone Boolean functions of n variables. In order to see this, recall that there is a one-to-one correspondence of $\mathsf{MI}(\mathcal{K})$ with Sperner families. It has already been mentioned that there is also a relationship between Sperner families and monotone Boolean functions. More precisely, let $\mathcal{S}$ be any Sperner family over $X = \{a_1, \ldots, a_n\}$ and define a Boolean function $f_\mathcal{S}$ via

$$f_\mathcal{S}(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } \exists S \in \mathcal{S} : S \subseteq \{a_i \mid x_i = 1\} \\ 0 & \text{otherwise} \end{cases} . \tag{5.11}$$

for all $x_1, \ldots, x_n \in \{0, 1\}$. In other words, the function $f_\mathcal{S}(x_1, \ldots, x_n)$ evaluates to 1 if there is a member $S \in \mathcal{S}$ such that the corresponding variables of the elements of S are 1. Observe that $f_\mathcal{S}$ is monotone for every Sperner family $\mathcal{S}$. Moreover, it can also be seen that for every monotone Boolean function f there is a uniquely determined Sperner family $\mathcal{S}$ such that $f = f_\mathcal{S}$. Consider now two monotone Boolean functions f_1, f_2 with $f_1 \sim f_2$. Then these correspond to two Sperner families $\mathcal{S}_1, \mathcal{S}_2$ over X, where $\mathcal{S}_2$ can be obtained from $\mathcal{S}_1$ by only permuting the elements of X. The only invariant between the Sperner families corresponding to the functions in $[f]_\sim$ is the number of sets in each, and the sizes of each set. In the context of minimal inconsistent sets, this means that there is a one-to-one correspondence between any MI-profile and an equivalence class $[f]_\sim$, leading to Equation (5.11).

For the specific case of $\mathcal{I}_{\mathsf{MI}^C}$ it has to be observed that the assignment of an MI-profile to the inconsistency value is not injective, i. e., there may be more than one MI-profile that is mapped to the same inconsistency value.

Example 14. Consider the knowledge bases $\mathcal{K}_3, \mathcal{K}_4 \in \mathbb{K}^f(5)$ defined via

$$\mathcal{K}_3 = \{a, \neg a, b, c, d\}$$
$$\mathcal{K}_4 = \{a, b, c, \neg a \vee \neg b \vee \neg c, \neg(a \wedge b \wedge c)\}$$

8 See also `http://oeis.org/A220880`.

n	$\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n)$	$\Psi(n)$
1	2	2
2	4	4
3	7	9
4	15	25
5	43	95

Table 4: Values of $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n)$ and $\Psi(n)$ for $n = 1, \ldots, 5$.

Here we have

$$\begin{aligned} \mathsf{MI}(\mathcal{K}_3) &= \{\{a, \neg a\}\} \\ \mathsf{MI}(\mathcal{K}_4) &= \{\{a, b, c, \neg a \vee \neg b \vee \neg c\}, \{a, b, c, \neg(a \wedge b \wedge c)\}\} \end{aligned}$$

and therefore

$$\begin{aligned} \mathsf{profile}(\mathcal{K}_3) &= (0, 1, 0, 0, 0) \\ \mathsf{profile}(\mathcal{K}_4) &= (0, 0, 0, 2, 0) \end{aligned}$$

yielding $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}_3) = \mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}_4) = 1/2$.

This behavior is the reason that $\Psi(n)$ is only an upper bound for $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n)$ in Theorem 1. As for $m(n)$ and $\Psi(n)$, no closed form for $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n)$ could be found in our investigation. Using a computational brute-force approach we could however determine the first five values for $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n)$ ($n = 1, \ldots, 5$) which are listed in Table 4 together with their corresponding upper bounds $\Psi(n)$. The measure $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$ is quite simplistic in its way to aggregate an MI-profile into a single inconsistency measure. A more elaborated measure is $\mathcal{I}_{D_f}$ (see Definition 5) where this aggregation is more fine-grained. For this measure, we obtain $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) \leq \mathcal{C}^f(\mathcal{I}_{D_f}, n) \leq \Psi(n)$, i.e., $\mathcal{I}_{D_f}$ is more expressive than $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$ but still bounded by $\Psi(n)$. Note again that Definition 5 represents only a single instance of a more general family of inconsistency measures presented in (Mu *et al.*, 2011a). Using an injective function h from the set of MI-profiles to real numbers one can instantiate this family with an instance $\mathcal{I}'_{D_f}$ where we could actually have $\mathcal{C}^f(\mathcal{I}'_{D_f}, n) = \Psi(n)$.[9] In any case, $\Psi(n)$ is always an upper bound for every measure following the paradigm of $\mathcal{I}_{D_f}$.

9 Note that injective functions of the form $h : \mathbb{N}^k \to \mathbb{R}$ do indeed exist for arbitrary (and infinite) k but require complex constructions. However, it is questionable whether there are instances that would lead to meaningful inconsistency measures.

4.3 *Knight's Inconsistency Measure and the Farey Series*

Let us now turn to another interesting relationship, namely that of the inconsistency measure $\mathcal{I}_\eta$ (Knight, 2002) with the *Farey series*[10,11]. The latter is defined as the series of numbers generated by the function $\Phi : \mathbb{N} \to \mathbb{N}$ defined via

$$\Phi(x) = |\{k/l \mid l = 1, \ldots, x, k = 0, \ldots, l\}|$$

for all $x \in \mathbb{N}$. In other words, $\Phi(x)$ is the number of *different* fractional expressions in $[0, 1]$ with maximal denominator x (where both nominator and denominator are natural numbers). For example, for $x = 3$ we have

$$\left\{\frac{0}{1}, \frac{1}{1}, \frac{0}{2}, \frac{1}{2}, \frac{2}{2}, \frac{0}{3}, \frac{1}{3}, \frac{2}{3}, \frac{3}{3}\right\} = \left\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\right\}$$

yielding $\Phi(3) = 5$. Let us now recall the measure $\mathcal{I}_\eta$ which is defined as

$$\mathcal{I}_\eta(\mathcal{K}) = 1 - \max\{\xi \mid \exists P \in \mathcal{P}(\mathsf{At}) : \forall \alpha \in \mathcal{K} : P(\alpha) \geq \xi\}$$

for every $\mathcal{K} \in \mathbb{K}$. So in order to determine $\mathcal{I}_\eta(\mathcal{K})$ we are seeking a probability function $P \in \mathcal{P}(\mathsf{At})$ that maximizes the minimum probability it assigns to formulas in $\mathcal{K}$. For the remainder of this section, let $P_\mathcal{K}$ be any probability function that maximizes the minimum probability of all formulas of $\mathcal{K}$, i.e., we have $\mathcal{I}_\eta(\mathcal{K}) = 1 - \xi_\mathcal{K}$ with $\xi_\mathcal{K} = \min\{P_\mathcal{K}(\alpha) \mid \alpha \in \mathcal{K}\}$.

A first observation one can make about $P_\mathcal{K}$ and $\mathcal{I}_\eta$—which also sets it apart from many other inconsistency measures—is that it does not care for the syntactic representation of formulas. More precisely, for any knowledge base $\mathcal{K}$, formulas ϕ, ϕ' with $\phi \equiv \phi'$ we have $\mathcal{I}_\eta(\mathcal{K} \cup \{\phi\}) = \mathcal{I}_\eta(\mathcal{K} \cup \{\phi'\})$ as $P(\phi) = P(\phi')$ for *every* probability function. Moreover, we also have $\mathcal{I}_\eta(\mathcal{K} \cup \{\phi\}) = \mathcal{I}_\eta(\mathcal{K} \cup \{\phi, \phi'\})$ for the same reason; adding syntactic variations of already present formulas does not change the inconsistency value. It follows that we can identify every formula $\phi \in \mathcal{K}$ with its set of models $\mathsf{Mod}(\phi)$ in all matters related to determining $\mathcal{I}_\eta(\mathcal{K})$. By abusing notation we therefore can rewrite $\mathcal{K}$ as

$$\mathcal{K}' = \{\mathsf{Mod}(\phi) \mid \phi \in \mathcal{K}\} \subseteq \mathfrak{P}(\mathsf{Int}(\mathsf{At}))$$

where $\mathfrak{P}(X)$ is the power set of a set X and At is the signature of the underlying propositional language. Assume that $\mathsf{At} = \{a_1, \ldots, a_n\}$, then $\mathsf{Int}(\mathsf{At})$ has 2^n elements and every $\mathsf{Mod}(\phi)$ is a subset of those.

Let us first consider the question of how $\mathcal{K}'$ and $\mathcal{K}$ could look like if we want to have $\xi_{\mathcal{K}'} = k/l$ for $l \in \{1, \ldots, 2^n\}$ and $k \in \{1, \ldots, l\}$. Consider an arbitrary set $\{\omega_1, \ldots, \omega_l\} \subseteq \mathsf{Int}(\mathsf{At})$ of interpretations (as $l \leq 2^n$ there

10 See `http://oeis.org/A005728` for more information on the Farey series.
11 The measure $\mathcal{I}_{LP_m}$ also has a relationship with the Farey series via $\mathcal{C}^v(\mathcal{I}_{LP_m}, n) = \Phi(n)$, but it is quite straightforwardly explained. See the proof of Theorem 1 for details.

are enough different interpretations) and define $P_{\mathcal{K}'}(\omega) = 1/l$ for all $\omega \in \{\omega_1, \ldots, \omega_l\}$ and $P_{\mathcal{K}'}(\omega') = 0$ for all remaining ω'. Then $P_{\mathcal{K}'}$ is indeed a probability function that assigns equal probability to all $\{\omega_1, \ldots, \omega_l\}$. Then define a formula ϕ in such a way that $\mathsf{Mod}(\phi)$ is any k-element subset of $\{\omega_1, \ldots, \omega_l\}$. Then we obtain $P_{\mathcal{K}}(\phi) = k/l$. Populating $\mathcal{K}$ with *all* ϕ that can be defined as such, we obtain a knowledge base $\mathcal{K}$ with $\xi_{\mathcal{K}} = k/l$.

Example 15. Consider the propositional signature $\mathsf{At}_3 = \{a_1, a_2, a_3\}$, i.e., we have $n = 3$ and $2^n = 8$. Choose $\xi_{\mathcal{K}'} = 5/6$ (note that $6 \in \{1, \ldots, 8\}$ and $5 \in \{1, \ldots, 6\}$) and consider the following 6 interpretations $\omega_1, \ldots, \omega_6 \in \mathsf{Int}(\mathsf{At}_3)$:

$$\omega_1 = a_1 a_2 a_3 \qquad \omega_2 = a_1 a_2 \overline{a_3} \qquad \omega_3 = a_1 \overline{a_2} a_3$$
$$\omega_4 = a_1 \overline{a_2} \overline{a_3} \qquad \omega_5 = \overline{a_1} a_2 a_3 \qquad \omega_6 = \overline{a_1} a_2 \overline{a_3}$$

The following sets $M_1, \ldots, M_6$ are all 5-element subsets of $\omega_1, \ldots, \omega_6$:

$$M_1 = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\} \qquad M_2 = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_6\}$$
$$M_3 = \{\omega_1, \omega_2, \omega_3, \omega_5, \omega_6\} \qquad M_4 = \{\omega_1, \omega_2, \omega_4, \omega_5, \omega_6\}$$
$$M_5 = \{\omega_1, \omega_3, \omega_4, \omega_5, \omega_6\} \qquad M_6 = \{\omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$$

Consider formulas ϕ_i with $\mathsf{Mod}(\phi_i) = M_i$ for $i = 1, \ldots, 6$. For example, we have $\phi_1 = a_1 \vee (a_2 \wedge a_3)$. For the knowledge base $\mathcal{K} = \{\phi_1, \ldots, \phi_6\}$ consider the probability function $P_{\mathcal{K}}$ with $P_{\mathcal{K}}(\omega_1) = \ldots = P_{\mathcal{K}}(\omega_6) = 1/6$ and $P(\omega) = 0$ for $\omega \in \mathsf{Int}(\mathsf{At}_3) \setminus \{\omega_1, \ldots \omega_6\}$. By construction we have $P_{\mathcal{K}}(\phi_1) = \ldots = P_{\mathcal{K}}(\phi_6) = 5/6$. Note also that there cannot be any other probability function that gives larger probability to all formulas.

So given a signature $\mathsf{At} = \{a_1, \ldots, a_n\}$ and any $l \in \{1, \ldots, 2^n\}$ and $k \in \{1, \ldots, l\}$ we can construct a knowledge base $\mathcal{K}$ such that $\xi_{\mathcal{K}} = k/l$ (and therefore $\mathcal{I}_\eta(\mathcal{K}) = 1 - k/l$). This gives us $\mathcal{C}^v(\mathcal{I}_\eta, n) \geq \Phi(2^n)$.

The remaining question is whether there are numbers $x \in [0, 1]$ that are not of the form $x = k/l$ with $l \in \{1, \ldots, 2^n\}$ and $k \in \{1, \ldots, l\}$ and for which a knowledge base $\mathcal{K}$ can be found such that $\xi_{\mathcal{K}} = x$. Knight already showed in (Knight, 2002) that $\xi_{\mathcal{K}}$ must always be a rational number in the unit interval, so it is clear that $x = p/q$ for some $p, q \in \mathbb{N}$ with $p \leq q$. So what about e.g. $x = 1/(2^n + 1)$? It can be shown (see the complete proof in the appendix) that due to combinatorial reasons a value such as $1/(2^n + 1)$ cannot be attained for $\xi_{\mathcal{K}}$ if the underlying signature has n elements. For example, the uniform probability function P with $P(\omega) = 1/2^n$ already yields $P(\phi) \geq 1/2^n$ for every formula $\phi \in \mathcal{K}$ as it has at least one model (note that if $\mathcal{K}$ contains a contradictory formula we always have $\xi_{\mathcal{K}} = 0$).

Although the expressivity of $\mathcal{I}_\eta$ is characterized by $\Phi(2^n)$ it has to be noted that $\Phi(2^n)$ increases quite rapidly which makes $\mathcal{I}_\eta$ a quite expressive inconsistency measure (see Section 5).

4.4 *Normal Forms for Knowledge Bases*

Many proofs of statements in Theorem 1 (in particular those showing infinite characteristics) involve the construction of particular families of knowledge bases that exhibit extreme inconsistency values, such as for $\mathcal{C}^f(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$ (for $n > 1$). Recall that $\mathcal{I}^{\Sigma}_{\text{dalal}}$ is defined by determining an interpretation $\omega \in \mathsf{Int}(\mathsf{At})$ such that each formula $\phi \in \mathcal{K}$ has minimal distance to ω (measured in the number of propositions that have to be flipped in ω order to obtain a model of ϕ). Then $\mathcal{I}^{\Sigma}_{\text{dalal}}$ is the sum of all these distances.

In the proof of $\mathcal{C}^f(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$ (for $n > 1$) the following family of knowledge bases $\mathcal{K}_i$ is used:

$$\mathcal{K}^1_i = \{a_1 \wedge \ldots \wedge a_i, \neg a_1 \wedge \ldots \wedge \neg a_i\}$$

for $i \in \mathbb{N}$. Note that each $\mathcal{K}^1_i$ consists of only two formulas but the number of mentioned propositions increases with increasing i. It can be seen that for *every* interpretation $\omega \in \mathsf{Int}(\mathsf{At})$ the sum of its distances to both formulas amounts to exactly i, i. e., $\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^1_i) = i$ and for $i \to \infty$ we obtain $\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^1_i) \to \infty$ and thus $\mathcal{C}^f(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$.

Constructions such as the above can be used to characterize *normal forms* of knowledge bases for inconsistency measures. For example, the above family $\mathcal{K}^1_i$ for $i \in \mathbb{N}$ exhaustively describes the image of $\mathcal{I}^{\Sigma}_{\text{dalal}}$, i. e.,

$$\mathsf{Im}\ \mathcal{I}^{\Sigma}_{\text{dalal}} = \{\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^1_0), \mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^1_1), \mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^1_2), \ldots\}$$

Note that $\mathcal{K}^1_0 = \emptyset$. Every other knowledge base can be transformed into one of these knowledge bases while retaining its inconsistency value. While this transformation, of course, does not maintain semantic equivalence (even in a paraconsistent context), it can be used for illustration purposes. In each $\mathcal{K}^1_i$ the inconsistency is boiled down to its essential core as it is measured by $\mathcal{I}^{\Sigma}_{\text{dalal}}$. Inspecting this normal form, instead of the original knowledge base where the inconsistency might be obfuscated, can lead to better understanding of the severity of the inconsistency. We leave a deeper investigation of this matter for future work.

4.5 *About the Distinction between* $\{\alpha, \beta\}$ *and* $\{\alpha \wedge \beta\}$

By studying Table 3 it can be observed that almost all inconsistency measures have trivial characteristic values, i. e., a value of ∞, wrt. $\mathcal{C}^l$ and $\mathcal{C}^p$, and the characteristics $\mathcal{C}^v$ and $\mathcal{C}^f$ seem to be much better suited for assessing the expressivity. The reason for this is that for many inconsistency measures some conjunctions $\alpha \wedge \beta$ can be replaced by two distinct formulas α and β without decreasing the inconsistency value, so large inconsistency values can be attained by either having few long formulas or many short formulas. As $\mathcal{C}^l$ and $\mathcal{C}^p$ only consider the formula-length as fixed (or the number of propositions per formula), arbitrary different inconsistency values can be attained by considering arbitrary large knowledge bases. The only exception, besides the drastic inconsistency measure $\mathcal{I}_d$, is the measure $\mathcal{I}^{\max}_{\text{dalal}}$. Recall that the measure $\mathcal{I}^{\max}_{\text{dalal}}$ is defined as the maximal distance of an optimally chosen $\omega \in \mathsf{Int}(\mathsf{At})$ to each formula of the knowledge base. If formulas are short, i. e., they each mention only few propositions, this distance is bounded, independently of the number of formulas in the knowledge base.

The important distinction between a set of formulas $\{\alpha, \beta\}$ and the conjunction $\alpha \wedge \beta$ has already been recognized within e. g. the fields of inconsistent-tolerant reasoning and belief revision (Konieczny *et al.*, 2005; Delgrande and Jin, 2012). For example, in the context of contracting from a knowledge base $\mathcal{K}_5 = \{a, b\}$ the inference a, the usually accepted result should be $\mathcal{K}_5 - a = \{b\}$. However, contracting from a knowledge base $\mathcal{K}_6 = \{a \wedge b\}$ the inference a would result in $\mathcal{K}_6 - a = \emptyset$. More generally, a conjunction $\alpha \wedge \beta$ establishes a relationship between the formulas α and β and stipulates that they have to appear together (if one does not appear then the other one should also not appear). For a more detailed discussion see (Konieczny *et al.*, 2005; Delgrande and Jin, 2012).

Our study on the characteristics $\mathcal{C}^l$ and $\mathcal{C}^p$ (for details see the proofs in the appendix) shows that many inconsistency measures do not recognize this difference and, moreover, behave quite incoherently in the general case of adding either separate formulas or a conjunction of the formulas to a knowledge base. Consider the following three properties for inconsistency measures. Let $\mathcal{I}$ be an inconsistency measure, $\mathcal{K} \in \mathbb{K}$, and $\phi, \psi \in \mathcal{L}(\mathsf{At})$ be arbitrary.

$\wedge$-Indifference $\mathcal{I}(\mathcal{K} \cup \{\alpha, \beta\}) = \mathcal{I}(\mathcal{K} \cup \{\alpha \wedge \beta\})$.

$\wedge$-Penalty $\mathcal{I}(\mathcal{K} \cup \{\alpha, \beta\}) \leq \mathcal{I}(\mathcal{K} \cup \{\alpha \wedge \beta\})$.

$\wedge$-Mitigation $\mathcal{I}(\mathcal{K} \cup \{\alpha, \beta\}) \geq \mathcal{I}(\mathcal{K} \cup \{\alpha \wedge \beta\})$.

Note that (Besnard, 2014) proposed $\wedge$-Indifference under the name *Adjunction Invariancy* as a desirable property. However, we do not aim to discuss which (if any) of these properties may be desirable.

But interestingly, only very few of the discussed measures satisfy any of them.

Theorem 3.

1. *The measures $\mathcal{I}_d$, $\mathcal{I}_c$, and $\mathcal{I}_{LP_m}$ satisfy $\wedge$-Indifference, $\wedge$-Penalty, and $\wedge$-Mitigation.*

2. *The measures $\mathcal{I}_\eta$, $\mathcal{I}_{hs}$, and $\mathcal{I}^{\max}_{dalal}$ satisfy $\wedge$-Penalty, but not $\wedge$-Mitigation.*

3. *The measures $\mathcal{I}^{hit}_{dalal}$ and $\mathcal{I}_{P_m}$ satisfy $\wedge$-Mitigation, but not $\wedge$-Penalty.*

4. *None of the measures $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$, $\mathcal{I}_{mc}$, $\mathcal{I}_p$, $\mathcal{I}^{\Sigma}_{dalal}$, $\mathcal{I}_{D_f}$, $\mathcal{I}_{mv}$, $\mathcal{I}_{nc}$ satisfies any of $\wedge$-Indifference, $\wedge$-Penalty, or $\wedge$-Mitigation.*

As a consequence, the inconsistency values for many measures change quite arbitrarily when a conjunction $\alpha \wedge \beta$ is replaced by it conjuncts α and β.

Example 16. Consider the knowledge base $\mathcal{K}_7$ given via

$$\mathcal{K}_7 = \{a, \neg a\}$$

and

$$\mathsf{MI}(\mathcal{K}_7) = \{\{a, \neg a\}\} \tag{5.12}$$
$$\mathsf{MI}(\mathcal{K}_7 \cup \{a, b\}) = \{\{a, \neg a\}\} \tag{5.13}$$
$$\mathsf{MI}(\mathcal{K}_7 \cup \{a \wedge b\}) = \{\{a, \neg a\}, \{a \wedge b, \neg a\}\} \tag{5.14}$$
$$\mathsf{MI}(\mathcal{K}_7 \cup \{a \wedge \neg a, \neg\neg a\}) = \{\{a, \neg a\}, \{\neg a, \neg\neg a\}, \{a \wedge \neg a\}\} \tag{5.15}$$
$$\mathsf{MI}(\mathcal{K}_7 \cup \{a \wedge \neg a \wedge \neg\neg a\}) = \{\{a, \neg a\}, \{a \wedge \neg a \wedge \neg\neg a\}\} \tag{5.16}$$

As one can see, the set $\mathsf{MI}(\mathcal{K})$ may change quite differently when adding separate formulas or conjunctions to $\mathcal{K}$. In Equations (5.13) and (5.14) the addition of a conjunction leads to more minimal inconsistent sets than the addition of separate formulas. In Equations (5.15) and (5.16) it is exactly the other way around. It follows that for measures $\mathcal{I}$ based on minimal inconsistent subsets—such as $\mathcal{I}_{\mathsf{MI}}$—there is no general relationship such as $\mathcal{I}(\mathcal{K} \cup \{\alpha, \beta\}) \leq \mathcal{I}(\mathcal{K} \cup \{\alpha \wedge \beta\})$ or $\mathcal{I}(\mathcal{K} \cup \{\alpha, \beta\}) \geq \mathcal{I}(\mathcal{K} \cup \{\alpha \wedge \beta\})$ for arbitrary knowledge bases $\mathcal{K}$ and formulas α and β.

5 EXPRESSIVITY ORDERS

Let us now come back to the original motivation of comparing inconsistency measures wrt. their expressivity. Definition 20 provides the basis for a comparative analysis of inconsistency measures wrt. their expressivity, which we address with the following definition.

Definition 25. An inconsistency measure $\mathcal{I}$ is *at least as expressive* as an inconsistency measure $\mathcal{I}'$ wrt. a characteristic $\mathcal{C}^\alpha$ ($\alpha \in \{f, v, l, p\}$), denoted by $\mathcal{I} \succeq_\alpha \mathcal{I}'$, if there is $n_0 \in \mathbb{N}$ such that for all $n > n_0$, $\mathcal{C}^\alpha(\mathcal{I}, n) \geq \mathcal{C}^\alpha(\mathcal{I}', n)$.

If both $\mathcal{I} \succeq_\alpha \mathcal{I}'$ and $\mathcal{I}' \succeq_\alpha \mathcal{I}$, we say that $\mathcal{I}$ and $\mathcal{I}'$ are *equally expressive* wrt. $\mathcal{C}^\alpha$ and denote this by $\mathcal{I} \sim_\alpha \mathcal{I}'$. If $\mathcal{I} \succeq_\alpha \mathcal{I}'$ but not $\mathcal{I} \sim_\alpha \mathcal{I}'$ we write $\mathcal{I} \succ_\alpha \mathcal{I}'$ ($\mathcal{I}$ is strictly more expressive than $\mathcal{I}'$). Note that the expressivity order $\succeq$ is not to be confused with the refinement order $\sqsubseteq$ sometimes used for pairwise comparisons of inconsistency measures, see e. g. (Thimm, 2016). The refinement order $\sqsubseteq$ is defined as $\mathcal{I}_1 \sqsubseteq \mathcal{I}_2$ iff $\mathcal{I}_2(\mathcal{K}) \geq \mathcal{I}_2(\mathcal{K}')$ implies $\mathcal{I}_1(\mathcal{K}) \geq \mathcal{I}_1(\mathcal{K}')$ for all $\mathcal{K}, \mathcal{K}'$. If $\mathcal{I}_1 \sqsubseteq \mathcal{I}_2$ this means that $\mathcal{I}_2$ is a refinement of $\mathcal{I}_1$. Note that $\succeq_\alpha$ compares measures in a quantitative way and also allows comparison of measures that induce totally different orders

$$\boxed{\mathcal{I}_{\mathsf{MI}}, \mathcal{I}_{\mathsf{MI}^C}, \mathcal{I}_{mc}, \mathcal{I}_p, \mathcal{I}^{\Sigma}_{\text{dalal}}, \mathcal{I}^{\text{hit}}_{\text{dalal}}, \mathcal{I}_{D_f}, \mathcal{I}_{P_m}, \mathcal{I}_{nc}} \succ_v \boxed{\mathcal{I}_\eta} \succ_v \boxed{\mathcal{I}_{hs}}$$
$$\succ_v \boxed{\mathcal{I}_{LP_m}} \succ_v \boxed{\mathcal{I}^{\max}_{\text{dalal}}} \succ_v \boxed{\mathcal{I}_c, \mathcal{I}_{mv}} \succ_v \boxed{\mathcal{I}_d}$$

$$\boxed{\mathcal{I}_c, \mathcal{I}_{LP_m}, \mathcal{I}^{\Sigma}_{\text{dalal}}, \mathcal{I}^{\max}_{\text{dalal}}, \mathcal{I}_{P_m}, \mathcal{I}_{mv}} \succ_f \boxed{\mathcal{I}_{\mathsf{MI}}} \succ_f \boxed{\mathcal{I}_{mc}} \succ_f \boxed{\mathcal{I}_p, \mathcal{I}_{hs}, \mathcal{I}^{\text{hit}}_{\text{dalal}}, \mathcal{I}_{nc}} \succ_f \boxed{\mathcal{I}_d}$$

$$\boxed{\mathcal{I}_c, \mathcal{I}_{LP_m}, \mathcal{I}^{\Sigma}_{\text{dalal}}, \mathcal{I}_{P_m}, \mathcal{I}_{mv}, \mathcal{I}_p, \mathcal{I}_{hs}, \mathcal{I}^{\text{hit}}_{\text{dalal}}, \mathcal{I}_{nc}, \mathcal{I}_{D_f}, \mathcal{I}_{\mathsf{MI}^C}, \mathcal{I}_\eta, \mathcal{I}_{\mathsf{MI}}, \mathcal{I}_{mc}} \succ_l \boxed{\mathcal{I}^{\max}_{\text{dalal}}} \succ_l \boxed{\mathcal{I}_d}$$

$$\boxed{\mathcal{I}_c, \mathcal{I}_{LP_m}, \mathcal{I}^{\Sigma}_{\text{dalal}}, \mathcal{I}_{P_m}, \mathcal{I}_{mv}, \mathcal{I}_p, \mathcal{I}_{hs}, \mathcal{I}^{\text{hit}}_{\text{dalal}}, \mathcal{I}_{nc}, \mathcal{I}_{D_f}, \mathcal{I}_{\mathsf{MI}^C}, \mathcal{I}_\eta, \mathcal{I}_{\mathsf{MI}}, \mathcal{I}_{mc}} \succ_p \boxed{\mathcal{I}^{\max}_{\text{dalal}}} \succ_p \boxed{\mathcal{I}_d}$$

Figure 1: Expressivity orders of the investigated inconsistency measures; measures within the same box are equally expressive wrt. the particular expressivity characteristic

on knowledge bases. However, it is also easy to see that $\mathcal{I}_1 \sqsubseteq \mathcal{I}_2$ implies $\mathcal{I}_2 \succeq_\alpha \mathcal{I}_1$ (for all $\alpha \in \{f, v, l, p\}$).

By exploiting the results from Theorem 1 we obtain the following simple corollary.

Corollary 1. *The expressivity orders wrt. α-characteristics* $\mathcal{C}^\alpha(\mathcal{I}, n)$ *(*$\alpha \in \{f, v, l, p\}$*) for the inconsistency measures* $\mathcal{I}_d$, $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^C}$, $\mathcal{I}_\eta$, $\mathcal{I}_c$, $\mathcal{I}_{LP_m}$, $\mathcal{I}_{mc}$, $\mathcal{I}_p$, $\mathcal{I}_{hs}$, $\mathcal{I}^{\Sigma}_{dalal}$, $\mathcal{I}^{\max}_{dalal}$, $\mathcal{I}^{hit}_{dalal}$, $\mathcal{I}_{D_f}$, $\mathcal{I}_{P_m}$, $\mathcal{I}_{mv}$, *and* $\mathcal{I}_{nc}$ *are as shown in Figure 1.*

The proof of the above corollary is omitted as the results follow directly from Theorem 1.

In Figure 1 the order $\succ_f$ does not show the placement of the measures $\mathcal{I}_{D_f}$, $\mathcal{I}_{\mathsf{MI}^C}$, and $\mathcal{I}_\eta$ as we only provided upper bounds for the corresponding characteristics in Theorem 1. However, we can give the following partial classification.

Corollary 2. *For* $\mathcal{I} \in \{\mathcal{I}_c, \mathcal{I}_{LP_m}, \mathcal{I}^{\Sigma}_{dalal}, \mathcal{I}^{\max}_{dalal}, \mathcal{I}_{P_m}, \mathcal{I}_{mv}\}$, $\mathcal{I} \succ_f \mathcal{I}_{D_f}$, $\mathcal{I} \succ_f \mathcal{I}_{\mathsf{MI}^C}$, $\mathcal{I} \succ_f \mathcal{I}_\eta$.

The proof of the above corollary is straightforward as, e. g., we provided a finite bound for $\mathcal{C}^f(\mathcal{I}_\eta, n)$ (for every n) while $\mathcal{C}^f(\mathcal{I}_c, n)$ is unbounded. Empirical evidence suggests also the following relationships, but a formal proof has yet to be found.

Conjecture 1. $\mathcal{I}_{D_f} \succ_f \mathcal{I}_{\mathsf{MI}^C} \succ_f \mathcal{I}_\eta \succ_f \mathcal{I}_{\mathsf{MI}}$.

Figure 1 shows that the measures $\mathcal{I}^{\Sigma}_{\text{dalal}}$ and $\mathcal{I}_{P_m}$ are the only measures that have maximal expressivity wrt. all four expressivity characteristics (among the considered inconsistency measures) and, as expected, the drastic inconsistency measure $\mathcal{I}_d$ is the least expressive one. One can also observe that for many measures their positioning in the orders $\succ_v$ and $\succ_p$ is complementary, i. e., if a measure has high expressivity wrt. $\mathcal{C}^f$ it has low expressivity wrt. $\mathcal{C}^v$ (consider e. g. $\mathcal{I}_c$ and $\mathcal{I}_p$). This is due to the fact that many measures measure only a specific aspect of inconsistency and usually belong either to the MI-based family of inconsistency measures—which focus on using minimal inconsistent subsets for measuring—or the variable-based family—which focus on conflicting propositions—, cf. (Hunter and Konieczny, 2008). Therefore, they are constrained in their expressivity if one of these dimensions is limited. For example, if the number of formulas in a knowledge base is restricted, so is the number of minimal inconsistent subsets.

Again, it should be noted that expressivity characteristics are meant to complement the investigation of rationality postulates, not to replace them. Rationality postulates are important to analyze the meaningfulness of the values of inconsistency measures, while our characteristics provide a quantitative assessment of their expressivity. However, we believe that the concept of expressivity characteristics and the results reported in this work will nurture general comparative analyses of inconsistency measures.

The expressivity characteristics considered in this paper each tackle one specific aspect of *size* of a knowledge base. Of course, one can also combine these characteristics to obtain hybrid versions via

$$\mathcal{C}^{\alpha,\alpha'}(\mathcal{I},n,m) = |\{\mathcal{I}(\mathcal{K}) \mid \mathcal{K} \in \mathbb{K}^{\alpha}(n) \cap \mathbb{K}^{\alpha'}(m)\}|$$

with $\alpha,\alpha' \in \{v,f,l,p\}$, $\alpha \neq \alpha'$, and $n,m > 0$. For example, $\mathcal{C}^{v,f}(\mathcal{I},n,m)$ is the number of different inconsistency values on knowledge bases which mention at most n propositions and consist of at most m formulas. A simple observation on these new characteristics is the following one.

Proposition 1. *Let $\alpha,\alpha' \in \{v,f,l,p\}$, $\alpha \neq \alpha'$, and $n,m > 0$. Then*

$$\mathcal{C}^{\alpha,\alpha'}(\mathcal{I},n,m) \leq \min\{\mathcal{C}^{\alpha}(\mathcal{I},n),\mathcal{C}^{\alpha'}(\mathcal{I},m)\}$$

The proof of the above proposition is straightforward. An investigation of these hybrid and other characteristics—and the resulting expressivity orders—is left for future work.

6 SUMMARY AND CONCLUSION

We conducted a focused but extensive comparative analysis of 16 inconsistency measures from the recent literature in terms of their expressivity. For that, we introduced 4 different expressivity characteristics and conducted an analytical evaluation of the considered measures wrt. these expressivity characteristics. Our findings revealed some interesting relationships of inconsistency measures to, e. g., set theory and monotone Boolean functions. Finally, the measures $\mathcal{I}^{\Sigma}_{\text{dalal}}$ (Grant and Hunter, 2013) and $\mathcal{I}_{P_m}$ (Jabbour and Raddaoui, 2013) have been proven to be maximally expressive wrt. all our characteristics.

Expressivity characteristics provide a novel evaluation method for assessing the quality of inconsistency measures. It has to be noted again, however, that high expressivity alone is not a sufficient criterion for doing this. It is straightforward to construct measures that exhibit maximal expressivity along all discussed dimensions, but fail to comply with the intuitions one expects from inconsistency measures. The use of rationality postulates—such as the ones presented and discussed in (Hunter and Konieczny, 2006; Mu *et al.*, 2011a; Besnard, 2014)—must still serve as first-level evaluation criterion. If measures satisfy the same (or a similar set of) rationality postulates, expressivity can be used to make further quality assessments.

To the best of our knowledge, our work is the most extensive comparative analysis of inconsistency measures so far. All inconsistency measures discussed in this paper have been implemented and an online interface to try out these measures is available[12].

APPENDIX: PROOFS OF TECHNICAL RESULTS

Theorem 1. *The α-characteristics $\mathcal{C}^{\alpha}(\mathcal{I},n)$ ($\alpha \in \{f,v,l,p\}$) for the inconsistency measures $\mathcal{I}_d$, $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^C}$, $\mathcal{I}_{\eta}$, $\mathcal{I}_c$, $\mathcal{I}_{mc}$, $\mathcal{I}_p$, $\mathcal{I}_{hs}$, $\mathcal{I}^{\Sigma}_{dalal}$, $\mathcal{I}^{\max}_{dalal}$, $\mathcal{I}^{hit}_{dalal}$, $\mathcal{I}_{D_f}$, $\mathcal{I}_{P_m}$, $\mathcal{I}_{mv}$, and $\mathcal{I}_{nc}$ are as shown in Table 5.*

Proof. Let $n > 0$ except in proofs regarding $\mathcal{C}^l$ where $n > 1$ is assumed (note that $\mathcal{C}^l(\mathcal{I},1) = 1$ for every measure $\mathcal{I}$ as every $\mathcal{K} \in \mathbb{K}^l(1)$ does not contain a negation and is therefore always consistent).

1. $\mathcal{C}^v(\mathcal{I}_d,n) = 2$
   By definition, $\mathcal{I}_d$ has co-domain $\{0,1\}$ and therefore $\mathcal{C}^v(\mathcal{I}_d,n) \leq 2$. For the knowledge bases $\mathcal{K}_8 = \{a\}$ and $\mathcal{K}_9 = \{a \wedge \neg a\}$ we get $\mathcal{I}_d(\mathcal{K}_8) = 0$ and $\mathcal{I}_d(\mathcal{K}_9) = 1$ and therefore $\mathcal{C}^v(\mathcal{I}_d,n) \geq 2$. As $\mathcal{K}_8$ and $\mathcal{K}_9$ use only one proposition the statement is true for all $n > 0$.

2. $\mathcal{C}^f(\mathcal{I}_d,n) = 2$
   Analogous to 1.

3. $\mathcal{C}^l(\mathcal{I}_d,n) = 2$
   Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_d,1) = 1$. Analogous to 1 but consider $\mathcal{K}_8 = \{a\}$ and $\mathcal{K}_7 = \{a, \neg a\}$.

4. $\mathcal{C}^p(\mathcal{I}_d,n) = 2$
   Analogous to 1.

12 http://tweetyproject.org/w/incmes/

	$\mathcal{C}^v(\mathcal{I}$,n)	$\mathcal{C}^f(\mathcal{I}$,n)	$\mathcal{C}^l(\mathcal{I}$,n)	$\mathcal{C}^p(\mathcal{I}$,n)
$\mathcal{I}_d$	2	2	2^*	2
$\mathcal{I}_{\mathsf{MI}}$	∞	$\binom{n}{\lfloor n/2 \rfloor}+1$	∞^*	∞
$\mathcal{I}_{\mathsf{MI}^C}$	∞	$\leq \Psi(n)^\ddagger$	∞^*	∞
$\mathcal{I}_\eta$	$\Phi(2^n)^\dagger$	$\leq \Phi(\binom{n}{\lfloor n/2 \rfloor})^\dagger$	∞^{**}	∞^*
$\mathcal{I}_c$	$n+1$	∞	∞^*	∞
$\mathcal{I}_{mc}$	∞	$\binom{n}{\lfloor n/2 \rfloor}^{**}$	∞^*	∞
$\mathcal{I}_p$	∞	$n+1$	∞^*	∞
$\mathcal{I}_{hs}$	2^n+1	$n+1$	∞^{**}	∞^*
$\mathcal{I}^{\Sigma}_{\text{dalal}}$	∞	∞^*	∞^*	∞
$\mathcal{I}^{\max}_{\text{dalal}}$	$n+2$	∞^*	$\lfloor (n+7)/3 \rfloor^{**}$	$n+2$
$\mathcal{I}^{\text{hit}}_{\text{dalal}}$	∞	$n+1$	∞^*	∞
$\mathcal{I}_{D_f}$	∞	$\leq \Psi(n)^\ddagger$	∞^*	∞
$\mathcal{I}_{P_m}$	∞	∞	∞^*	∞
$\mathcal{I}_{mv}$	$n+1$	∞^*	∞^*	∞
$\mathcal{I}_{nc}$	∞	$n+1$	∞^*	∞

Table 5: Characteristics of inconsistency measures ($n \geq 1$)
*only holds for $n > 1$
**only holds for $n > 3$
$^\dagger\Phi(x)$ is the number of fractions in the Farey series of order x and can be defined as $\Phi(x) = |\{k/l \mid l = 1, \ldots, x, k = 0, \ldots, l\}|$, see e. g. `http://oeis.org/A005728`
$^\ddagger\Psi(n)$ is the number of profiles of monotone Boolean functions of n variables, see e. g. `http://oeis.org/A220880`

5. $\mathcal{C}^v(\mathcal{I}_{\mathsf{MI}}, n) = \infty$
Consider for $i \in \mathbb{N}$ the knowledge bases $\mathcal{K}^2_i = \{\neg a, a, a \wedge a, a \wedge a \wedge a, \ldots, \bigwedge_{j=1}^{i} a\}$. Then $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}^2_i) = i$ and $\lim_{i\to\infty} \mathcal{I}_{\mathsf{MI}}(\mathcal{K}^2_i) = \infty$. As each $\mathcal{K}^2_i$ only uses one proposition the statement is true for every $n > 0$.

6. $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n) = \binom{n}{\lfloor n/2 \rfloor} + 1$
Note that for every inconsistent knowledge base $\mathcal{K}$ the set $\mathsf{MI}(\mathcal{K})$ is a *Sperner family* of $\mathcal{K}$, i. e. a set S of subsets from a set T for which for no two $X, Y \in S$ it holds $X \subseteq Y$. According to Sperner's theorem the maximal cardinality (which is also attained) of any Sperner family of a set T with $|T| = n$ is $\binom{n}{\lfloor n/2 \rfloor}$ (Sperner, 1928). If $\mathcal{K}$ is consistent we have $\mathsf{MI}(\mathcal{K}) = \emptyset$ and $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}) = 0$, yielding $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n) \leq \binom{n}{\lfloor n/2 \rfloor} + 1$. To show $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n) \geq \binom{n}{\lfloor n/2 \rfloor} + 1$ we show that every Sperner family can be represented through $\mathsf{MI}(\mathcal{K})$ of a knowledge base $\mathcal{K}$. Let $T = \{\alpha_1, \ldots, \alpha_n\}$ be a set and define a propositional signature

$\mathsf{At} = \{a_1, \ldots, a_n\}$. Let S be any Sperner family of T with cardinality $\binom{n}{\lfloor n/2 \rfloor}$. Define a knowledge base $\mathcal{K}^S_n = \{\phi_1, \ldots, \phi_n\}$ via

$$\phi_i = a_i \wedge \bigwedge_{M \in S, \alpha_i \in M} \bigvee_{\alpha_j \in M \setminus \{\alpha_i\}} \neg a_j$$

for $i = 1, \ldots, n$. Informally, every ϕ_i states that a_i is accepted and for each set M in S which contains α_i at least one of the other elements must not be accepted. If follows that $M = \{\alpha_{k_1}, \ldots, \alpha_{k_m}\}$ (for some $k_1, \ldots, k_m \in \{1, \ldots, n\}$) is an element of S if and only if the set $\{\phi_{k_1}, \ldots, \phi_{k_m}\}$ is a minimal inconsistent set. It follows $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}^S_n) = |\mathsf{MI}(\mathcal{K}^S_n)| = |S| = \binom{n}{\lfloor n/2 \rfloor}$. As removing any element from a Sperner family still yields a Sperner family, every value between 1 and $\binom{n}{\lfloor n/2 \rfloor}$ can be attained. Together with the fact that 0 is also a possible value of $\mathcal{I}_{\mathsf{MI}}$ we obtain $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n) \geq \binom{n}{\lfloor n/2 \rfloor} + 1$ and thus $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}}, n) = \binom{n}{\lfloor n/2 \rfloor} + 1$.

7. $\mathcal{C}^l(\mathcal{I}_{\mathsf{MI}}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{\mathsf{MI}}, 1) = 1$. Consider the family of knowledge bases $\mathcal{K}^3_i = \{a_1, \ldots, a_i, \neg a_1, \ldots, \neg a_i\}$ for $i \in \mathbb{N}$. Then $\mathcal{I}_{\mathsf{MI}}(\mathcal{K}^3_i) = i$ and $\mathcal{C}^l(\mathcal{I}_{\mathsf{MI}}, n) = \infty$ as only formulas of maximum length two have been used.

8. $\mathcal{C}^p(\mathcal{I}_{\mathsf{MI}}, n) = \infty$
Analogous to 7 (note that every formula in $\mathcal{K}^3_i$ mentions only one proposition).

9. $\mathcal{C}^v(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}^2_i$ from 5. Observe that $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}^2_i) = i/2$ and therefore $\mathcal{C}^v(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) = \infty$.

10. $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) \leq \Psi(n)$
Consider the vector $\mathsf{profile}(\mathcal{K}) = (\mathsf{MI}^0(\mathcal{K}), \ldots, \mathsf{MI}^n(\mathcal{K}))$, called MI-profile of $\mathcal{K}$ in the following, where $\mathsf{MI}^i(\mathcal{K})$ is the set if i-size minimal inconsistent subsets of $\mathcal{K}$. Note that every MI-profile induces the inconsistency value wrt. $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$ of its corresponding knowledge base by $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}) = \sum_{i=1}^n |\mathsf{MI}^i(\mathcal{K})| \cdot 1/i$. Furthermore, note that two distinct MI-profiles may yield the same inconsistency value, e. g. $(1, 0, 0)$ and $(0, 2, 0)$ yield the same inconsistency value 1. It follows

$$\begin{aligned}\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) \leq |\{(|\mathsf{MI}^0(\mathcal{K})|, \ldots, |\mathsf{MI}^n(\mathcal{K})|) \mid \\ (\mathsf{MI}^0(\mathcal{K}), \ldots, \mathsf{MI}^n(\mathcal{K})) \text{ is an} \\ \text{MI-profile for some } \mathcal{K} \in \mathbb{K}^f(n)\}|\end{aligned}$$

As discussed in 6, for every knowledge base $\mathcal{K}$ the set $\mathsf{MI}(\mathcal{K})$ is a Sperner family. It is well-known that there is an equivalence between

Sperner families and monotone boolean functions, cf. (Stephen and Yusun, 2012). In (Stephen and Yusun, 2012) the number of inequivalent monotone boolean functions has been investigated, see also `http://oeis.org/A220880`. These numbers are the same of inequivalent Sperner families as well. Here, two Sperner families $\mathsf{MI}(\mathcal{K})$ and $\mathsf{MI}(\mathcal{K}')$ are equivalent if they yield the same MI-profiles. The number of different MI-profiles is also the number on the right-hand side of the above equation, thus showing the claim $\mathcal{C}^f(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) \leq \Psi(n)$ where $\Psi(n)$ is the number of inequivalent monotone Boolean functions on n variables.

11. $\mathcal{C}^l(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, 1) = 1$. Consider then the family of knowledge bases $\mathcal{K}_i^3$ from 7 and observe than $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}_i^3) = i/2$ and therefore $\mathcal{C}^l(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) = \infty$.

12. $\mathcal{C}^p(\mathcal{I}_{\mathsf{MI}^\mathsf{C}}, n) = \infty$
Analogous to 11 (note that every formula in $\mathcal{K}_i^3$ mentions only one proposition).

13. $\mathcal{C}^v(\mathcal{I}_\eta, n) = \Phi(2^n)$
We first show $\mathcal{C}^v(\mathcal{I}_\eta, n) \leq \Phi(2^n)$. In (Knight, 2002) it has already been shown that $\mathcal{I}_\eta(\mathcal{K}) \in [0,1] \cap \mathbb{Q}$ for every $\mathcal{K}$ (Definition 2.7 and Theorem 2.28). Hence, assume $\eta = k/l$ and $\mathcal{I}_\eta(\mathcal{K}) = 1 - \eta$ for $k, l \in \mathbb{N}$ and $k \leq l$. We also assume for now that $\mathcal{K}$ contains no contradictory formula. Furthermore, we assume that $\mathcal{K}$ contains no free formulas (as $\mathcal{I}_\eta$ satisfies independence they have no influence on the inconsistency value, cf. (Thimm, 2013b)). Let P be a probability function such that $P(\phi) \geq k/l$ for all $\phi \in \mathcal{K}$. It can be assumed that there is no $\omega \in \mathsf{Int}(\mathsf{At})$ such that $P(\omega) > 0$ but $\omega \not\models \phi$ for every $\phi \in \mathcal{K}$ (otherwise one could set $P(\omega) = 0$ and distribute the "probability mass" $P(\omega)$ on the remaining interpretations which have already a positive probability; this cannot change the fact that $P(\phi) \geq k/l$ for all $\phi \in \mathcal{K}$). So it holds that for all $\omega \in \mathsf{Int}(\mathsf{At})$ we have that $P(\omega) > 0$ implies $\omega \models \phi$ for some $\phi \in \mathcal{K}$. Define $F_\mathcal{K}(\omega) = \{\phi \in \mathcal{K} \mid \omega \models \phi\}$ for all $\omega \in \mathsf{Int}(\mathsf{At})$, i.e., $F_\mathcal{K}(\omega)$ is the set of formulas in $\mathcal{K}$ that are satisfied by ω. We can furthermore assume that for all $\omega, \omega' \in \mathsf{Int}(\mathsf{At})$ with $P(\omega) > 0$ and $P(\omega') > 0$ we have $F_\mathcal{K}(\omega) \not\subseteq F_\mathcal{K}(\omega')$ (otherwise we could set $P(\omega) = 0$ and add the probability mass $P(\omega)$ to $P(\omega')$, without decreasing the probabilities of the formulas). Assume furthermore, that among all probability functions that satisfy the above constraints, P is one such that $|\{\omega | P(\omega) > 0\}|$ is minimal.

Now consider the case $|\{\omega \mid P(\omega) > 0\}| = 2$, i.e., there are two interpretations ω_1, ω_2 that receive positive probability. For every formula $\phi \in \mathcal{K}$ it holds that either $\phi \in F_\mathcal{K}(\omega_1)$, or $\phi \in F_\mathcal{K}(\omega_2)$, or $\phi \in F_\mathcal{K}(\omega_1) \cap F_\mathcal{K}(\omega_2)$. Note that the latter case cannot be possible for all $\phi \in \mathcal{K}$ as otherwise $F_\mathcal{K}(\omega_1) = F_\mathcal{K}(\omega_2)$. Furthermore, there is one formula $\phi' \in \mathcal{K}$ with $P(\phi') = P(\omega_1)$ and one formula $\phi'' \in \mathcal{K}$ with $P(\phi'') = P(\omega_2)$, otherwise we would have $F_{\mathcal{K}}(\omega_1) \subseteq F_{\mathcal{K}}(\omega_2)$ or $F_{\mathcal{K}}(\omega_2) \subseteq F_{\mathcal{K}}(\omega_1)$. As $P(\phi)$ has to be maximal for all $\phi \in \mathcal{K}$ we can conclude $P(\omega_1) = P(\omega_2)$ and therefore $\eta = 1/2$.

Now consider the case $|\{\omega \mid P(\omega) > 0\}| = 3$, i.e., there are three interpretations $\omega_1, \omega_2, \omega_3$ that receive positive probability. For each $\phi \in \mathcal{K}$ let $\Delta_P(\phi) = \{\omega \in \mathsf{Int}(\mathsf{At}) \mid P(\omega) > 0, \omega \models \phi\}$, i.e., $\Delta_P(\phi)$ is such that $P(\phi) = \sum_{\omega \in \Delta_P(\phi)} P(\omega)$. Note that it cannot be the case that $|\Delta_P(\phi)| = 3$ for any $\phi \in \mathcal{K}$ (otherwise ϕ would be free in $\mathcal{K}$) or that $|\Delta_P(\phi)| = 0$ (then ϕ would be self-contradictory). Consider the following sub-cases:

a) for all $\phi \in \mathcal{K}$ we have $|\Delta_P(\phi)| = 1$:
Then for all $\phi \in \mathcal{K}$ we have $P(\phi) = P(\omega)$ for some $\omega \in \{\omega_1, \omega_2, \omega_3\}$ and as there are not subset relations between any $F_{\mathcal{K}}(\omega_1)$, $F_{\mathcal{K}}(\omega_2)$, and $F_{\mathcal{K}}(\omega_3)$, it follows that $P(\omega_1) = P(\omega_2) = P(\omega_3) = 1/3$ maximizes each probability and we have $\eta = 1/3$.

b) for all $\phi \in \mathcal{K}$ we have $|\Delta_P(\phi)| = 2$:
Then for all $\phi \in \mathcal{K}$ we have $P(\phi) = P(\omega) + P(\omega')$ for some $\omega, \omega' \in \{\omega_1, \omega_2, \omega_3\}$ with $\omega \neq \omega'$ and as there are not subset relations between any $F_{\mathcal{K}}(\omega_1)$, $F_{\mathcal{K}}(\omega_2)$, and $F_{\mathcal{K}}(\omega_3)$, it follows that $P(\omega_1) = P(\omega_2) = P(\omega_3) = 1/3$ maximizes each probability and we have $\eta = 2/3$.

c) otherwise:
Let $\phi_1 \in \mathcal{K}$ with $|\Delta_P(\phi_1)| = 1$. Without loss of generality assume $\Delta_P(\phi_1) = \{\omega_1\}$. As $F_{\mathcal{K}}(\omega_2) \not\subseteq F_{\mathcal{K}}(\omega_1)$ there is $\phi_2 \in F_{\mathcal{K}}(\omega_2)$ with $\phi_2 \notin F_{\mathcal{K}}(\omega_1)$. Consider the case that for all $\phi \in F_{\mathcal{K}}(\omega_3)$ either $\phi \in F_{\mathcal{K}}(\omega_2)$ or $\phi \in F_{\mathcal{K}}(\omega_2)$. Then P' defined via $P'(\omega_1) = 0.5$, $P'(\omega_2) = 0.5$, and $P'(\omega) = 0$ for all other ω yields $P'(\phi) \geq 0.5$ for all $\phi \in \mathcal{K}$. Assuming P obtains a larger probability for all formulas it must hold $P(\omega) > 0.5$ (in order to have $P(\phi_1) > 0.5$), but then $P(\phi_2) < 0.5$. So we have a contradiction since P is supposed to be minimal wrt. $|\{\omega | P(\omega) > 0\}|$. It follows that there is $\phi_3 \in F_{\mathcal{K}}(\omega_3)$ with $\phi_3 \notin F_{\mathcal{K}}(\omega_1)$ and $\phi_3 \notin F_{\mathcal{K}}(\omega_2)$, so $P(\phi_3) = P(\omega_3)$. Similarly, it can be assumed that $\phi_2 \notin F_{\mathcal{K}}(\omega_3)$ as well. As $P(\phi_1) = P(\omega_1)$, $P(\phi_2) = P(\omega_2)$, and $P(\phi_3) = P(\omega_3)$ it follows that $P(\omega_1) = P(\omega_2) = P(\omega_3) = 1/3$ maximizes each probability and we have $\eta = 1/3$.

So for $|\{\omega \mid P(\omega) > 0\}| = 3$ we have that $\eta \in \{1/3, 2/3\}$. Inductively it follows that for $|\{\omega \mid P(\omega) > 0\}| = h$ we have $\eta \in \{1/h, \ldots, (h-1)/h\}$. As a signature with n propositions has 2^n different interpretations, and together with the cases of a consistent knowledge base (inconsistency value 0) and one that contains a contradictory formula (inconsistency value 1) we obtain $\mathcal{C}^v(\mathcal{I}_\eta, n) \leq |\{k/l \mid l = 1, \ldots, 2^n, k = 0, \ldots, l\}| = \Phi(2^n)$.

We now show $\mathcal{C}^v(\mathcal{I}_\eta, n) \geq \Phi(2^n)$. For that let $\eta = k/l$ for $l \in \{1,\ldots,2^n\}$ and $k \in \{1,\ldots,l\}$. Let $X = \{\omega_1,\ldots,\omega_l\} \subseteq \mathsf{Int}(\mathsf{At})$ be any set of l different interpretations. Define P via $P(\omega) = 1/l$ if $\omega \in X$ and $P(\omega) = 0$ otherwise. Define $\hat{\mathcal{K}}$ via $\phi \in \hat{\mathcal{K}}$ if and only if

$$\phi = \bigvee_{\omega \in X_k} \phi_\omega$$

where X_k is a k-element subset of X and ϕ_ω is the complete conjunction that has $\omega \in \mathsf{Int}(\mathsf{At})$ as its only model (note that $\hat{\mathcal{K}}$ contains $\binom{l}{k}$ formulas, one for each k-element subset of X). Observe that for all $\phi \in \hat{\mathcal{K}}$ we have

$$P(\phi) = P(\bigvee_{\omega \in X_k} \phi_\omega) = \sum_{\omega \in X_k} P(\phi_\omega) = k/l$$

and that this is obviously the maximal possible value for $\hat{\mathcal{K}}$. It follows $\mathcal{I}_\eta(\hat{\mathcal{K}}) = 1 - k/l$ and therefore $\mathcal{C}^v(\mathcal{I}_\eta, n) \geq \Phi(2^n)$.

14. $\mathcal{C}^f(\mathcal{I}_\eta, n) \leq \Phi(\binom{n}{\lfloor n/2 \rfloor})$
Analogous to 13. However, note that that maximal number of interpretations that may receive a positive probability is bounded by the number of different $F_\mathcal{K}(\omega)$ for $\omega \in \mathsf{Int}(\mathsf{At})$. As the set of $F_\mathcal{K}(\omega)$ with $P(\omega) > 0$ form a Sperner family (no two elements have a subset relation) the maximal cardinality of this set is $\binom{n}{\lfloor n/2 \rfloor}$, cf. (Sperner, 1928).

15. $\mathcal{C}^l(\mathcal{I}_\eta, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_\eta, 1) = 1$. For $n = 2$ observe that either $\mathcal{I}_\eta(\mathcal{K}) = 0$ (for consistent $\mathcal{K}$) or $\mathcal{I}_\eta(\mathcal{K}) = 0.5$. For the latter, note that if $\mathcal{K}$ can only be inconsistent if and only if there is at least one (possibly more) $a \in \mathsf{At}$ such that $a, \neg a \in \mathcal{K}$ (or semantically equivalent formulas). Then any probability function P with $P(\phi)$ maximal for all $\phi \in \mathcal{K}$ has to satisfy $P(a) = P(\neg a) = 0.5$. Therefore we have $\mathcal{C}^l(\mathcal{I}_\eta, 2) = 2$. For $n = 3$ we additionally have the case that a three-element minimal inconsistent subset $\{\neg a_1, \neg a_2, a_1 \vee a_2\}$ may occur with corresponding inconsistency value $1/3$, thus $\mathcal{C}^l(\mathcal{I}_\eta, n) = 3$. For $n > 3$ consider the family of knowledge bases $\mathcal{K}_i^4 = \{\neg a_1 \vee a_2, \neg a_2 \vee a_3, \ldots, \neg a_{i-1} \vee a_i, \neg a_i \wedge a_1\}$. Note that $\mathcal{K}_i^4$ is a minimal inconsistent set. By Theorem 2.12 of (Knight, 2002) it follows $\mathcal{I}_\eta(\mathcal{K}_i^4) = 1/|\mathcal{K}_i^4| = 1/i$ and therefore $\mathcal{C}^l(\mathcal{I}_\eta, n) = \infty$.

16. $\mathcal{C}^p(\mathcal{I}_\eta, n) = \infty$
First, for $n = 1$ observe that either $\mathcal{I}_\eta(\mathcal{K}) = 0$ (for consistent $\mathcal{K}$), $\mathcal{I}_\eta(\mathcal{K}) = 1$ (for $\mathcal{K}$ containing a contradictory formula), or $\mathcal{I}_\eta(\mathcal{K}) = 0.5$. For the latter, note that if $\mathcal{K}$ can only be inconsistent without containing a contradictory formula if and only if there is at least one (possibly more) $a \in \mathsf{At}$ such that $a, \neg a \in \mathcal{K}$ (or semantically equivalent formulas). Then any probability function P with $P(\phi)$

maximal for all $\phi \in \mathcal{K}$ has to satisfy $P(a) = P(\neg a) = 0.5$. For $n > 1$ consider the family of knowledge bases $\mathcal{K}_i^4$ from 15. Note that $\mathcal{K}_i^4$ is a minimal inconsistent set. By Theorem 2.12 of (Knight, 2002) it follows $\mathcal{I}_\eta(\mathcal{K}_i^4) = 1/|\mathcal{K}_i^4| = 1/i$ and therefore $\mathcal{C}^p(\mathcal{I}_\eta, n) = \infty$.

17. $\mathcal{C}^v(\mathcal{I}_c, n) = n + 1$
Consider the propositional signature $\mathsf{At} = \{a_1, \dots, a_n\}$ and for each $i = 0, \dots, n$ consider the knowledge base $\mathcal{K}_i^5 = \{a_1 \wedge \neg a_1, \dots, a_i \wedge \neg a_i\}$ (with $\mathcal{K}_0^5 = \emptyset$). Then $\mathcal{I}_c(\mathcal{K}_i^5) = i$ as every $a_1, \dots, a_i$ has to be set to B in every model of $\mathcal{K}_i^5$. Together with the fact that every model can assign the value B to at most $|\mathsf{At}| = n$ different propositions we have $\mathcal{C}^v(\mathcal{I}_c, n) = n + 1$.

18. $\mathcal{C}^f(\mathcal{I}_c, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}_i^6 = \{a_1 \wedge \ldots \wedge a_i \wedge \neg a_1 \wedge \ldots \wedge \neg a_i\}$. Then $\mathcal{I}_c(\mathcal{K}_i^6) = i$ for $i > 0$ and $\lim_{i \to \infty} \mathcal{I}_c(\mathcal{K}_i^6) = \infty$. As each $\mathcal{K}_i^6$ has only one formula the statement is true for every $n > 0$.

19. $\mathcal{C}^l(\mathcal{I}_c, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_c, 1) = 1$. Consider the family of knowledge bases $\mathcal{K}_i^3 = \{a_1, \dots, a_i, \neg a_1, \dots, \neg a_i\}$ for $i \in \mathbb{N}$. Then $\mathcal{I}_c(\mathcal{K}_i^3) = i$ and $\mathcal{C}^l(\mathcal{I}_c, n) = \infty$ as only formulas of maximum length two have been used.

20. $\mathcal{C}^p(\mathcal{I}_c, n) = \infty$
Analogous to 19 (note that every formula in $\mathcal{K}_i^3$ mentions only one proposition).

21. $\mathcal{C}^v(\mathcal{I}_{mc}, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}_i^7 = \{a \wedge \neg a, a \wedge a \wedge \neg a \wedge \neg a, \ldots \bigwedge_{j=1}^{i} a \wedge \neg a\}$ and observe $\mathcal{I}_{mc}(\mathcal{K}_i^7) = |\mathsf{SC}(\mathcal{K}_i^7)| = i + 1$ (all formulas in $\mathcal{K}_i^7$ are self-contradicting and only the empty subset is a maximal consistent subset). It follows $\mathcal{C}^v(\mathcal{I}_{mc}, n) = \infty$.

22. $\mathcal{C}^f(\mathcal{I}_{mc}, n) = \binom{n}{\lfloor n/2 \rfloor}$
Note first, that if $\mathcal{K}$ contains only self-contradictory formulas we have $\mathtt{MC}(\mathcal{K}) = \emptyset$. Otherwise, analogously to 6, observe that for every other consistent or inconsistent knowledge base $\mathcal{K}$ the set $\mathtt{MC}(\mathcal{K})$ is a *Sperner family* of $\mathcal{K}$, i. e. a set S of subsets from a set T for which for no two $X, Y \in S$ it holds $X \subseteq Y$. According to Sperner's theorem the maximal cardinality (which is also attained) of any Sperner family of a set T with $|T| = n$ is $\binom{n}{\lfloor n/2 \rfloor}$ (Sperner, 1928). Hence we have $0 \leq |\mathtt{MC}(\mathcal{K})| \leq \binom{n}{\lfloor n/2 \rfloor}$. Also analogously to 6 observe that every value can be attained by some knowledge base. For that let S be any Sperner family of cardinality $\lfloor n/2 \rfloor$ of a set $T = \{\alpha_1, \dots, \alpha_n\}$. Let

$\mathsf{At} = \{a_1, \ldots, a_n\}$ and define a knowledge base $\hat{\mathcal{K}}_n^S = \{\phi_1, \ldots, \phi_n\}$ with

$$\phi_i = \bigvee_{\alpha_i \in M \in S} (\bigwedge_{\alpha_j \in M} a_j \wedge \bigwedge_{\alpha_j \notin M} \neg a_j)$$

for $i = 1, \ldots, n$. Informally, every ϕ_i lists all sets $M \in S$ that include α_i. Then a set $M \subseteq \hat{\mathcal{K}}_n^S$ is a maximal consistent subset if and only if it corresponds to an element of S. As removing any element from a Sperner family still yields a Sperner family, we have $\{|\mathtt{MC}(\mathcal{K})| \mid \mathcal{K} \in \mathbb{K}^f(n)\} = \{i \in \mathbb{N} \mid 0 \leq i \leq \binom{n}{\lfloor n/2 \rfloor}\}$.

Note that for $|\mathcal{K}| = n$ it holds $0 \leq |\mathsf{SC}(\mathcal{K})| \leq n$ (and every value can be attained). However, we cannot simply obtain the value $\mathcal{C}^f(\mathcal{I}_{mc}, n)$ by adding the upper bounds of $|\mathtt{MC}(\mathcal{K})|$ and $|\mathsf{SC}(\mathcal{K})|$ as these two values are dependent. Observe that if $\mathcal{K}$ with $|\mathcal{K}| = n$ contains a self-contradictory formula ϕ then ϕ cannot be part of any maximal consistent subset of $\mathcal{K}$, i e., we have $\mathtt{MC}(\mathcal{K}) = \mathtt{MC}(\mathcal{K} \setminus \{\phi\})$. In general, we have that if $\mathcal{K}$ contains k contradictory formulas then we have $|\mathtt{MC}(\mathcal{K})| \leq \binom{(n-k)}{\lfloor (n-k)/2 \rfloor}$. Define $c_{mc}^{n,k} = \binom{(n-k)}{\lfloor (n-k)/2 \rfloor}$ and then we obtain the following characterization of $\mathcal{C}^f(\mathcal{I}_{mc}, n)$:

$$\mathcal{C}^f(\mathcal{I}_{mc}, n) = \max\{c_{mc}^{n,0}, c_{mc}^{n,1} + 1, c_{mc}^{n,2} + 2, \ldots, c_{mc}^{n,n} + n\}$$

That is, the value $\mathcal{C}^f(\mathcal{I}_{mc}, n)$ is either $c_{mc}^{n,0}$ (considering no self-contradictory formulas) or $c_{mc}^{n,1} + 1$ (considering one self-contradictory formula), etc.. Observe that the first element of the above maximum is dominant for $n > 3$. For $n = 1$ we obtain $\mathcal{C}^f(\mathcal{I}_{mc}, n) = 2$ (a knowledge base with one formula is either consistent, i. e., $\mathtt{MC}(\mathcal{K}) = \{\mathcal{K}\}$, $\mathsf{SC}(\mathcal{K}) = \emptyset$, and thus $\mathcal{I}_{mc}(\mathcal{K}) = 0$; or it is inconsistent with $\mathtt{MC}(\mathcal{K}) = \{\emptyset\}$, $\mathsf{SC}(\mathcal{K}) = \mathcal{K}$, and thus $\mathcal{I}_{mc}(\mathcal{K}) = 1$), note that either the empty set or the whole set are the only possible maximal consistent subsets. For $n = 2$ we obtain $\mathcal{C}^f(\mathcal{I}_{mc}, n) = 3$: either $\mathcal{K}$ is consistent ($\mathcal{I}_{mc}(\mathcal{K}) = 0$), or it contains one self-contradictory formula ($\mathcal{I}_{mc}(\mathcal{K}) = 1$), or it contains two contradictory formulas ($\mathcal{I}_{mc}(\mathcal{K}) = 2$), note that the maximal number of consistent subsets of $\mathcal{K}$ is 2 (for the case that $\mathcal{K}$ is a two-element minimal inconsistent set), but then there cannot be self-contradictory formulas and we have $\mathcal{I}_{mc}(\mathcal{K}) = 1$. For $n = 3$ we obtain $\mathcal{C}^f(\mathcal{I}_{mc}, n) = 4$ (for a consistent knowledge base and knowledge bases with 1 to 3 self-contradictory formulas and one maximal consistent subset), note that the maximal number of consistent subsets of $\mathcal{K}$ is 3, e. g. all two-element subsets, but then $\mathcal{K}$ cannot contain any self-contradictory formula and we have $\mathcal{I}_{mc}(\mathcal{K}) = 2$ which we can also obtain by having two self-contradictory formulas. For $n > 3$ and $k = 1, \ldots, n$ we have

$$\begin{aligned}
c_{mc}^{n,0} = \binom{n}{\lfloor n/2 \rfloor} &= \binom{n-1}{\lfloor n/2 \rfloor} + \underbrace{\binom{n-1}{\lfloor n/2 \rfloor - 1}}_{\geq n \text{ as } n>3} \\
&\geq \binom{n-1}{\lfloor n/2 \rfloor} + n \\
&= \binom{n-1}{\lfloor (n-1)/2 \rfloor} + n \\
&\geq \binom{n-k}{\lfloor (n-k)/2 \rfloor} + n \\
&\geq \binom{n-k}{\lfloor (n-k)/2 \rfloor} + k = c_{mc}^{n,k} + k
\end{aligned}$$

Note that $\binom{n-1}{\lfloor n/2 \rfloor} = \binom{n-1}{\lfloor (n-1)/2 \rfloor}$ holds as for odd n we have $\binom{n-1}{\lfloor n/2 \rfloor} = \binom{n-1}{(n-1)/2} = \binom{n-1}{\lfloor (n-1)/2 \rfloor}$ and for even n we have $\binom{n-1}{\lfloor n/2 \rfloor} = \binom{n-1}{n/2}$ which (as $n-1$ is odd) is the same as $\binom{n-1}{n/2-1} = \binom{n-1}{\lfloor (n-1)/2 \rfloor}$. Hence, for $n > 3$ we obtain $\mathcal{C}^f(\mathcal{I}_{mc}, n) = \binom{n}{\lfloor n/2 \rfloor}$.

23. $\mathcal{C}^l(\mathcal{I}_{mc}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{mc}, 1) = 1$. Note furthermore that for $n = 2$ only literals are allowed as formulas in $\mathcal{K}$. Consider the family of knowledge bases $\mathcal{K}_i^3 = \{a_1, \neg a_1, \ldots, a_i, \neg a_i\}$ and observe $|\mathcal{K}_i^3| = 2i$ and $\mathcal{I}_{mc}(\mathcal{K}_i^3) = 2^i$ (every interpretation ω corresponds to a maximal consistent subset of $\mathcal{K}_i^3$, i.e., the union of all a_i with $\omega(a_i) = \mathsf{true}$ and $\neg a_i$ with $\omega(a_i) = \mathsf{false}$; adding any other formula from $\mathcal{K}_i^3$ makes this set inconsistent). As only formulas of maximum length 2 are used in $\mathcal{K}_i^3$ it follows $\mathcal{C}^l(\mathcal{I}_{mc}, n) = \infty$.

24. $\mathcal{C}^p(\mathcal{I}_{mc}, n) = \infty$
Consider the family knowledge bases $\mathcal{K}_i^7$ from 21 and observe that every formula mentions only one proposition. It follows $\mathcal{C}^p(\mathcal{I}_{mc}, n) = \infty$.

25. $\mathcal{C}^v(\mathcal{I}_p, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}_i^2$ from 5. Then $\mathcal{I}_p(\mathcal{K}_i^2) = i$ and $\lim_{i\to\infty} \mathcal{I}_p(\mathcal{K}_i^2) = \infty$. As each $\mathcal{K}_i^2$ only uses one proposition the statement is true for every $n > 0$.

26. $\mathcal{C}^f(\mathcal{I}_p, n) = n + 1$
For $k < n$ let $M_k = \{a_1, \ldots, a_{k-1}, \neg a_1 \vee \ldots \neg a_{k-1}\}$. Note that M_k is a minimal inconsistent set. Consider $\mathcal{K}_{n,k} = M_k \cup \{a_{k+1}, \ldots, a_n\}$. Then $\mathcal{K}_{n,k}$ has exactly one minimal inconsistent subset (M_k) and $\mathcal{I}_p(\mathcal{K}_{n,k}) = k$. Hence, for $k = 1, \ldots, n$ every value in $\{1, \ldots, n\}$ is attained for

$\mathcal{I}_p(\mathcal{K}_{n,k})$. Together with $\mathcal{I}_p(\mathcal{K}) = 0$ for any consistent $\mathcal{K}$ of size n we have $\mathcal{C}^f(\mathcal{I}_p, n) = n + 1$.

27. $\mathcal{C}^l(\mathcal{I}_p, n) = \infty$
Analogous to 19.

28. $\mathcal{C}^p(\mathcal{I}_p, n) = \infty$
Analogous to 20.

29. $\mathcal{C}^v(\mathcal{I}_{hs}, n) = 2^n + 1$
Any hitting set can be of maximal size 2^n as there are that many interpretations in a language with n propositions, and as a hitting set may not be defined in case of a contradictory formula we get $\mathcal{C}^v(\mathcal{I}_{hs}, n) \leq 2^n + 1$. Let now At be a propositional signature with $|\mathsf{At}| = n$ and consider the knowledge base $\mathcal{K}^8_n = \bigcup_{\omega \in \mathsf{Int}(\mathsf{At})} \{\phi_\omega\}$ where ϕ_ω is any formula with $\mathsf{Mod}(\phi_\omega) = \{\omega\}$. Then $\mathcal{K}^8_n$ contains 2^n formulas and each of them is satisfied by only one interpretation. We get $\mathcal{I}_{hs}(\mathcal{K}^8_n) = 2^n - 1$ and removing any formula from $\mathcal{K}^8_n$ reduces the value by one, so all values $0, \ldots, 2^n - 1$ are attained. Taking the case of a knowledge base with a contradictory formula into account, we obtain $\mathcal{C}^v(\mathcal{I}_{hs}, n) \geq 2^n + 1$ and thus $\mathcal{C}^v(\mathcal{I}_{hs}, n) = 2^n + 1$.

30. $\mathcal{C}^f(\mathcal{I}_{hs}, n) = n + 1$
For $\mathcal{K}$ with $|\mathcal{K}| = n$ any hitting set can be of maximal size n, as only formulas in $\mathcal{K}$ need to be hit. Considering the case of a knowledge base with a contradictory formula with get $\mathcal{C}^f(\mathcal{I}_{hs}, n) \leq n + 1$. For $\mathcal{C}^f(\mathcal{I}_{hs}, n) \geq n + 1$ consider a knowledge base with n pairwise inconsistent formulas, such as in 29 (note that the signature can be arbitrarily large). Therefore we get $\mathcal{C}^f(\mathcal{I}_{hs}, n) = n + 1$.

31. $\mathcal{C}^l(\mathcal{I}_{hs}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{hs}, 1) = 1$. For $n = 2$ or $n = 3$ consider the interpretations ω_1, ω_2 with $\omega_1(a) = \mathsf{true}$ and $\omega_2(a) = \mathsf{false}$ for all $a \in \mathsf{At}$. As formulas of maximal length 2 are either a, or $\neg a$, and formulas of length 3 are either $a \wedge b$ or $a \vee b$ for $a, b \in \mathsf{At}$, either ω_1 or ω_2 is a model of each formula. Therefore, $\mathcal{I}_{hs}(\mathcal{K}) = 1$ or $\mathcal{I}_{hs}(\mathcal{K}) = 0$ and $\mathcal{C}^l(\mathcal{I}_{hs}, 2) = \mathcal{C}^l(\mathcal{I}_{hs}, 3) = 2$. For $n > 3$ consider the signature $\mathsf{At}_m = \{a_1, \ldots, a_m\}$ and the knowledge base $\mathcal{K}^9_m = \{a \wedge b, \neg a \wedge b, a \wedge \neg b \mid a, b \in \mathsf{At}_m, a \neq b\}$. Observe that for $m \to \infty$ we have $\mathcal{I}_{hs}(\mathcal{K}^9_m) \to \infty$ and therefore $\mathcal{C}^l(\mathcal{I}_{hs}, n) = \infty$.

32. $\mathcal{C}^p(\mathcal{I}_{hs}, n) = \infty$
First, for $n = 1$ observe that either $\mathcal{I}_{hs}(\mathcal{K}) = 0$ (for consistent $\mathcal{K}$), $\mathcal{I}_{hs}(\mathcal{K}) = \infty$ (for $\mathcal{K}$ containing a contradictory formula), or $\mathcal{I}_{hs}(\mathcal{K}) = 1$. For the latter, note that given a signature $\mathsf{At}_m = \{a_1, \ldots, a_m\}$ the two interpretations ω_1, ω_2 with $\omega_1(a) = \mathsf{true}$ and $\omega_2(a) = \mathsf{false}$, for all $a \in \mathsf{At}_m$, form a hitting set for every knowledge base where the formulas mention at most one proposition. It follows $\mathcal{C}^p(\mathcal{I}_{hs}, 1) =$

3. For $n > 1$ consider the signature $\mathsf{At}_m = \{a_1, \ldots, a_m\}$ and the knowledge base $\mathcal{K}^9_m = \{a \wedge b, \neg a \wedge b, a \wedge \neg b, \neg a \wedge \neg b \mid a, b \in \mathsf{At}_m, a \neq b\}$. Observe that for $m \to \infty$ we have $\mathcal{I}_{hs}(\mathcal{K}^9_m) \to \infty$ and therefore $\mathcal{C}^p(\mathcal{I}_{hs}, n) = \infty$.

33. $\mathcal{C}^v(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$
Consider the family of knowledge bases

$$\mathcal{K}^{10}_i = \{a, \neg a, a \wedge a, \neg a \wedge \neg a, \ldots, \bigwedge_{j=1}^{i} a, \bigwedge_{j=1}^{i} \neg a\}$$

for $i \in \mathbb{N}$. Then $\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^{10}_i) = i$ as there are only two interpretations ω_1 and ω_2 with $\omega_1(a) = \mathsf{true}$ and $\omega_2(a) = \mathsf{false}$ and for both interpretations it is the case that the sets of models of half of the formulas in $\mathcal{K}^{10}_i$ have a distance of one to each of them (note that $|\mathcal{K}^{10}_i| = 2i$). Therefore we have $\mathcal{C}^v(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$.

34. $\mathcal{C}^f(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$
For $\mathcal{K}$ with $|\mathcal{K}| = 1$ we have that either $\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}) = 0$ or $\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}) = \infty$ and therefore $\mathcal{C}^f(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = 2$. For $n > 1$ consider the family of knowledge bases $\mathcal{K}^1_i = \{a_1 \wedge \ldots \wedge a_i, \neg a_1 \wedge \ldots \wedge \neg a_i\}$ for $i \in \mathbb{N}$ and observe $\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^1_i) = i$. Therefore we have $\mathcal{C}^f(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$.

35. $\mathcal{C}^l(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}^{\Sigma}_{\text{dalal}}, 1) = 1$. Then for $i \in \mathbb{N}$ consider the family of knowledge bases $\mathcal{K}^3_i = \{a_1, \ldots, a_i, \neg a_1, \ldots, \neg a_i\}$ and observe $\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}^3_i) = i$. Hence, we obtain $\mathcal{C}^l(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$.

36. $\mathcal{C}^p(\mathcal{I}^{\Sigma}_{\text{dalal}}, n) = \infty$
Analogous to 33 (note that in every formula of $\mathcal{K}^0_i$ only one proposition is used).

37. $\mathcal{C}^v(\mathcal{I}^{\max}_{\text{dalal}}, n) = n + 2$
For every consistent knowledge base $\mathcal{K}$ we have $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}) = 0$ and for the knowledge base $\mathcal{K}^9 = \{a \wedge \neg a\}$ we have $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}^9) = \infty$. Furthermore, for the signature $\mathsf{At}_i = \{a_1, \ldots, a_i\}$ and $i = 1, \ldots, n$ consider the family of knowledge bases $\mathcal{K}^{11}_i = \{\phi_\omega \mid \omega \in \mathsf{Int}(\mathsf{At}_i)\}$ where ϕ_ω is any formula with $\mathsf{Mod}(\phi_\omega) = \{\omega\}$. Observe that for every $\omega \in \mathsf{Int}(\mathsf{At}_i)$ there is one formula $\phi \in \mathcal{K}^{11}_i$ with $d_{\text{d}}(\mathsf{Mod}(\phi), \omega) = i$ and therefore $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}^{11}_i) = i$. Note also that $d_{\text{d}}(\omega, \omega') \leq i$ for every pair $\omega, \omega' \in \mathsf{Int}(\mathsf{At}_i)$ as ω and ω' can differ in at most i propositions. Hence, we have $\mathcal{C}^v(\mathcal{I}^{\max}_{\text{dalal}}, n) = n + 2$.

38. $\mathcal{C}^f(\mathcal{I}^{\max}_{\text{dalal}}, n) = \infty$
First, for $\mathcal{K}$ with $|\mathcal{K}| = 1$ observe that either $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}) = 0$ or $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}) = \infty$ ($\mathcal{K}$ can only be inconsistent if it contains a contradictory formula and then the Dalal distance between the set of models

of this formula (which is the empty set) to any interpretation is ∞). Therefore we have $\mathcal{C}^f(\mathcal{I}^{\max}_{\text{dalal}}, 1) = 2$. For $n > 1$ consider the family of knowledge bases $\mathcal{K}^1_i = \{a_1 \wedge \ldots \wedge a_i, \neg a_1 \wedge \ldots \wedge \neg a_i\}$ with $i \in \mathbb{N}$. Then we have $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}^1_i) = \lceil i/2 \rceil$ and therefore $\mathcal{C}^f(\mathcal{I}^{\max}_{\text{dalal}}, n) = \infty$.

39. $\mathcal{C}^l(\mathcal{I}^{\max}_{\text{dalal}}, n) = \lfloor (n+7)/3 \rfloor$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}^{\max}_{\text{dalal}}, 1) = 1$. For $n = 2$ observe that $\mathcal{K}$ only contains propositions or negations of propositions and therefore, the set of models of every formula of $\mathcal{K}$ has at most distance 1 to any interpretation. Therefore, $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}) = 0$ or $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}) = 1$ and we have $\mathcal{C}^l(\mathcal{I}^{\max}_{\text{dalal}}, 2) = 2$. For $n = 3$ observe that formulas may only have the form a, $\neg a$, $a \wedge b$, or $a \vee b$ for $a, b \in \mathsf{At}$. Note that only the models of a formula $a \wedge b$ may have distance 2 to some interpretation. However, as a conjunction cannot contain a negation, the minimal maximal distance of any formula from $\mathcal{K}$ is 1 as the models of e. g. $\neg a$ has only distance 1 to any models of $a \wedge b$. As there can also be no contradictory formula (for that a length of 4 of a formula is required) we get $\mathcal{C}^l(\mathcal{I}^{\max}_{\text{dalal}}, 3) = 2$. For $n = 4$ consider the interpretation ω with $\omega(a) = \mathsf{true}$ for every $a \in \mathsf{Int}(\mathsf{At})$. As every conjunction can contain at most one negation, the distance of the models of every formula to ω is also maximally 1. Additionally, we have self-contradictory formulas which yield in total $\mathcal{C}^l(\mathcal{I}^{\max}_{\text{dalal}}, 4) = 3$. Assume $n > 4$ with $n = 2 + 3k$ with $k > 0$ and let ϕ_ω be the complete conjunction that has $\omega \in \mathsf{Int}(\mathsf{At})$ as its only model. Then observe that $\mathcal{K}^{12}_k = \{\phi_\omega \mid \omega \in \mathsf{Int}(\{a_0, \ldots, a_k\})\}$ has only formulas of maximal length n and $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}^{12}_k) = k + 1$ (and that smaller inconsistency values can be attained by removing some formula in $\mathcal{K}^{12}_i$). Observe further that for $n \in \{3 + 3k, 4 + 3k\}$ the maximal distance cannot be larger than for $n = 2 + 3k$. Together with consistent knowledge bases and knowledge bases containing self-contradictory formulas we obtain $\mathcal{C}^l(\mathcal{I}^{\max}_{\text{dalal}}, n) = \lfloor (n-2)/3 \rfloor + 3 = \lfloor (n+7)/3 \rfloor$ which holds for $n > 3$.

40. $\mathcal{C}^p(\mathcal{I}^{\max}_{\text{dalal}}, n) = n + 2$
Consider the signature $\mathsf{At}_i = \{a_1, \ldots, a_i\}$ for $i = 1, \ldots, n$ and the family of knowledge bases $\mathcal{K}^{11}_i = \{\phi_\omega \mid \omega \in \mathsf{Int}(\mathsf{At}_i)\}$. Note that every $\phi \in \mathcal{K}^{11}_i$ mentions exactly i propositions and, with the same argumentation as in 37, we have $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}^{11}_i) = i$. Furthermore, for a consistent knowledge base $\mathcal{K}$ we have $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}) = 0$ and for the knowledge base $\mathcal{K}^9 = \{a \wedge \neg a\}$ we have $\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}^9) = \infty$ and therefore $\mathcal{C}^p(\mathcal{I}^{\max}_{\text{dalal}}, n) = n + 2$.

41. $\mathcal{C}^v(\mathcal{I}^{\text{hit}}_{\text{dalal}}, n) = \infty$
For every $i \in \mathbb{N}$ consider the knowledge base $\mathcal{K}^{10}_i = \{a, \neg a, a \wedge a, \neg a \wedge \neg a, \ldots, \bigwedge_{j=1}^{i} a, \bigwedge_{j=1}^{i} \neg a\}$. Then $\mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K}^{10}_i) = i$ (consider e. g. the interpretation ω with $\omega(a) = \mathsf{true}$, then the models of half of the formulas of $\mathcal{K}^{10}_i$ have distance one to ω and $|\mathcal{K}^{10}_i| = 2i$). As only one

proposition is necessary to create any inconsistency value we have $\mathcal{C}^v(\mathcal{I}^{\text{hit}}_{\text{dalal}}, n) = \infty$.

42. $\mathcal{C}^f(\mathcal{I}^{\text{hit}}_{\text{dalal}}, n) = n + 1$
For every consistent knowledge base $\mathcal{K}$ we have $\mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K}) = 0$ and for the family of knowledge base $\mathcal{K}^5_i = \{a_1 \wedge \neg a_1, \ldots, a_i \wedge \neg a_i\}$ (for $i = 1, \ldots, n$) we have $\mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K}^5_i) = i$. Note that n is also the maximal value of $\mathcal{I}^{\text{hit}}_{\text{dalal}}$ as this is the maximal number of formulas in any $\mathcal{K}$. Hence, we have $\mathcal{C}^f(\mathcal{I}^{\text{hit}}_{\text{dalal}}, n) = n + 1$.

43. $\mathcal{C}^l(\mathcal{I}^{\text{hit}}_{\text{dalal}}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}^{\text{hit}}_{\text{dalal}}, 1) = 1$. Consider the family of knowledge bases $\mathcal{K}^3_i = \{a_1, \neg a_1, \ldots, a_i, \neg a_i\}$. Then $\mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K}^3_i) = i$ and therefore $\mathcal{C}^l(\mathcal{I}^{\text{hit}}_{\text{dalal}}, n) = \infty$ as only formula of length 2 are necessary to produce any inconsistency value.

44. $\mathcal{C}^p(\mathcal{I}^{\text{hit}}_{\text{dalal}}, n) = \infty$
Analogous to 41.

45. $\mathcal{C}^v(\mathcal{I}_{D_f}, n) = \infty$
Consider for $i \in \mathbb{N}$ the knowledge bases $\mathcal{K}^2_i = \{\neg a, a, a \wedge a, a \wedge a \wedge a, \ldots, \bigwedge_{j=1}^{i} a\}$. Then $\mathsf{MI}^{(1)}(\mathcal{K}^2_i) = \mathsf{MI}^{(3)}(\mathcal{K}^2_i) = \mathsf{MI}^{(4)}(\mathcal{K}^2_i) = \ldots = \mathsf{MI}^{(|\mathcal{K}^2_i|)}(\mathcal{K}^2_i) = \emptyset$ and $\mathsf{MI}^{(2)}(\mathcal{K}^2_i) = \{\{\neg a, a\}, \{\neg a, a \wedge a\}, \{\neg a, a \wedge a \wedge a\}, \ldots, \{\neg a, \bigwedge_{j=1}^{i} a\}\}$. Therfore $|\mathsf{MI}^{(1)}(\mathcal{K}^2_i)| = |\mathsf{MI}^{(3)}(\mathcal{K}^2_i)| = |\mathsf{MI}^{(4)}(\mathcal{K}^2_i)| = \ldots = |\mathsf{MI}^{(|\mathcal{K}^2_i|)}(\mathcal{K}^2_i)| = 0$ and $|\mathsf{MI}^{(2)}(\mathcal{K}^2_i)| = i$. Furthermore, note that $\mathsf{CN}^{(2)}$ is comprised of every two-element subset of $\mathcal{K}^2_i \setminus \{\neg a\}$ and therefore $|\mathsf{CN}^{(2)}| = \binom{i}{2}$. It follows $R_1(\mathcal{K}^2_i) = R_3(\mathcal{K}^2_i) = R_4(\mathcal{K}^2_i) = \ldots = R_{|\mathcal{K}^2_i|-1}(\mathcal{K}^2_i) = 0$ and $R_2(\mathcal{K}^2_i) = i/(i + \binom{i}{2})$. We obtain $\mathcal{I}_{D_f}(\mathcal{K}^2_i) = 1 - R_2(\mathcal{K}^2_i)/2 = 1 - i/(2i + 2\binom{i}{2}) = 1 - i/(2i + i(i-1)) = 1 - 1/(i+1)$. As $i \in \mathbb{N}$ we obtain $\mathcal{C}^v(\mathcal{I}_{D_f}, n) = \infty$.

46. $\mathcal{C}^f(\mathcal{I}_{D_f}, n) \leq \Psi(n)$
Analogous to 10. Observe, that each MI-profile $(\mathsf{MI}^0(\mathcal{K}), \ldots, \mathsf{MI}^n(\mathcal{K}))$ also uniquely determines the corresponding CN-profile $(\mathsf{CN}^0(\mathcal{K}), \ldots, \mathsf{CN}^n(\mathcal{K}))$. That is, there are no two knowledge bases $\mathcal{K}$ and $\mathcal{K}'$ that have the same MI-profile but different CN-profiles. Then there cannot be more R-profiles $(R_1(\mathcal{K}), \ldots, R_n(\mathcal{K}))$ then there are MI-profiles and, thus, the number of inequivalent monotone Boolean functions on n variables is also an upper bound for $\mathcal{C}^f(\mathcal{I}_{D_f}, n)$.

47. $\mathcal{C}^l(\mathcal{I}_{D_f}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{D_f}, 1) = 1$. Consider then the family of knowledge bases $\mathcal{K}^3_i$ from 7. Observe that $\mathsf{MI}^2(\mathcal{K}^3_i) = i$

and $\mathsf{MI}^j(\mathcal{K}_i^3) = 0$ for $j \neq 2$. Furthermore, it is $\mathsf{CN}^2 = \binom{2i}{2} - i$ and therefore $R_2(\mathcal{K}_i^3) = i/\binom{2i}{2}$.Then

$$\mathcal{I}_{D_f}(\mathcal{K}_i^3) = 1 - \frac{i}{2\binom{2i}{2}}$$

Hence, we have $\mathcal{C}^l(\mathcal{I}_{D_f}, n) = \infty$.

48. $\mathcal{C}^p(\mathcal{I}_{D_f}, n) = \infty$
Analogous to 45.

49. $\mathcal{C}^v(\mathcal{I}_{P_m}, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}_i^2 = \{\neg a, a, a \wedge a, \ldots, \bigwedge_{j=1}^{i} a\}$ for $i \in \mathbb{N}$. Then $\mathcal{K}_i^2$ contains one minimal proof for $\neg a$ and i minimal proofs for a (each formula other than $\neg a$ is a minimal proof for a). Then we have $\mathcal{I}_{D_f}(\mathcal{K}_i^2) = i$ and therefore $\mathcal{C}^v(\mathcal{I}_{D_f}, n) = \infty$.

50. $\mathcal{C}^f(\mathcal{I}_{P_m}, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}_i^6 = \{a_1 \wedge \ldots \wedge a_i \wedge \neg a_1 \wedge \ldots \wedge \neg a_i\}$ for $i \in \mathbb{N}$. Then $\mathcal{K}_i^6$ contains one minimal proof for each a_j and one minimal proof for each $\neg a_j$ for $j = 1, \ldots, i$. Then we have $\mathcal{I}_{D_f}(\mathcal{K}_i^6) = i$ and therefore $\mathcal{C}^f(\mathcal{I}_{D_f}, n) = \infty$.

51. $\mathcal{C}^l(\mathcal{I}_{P_m}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{D_f}, 1) = 1$. Consider the family of knowledge bases $\mathcal{K}_i^3 = \{a_1, \ldots, a_i, \neg a_1, \ldots, \neg a_i\}$ for $i \in \mathbb{N}$. Then $\mathcal{K}_i^3$ contains one minimal proof for each a_j and one minimal proof for each $\neg a_j$ for $j = 1, \ldots, i$. Then we have $\mathcal{I}_{D_f}(\mathcal{K}_i^3) = i$ and therefore $\mathcal{C}^l(\mathcal{I}_{D_f}, n) = \infty$ as only formulas of maximal size 2 have been used.

52. $\mathcal{C}^p(\mathcal{I}_{P_m}, n) = \infty$
Analogous to 51 (note that $\mathcal{K}_i^3$ mentions only one proposition in each formula).

53. $\mathcal{C}^v(\mathcal{I}_{mv}, n) = n + 1$
Consider the propositional signature $\mathsf{At} = \{a_1, \ldots, a_n\}$ and for each $i = 0, \ldots, n$ consider the knowledge base $\mathcal{K}_i^{13} = \{a_1 \wedge \neg a_1, \ldots, a_i \wedge \neg a_i, a_{i+1}, \ldots, a_n\}$ (with $\mathcal{K}_0^{13} = \emptyset$). Then $\mathcal{I}_{mv}(\mathcal{K}_i^{13}) = i/n$ and we obtain $\mathcal{C}^v(\mathcal{I}_{mv}, n) = n + 1$.

54. $\mathcal{C}^f(\mathcal{I}_{mv}, n) = \infty$
If $|\mathcal{K}| = 1$ then observe that either $\mathcal{I}_{mv}(\mathcal{K}) = 0$ or $\mathcal{I}_{mv}(\mathcal{K}) = 1$ (if the knowledge base is inconsistent then all propositions appearing in $\mathcal{K}$ also appear in the only formula of $\mathcal{K}$ and are thus part of a minimal inconsistent subset) , so we have $\mathcal{C}^f(\mathcal{I}_{mv}, 1) = 2$. For

$n = |\mathcal{K}| > 1$ consider for $i \in \mathbb{N}$ the family of knowledge bases $\mathcal{K}_i^{14} = \{a_1 \wedge \ldots \wedge a_{i-1}, a_i \wedge \neg a_i\}$ and observe $\mathcal{I}_{mv}(\mathcal{K}_i^{14}) = 1/i$. Hence, we have $\mathcal{C}^f(\mathcal{I}_{mv}, n) = \infty$ for $n > 1$.

55. $\mathcal{C}^l(\mathcal{I}_{mv}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{mv}, 1) = 1$. For $n > 1$ consider $\mathcal{K}_i^{15} = \{\neg a_1, a_1, a_2, \ldots, a_i\}$ with $i > 1$ and note that $\{\neg a_1, a_1\}$ is the only minimal inconsistent subset of $\mathcal{K}_i^{15}$. Then $\mathcal{I}_{mv}(\mathcal{K}_i^{15}) = 1/i$ and therefore $\mathcal{C}^l(\mathcal{I}_{mv}, n) = \infty$.

56. $\mathcal{C}^p(\mathcal{I}_{mv}, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}_i^{16} = \{a_1 \wedge \neg a_1, a_2, \ldots, a_i\}$ for $i \in \mathbb{N}$. Then $\mathcal{I}_{mv}(\mathcal{K}_i^{16}) = 1/i$ and therefore $\mathcal{C}^p(\mathcal{I}_{mv}, n) = \infty$.

57. $\mathcal{C}^v(\mathcal{I}_{nc}, n) = \infty$
Consider the family of knowledge bases $\mathcal{K}_i^2$ from 5. Then $\mathcal{I}_{nc}(\mathcal{K}_i^2) = i$ and $\lim_{i\to\infty} \mathcal{I}_{nc}(\mathcal{K}_i^2) = \infty$. As each $\mathcal{K}_i^2$ only uses one proposition the statement is true for every $n > 0$.

58. $\mathcal{C}^f(\mathcal{I}_{nc}, n) = n + 1$
Obviously, $\mathcal{I}_{nc}(\mathcal{K}) \leq |\mathcal{K}| \in \{0, \ldots, n\}$ for every knowledge base. Consider the family of knowledge bases $\mathcal{K}_{n,k}$ from 26. Then $\mathcal{I}_p(\mathcal{K}_{n,k}) = k - 1$ for $k = 1, \ldots, n$. Furthermore, a knowledge base $\mathcal{K}$ containing at least one contradictory formula has $\mathcal{I}_{nc}(\mathcal{K}) = |\mathcal{K}| = n$. Hence, we have $\mathcal{C}^f(\mathcal{I}_{nc}, n) = n + 1$.

59. $\mathcal{C}^l(\mathcal{I}_{nc}, n) = \infty$
Note that $n > 1$ is assumed as trivially $\mathcal{C}^l(\mathcal{I}_{nc}, 1) = 1$. For $n > 1$ consider $\mathcal{K}_i^{15} = \{\neg a_1, a_1, a_2, \ldots, a_i\}$ with $i > 1$ and note that $\{\neg a_1, a_1\}$ is the only minimal inconsistent subset of $\mathcal{K}_i^{15}$. Then $\mathcal{I}_{nc}(\mathcal{K}_i^{15}) = |\mathcal{K}_i| - 1 = i$ and therefore $\mathcal{C}^l(\mathcal{I}_{nc}, n) = \infty$.

60. $\mathcal{C}^p(\mathcal{I}_{nc}, n) = \infty$
Analogous to 57.

□

Theorem 3.

1. *The measures $\mathcal{I}_d$ and $\mathcal{I}_c$ satisfy $\wedge$-Indifference, $\wedge$-Penalty, and $\wedge$-Mitigation.*
2. *The measures $\mathcal{I}_\eta$, $\mathcal{I}_{hs}$, and $\mathcal{I}_{dalal}^{\max}$ satisfy $\wedge$-Penalty, but not $\wedge$-Mitigation.*
3. *The measures $\mathcal{I}_{dalal}^{hit}$ and $\mathcal{I}_{P_m}$ satisfy $\wedge$-Mitigation, but not $\wedge$-Penalty.*
4. *None of the measures $\mathcal{I}_{\mathsf{MI}}$, $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$, $\mathcal{I}_{mc}$, $\mathcal{I}_p$, $\mathcal{I}_{dalal}^{\Sigma}$, $\mathcal{I}_{D_f}$, $\mathcal{I}_{mv}$, $\mathcal{I}_{nc}$ satisfies any of $\wedge$-Indifference, $\wedge$-Penalty, or $\wedge$-Mitigation.*

Proof. Let $\mathcal{K}$ be some arbitrary knowledge base and $\alpha, \beta \in \mathcal{L}(\mathsf{At})$ formulas.

1. Note first that if an inconsistency measure $\mathcal{I}$ satisfies $\wedge$-Indifference it also satisfies $\wedge$-Penalty and $\wedge$-Mitigation. Therefore, we only show that $\mathcal{I}_d$ and $\mathcal{I}_c$ satisfy $\wedge$-Indifference.

   a) $\mathcal{I}_d$: the statement follows directly from that fact that every knowledge base $\mathcal{K} \cup \{\alpha, \beta\}$ is inconsistent if and only if $\mathcal{K} \cup \{\alpha \wedge \beta\}$ is inconsistent.

   b) $\mathcal{I}_c$: let v be a three-valued interpretation $v : \mathsf{At} \to \{T, F, B\}$ with $v \models^3 \mathcal{K} \cup \{\alpha, \beta\}$. Then $v(\alpha), v(\beta) \in \{T, B\}$. From Table 2 it follows $v(\alpha \wedge \beta) \in \{T, B\}$ and therefore $v \models^3 \mathcal{K} \cup \{\alpha \wedge \beta\}$ as well. Note that the converse holds as well. Then we have

$$\begin{aligned}\mathcal{I}_c(\mathcal{K} \cup \{\alpha, \beta\}) &= \min\{|v^{-1}(B)| \mid v \models^3 \mathcal{K} \cup \{\alpha, \beta\}\} \\ &= \min\{|v^{-1}(B)| \mid v \models^3 \mathcal{K} \cup \{\alpha \wedge \beta\}\} \\ &= \mathcal{I}_c(\mathcal{K} \cup \{\alpha \wedge \beta\})\end{aligned}$$

2. a) $\mathcal{I}_\eta$: let P be a probability distribution $P\mathsf{Int}(\mathsf{At}) \to [0, 1]$ with $P(\phi) \geq \hat{\xi}$ for all $\phi \in \mathcal{K} \cup \{\alpha \wedge \beta\}$ with $\hat{\xi}$ being maximal. Then $P(\alpha) \geq \hat{\xi}$ and $P(\beta) \geq \hat{\xi}$ as well. It follows

$$\begin{aligned}\hat{\xi} &= \max\{\xi \mid \exists P \in \mathcal{P}(\mathsf{At}) : \forall \phi \in \mathcal{K} \cup \{\alpha \wedge \beta\} : P(\phi) \geq \xi\} \\ &\leq \max\{\xi \mid \exists P \in \mathcal{P}(\mathsf{At}) : \forall \phi \in \mathcal{K} \cup \{\alpha, \beta\} : P(\phi) \geq \xi\}\end{aligned}$$

   and therefore $\mathcal{I}_\eta(\mathcal{K} \cup \{\alpha \wedge \beta\}) \geq \mathcal{I}_\eta(\mathcal{K} \cup \{\alpha, \beta\})$. To see that $\mathcal{I}_\eta$ does not satisfy $\wedge$-Mitigation consider $\mathcal{K}_{10} = \{\neg a \wedge b, a \wedge \neg b\}$ and

$$\begin{aligned}\mathcal{I}_\eta(\mathcal{K}_{10} \cup \{a \wedge b\}) &= 2/3 \\ \mathcal{I}_\eta(\mathcal{K}_{10} \cup \{a, b\}) &= 1/2\end{aligned}$$

   b) $\mathcal{I}_{hs}$: let H be a hitting set of $\mathcal{K} \cup \{\alpha \wedge \beta\}$ and let $\omega \in H$ be such that $\omega \models \alpha \wedge \beta$. Then $\omega \models \alpha$ and $\omega \models \beta$ as well and H is a hitting set of $\mathcal{K} \cup \{\alpha, \beta\}$. So we have that every hitting set of $\mathcal{K} \cup \{\alpha \wedge \beta\}$ is also a hitting set of $\mathcal{K} \cup \{\alpha, \beta\}$ and therefore

$$\begin{aligned}&\{H \mid H \text{ is a hitting set of } \mathcal{K} \cup \{\alpha \wedge \beta\}\} \\ \subseteq &\{H \mid H \text{ is a hitting set of } \mathcal{K} \cup \{\alpha, \beta\}\}\end{aligned}$$

   and

$$\begin{aligned}\mathcal{I}_{hs}(\mathcal{K} \cup \{\alpha \wedge \beta\}) &= \min\{|H| \mid H \text{ is a hitting set of } \mathcal{K} \cup \{\alpha \wedge \beta\}\} \\ &\geq \min\{|H| \mid H \text{ is a hitting set of } \mathcal{K} \cup \{\alpha, \beta\}\} \\ &= \mathcal{I}_{hs}(\mathcal{K} \cup \{\alpha, \beta\})\end{aligned}$$

To see that $\mathcal{I}_{hs}$ does not satisfy $\wedge$-Mitigation consider $\mathcal{K}_{10} = \{\neg a \wedge b, a \wedge \neg b\}$ and

$$\begin{aligned}\mathcal{I}_{hs}(\mathcal{K}_{10} \cup \{a \wedge b\}) &= 2\\ \mathcal{I}_{hs}(\mathcal{K}_{10} \cup \{a, b\}) &= 1\end{aligned}$$

c) $\mathcal{I}^{\max}_{\text{dalal}}$: Let $\omega \in \mathsf{Int}(\mathsf{At})$ be arbitrary. As $\mathsf{Mod}(\alpha \wedge \beta) \subseteq \mathsf{Mod}(\alpha)$ and $\mathsf{Mod}(\alpha \wedge \beta) \subseteq \mathsf{Mod}(\beta)$ we have

$$\begin{aligned}d_{\text{d}}(\mathsf{Mod}(\alpha \wedge \beta), \omega) &\geq d_{\text{d}}(\mathsf{Mod}(\alpha), \omega)\\ d_{\text{d}}(\mathsf{Mod}(\alpha \wedge \beta), \omega) &\geq d_{\text{d}}(\mathsf{Mod}(\beta), \omega)\end{aligned}$$

It follows

$$\max_{\phi \in \mathcal{K} \cup \{\alpha \wedge \beta\}} d_{\text{d}}(\mathsf{Mod}(\phi), \omega) \geq \max_{\phi \in \mathcal{K} \cup \{\alpha, \beta\}} d_{\text{d}}(\mathsf{Mod}(\phi), \omega)$$

and

$$\begin{aligned}\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K} \cup \{\alpha \wedge \beta\}) &= \min\{\max_{\phi \in \mathcal{K} \cup \{\alpha \wedge \beta\}} d_{\text{d}}(\mathsf{Mod}(\phi), \omega) \mid \omega \in \mathsf{Int}(\mathsf{At})\}\\ &\geq \min\{\max_{\phi \in \mathcal{K} \cup \{\alpha, \beta\}} d_{\text{d}}(\mathsf{Mod}(\phi), \omega) \mid \omega \in \mathsf{Int}(\mathsf{At})\}\\ &= \mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K} \cup \{\alpha, \beta\})\end{aligned}$$

To see that $\mathcal{I}^{\max}_{\text{dalal}}$ does not satisfy $\wedge$-Mitigation consider $\mathcal{K}_{11} = \{\neg a \wedge b, a \wedge \neg b, \neg a \wedge \neg b\}$ and

$$\begin{aligned}\mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}_{11} \cup \{a \wedge b\}) &= 2\\ \mathcal{I}^{\max}_{\text{dalal}}(\mathcal{K}_{11} \cup \{a, b\}) &= 1\end{aligned}$$

3. a) $\mathcal{I}^{\text{hit}}_{\text{dalal}}$: Let $\omega \in \mathsf{Int}(\mathsf{At})$ be arbitrary. Note that $\omega \not\models \alpha \wedge \beta$ if and only if either $\omega \not\models \alpha$, $\omega \not\models \beta$, or both. Therefore, $d_{\text{d}}(\mathsf{Mod}(\alpha \wedge \beta), \omega) > 0$ if and only if either $d_{\text{d}}(\mathsf{Mod}(\alpha), \omega) > 0$, $d_{\text{d}}(\mathsf{Mod}(\beta), \omega) > 0$, or both. Then it holds

$$\begin{aligned}&|\{\phi \in \mathcal{K} \cup \{\alpha \wedge \beta\} \mid d_{\text{d}}(\mathsf{Mod}(\phi), \omega) > 0\}|\\ \leq &|\{\phi \in \mathcal{K} \cup \{\alpha, \beta\} \mid d_{\text{d}}(\mathsf{Mod}(\phi), \omega) > 0\}|\end{aligned}$$

and $\mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K} \cup \{\alpha \wedge \beta\}) \leq \mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K} \cup \{\alpha, \beta\})$. To see that $\mathcal{I}^{\text{hit}}_{\text{dalal}}$ does not satisfy $\wedge$-Penalty consider $\mathcal{K}_{12} = \{\neg a, \neg b\}$ and

$$\begin{aligned}\mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K}_{12} \cup \{a \wedge b\}) &= 1\\ \mathcal{I}^{\text{hit}}_{\text{dalal}}(\mathcal{K}_{12} \cup \{a, b\}) &= 2\end{aligned}$$

b) $\mathcal{I}_{P_m}$: observe that if π is a minimal proof for some γ in $\mathcal{K} \cup \{\alpha \wedge \beta\}$ it holds that, if $\alpha \wedge \beta \in \pi$ then either

i. $\pi \setminus \{\alpha \wedge \beta\} \cup \{\alpha\}$, or

ii. $\pi \setminus \{\alpha \wedge \beta\} \cup \{\beta\}$,

iii. both of the above, or

iv. $\pi \setminus \{\alpha \wedge \beta\} \cup \{\alpha, \beta\}$

is a minimal proof (are minimal proofs) for γ in $\mathcal{K} \cup \{\alpha, \beta\}$. In any case, for every γ it holds $|P_m^{\mathcal{K}\cup\{\alpha\wedge\beta\}}(\gamma)| \leq |P_m^{\mathcal{K}\cup\{\alpha,\beta\}}(\gamma)|$ and therefore $\mathcal{I}_{P_m}(\mathcal{K} \cup \{\alpha \wedge \beta\}) \leq \mathcal{I}_{P_m}(\mathcal{K} \cup \{\alpha, \beta\})$. To see that $\mathcal{I}_{P_m}$ does not satisfy $\wedge$-Penalty consider $\mathcal{K}_{13} = \{\neg a\}$ and

$$\mathcal{I}_{P_m}(\mathcal{K}_{13} \cup \{a \wedge \neg\neg a\}) = 1$$
$$\mathcal{I}_{P_m}(\mathcal{K}_{13} \cup \{a, \neg\neg a\}) = 2$$

4. a) $\mathcal{I}_{\mathsf{MI}}$: consider Example 16 with $\mathcal{K}_7 = \{a, \neg a\}$. Here we have

$$\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_7 \cup \{a, b\}) = 1 < 2 = \mathcal{I}_{\mathsf{MI}}(\mathcal{K}_7 \cup \{a \wedge b\})$$

and

$$\mathcal{I}_{\mathsf{MI}}(\mathcal{K}_7 \cup \{a \wedge \neg a, \neg\neg a\}) = 3 > 2 = \mathcal{I}_{\mathsf{MI}}(\mathcal{K}_7 \cup \{a \wedge \neg a \wedge \neg\neg a\})$$

b) $\mathcal{I}_{\mathsf{MI}^\mathsf{C}}$: consider Example 16 with $\mathcal{K}_7 = \{a, \neg a\}$. Here we have

$$\mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}_7 \cup \{a, b\}) = 1/2 < 1 = \mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}_7 \cup \{a \wedge b\})$$

and

$$\mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}_7 \cup \{a \wedge \neg a, \neg\neg a\}) = 2 > 3/2 = \mathcal{I}_{\mathsf{MI}^\mathsf{C}}(\mathcal{K}_7 \cup \{a \wedge \neg a \wedge \neg\neg a\})$$

c) $\mathcal{I}_{mc}$: consider Example 16 with $\mathcal{K}_7 = \{a, \neg a\}$. Here we have

$$\mathcal{I}_{mc}(\mathcal{K}_7 \cup \{a, \neg a\}) = 1 < 2 = \mathcal{I}_{mc}(\mathcal{K}_7 \cup \{a \wedge \neg a\})$$

and for $\mathcal{K}_9 = \{a, b\}$ we have

$$\mathcal{I}_{mc}(\mathcal{K}_9 \cup \{\neg a, \neg b\}) = 2 > 1 = \mathcal{I}_{mc}(\mathcal{K}_9 \cup \{\neg a \wedge \neg b\})$$

d) $\mathcal{I}_p$: consider $\mathcal{K}_{10} = \{\neg a, a, \neg b, b\}$. Here we have

$$\mathcal{I}_p(\mathcal{K}_{10} \cup \{a, b\}) = 4 < 5 = \mathcal{I}_p(\mathcal{K}_{10} \cup \{a \wedge b\})$$

and for $\mathcal{K}_5 = \{a, b\}$ we have

$$\mathcal{I}_p(\mathcal{K}_5 \cup \{\neg a, \neg b\}) = 4 > 3 = \mathcal{I}_p(\mathcal{K}_5 \cup \{\neg a \wedge \neg b\})$$

e) $\mathcal{I}^{\Sigma}_{\text{dalal}}$: consider $\mathcal{K}_5 = \{a, b\}$. Here we have

$$\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}_5 \cup \{\neg a, a\}) = 1 < \infty = \mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}_5 \cup \{\neg a \wedge a\})$$

and for $\mathcal{K}_{14} = \{a, b, a \wedge b, \neg a \wedge \neg b\}$ we have

$$\mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}_{14} \cup \{\neg a, \neg b\}) = 4 > 2 = \mathcal{I}^{\Sigma}_{\text{dalal}}(\mathcal{K}_{14} \cup \{\neg a \wedge \neg b\})$$

f) $\mathcal{I}_{D_f}$: consider Example 16 with $\mathcal{K}_7 = \{a, \neg a\}$. Here we have

$$\mathcal{I}_{D_f}(\mathcal{K}_7 \cup \{a, b\}) = 1/6 < 1/3 = \mathcal{I}_{D_f}(\mathcal{K}_7 \cup \{a \wedge b\})$$

and for $\mathcal{K}_{15} = \{a \wedge b, a \wedge c, a \wedge d\}$ we have

$$\mathcal{I}_{D_f}(\mathcal{K}_{15} \cup \{\neg a, \neg a \wedge e\}) = 3/10 > 1/4 = \mathcal{I}_{D_f}(\mathcal{K}_{15} \cup \{\neg a \wedge \neg a \wedge e\})$$

g) $\mathcal{I}_{mv}$: consider $\mathcal{K}_8 = \{a\}$. Here we have

$$\mathcal{I}_{D_f}(\mathcal{K}_8 \cup \{\neg a, b\}) = 1 < 2 = \mathcal{I}_{D_f}(\mathcal{K}_8 \cup \{\neg a \wedge b\})$$

and for $\mathcal{K}_6 = \{a \wedge b\}$ we have

$$\mathcal{I}_{D_f}(\mathcal{K}_6 \cup \{\neg a, a\}) = 2 > 1 = \mathcal{I}_{D_f}(\mathcal{K}_6 \cup \{\neg a \wedge a\})$$

h) $\mathcal{I}_{nc}$: consider $\mathcal{K}_8 = \{a\}$. Here we have

$$\mathcal{I}_{nc}(\mathcal{K}_8 \cup \{\neg a, a\}) = 1 < 2 = \mathcal{I}_{nc}(\mathcal{K}_8 \cup \{\neg a \wedge a\})$$

and for $\mathcal{K}_7 = \{a, \neg a\}$ from Example 16 we have

$$\mathcal{I}_{nc}(\mathcal{K}_7 \cup \{b, c\}) = 3 > 2 = \mathcal{I}_{nc}(\mathcal{K}_7 \cup \{b \wedge c\})$$

$\square$

APPENDIX: LIST OF KNOWLEDGE BASES

$$\begin{aligned}
\mathcal{K}_1 &= \{a, b \vee c, \neg a \wedge \neg b, d\} \\
\mathcal{K}_2 &= \{a, \neg a, b, \neg b\} \\
\mathcal{K}_3 &= \{a, \neg a, b, c, d\} \\
\mathcal{K}_4 &= \{a, b, c, \neg a \vee \neg b \vee \neg c, \neg(a \wedge b \wedge c)\} \\
\mathcal{K}_5 &= \{a, b\} \\
\mathcal{K}_6 &= \{a \wedge b\} \\
\mathcal{K}_7 &= \{a, \neg a\} \\
\mathcal{K}_8 &= \{a\}
\end{aligned}$$

$$
\begin{aligned}
\mathcal{K}_9 &= \{a \wedge \neg a\} \\
\mathcal{K}_{10} &= \{\neg a \wedge b, a \wedge \neg b\} \\
\mathcal{K}_{11} &= \{\neg a \wedge b, a \wedge \neg b, \neg a \wedge \neg b\} \\
\mathcal{K}_{12} &= \{\neg a, \neg b\} \\
\mathcal{K}_{13} &= \{\neg a\} \\
\mathcal{K}_{14} &= \{a, b, a \wedge b, \neg a \wedge \neg b\} \\
\mathcal{K}_{15} &= \{a \wedge b, a \wedge c, a \wedge d\} \\
\mathcal{K}_i^1 &= \{a_1 \wedge \ldots \wedge a_i, \neg a_1 \wedge \ldots \wedge \neg a_i\} && i \in \mathbb{N} \\
\mathcal{K}_i^2 &= \{\neg a, a, a \wedge a, a \wedge a \wedge a, \ldots, \bigwedge_{j=1}^{i} a\} && i \in \mathbb{N} \\
\mathcal{K}_i^3 &= \{a_1, \ldots, a_i, \neg a_1, \ldots, \neg a_i\} && i \in \mathbb{N} \\
\mathcal{K}_i^4 &= \{\neg a_1 \vee a_2, \neg a_2 \vee a_3, \ldots, \neg a_{i-1} \vee a_i, a_i \wedge \neg a_1\} && i \in \mathbb{N} \\
\mathcal{K}_i^5 &= \{a_1 \wedge \neg a_1, \ldots, a_i \wedge \neg a_i\} && i \in \mathbb{N} \\
\mathcal{K}_i^6 &= \{a_1 \wedge \ldots \wedge a_i \wedge \neg a_1 \wedge \ldots \wedge \neg a_i\} && i \in \mathbb{N} \\
\mathcal{K}_i^7 &= \{a \wedge \neg a, a \wedge a \wedge \neg a \wedge \neg a, \ldots \bigwedge_{j=1}^{i} a \wedge \neg a\} && i \in \mathbb{N} \\
\mathcal{K}_i^8 &= \{a_1 \wedge \neg a_1, a_2, \ldots, a_i\} && i \in \mathbb{N} \\
\mathcal{K}_i^9 &= \{a \wedge b, \neg a \wedge b, a \wedge \neg b \mid a, b \in \mathsf{At}_i, a \neq b\} && i \in \mathbb{N} \\
\mathcal{K}_i^{10} &= \{a, \neg a, a \wedge a, \neg a \wedge \neg a, \ldots, \bigwedge_{j=1}^{i} a, \bigwedge_{j=1}^{i} \neg a\} && i \in \mathbb{N} \\
\mathcal{K}_i^{11} &= \{\phi_\omega \mid \omega \in \mathsf{Int}(\mathsf{At}_i)\} && i \in \mathbb{N} \\
\mathcal{K}_i^{12} &= \{\neg a_1, a_1, a_2, \ldots, a_i\} && i \in \mathbb{N} \\
\mathcal{K}_i^{13} &= \{a_1 \wedge \neg a_1, \ldots, a_i \wedge \neg a_i, a_{i+1}, \ldots, a_n\} && i \in \mathbb{N} \\
\mathcal{K}_i^{14} &= \{a_1 \wedge \ldots \wedge a_{i-1}, a_i \wedge \neg a_i\} && i \in \mathbb{N} \\
\mathcal{K}_i^{15} &= \{\neg a_1 \wedge a_1 \wedge a_2 \wedge \ldots \wedge a_i\} && i \in \mathbb{N} \\
\mathcal{K}_{i,j}^{16} &= \{a_1, \ldots, a_i, \neg a_1, \ldots, \neg a_i\} && i, j \in \mathbb{N}
\end{aligned}
$$

TWEETY: A COMPREHENSIVE COLLECTION OF JAVA LIBRARIES FOR LOGICAL ASPECTS OF ARTIFICIAL INTELLIGENCE AND KNOWLEDGE REPRESENTATION

Matthias Thimm. Tweety - A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation. In Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14). Vienna, Austria, July 2014.

Abstract

This paper presents Tweety, an open source project for scientific experimentation on logical aspects of artificial intelligence and particularly knowledge representation. Tweety provides a general framework for implementing and testing knowledge representation formalisms in a way that is familiar to researchers used to logical formalizations. This framework is very general, widely applicable, and can be used to implement a variety of knowledge representation formalisms from classical logics, over logic programming and computational models for argumentation, to probabilistic modeling approaches. Tweety already contains over 15 different knowledge representation formalisms and allows easy computation of examples, comparison of algorithms and approaches, and benchmark tests. This paper gives an overview on the technical architecture of Tweety and a description of its different libraries. We also provide two case studies that show how Tweety can be used for empirical evaluation of different problems in artificial intelligence.

1 INTRODUCTION

Knowledge Representation and Reasoning (KR) (Brachman and Levesque, 2004) is an important subfield in Artificial Intelligence (AI) that deals with issues regarding formalizing knowledge in such a way that machines can read, understand, and reason with it. Nowadays, KR has a lot of applications within, e.g., the semantic web (Antoniou and van Harmelen, 2004), as a lot of work on description logics (Baader *et al.*, 2003) and ontologies originate from this field (at least the technical or computer-science-oriented perspectives on those). Apart from that, more fundamental work in KR deals with issues regarding uncertainty of beliefs, dynamics of belief, and defeasible reasoning. Many branches of research in knowledge representation and reasoning is theoretical in nature and researchers usually do not put effort in implementation and empirical evaluation. To address this issue we present in this field study the Tweety libraries for logical aspects of artificial intelligence and knowledge representation.

Approaches to knowledge representation follow almost always a specific pattern. Starting from a formal syntax one can build formulas which are collected in knowledge bases. Using knowledge bases one can derive new information using either the underlying semantics of the language or a specific reasoner. For example, propositional logic is the most basic form for knowledge representation. Given some set of propositions (or atoms) one can build complex formulas using disjunction, conjunction, or negation. A set of propositional formulas, i. e., a knowledge base, can be used to derive new propositional formulas as conclusions. For instance, this can be done using the standard model-theoretic semantics of propositional logic or more sophisticated reasoning techniques such as paraconsistent reasoning. Most logical approaches to knowledge representation such as first-order logic, description logics, defeasible logics, default logics, probabilistic logics, fuzzy logics, etc. follow this pattern. Moreover, many other formalisms which are not so obviously rooted in logic such as abstract argumentation or Bayes nets can also be cast into this framework. For example, for abstract argumentation frameworks (Dung, 1995), a knowledge base is given by a conjunction of attack statements between arguments and different kinds of semantics such as grounded or stable semantics determine how sets of arguments can be derived from a knowledge base.

The Tweety libraries support the implementation of such approaches by providing a couple of abstract classes and interfaces for components such as `Formula`, `BeliefBase`, and `Reasoner`. Furthermore, many strictly logic-based approaches to knowledge representation can also utilize further classes such as `Predicate`, `Atom`, and `Variable`, to name just a few. Currently, Tweety already contains implementations of over 15 different approaches to knowledge representation such as propositional logic, first-order logic, several approaches to probabilistic logics, and several approaches to computational models of argumentation.

In this paper, besides giving an overview on the technical details of Tweety and its libraries, we also report on two case studies that use Tweety as a framework for experimentation and empirical evaluation. The first study is on inconsistency measurement for probabilistic logics (Thimm, 2011, 2013b). In general, probabilistic logics are concerned with using quantitative uncertainty for non-monotonic reasoning. Naturally, these approaches are computationally hard and not easy to understand, as the underlying reasoning mechanisms are quite complicated. Consequently, implementations serve well to understand examples and to (in-)validate conjectures. Our second case study is about strategic argumentation in multiagent systems (Thimm and Garcia, 2010; Rienstra *et al.*, 2013). Similarly, when defining agent models and negotiation strategies in such an environment, effects that occur on a larger scale are hard to predict by hand. Moreover, just an analytical evaluation of different negotiation strategies is often also simply too weak to provide meaningful insights (Rienstra *et al.*, 2013). For that reason Tweety can also be used as a tool for empirical evaluation as it has been done in (Rienstra *et al.*, 2013) to provide average performance results on a series of experiment runs in random settings. Further works that use Tweety for implementing knowledge representation formalisms or for empirical evaluation are, e. g., (Thimm and Kern-Isberner, 2008; Thimm and Garcia, 2010; Krümpelmann *et al.*, 2011; Thimm, 2011; Kern-Isberner and Thimm, 2012; Thimm, 2012, 2013b; Rienstra *et al.*, 2013; Thimm, 2013a; Krümpelmann and Kern-Isberner, 2012).

The rest of this paper is organized as follows. Section 2 gives an overview on the architecture of Tweety and Section 3 presents some technical details on its different libraries. In Section 4 two case studies are presented that show how Tweety can be used for evaluation in scientific research. Section 5 concludes with a summary and pointers to future work.

2 TECHNICAL OVERVIEW

Tweety is organized as a modular collection of Java libraries with a clear dependence structure. The programming language Java has been chosen as it is easy to understand, commonly used, and platform-independent. Each knowledge representation formalism has a dedicated Tweety library (ranging from a library on propositional logic to libraries on computational models of argumentation) which provides implementations for both syntactic and semantic constructs of the given formalism as well as reasoning capabilities. Several libraries provide basic functionalities that can be used in other projects. Among those is the *Tweety Core* library which contains abstract classes and interfaces for all kinds of knowledge representation formalisms. Furthermore, the library *Math* contains classes for dealing with mathematical sub-problems that often occur, in particular, in probabilistic approaches to reasoning. Most other Tweety projects deal with specific approaches to knowledge representation. In the next section, we have a closer look on the individual libraries.

Each Tweety library is organized as a Maven[1] project (Maven is a tool for organizing dependencies between projects, building, and deploying). Most libraries can be used right away as they only have dependencies to other Tweety libraries. Some libraries provide bridges to third-party libraries such as numerical optimization solvers which are not automatically found by Maven and have to be installed beforehand. However, all necessary third-party libraries can be installed by executing a single install file located within the Tweety distribution.

In order to use and develop with Tweety we recommend using the Eclipse IDE[2] and its Maven plugin[3]. As all Tweety libraries are organized as Maven projects they can all be easily imported and used for other projects within

1 http://maven.apache.org
2 http://www.eclipse.org
3 http://maven.apache.org/eclipse-plugin.html

Eclipse. Furthermore, pre-compiled JARs for every library can be downloaded from the Tweety homepage[4] and directly used in other projects. A third way of using the functionalities of Tweety is by using its *Command Line Interface* which is currently under development.

Currently, Tweety contains 161 Java packages which themselves contain 794 Java classes in overall 87421 lines of code. The average cyclomatic complexity number per method (CCN) (McCabe, 1976) is 2.87. This means, that every method roughly contains two to three if-else statements thus reducing complexity on a method-basis and emphasizing the modular nature of Tweety. Furthermore, the average code-to-comment ratio in Tweety is 2.13, meaning that for roughly every two lines of code one line of comment is given.

3 LIBRARIES

In the following we give a detailed description of the currently available libraries within Tweety. An overview of these libraries is given in Table 1 which provides both the name of a library and its Java root package name. Furthermore, the final column lists references to original literature and the implemented reasoning mechanisms and solvers. There, a dagger (†) indicates that a the particular reasoning mechanism has been directly implemented from the original literature, a double dagger (‡) means that a wrapper for the existing original implementation is provided, and an asterisk (*) refers to related literature.

3.1 *General Libraries*

The *General* libraries of Tweety provide basic functionalities and utility classes for all other Tweety classes.

Tweety Core

The *Tweety Core* library contains abstract classes and interfaces for various knowledge representation concepts. Among the most important ones are

`Formula` A formula of a representation formalism.

`BeliefBase` Some structure containing beliefs.

`BeliefSet` A set of beliefs, i. e., a set of formulas, it is the most commonly used class derived from `BeliefBase`. Please note that we follow the Java guideline for naming a class containing a set of beliefs a *belief set* (it contains a finite unordered set of elements), opposed to the naming convention in belief dynamics where a *belief set* is usually deductively closed. In terms of belief dynamics research the class `BeliefSet` actually represents a *belief base*.

4 `http://www.mthimm.de/projects/tweety/`

Library	Project root package	Notes
General Libraries		
Tweety Core	`net.sf.tweety`	
Command Line Interface	`nst.cli`	
Plugin	`nst.plugin`	*JSPF[a]
Math	`nst.math`	‡lpsolve[b], ‡OpenOpt[c], ‡Apache simplex[d], ‡Choco[e], †Gradient descent
Graphs	`nst.graphs`	
Logic Libraries		
Logic Commons	`nst.logics.commons`	†(Grant and Hunter, 2006), †(Hunter and Konieczny, 2006)
Propositional Logic	`nst.logics.pl`	‡Sat4j[f]
First-Order Logic	`nst.logics.fol`	*(Brachman and Levesque, 2004)
Conditional Logic	`nst.logics.cl`	†(Pearl, 1990), †(Kern-Isberner, 2001)
Relational Conditional Logic	`nst.logics.rcl`	†(Kern-Isberner and Thimm, 2012)
Prob. Conditional Logic	`nst.logics.pcl`	†(Rödder, 2000), †(Thimm, 2013b), †(Potyka, 2014)
Rel. Prob. Conditional Logic	`nst.logics.rpcl`	†(Kern-Isberner and Thimm, 2010)
Markov Logic	`nst.logics.ml`	†(Richardson and Domingos, 2006) ‡Alchemy[g]
Epistemic Logic	`nst.logics.el`	*(Fagin *et al.*, 1995)
Description Logic	`nst.logics.dl`	*(Baader *et al.*, 2003)
Logic Translators	`nst.logics.translators`	
Logic Programming Libraries		
Answer Set Programming	`nst.lp.asp`	‡Clingo[h], ‡DLV[i]
Dynamics in ASP	`nst.lp.asp.beliefdynamics`	†(Krümpelmann and Kern-Isberner, 2012)
Nested Logic Programming	`nst.lp.nlp`	*(Lifschitz *et al.*, 1999)
Argumentation Libraries		
Abstract Argumentation	`nst.arg.dung`	†(Dung, 1995), †(Baroni *et al.*, 2011)
Deductive Argumentation	`nst.arg.deductive`	†(Besnard and Hunter, 2001), †(Besnard and Hunter, 2006)
Structured AFs	`nst.arg.saf`	†(Thimm and Garcia, 2010)
DeLP	`nst.arg.delp`	†(Garcia and Simari, 2004), †(Stolzenburg *et al.*, 2003)
Logic Prog. Argumentation	`nst.arg.lp`	†(Schweimeier and Schroeder, 2003)
Probabilistic Argumentation	`nst.arg.prob`	†(Thimm, 2012)
Agent Libraries		
Agents	`nst.agents`	*(Weiss, 2013)
Dialogues	`nst.agents.dialogues`	†(Thimm and Garcia, 2010), †(Rienstra *et al.*, 2013)
Other Libraries		
Action and Change	`nst.action`	†(Gelfond and Lifschitz, 1998)
Belief Dynamics	`nst.beliefdynamics`	†(Hansson, 2001), †(Fermé and Hansson, 1999), †(Krümpelmann *et al.*, 2011)
Machine Learning	`nst.machinelearning`	‡LIBSVM[j]
Preferences	`nst.preferences`	*(Walsh, 2007), †(Thimm, 2013a)

Table 1: Overview on the Tweety libraries (the prefix `nst` stands for `net.sf.tweety`)

a `https://code.google.com/p/jspf/`
b `http://lpsolve.sourceforge.net`
c `http://openopt.org`
d `http://commons.apache.org/math`
e `http://www.emn.fr/z-info/choco-solver/`
f `http://www.sat4j.org`
g `http://alchemy.cs.washington.edu`
h `http://potassco.sourceforge.net`
i `http://www.dlvsystem.com`
j `http://www.csie.ntu.edu.tw/~cjlin/libsvm/`

`Signature` The signature of a representation formalism.

`Interpretation` An interpretation that evaluates the truth of formulas.

`Reasoner` Implements a specific reasoning strategy to answer queries for a representation formalism.

`Parser`, `Writer` For reading/writing formulas and belief sets.

Most other Tweety libraries provide specific implementations of the above abstract classes and interfaces for their specific representation formalisms. For example, the library *Propositional Logic* implements `Formula` by `PropositionalFormula` (which is recursively defined using conjunction, disjunction, and negation) and `Interpretation` by `PossibleWorld`. In this way, the classical approach to formally define a logical language via syntax and semantics has a one-to-one correspondence with its implementation in Tweety.

Besides the above mentioned abstract classes and interfaces, *Tweety Core* provides abstract implementations of several other knowledge representation concepts and several utility classes for working with sets, subsets, vectors, and general rules.

Plugin

The *Plugin* library provides classes for implementing Tweety plugins that can be used by, e. g., the *Command Line Interface*. This library makes use of the *Java Simple Plugin Framework (JSPF)*[5]. Using these classes one can encapsulate the functionalities of a specific knowledge representation formalism and expose them in the same way to user interfaces. The most important class is the abstract class `AbstractTweetyPlugin` which is the basis for developing plugins. Please note that the *Plugin* library is currently in an experimental phase.

Command Line Interface

All Tweety libraries can be accessed programmatically in Java through their API (*Application Programming Interface*). However, for non-programmers this way of utilizing the libraries is not very convenient. Using the *Plugin* library the *Command Line Interface* library provides a general command line interface for many Tweety libraries. Every library can expose its functionality through a Tweety plugin that can be plugged into the command line interface and accessed in a uniform way. Please note that the *Command Line Interface* library is currently in an experimental phase.

5 `https://code.google.com/p/jspf/`

Math

Many algorithms for knowledge representation and reasoning are based on mathematical methods such as optimization techniques. The *Math* library encapsulates those mathematical methods and exposes them through simple interfaces to other libraries for realizing these algorithms. At the core, the *Math* library contains classes for representing mathematical terms (such as `Constant`, `Variable`, `Product`, `Logarithm`) and statements (such as `Equation`). Using these constructs one can represent, e. g., constraint satisfaction problems (`ConstraintSatisfactionProblem`) and optimization problems (`OptimizationProblem`). Through the `Solver` interface the *Math* library provides bridges to several third-party solvers such as the *ApacheCommons* Simplex-algorithm[6], the *OpenOpt* solvers[7], or Choco[8].

Graphs

The *Graphs* library contains a simple graph implementation with utility functions as it can be used, e. g., to represent abstract argumentation frameworks (see *Abstract Argumentation* library).

3.2 *Logic Libraries*

The *Tweety Logic* libraries (under the package `net.sf.tweety.logics`) provide implementations for various knowledge representation formalisms based on classical logics (propositional logic and first-order logic) and non-classical logics such as conditional logic, probabilistic logics, epistemic logics, or description logic. Each library follows a strict approach in defining the formalism by implementing the abstract classes and interfaces `Formula`, `BeliefBase`, `Interpretation`,... from the *Tweety Core* library. Each library contains a sub-package `syntax` which contains the elements to construct formulas of the formalism and a sub-package `semantics` which contains elements for realizing the semantics of the formalism. Besides these two common sub-packages many libraries also contain parsers for reading formulas from file and reasoner that implement a specific reasoning approach.

Logic Commons

The *Logic Commons* library contains abstract classes and interfaces which further refine the general `Formula` interface from the *Tweety Core* library. Among these refinements are several concepts that are shared among a great number of knowledge representation formalism such as `Predicate`, `Variable` or `Atom`.

6 `http://commons.apache.org/math`
7 `http://openopt.org`
8 `http://www.emn.fr/z-info/choco-solver/`

Propositional Logic

The *Propositional Logic* library provides an implementation of classical propositional logic. Propositional formulas can be constructed using, e. g., classes `Conjunction` or `Disjunction` and propositional formulas can be put into a knowledge base of type `PlBeliefSet`. Currently, the *Propositional Logic* library supports two different reasoners. The first is a simple brute force approach that directly follows the definition of classical entailment, i. e., in order to prove a given propositional formula wrt. a given set of propositional formulas all possible worlds are enumerated and tested. Obviously, this reasoner only works for small examples but is useful when one is interested in all models of a knowledge base. The second supported reasoner incorporates the Sat4j reasoner[9]. Other SAT-solvers can be added in a straightforward way.

First-Order Logic

This library contains an implementation of first-order logic as a knowledge representation formalism. Both the *Propositional Logic* library and the *First-Order Logic* library are used by many other libraries of knowledge representation formalisms.

Conditional Logic

The *Conditional Logic* library extends the *Propositional Logic* library by *conditionals*, i. e., non-classical rules of the form $(B \mid A)$ ("A usually implies B"), cf. (Nute and Cross, 2002). In the literature, several different semantics and reasoning approaches for conditional logics have been proposed and this library can be used to easily compare their reasoning behavior. Currently, the *Conditional Logic* library implements interpretations in the form of ranking functions (Spohn, 1988) and conditional structures (Kern-Isberner, 2001), and provides reasoner based on z-ranking (Goldszmidt and Pearl, 1996) and c-representations (Kern-Isberner, 2001).

Relational Conditional Logic

Similar to the *Conditional Logic* library the *Relational Conditional Logic* extends the *First-Order Logic* libraries with relational conditionals (i. e. conditionals that may contain first-order formulas), cf. (Delgrande, 1998; Kern-Isberner and Thimm, 2012). Currently, this library contains an implementation of the relational c-representation reasoning approach of (Kern-Isberner and Thimm, 2012).

9 `http://www.sat4j.org`

Probabilistic Conditional Logic

This library further extends the *Conditional Logic* library by extending conditionals to *probabilistic conditionals* of the form $(B \mid A)[p]$ ("A usually implies B with probability p"), cf. (Rödder, 2000). Besides a naive implementation of probabilistic reasoning based on the principle of maximum entropy (Paris, 1994) this library also contains several classes for analyzing and repairing inconsistent sets of probabilistic conditionals, cf. (Thimm, 2011, 2013b). We will discuss this package in more detail in Section 4.

Relational Probabilistic Conditional Logic

By combining both the *Relational Conditional Logic* and *Probabilistic Conditional Logic* libraries the *Relational Probabilistic Conditional Logic* library introduces relational conditionals with probabilities, cf. (Kern-Isberner and Thimm, 2010). It implements both the *averaging* and *aggregating* semantics from (Kern-Isberner and Thimm, 2010) and also allows for lifted inference as proposed in (Thimm, 2011).

Markov Logic

This library builds on the *First-Order Logic* library to implement *Markov Logic*, an extension of first-order logic with weights to allow for probabilistic reasoning, cf. (Richardson and Domingos, 2006). It provides several propriety sampling-based reasoner and a bridge to the Alchemy reasoner[10].

Epistemic Logic

This library extends the *Propositional Logic* library with modal operators for epistemic logic and its semantics with accessibility relations and Kripke models. Please note that the *Epistemic Logic* library is currently in an experimental phase.

Description Logic

The *Description Logic* library provides a general description logic implementation (Baader *et al.*, 2003) based on the *First-Order Logic* library. Please note that the *Description Logic* library is currently in an experimental phase.

Logic Translators

This library provides the abstract class `Translator` that provides basic functionalities to implement translators between different knowledge representation formalisms. Currently, the *Logic Translators* library contains translators between first-order logic and answer set programming, between nested logic programming and answer set programming, and between propositional logic and first-order logic.

---

10 http://alchemy.cs.washington.edu

3.3 *Logic Programming Libraries*

The *Logic Programming* libraries (located under the package `net.sf.tweety.lp`) provide implementations of knowledge representation formalisms based on logic programming.

Answer Set Programming

The *Answer Set Programming* library provides classes for representing extended logic programs (Gelfond and Leone, 2002). Answer set programs are logic programs of the form $A \leftarrow B_1, \ldots, B_m$ with first-order literals $A, B_1, \ldots, B_m$ where the body literals $B_1, \ldots, B_m$ may also have a default negation `not`. This library provides bridges to several established solvers such as DLV[11], DLV Complex[12], and Clingo[13].

Dynamics in Answer Set Programming

This library extends the *Answer Set Programming* library by introducing revision and update approaches. The library contains implementations of the approaches introduced in (Krümpelmann and Kern-Isberner, 2012; Delgrande *et al.*, 2007) and also revision approaches based on argumentation.

Nested Logic Programming

This library contains an implementation of nested logic programs which allow for complex first-order formulas to appear in logic programming rules (Lifschitz *et al.*, 1999).

3.4 *Argumentation Libraries*

The argumentation libraries (under the package `net.sf.tweety.arg`) are one of the most mature libraries of Tweety and contain a wide variety of implementations of different approaches to computational argumentation.

Abstract Argumentation

This library implements abstract argumentation as proposed in (Dung, 1995). An abstract argumentation framework is a directed graph (A, Att) where A is interpreted as a set of arguments and an edge $(A, A') \in Att$ is an attack of A on A'. The library provides implementations of the mostly used semantics and their corresponding reasoner, both in terms of extensions (an extension is a set of arguments that is regarded as accepted by a semantics) and labelings (a labeling is a function with a three-valued truth assignment to each argument). Several utility classes for generating random argumentation frameworks complement this library.

11 `http://www.dlvsystem.com`
12 `https://www.mat.unical.it/dlv-complex`
13 `http://potassco.sourceforge.net`

Deductive Argumentation

The *Deductive Argumentation* library provides an implementation of the approach proposed in (Besnard and Hunter, 2001). In deductive argumentation, an argument is composed of a set of propositional formulas that derive the claim of the argument. Attack between arguments is derived from classical unsatisfiability.

Structured Argumentation Frameworks

This library implements the approach of structured argumentation frameworks as proposed in (Thimm and Garcia, 2010). In structured argumentation frameworks arguments are composed of subarguments and a conclusion.

Defeasible Logic Programming

This library provides an implementation of Defeasible Logic Programming (DeLP) (Garcia and Simari, 2004). In DeLP knowledge bases contain strict and defeasible rules and facts, similar to knowledge representation formalisms for logic programming. Defeasible rules can be collected in arguments and compared by generalized specificity (Stolzenburg *et al.*, 2003).

Logic Programming Argumentation

This library provides an implementation of the argumentation approach of (Schweimeier and Schroeder, 2003) which is also based on logic programming techniques.

Probabilistic Argumentation

The *Probabilistic Argumentation* library extends the *Abstract Argumentation* library with non-classical semantics based on probabilistic assessments (Thimm, 2012).

3.5 *Agent Libraries*

The agent libraries (located under the package `net.sf.tweety.agents`) provide a framework for analyzing and simulating interactions between agents.

Agents

This general library contains an abstract formalization of agents and multi-agent systems. Classes such as `Agent`, `Environment`, `MultiAgentSystem`, and `Protocol` can be used to set up and simulate a system of agents within an environment. This library has a specific focus on the simulation aspect and provides classes such as `MultiAgentSystemGenerator` and `GameSimulator` that allow the automatic generation of test scenarios and their evaluation.

Dialogues

The library *Dialogues* extends the *Agents* library with the capability of simulating dialogues between agents, as they are investigated in the context of argumentation in multi-agent systems (Karunatillake *et al.*, 2009). It also provides an implementation of agents with an opponent model as proposed in (Rienstra *et al.*, 2013). We will discuss this package in more detail in Section 4.

3.6 *Other Libraries*

The above discussed libraries constitute the core of Tweety by providing implementations of several knowledge representation formalisms. This collection is complemented by some further libraries that relate either to topics that do not strictly belong to the field of knowledge representation (such as the *Machine Learning* library) or can be applied across several different knowledge representation formalisms (such as the *Belief Dynamics* library).

Action and Change

The *Action and Change* library implements several action languages and their dynamics from (Gelfond and Lifschitz, 1998).

Belief Dynamics

This library provides a general implementation for various approaches to belief (base) revision and update (Hansson, 2001). It provides interfaces and several implementations of many concepts used in belief dynamics such as `BaseRevisionOperator`, `BaseContractionOperator`, `IncisionFunction`, and `LeviBaseRevisionOperator`. Those classes are defined in such a general way that they can be used not only to implement belief dynamics for propositional logic but also for other knowledge representation formalisms implementing the corresponding Tweety interfaces. This library contains also specific revision approaches such as selective revision (Fermé and Hansson, 1999) and argumentative selective revision (Krümpelmann *et al.*, 2011).

Machine Learning

The *Machine Learning* library provides several abstract concepts that can be used in a machine learning context such as `Observation`, `Classifier`, and `CrossValidator`. It contains also an implementation of support vector machines utilizing LIBSVM[14].

Preferences

This library contains classes for representing preference orders and approaches for aggregating them (Walsh, 2007). It also contains an implementation of the dynamic preference aggregation approach proposed in (Thimm, 2013a).

4 CASE STUDIES AND EVALUATION

In this section we discuss two case studies that make use of Tweety as a platform for experimentation and empirical evaluation. The first case study is on inconsistency handling for probabilistic logics (Thimm, 2011, 2013b) while the second study is on strategic argumentation in multi-agent systems (Thimm and Garcia, 2010; Rienstra *et al.*, 2013).

4.1 *Inconsistency Handling for Probabilistic Logics*

In order to motivate the work described in this section we give a brief introduction into probabilistic conditional logic and its inconsistency measures, cf. (Thimm, 2011, 2013b).

For propositional formulas ϕ, ψ and a real-value $p \in [0,1]$ we call $(\phi \mid \psi)[p]$ a *probabilistic conditional*. A probabilistic conditional $(\phi \mid \psi)[p]$ represents a specific form of defeasible rule and has the intuitive meaning "if ψ is true then ϕ is true with probability p". A (probabilistic conditional) *knowledge base* $\mathcal{K}$ is a set of probabilistic conditionals. Semantics are given to probabilistic conditionals by probability functions $P : \Omega \to [0,1]$ with Ω being the set of interpretations (possible worlds) of the underlying propositional logic (We assume that the set of propositions is finite and so is the set of possible worlds). A probability function P *satisfies* a conditional $(\phi \mid \psi)[p]$ if and only if $P(\phi \wedge \psi) = pP(\psi)$ (the probability of a formula is defined to be the sum of the probabilities of all possible worlds satisfying it). Note that this follows the definition of conditional probability ($P(\phi \mid \psi) = P(\phi \wedge \phi)/P(\psi) = p$) as long as $P(\psi) \neq 0$. In order to avoid a case differentiation for $P(\psi) = 0$ we use the above definition, cf. (Paris, 1994). A probability function P satisfies a knowledge base $\mathcal{K}$ if and only if it satisfies all its probabilistic conditionals. A knowledge base is consistent if such a probability function exists.

14 `http://www.csie.ntu.edu.tw/~cjlin/libsvm/`

Example 1. Consider $\mathcal{K} = \{(f \mid b)[0.9], (b \mid p)[1], (f \mid p)[0.01]\}$ with the intuitive meaning that birds (b) usually (with probability 0.9) fly (f), that penguins (p) are always birds, and that penguins usually do not fly (only with probability 0.01). The knowledge base $\mathcal{K}$ is consistent as a probability function satisfying it can easily be constructed, cf. (Thimm, 2013b). Note that, e. g., the knowledge base $\mathcal{K} = \{(x \mid y)[0.9], (y \mid \top)[0.9], (x \mid \top)[0.2]\}$ is inconsistent ($\top$ is a logical tautology): considering just the conditionals $(x \mid y)[0.9]$ and $(y \mid \top)[0.9]$ we obtain that x has to be at the least probability 0.81 which is inconsistent with stating that x has probability 0.2.

In order to deal with inconsistent knowledge bases the work (Thimm, 2013b) proposes inconsistency measures as a tool for analyzing inconsistencies. An inconsistency measure is a function $\mathcal{I}$ that takes a knowledge base $\mathcal{K}$ and computes an inconsistency value $\mathcal{I}(\mathcal{K}) \in [0, \infty)$ with the intuitive meaning that a larger value indicates a more severe inconsistency (and $\mathcal{I}(\mathcal{K}) = 0$ means that $\mathcal{K}$ is consistent). See (Thimm, 2013b) for more details, some rationality postulates on inconsistency measurement, and specific approaches.

The inconsistency measurement framework for probabilistic logics has been implemented in the *Probabilistic Conditional Logic* library of Tweety (sub-package `net.sf.tweety.logics.pcl.analysis`). However, as inconsistency measurement is a broader topic that can also be used in other knowledge representation formalisms such as classical logics (Grant and Hunter, 2006), the concept *inconsistency measure* is already implemented in the *Logic Commons* library (sub-package `net.sf.tweety.logics.commons.analysis`) as a very general interface (only a simplified version is shown):

```
public interface InconsistencyMeasure
            <T extends BeliefBase> {
  public Double inconsistencyMeasure
                        (T beliefBase);
}
```

The interface above is parametrized by the specific type of belief base using Java Generics. All types of belief bases used within Tweety, such as propositional belief sets (`PlBeliefSet`) or probabilistic conditional knowledge bases (`PclBeliefSet`), are derived from `BeliefBase`. The package `net.sf.tweety.logics.commons.analysis` provides several generally applicable implementations of the above interface such as (only a simplified version is shown):

```
public class MiInconsistencyMeasure
  <S extends Formula,T extends BeliefSet<S>>
  implements InconsistencyMeasure<T> {

  private BeliefSetConsistencyTester<S,T> consTester;

  public MiInconsistencyMeasure(BeliefSetConsistencyTester<S,T>
      consTester){
    this.consTester = consTester;
```

```
  }

  @Override
  public Double inconsistencyMeasure
    (T beliefSet) {
      return new Double(this.consTester.minimalInconsistentSubsets
          (beliefSet).size());
  }
}
```

The above measure is an implementation of the MI-inconsistency measure (Grant and Hunter, 2006) and is applicable for all kinds of logics that provide an implementation of an `BeliefSetConsistencyTester`. This measure takes the number of minimal inconsistent subsets of a knowledge as an assessment of its inconsistency. Another example, particularly for the case of probabilistic conditional logic, is the `DistanceMinimizationInconsistencyMeasure` that makes use of the *Math* library. This measure assesses the grade of inconsistency by measuring how much the probabilities of the conditionals have to be modified in order to obtain a consistent knowledge base, cf. (Thimm, 2013b). This problem is solved by optimization techniques that can be found in the *Math* library.

For probabilistic conditional logic, determining whether a knowledge base is inconsistent and assessing its inconsistency value is not easily done by hand. Using the implementations of various inconsistency measures in Tweety, we were able to compute and compare inconsistency values for various knowledge bases, cf. (Thimm, 2011). Furthermore, as the interfaces and abstract classes provided by Tweety are very general and force the programmer to work as abstract as possible, even the implementation of such specific concepts such as inconsistency measures yield very generally applicable classes that are easily adapted to other approaches.

4.2 *Strategic Argumentation*

Our second case study is about strategic argumentation in multi-agent systems. In the works (Thimm and Garcia, 2010; Rienstra *et al.*, 2013) we investigated systems of agents that are engaged in dialogues and aim at resolving contradiction by exchange of arguments. We give a brief introduction into the topic now, but simplify the formalization for the sake of readability.

We consider two agents PRO (proponent) and OPP (opponent) engaged in a dialogue about a specific argument A (we use the terminology of abstract argumentation frameworks as mentioned earlier). The proponent has the goal to establish that A is acceptable and the opponent has the goal to establish that A is not acceptable. Both agents have only access to a subset of all available arguments and are, in general, ignorant or uncertain about the arguments the other agent has access to. Both agents take turn in forwarding a set of arguments. In (Rienstra *et al.*, 2013) several different belief states with *opponent models* were proposed and discussed that help an agent to act strategically in these kinds of dialogues. The first type T_1 of belief state is a tuple (B, E) where B is the set of arguments a particular agent (either PRO or OPP) has access to, and E is the opponent model which is itself a belief state of type T_1[15]. This type of belief state therefore models what an agent thinks another agent beliefs, etc.. The second T_2 and third T_3 types of belief state extend the first type by introducing uncertainty on the set of arguments believed by the other agent and uncertainty about the arguments themselves. The second type of belief state T_2 is a tuple (B, P) where B is again the set of arguments a particular agent has access to and P is a probability distribution over some set $\{K_1, \ldots, K_n\}$ where each K_i $(i = 1, \ldots, n)$ is again a belief state of type T_2. For a formalization of the belief state of type T_3 see (Rienstra *et al.*, 2013). In (Rienstra *et al.*, 2013) it has been analytically shown that the expressiveness of the three models is increasing from T_1 to T_3. However, in order to understand the differences between the three models examples have to be created and computed with different belief states. In the setting of strategic argumentation, this is a hard task to do by hand. In a system with at least two agents where both agents are equipped with a non-trivial belief state that changes with every action, running through a complete example by hand is a tedious task.

The complete setting of (Rienstra *et al.*, 2013) has been implemented in the *Dialogues* library which makes heavy use of the general agent classes from the *Agents* library and, of course, the knowledge representation formalism from the *Abstract Argumentation* library. The central class of the implementation is the `ArguingAgent` class (we only show an excerpt):

```java
public class ArguingAgent extends Agent {
  private BeliefState beliefState;
  private AgentFaction faction;

  @Override
  public Executable next(Collection<? extends Perceivable>
      percepts) {
    // [env = the environment object]
    this.beliefState.update(env.getDialogueTrace());
    return this.beliefState.move(env);
  }
}
```

The central attributes of an arguing agent are its belief state and its faction, e. g., either PRO or OPP. The method `next(...)` (derived from the superclass `Agent`) determines the agent's behavior on receiving some perception from the environment and returns some action (of type `Executable`). Here, the agent first updates its belief state with the current dialogue trace (a sequence of sets of arguments advanced so far) and then returns its own move (a set of arguments). The three different belief state types have been implemented in the classes `T1BeliefState`, `T2BeliefState`, and `T3BeliefState`. Arguing agents are organized in a `GroundedGameSystem` which is of type

15 Note that this model has originally been proposed in (Oren and Norman, 2009)

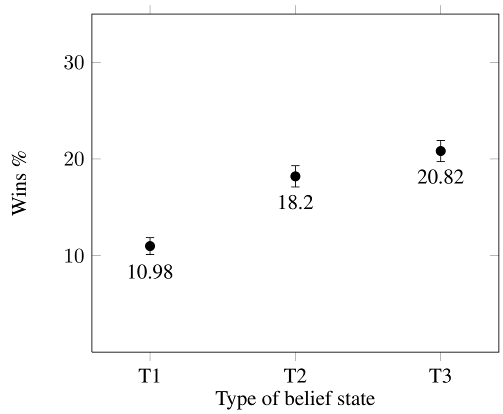


Figure 1: Average performance of T_1, T_2, and T_3 belief state models after 5000 simulation runs (with Binomial proportion confidence intervals)

`MultiAgentSystem<ArguingAgent>` and models an argumentation dialogue ("grounded" refers to the grounded semantics used for this type of game). On top of this implementation of the actual dialogue system a simulation framework was implemented that allows the (random) generation of the above multi-agent systems and measures the performance of the individual agents over a series of runs. The central class for the simulation framework is the `GroundedGameGenerator` which implements the interface `MultiAgentSystemGenerator` and is able to generate (random) multi-agent systems of arguing agents.

In (Rienstra *et al.*, 2013), for evaluating performance we generated a random abstract argumentation theory with 10 arguments, ensuring that the argument under consideration is in its grounded extension, i. e., under perfect information the proponent should win the dialogue. However, from these 10 arguments only 50 % are known by the proponent but 90 % by the opponent. We used a proponent without opponent model and generated an belief state of type T_3 for the opponent. From this T_3 belief state we derived T_2 and T_1 belief states by ignoring the added expressivity. For each belief state we simulated a dialogue against the same opponent and counted the number of wins. We repeated the experiment 5000 times, Figure 1 shows our results, cf. (Rienstra *et al.*, 2013). There, it can be seen that increasing the complexity of the belief state yields better overall performance (thus confirming the analytical evaluation). However, this empirical evaluation sheds also more light on the importance of the added expressivity for strategic argumentation. While T_2 is significantly better that T_1, the difference between T_3 and T_2 is nearly marginal. These kinds of nuances are very difficult to discover when considering only analytical evaluation.

5 SUMMARY AND FUTURE WORK

In this paper we presented Tweety, a comprehensive collection of Java libraries for logical aspects of artificial intelligence and knowledge representation. We gave an overview on the technical aspects and provided details on its individual packages. Finally, we presented two case studies that make use of Tweety as a framework for experimentation and empirical evaluation.

Tweety is an open source project[16,17] and can therefore be used and extended by everyone. In particular, instantiating the abstract Tweety classes for a particular formalism is simple. Although Tweety is implemented in a object-oriented programming language it follows a strict declarative formal way to define concepts from theoretical knowledge representation research. Tweety is available under the GNU General Public License version 3.0. In order to contribute to the main Tweety repository contact the author.

To the best of our knowledge, Tweety is the first attempt to provide a general-purpose framework for a broad variety of knowledge representation formalisms. However, there exist also more specialized frameworks for specific approaches or areas, such as the OWLAPI[18] for working with OWL ontologies, KReator[19] for relational probabilistic knowledge representation, or bcontractor[20] for belief dynamics.

Current and future work on Tweety is mainly concerned with extending the general infrastructure and improving usability. In particular, current work is about implementation of the plugin architecture for all libraries, a command line interface, and a web front-end. The ultimate goal there is to have several standardized user interfaces that are apt to work with any kind of knowledge representation mechanism and thus remove the burden of designing and implementing user interfaces from the researcher.

16 `http://www.mthimm.de/projects/tweety/`
17 The source code of Tweety is hosted at SourceForge: `http://tweety.svn.sourceforge.net`.
18 `http://owlapi.sourceforge.net`
19 `http://kreator-ide.sourceforge.net`
20 `https://code.google.com/p/bcontractor/`

BIBLIOGRAPHY

T. Alsinet, C. I. Chesñevar, L. Godo, and G. R. Simari. A logic programming framework for possibilistic argumentation: Formalization and logical properties. *Fuzzy Sets and Systems*, 159(10):1208–1228, 2008. (Cited on pages 16 and 28.)

Leila Amgoud and Philippe Besnard. Logical limits of abstract argumentation frameworks. *Journal of Applied Non-Classical Logics*, 23(3):229–267, 2013. (Cited on page 4.)

Leila Amgoud, Nicolas Maudet, and Simon Parsons. Modelling dialogues using argumentation. In *Proceedings of the Fourth International Conference on Agents and Multiagent Systems (AAMAS'00)*, pages 31–38, 2000. (Cited on page 5.)

Leila Amgoud, Claudette Cayrol, Marie-Christine Lagasquie-Schiex, and P. Livet. On Bipolarity in Argumentation Frameworks. *International Journal of Intelligent Systems*, 23(10):1062–1093, 2008. (Cited on page 4.)

Grigoris Antoniou and Frank van Harmelen. *A Semantic Web Primer*. MIT Press, Cambridge, MA, USA, 2004. (Cited on pages 1 and 151.)

Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, New York, NY, USA, 2003. (Cited on pages 1, 151, 155, and 159.)

Franz Baader, Ian Horrocks, and Ulrike Sattler. Description Logics as Ontology Languages for the Semantic Web. In *Mechanizing Mathematical Reasoning*, volume 2605 of *Lecture Notes in Computer Science*, pages 228–248. Springer-Verlag, 2005. (Cited on page 1.)

Pietro Baroni and Massimiliano Giacomin. A systematic classification of argumentation frameworks where semantics agree. In *Proceedings of the Second International Conference on Computational Models of Argument (COMMA'08)*, pages 37–48, 2008. (Cited on page 98.)

P. Baroni, M. Giacomin, and G. Guida. Scc-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence*, 168(1–2):162–210, 2005. (Cited on pages 15 and 18.)

P. Baroni, P. E. Dunne, and M. Giacomin. On extension counting problems in argumentation frameworks. In *Proc. of the Third Int. Conf. on Computational Models of Argument (COMMA'10)*, pages 63–74, 2010. (Cited on page 15.)

Pietro Baroni, Martin Caminada, and Massimiliano Giacomin. An Introduction to Argumentation Semantics. *The Knowledge Engineering Review*, 26(4):365–410, 2011. (Cited on pages 4 and 155.)

Pietro Baroni, Massimiliano Giacomin, and Guillermo R. Simari. Belief revision and argumentation: a reasoning process view. In Eduardo Fermé, Guillermo R. Simari, and Dov Gabbay, editors, *Trends in Belief Revision and Argumentation Dynamics*. College Publications, 2013. (Cited on page 5.)

Pietro Baroni, Massimiliano Giacomin, and Paolo Vicig. On rationality conditions for epistemic probabilities in abstract argumentation. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*, 2014. (Cited on page 5.)

Ringo Baumann. What Does it Take to Enforce an Argument? Minimal Change in Abstract Argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, 2012. (Cited on page 5.)

Harald Beck, Minh Dao-Tran, Thomas Eiter, and Michael Fink. LARS: A Logic-based Framework for Analyzing Reasoning over Streams. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI'15)*, 2015. (Cited on page 77.)

Trevor J.M. Bench-Capon. Persuasion in Practical Argument Using Value Based Argumentation Frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003. (Cited on page 4.)

P. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128(1-2):203–235, 2001. (Cited on pages 13, 155, and 161.)

P. Besnard and A. Hunter. Knowledgebase compilation for efficient logical argumentation. In *Proceedings of the 10th International Conference on Knowledge Representation (KR'06)*, pages 123–133. AAAI Press, 2006. (Cited on page 155.)

Philippe Besnard and Anthony Hunter. *Elements of Argumentation*. The MIT Press, 2008. (Cited on pages 4 and 5.)

Philippe Besnard. Revisiting Postulates for Inconsistency Measures. In Eduardo Fermé and João Leite, editors, *Logics in Artificial Intelligence*, volume 8761 of *Lecture Notes in Computer Science*, pages 383–396. Springer International Publishing, 2014. (Cited on pages 6, 12, 66, 68, 72, 98, 100, 114, 126, and 130.)

Jean-Yves Béziau, Walter Carnielli, and Dov Gabbay, editors. *Handbook of Paraconsistency*. College Publications, London, 2007. (Cited on page 3.)

Stefano Bistarelli, Fabio Rossi, and Francesco Santini. Benchmarking hard problems in random abstract afs: The stable semantics. In *Proceedings of* *the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*, 2014. (Cited on page 5.)

Elizabeth Black and Anthony Hunter. A generative inquiry dialogue system. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'07)*. IEEE Press, 2007. (Cited on page 5.)

Ronald J. Brachman and Hector J. Levesque. *Knowledge Representation and Reasoning*. The Morgan Kaufmann Series in Artificial Intelligence. Morgan Kaufmann Publishers, 2004. (Cited on pages 1, 151, and 155.)

Gerhard Brewka, Stefan Ellmauthaler, Hannes Strass, Johannes Peter Wallner, and Stefan Woltran. Abstract Dialectical Frameworks Revisited. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, 2013. (Cited on page 5.)

Gerhard Brewka, Sylwia Polberg, and Stefan Woltran. Generalizations of dung frameworks and their role in formal argumentation. *IEEE Intelligent Systems*, 29(1), 2014. (Cited on page 4.)

H. K. Büning and O. Kullmann. Minimal Unsatisfiability and Autarkies. In *Handbook of Satisfiability*, pages 339–401. IOS Press, 2009. (Cited on page 88.)

Martin Caminada and Leila Amgoud. On the Evaluation of Argumentation Formalisms. *Artificial Intelligence*, 171(5-6):286–310, 2007. (Cited on page 4.)

Martin Caminada and Dov M. Gabbay. A Logical Account of Formal Argumentation. *Studia Logica*, 93(2–3):109–145, 2009. (Cited on page 7.)

M. Caminada. Semi-stable semantics. In *Proc. of the First Int. Conf. on Computational Models of Argument (COMMA'06)*, pages 121–130, 2006. (Cited on pages 15 and 17.)

David Carmel and Shaul Markovitch. Learning and using opponent models in adversary search. Technical Report CIS9609, Technion, 1996. (Cited on page 35.)

L. Cholvy and A. Hunter. Information fusion in logic: A brief overview. In *Qualitative and Quantitative Practical Reasoning (ECSQARU'97/FAPR'97)*, volume 1244 of *Lecture Notes in Computer Science*, pages 86–95. Springer, 1997. (Cited on pages 3 and 11.)

Petr Cintula, Petr Hájek, and Carles Noguera, editors. *Handbook of Mathematical Fuzzy Logic*, volume 1. College Publications, 2011. (Cited on page 3.)

Stephen A. Cook. The Complexity of Theorem-Proving Procedures. In *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing*, pages 151–158, 1971. (Cited on page 68.)

Sylvie Coste-Marquis, Sebastien Konieczny, Jean-Guy Mailly, and Pierre Marquis. On the Revision of Argumentation Systems: Minimal Change of Arguments Statuses. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14)*, 2014. (Cited on page 5.)

Lionel Daniel. *Paraconsistent Probabilistic Reasoning*. PhD thesis, L'École Nationale Supérieure des Mines de Paris, December 2009. (Cited on page 66.)

Glauber De Bona and Marcelo Finger. Measuring Inconsistency in Probabilistic Logic: Rationality Postulates and Dutch Book Interpretation. *Artificial Intelligence*, online first, doi:10.1016/j.artint.2015.06.006, 2015. (Cited on page 6.)

L. De Raedt and K. Kersting. Probabilistic inductive logic programming. In *Probabilistic Inductive Logic Programming*, pages 1–27. Springer, 2008. (Cited on page 47.)

Luc De Raedt. *Logical and Relational Learning*. Springer, 2008. (Cited on page 10.)

Richard Dedekind. Über Zerlegungen von Zahlen durch ihre grössten gemeinsamen Teiler. In *Gesammelte Werke 2*, pages 103–148. 1897. (Cited on page 120.)

James Delgrande and Yi Jin. Parallel Belief Revision: Revising by Sets of Formulas. *Artificial Intelligence*, 176:2223–2245, 2012. (Cited on page 126.)

James P. Delgrande, Torsten Schaub, and Hans Tompits. A preference-based framework for updating logic programs. In *Proceedings of the Ninth International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'07)*, 2007. (Cited on page 160.)

James P. Delgrande. On first-order conditional logics. *Artificial Intelligence*, 105(1-2):105–137, 1998. (Cited on page 158.)

Xi Deng, Volker Haarslev, and Nematollaah Shiri. Measuring Inconsistencies in Ontologies. In *Proceedings of the 4th European conference on The Semantic Web (ESWC '07)*, pages 326–340, Berlin, Heidelberg, 2007. Springer International Publishing. (Cited on page 66.)

Dragan Doder and Stefan Woltran. Probabilistic argumentation frameworks – a logical approach. In Umberto Straccia and Andrea Calì, editors, *Scalable Uncertainty Management*, volume 8720 of *Lecture Notes in Computer Science*, pages 134–147. Springer International Publishing, 2014. (Cited on page 5.)

Dragan Doder, Miodrag Raskovic, Zoran Markovic, and Zoran Ognjanovic. Measures of Inconsistency and Defaults. *International Journal of Approximate Reasoning*, 51:832–845, 2010. (Cited on pages 6, 66, 98, 100, 105, and 106.)

P. Domingos and D. Lowd. *Markov Logic: An Interface Layer for Artificial Intelligence*. Morgan and Claypool, 2009. (Cited on pages 10, 47, 49, and 51.)

Pierpaolo Dondio. Multi-valued and probabilistic argumentation frameworks. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*, 2014. (Cited on page 5.)

P. M. Dung and P. M. Thang. Towards (probabilistic) argumentation for jury-based dispute resolution. In *Proc. of the Third Int. Conf. on Computational Models of Argument (COMMA'10)*, pages 171–182, 2010. (Cited on pages 16 and 28.)

Phan Minh Dung. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. *Artificial Intelligence*, 77(2):321–358, 1995. (Cited on pages 4, 13, 15, 17, 18, 28, 32, 33, 152, 155, and 160.)

Paul E. Dunne and Michael Wooldridge. Complexity of Abstract Argumentation. In Iyad Rahwan and Guillermo R. Simari, editors, *Argumentation in Artificial Intelligence*, pages 85–104. Springer-Verlag, 2009. (Cited on page 5.)

P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457–486, 2011. (Cited on pages 5, 15, 16, and 28.)

Patricia Everaere, Sebastien Konieczny, and Pierre Marquis. Belief merging versus judgment aggregation. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, 2015. (Cited on page 11.)

Ronald Fagin, Joseph Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1995. (Cited on page 155.)

Bettina Fazzinga, Sergio Flesca, and Francesco Parisi. On the complexity of probabilistic abstract argumentation. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, 2013. (Cited on page 5.)

Eduardo Fermé and Sven Ove Hansson. Selective revision. *Studia Logica*, 63(3):331–342, 1999. (Cited on pages 155 and 162.)

M. Finthammer, C. Beierle, J. Fisseler, G. Kern-Isberner, and J.I. Baumbach. Using probabilistic relational learning to support bronchial carcinoma diagnosis based on ion mobility spectrometry. *International Journal for Ion Mobility Spectrometry*, 13:83–93, 2010. (Cited on page 47.)

J. Fisseler. Toward markov logic with conditional probabilities. In D. C. Wilson and H. C. Lane, editors, *Proceedings of the Twenty-First International FLAIRS Conference*, pages 643–648. AAAI Press, 2008. (Cited on pages 49 and 51.)

Dov M. Gabbay, C. J. Hogger, and J. A. Robinson. *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*. Clarendon Press, 1994. (Cited on page 2.)

Dov Gabbay. Equational approach to argumentation networks. *Argument and Computation*, 3(2–3):87–142, 2012. (Cited on page 5.)

A. Garcia and Guillermo R. Simari. Defeasible Logic Programming: An Argumentative Approach. *Theory and Practice of Logic Programming*, 4(1–2):95–138, 2004. (Cited on pages 5, 13, 155, and 161.)

Michael Gelfond and Nicola Leone. Logic programming and knowledge representation - the a-prolog perspective. *Artificial Intelligence*, 138(1-2):3–38, June 2002. (Cited on pages 3, 13, 48, and 160.)

Michael Gelfond and Vladimir Lifschitz. Action languages. *Electronic Transactions on AI*, 2:193–210, 1998. (Cited on pages 155 and 162.)

Lise Getoor and Ben Taskar, editors. *Introduction to Statistical Relational Learning*. MIT Press, 2007. (Cited on page 10.)

Moises Goldszmidt and Judea Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, 84:57–112, 1996. (Cited on page 158.)

Matthias Grabmair and Kevin D. Ashley. Argumentation with value judgments - an example of hypothetical reasoning. In Radboud Winkels, editor, *JURIX*, volume 223 of *Frontiers in Artificial Intelligence and Applications*, pages 67–76. IOS Press, 2010. (Cited on page 31.)

John Grant and Anthony Hunter. Measuring Inconsistency in Knowledgebases. *Journal of Intelligent Information Systems*, 27:159–184, 2006. (Cited on pages 5, 66, 72, 155, 164, and 165.)

John Grant and Anthony Hunter. Analysing Inconsistent First-Order Knowledgebases. *Artificial Intelligence*, 172(8-9):1064–1093, May 2008. (Cited on pages 6, 66, and 98.)

John Grant and Anthony Hunter. Measuring Consistency Gain and Information Loss in Stepwise Inconsistency Resolution. In W. Liu, editor, *Proceedings of the 11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'11)*, number 6717 in Lecture Notes in Artificial Intelligence, pages 362–373. Springer International Publishing, 2011. (Cited on pages 6, 66, 68, 69, 98, 100, 104, 105, and 108.)

J. Grant and A. Hunter. Distance-based Measures of Inconsistency. In *Proceedings of the 12th Europen Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'13)*, volume 7958 of *Lecture Notes in Computer Science*, pages 230–241. Springer International Publishing, 2013. (Cited on pages 6, 12, 66, 68, 97, 98, 100, 110, 111, and 130.)

Davide Grossi and Wiebe van der Hoek. Uncertainty in abstract argument games: A case study on the game for the grounded extension. Technical Report ULCS-12-001, University of Liverpool, Department of Computer Science, 2012. (Cited on page 5.)

Davide Grossi and Wiebe van der Hoek. Audience-based uncertainty in abstract argument games. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, August 2013. (Cited on page 5.)

Christos Hadjinikolis, Sanjay Modgil, Elizabeth Black, Peter McBurney, and Michael Luck. Investigating strategic considerations in persuasion dialogue games. In *STAIRS 2012 - Proceedings of the Sixth Starting AI Researchers' Symposium*, pages 137–148, 2012. (Cited on page 45.)

Christos Hadjinikolis, Yiannis Siantos, Sanjay Modgil, Elizabeth Black, and Peter McBurney. Opponent Modelling in Persuasion Dialogues. In Francesca Rossi, editor, *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, pages 164–170, 2013. (Cited on page 9.)

Joseph Y. Halpern. An Analysis of First-Order Logics of Probability. *Artificial Intelligence*, 46:311–350, 1990. (Cited on page 10.)

Sven Ove Hansson. *A Textbook of Belief Dynamics*. Kluwer Academic Publishers, Norwell, MA, USA, 2001. (Cited on pages 3, 13, 47, 66, 97, 155, and 162.)

A. Hunter and S. Konieczny. Approaches to Measuring Inconsistent Information. In *Inconsistency Tolerance*, volume 3300 of *Lecture Notes in Computer Science*, pages 189–234. Springer International Publishing, 2004. (Cited on pages 5, 6, 66, 68, and 98.)

A. Hunter and S. Konieczny. Shapley inconsistency values. In *Proceedings of the 10th International Conference on Knowledge Representation (KR'06)*, pages 249–259. AAAI Press, 2006. (Cited on pages 98, 113, 114, 130, and 155.)

Anthony Hunter and Sébastien Konieczny. Measuring Inconsistency through Minimal Inconsistent Sets. In Gerhard Brewka and Jérôme Lang, editors, *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'2008)*, pages 358–366, Sydney, Australia, September 2008. AAAI Press, Menlo Park, California. (Cited on pages 6, 66, 68, 98, 100, 102, and 129.)

Anthony Hunter and Sebastien Konieczny. On the Measure of Conflicts: Shapley Inconsistency Values. *Artificial Intelligence*, 174(14):1007–1026, July 2010. (Cited on pages 6, 66, 68, 73, 74, 90, 98, 100, 108, and 109.)

Anthony Hunter and Matthias Thimm. Probabilistic argument graphs for argumentation lotteries. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*, September 2014. (Cited on pages 5 and 14.)

Anthony Hunter and Matthias Thimm. Probabilistic argumentation with epistemic extensions. In *Proceedings of the International Workshop on Defeasible and Ampliative Reasoning (DARe'14)*, August 2014. (Cited on pages 8 and 14.)

Anthony Hunter and Matthias Thimm. Probabilistic argumentation with epistemic extensions and incomplete information. Technical report, ArXiv, May 2014. (Cited on pages 8 and 14.)

Anthony Hunter and Matthias Thimm. Probabilistic argumentation with incomplete information. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 1033–1034, August 2014. (Cited on pages 5, 8, and 14.)

A. Hunter, S. Parsons, and M. Wooldridge. Measuring inconsistency in multi-agent systems. *Künstliche Intelligenz*, 28:169–178, 2014. (Cited on page 100.)

Anthony Hunter. Measuring Inconsistency in Knowledge via Quasi-Classical Models. In *Proceedings of the 18th American National Conference on Artificial Intelligence (AAAI'02)*, pages 68–73. MIT Press, 2002. (Cited on pages 66 and 68.)

Anthony Hunter. Some foundations for probabilistic abstract argumentation. In Verheij et al. (Verheij *et al.*, 2012), pages 117–128. (Cited on page 43.)

Anthony Hunter. A Probabilistic Approach to Modelling Uncertain Logical Arguments. *International Journal of Approximate Reasoning*, 54(1):47–81, 2013. (Cited on page 5.)

Anthony Hunter. Probabilistic strategies in dialogical argumentation. In *Proceedings of the 8th International Conference on Scalable Uncertainty Management (SUM'14)*, 2014. (Cited on page 5.)

Robert Isele, Jürgen Umbrich, Chris Bizer, and Andreas Harth. LDSpider: An Open-Source Crawling Framework for the Web of Linked Data. In *Proceedings of 9th International Semantic Web Conference (ISWC'10)*, 2010. (Cited on pages 75 and 89.)

S. Jabbour and B. Raddaoui. Measuring inconsistency through minimal proofs. In *Proceedings of the 12th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, ECSQARU'13, pages 290–301. Springer, 2013. (Cited on pages 12, 97, 98, 99, 100, 112, 113, and 130.)

Said Jabbour, Yue Ma, and Badran Raddaoui. Inconsistency Measurement Thanks to MUS Decomposition. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'14)*, pages 877–884, 2014. (Cited on page 72.)

Said Jabbour, Yue Ma, Badran Raddaoui, and Lakhdar Sais. Prime Implicates Based Inconsistency Characterization. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 1037–1038, 2014. (Cited on pages 6, 66, 68, 72, 98, and 100.)

Manfred Jaeger. *Default Reasoning about Probabilities*. PhD thesis, Universität des Saarlandes, 1995. (Cited on page 10.)

J. Janssen, M. D. Cock, and D. Vermeir. Fuzzy argumentation frameworks. In *Proc. of the 12th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'08)*, pages 513–520, 2008. (Cited on pages 15, 16, and 28.)

Edwin T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003. (Cited on page 20.)

Phillip Judson. *Knowledge-based Expert Systems in Chemistry*. RSC Theoretical and Computational Chemistry Series. Springer-Verlag, 2009. (Cited on page 1.)

Nishan C. Karunatillake, Nicholas R. Jennings, Iyad Rahwan, and Peter McBurney. Dialogue Games that Agents play within a Society. *Artificial Intelligence*, 173(9-10):935–981, April 2009. (Cited on pages 5 and 162.)

Gabriele Kern-Isberner and Matthias Thimm. Novel semantical approaches to relational probabilistic conditionals. In Fangzhen Lin, Ulrike Sattler, and Mirosław Truszczyński, editors, *Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR'10)*, pages 382–392. AAAI Press, May 2010. (Cited on pages 155 and 159.)

Gabriele Kern-Isberner and Matthias Thimm. A ranking semantics for first-order conditionals. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, August 2012. (Cited on pages 153, 155, and 158.)

Gabriele Kern-Isberner. *Conditionals in Nonmonotonic Reasoning and Belief Revision*. Number 2087 in Lecture Notes in Computer Science. Springer-Verlag, 2001. (Cited on pages 3, 155, and 158.)

Kristian Kersting. Lifted probabilistic inference. In *Proceedings of 20th European Conference on Artificial Intelligence (ECAI'12)*. IOS Press, August 2012. (Cited on page 51.)

K. M. Knight. Measuring inconsistency. *Journal of Philosophical Logic*, 31:77–98, 2001. (Cited on pages 66 and 68.)

Kevin M. Knight. *A Theory of Inconsistency*. PhD thesis, University Of Manchester, 2002. (Cited on pages 67, 69, 74, 98, 100, 106, 107, 116, 123, 124, 133, 135, and 136.)

J. Kohlas. Probabilistic argumentation systems: A new way to combine logic with probability. *Journal of Applied Logic*, 1(3-4):225–253, 2003. (Cited on pages 16 and 28.)

S. Konieczny and R. Pino Pérez. On the logic of merging. In *Proceedings of the Sixth International Conference on Principles of Knowledge Representationa and Reasoning (KR'98)*. Morgan Kaufmann, 1998. (Cited on pages 3 and 11.)

Sébastien Konieczny, Jérôme Lang, and Pierre Marquis. Reasoning under inconsistency: the forgotten connective. In *Proceedings of the 19th International Joint Conference of Artificial Intelligence (IJCAI'05)*, pages 484–489, 2005. (Cited on page 126.)

P. Krause, S. Ambler, M. Elvang-Goransson, and J. Fox. A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, 11(1):113–131, 1995. (Cited on pages 16 and 28.)

Patrick Krümpelmann and Gabriele Kern-Isberner. Belief base change operations for answer set programming. In Luis Farinas Cerro, Andreas Herzig, and Jerome Mengin, editors, *Proceedings of the 13th European conference on Logics in Artificial Intelligence (JELIA'12)*, pages 294–306. Springer, 2012. (Cited on pages 153, 155, and 160.)

Patrick Krümpelmann, Matthias Thimm, Marcelo A. Falappa, Alejandro J. Garcia, Gabriele Kern-Isberner, and Guillermo R. Simari. Selective revision by deductive argumentation. In Sanjay Modgil, Nir Oren, and Francesca Toni, editors, *Theory and Applications of Formal Argumentation, Proceedings of the First International Workshop (TAFA'11, revised selected papers)*, volume 7132 of *Lecture Notes in Artificial Intelligence*, pages 147–162. Springer, July 2011. (Cited on pages 153, 155, and 162.)

S. Kullback and R.A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, 1951. (Cited on page 54.)

D. Lawrence. *Genetic Algorithms and Simulated Annealing*. Pitman Publishing, 1987. (Cited on page 79.)

H. Li, N. Oren, and T. J. Norman. Probabilistic argumentation frameworks. In *Proceedings of the First International Workshop on the Theory and Applications of Formal Argumentation (TAFA'11)*, 2011. (Cited on pages 5, 15, 16, and 28.)

Mark H. Liffiton and Ammar Malik. Enumerating Infeasibility: Finding Multiple MUSes Quickly. In *Proceedings of the 10th International Conference on Integration of Artificial Intelligence and Operations Research techniques in Constraint Programming (CPAIOR'13)*, pages 160–175, 2013. (Cited on pages 84 and 88.)

Mark H. Liffiton and Karem A. Sakallah. On Finding All Minimally Unsatisfiable Subformulas. In *Proceedings of the 8th International Conference on Theory and Applications of Satisfiability Testing (SAT'05)*, pages 173–186. Springer International Publishing, 2005. (Cited on page 88.)

Vladimir Lifschitz, Lappoon R. Tang, and Hudson Turner. Nested expressions in logic programs. *Annals of Mathematics and Artificial Intelligence*, 25:369–389, 1999. (Cited on pages 155 and 160.)

Eliezer L. Lozinskii. Information and Evidence in Logic Systems. *Journal of Experimental and Theoretical Artificial Intelligence*, 6:163–193, 1994. (Cited on page 6.)

T. Lukasiewicz and G. Kern-Isberner. Probabilistic Logic Programming under Maximum Entropy. In *Proceedings ECSQARU-99*, volume 1638, pages 279–292. Springer Lecture Notes in Artificial Intelligence, 1999. (Cited on page 3.)

Yue Ma, Guilin Qi, Pascal Hitzler, and Zuoquan Lin. Measuring Inconsistency for Description Logics Based on Paraconsistent Semantics. In *Proceedings of the 9th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'07)*, pages 30–41, Berlin, Heidelberg, 2007. Springer International Publishing. (Cited on pages 6, 66, 98, and 109.)

Yue Ma, Guilin Qi, Guohui Xiao, Pascal Hitzler, and Zuoquan Lin. An Anytime Algorithm for Computing Inconsistency Measurement. In *Knowledge Science, Engineering and Management*, number 5914 in Lecture Notes in Computer Science, pages 29–40. Springer International Publishing, 2009. (Cited on pages 6, 12, 66, 68, 87, 98, and 113.)

Yue Ma, Guilin Qi, Guohui Xiao, Pascal Hitzler, and Zuoquan Lin. Computational Complexity and Anytime Algorithm for Inconsistency Measurement. *International Journal of Software and Informatics*, 4(1):3–21, 2010. (Cited on page 87.)

Y. Ma, G. Qi, and P. Hitzler. Computing inconsistency measure based on paraconsistent semantics. *Journal of Logic and Computation*, 21(6):1257–1281, December 2011. (Cited on page 109.)

Jianbing Ma, Weiru Liu, and Paul Miller. A Characteristic Function Approach to Inconsistency Measures for Knowledge Bases. In *Proceedings of the 6th International Conference on Scaleable Uncertainty Management (SUM'12)*, pages 473–485, 2012. (Cited on pages 6, 66, and 98.)

D. Makinson. *Bridges from Classical to Nonmonotonic Logic*. College Publications, 2005. (Cited on pages 66 and 97.)

P.-A. Matt and F. Toni. Game-Theoretic Measure of Argument Strength for Abstract Argumentation. In *Proc. of the 11th European Conf. on Logics in Artificial Intelligence (JELIA'08)*, 2008. (Cited on pages 19 and 29.)

Kevin McAreavey, Weiru Liu, and Paul Miller. Computational Approaches to Finding and Measuring Inconsistency in Arbitrary Knowledge Bases. *International Journal of Approximate Reasoning*, 55:1659–1693, 2014. (Cited on pages 6, 12, 66, 68, 72, 84, 88, 98, 100, and 113.)

Thomas J. McCabe. A complexity measure. *IEEE Transactions on Software Engineering*, pages 308–320, 1976. (Cited on page 154.)

W. Meijs. *Probabilistic Measures of Coherence*. Haveka, 2005. (Cited on page 61.)

Sanjay Modgil and Trevor J.M Bench-Capon. Metalevel Argumentation. *Journal of Logic and Computation*, 21(6):959–1003, 2011. (Cited on page 4.)

Sanjay Modgil and Martin Caminada. Proof Theories and Algorithms for Abstract Argumentation Frameworks. In Guillermo Simari and Iyad Rahwan, editors, *Argumentation in Artificial Intelligence*, chapter 6, pages 105–129. Springer US, Boston, MA, 2009. (Cited on pages 36 and 37.)

Sanjay Modgil. Labellings and Games for Extended Argumentation Frameworks. In *Proc. of the 21st Int. Joint Conf. on Artificial Intelligence (IJCAI'09)*, pages 873–878, 2009. (Cited on page 27.)

K. Mu, W. Liu, Z. Jin, and D. A. Bell. A Syntax-based Approach to Measuring the Degree of Inconsistency for Belief Bases. *International Journal of Approximate Reasoning*, 52(7):978–999, 2011. (Cited on pages 6, 66, 68, 98, 100, 103, 114, 122, and 130.)

Kedian Mu, Weiru Liu, and Zhi Jin. A general framework for measuring inconsistency through minimal inconsistent sets. *Knowledge and Information Systems*, 27:85–114, 2011. (Cited on pages 66 and 68.)

Kedian Mu, Kewen Wang, and Lian Wen. Approaches to Measuring Inconsistency for Stratified Knowledge Bases. *International Journal of Approximate Reasoning*, 55:529–556, 2014. (Cited on pages 66 and 72.)

David Picado Muiño. Measuring and Repairing Inconsistency in Probabilistic Knowledge Bases. *International Journal of Approximate Reasoning*, 52(6):828–840, September 2011. (Cited on pages 2 and 66.)

M. Niepert. Symmetry-aware marginal density estimation. In *Proceedings of the 27th Conference on Artificial Intelligence (AAAI'13)*, 2013. (Cited on page 60.)

Donald Nute and Charles Cross. Conditional logic. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume 4, pages 1–98. Kluwer Academic Publishers, second edition edition, 2002. (Cited on pages 3 and 158.)

Donald Nute. Defeasible logic. In Dov M. Gabbay, Hogger C. J., and Robinson J. A., editors, *Handbook of logic in artificial intelligence and logic programming (vol. 3)*, pages 353–395. Oxford University Press, Inc., New York, NY, USA, 1994. (Cited on page 3.)

Nir Oren and Timothy J. Norman. Arguing using opponent models. In *Proceedings of the 6th international conference on Argumentation in Multi-Agent Systems*, ArgMAS'09, pages 160–174, Berlin, Heidelberg, 2009. Springer-Verlag. (Cited on pages 9, 31, 32, 35, 36, and 166.)

Nir Oren, Timothy J. Norman, and Alun Preece. Arguing with confidential information. In *Proceedings of the 18th European Conference on Artificial Intelligence*, pages 280–284, Riva del Garda, Italy, August 2006. (Cited on page 43.)

Nir Oren, Katie Atkinson, and Hengfei Li. Group persuasion through uncertain audience modelling. In Verheij et al. (Verheij *et al.*, 2012), pages 350–357. (Cited on page 43.)

Christos H. Papadimitriou. *Computational Complexity*. Addison Wesley, 1994. (Cited on page 11.)

T. Pápai, S. Ghosh, and H. Kautz. Combining subjective probabilities and data in training markov logic networks. In *Proceedings of the 2012 European conference on Machine Learning and Knowledge Discovery in Databases (ECML+PKDD'12)*, pages 90–105, Berlin, Heidelberg, 2012. Springer-Verlag. (Cited on page 60.)

Jeff B. Paris. *The Uncertain Reasoner's Companion – A Mathematical Perspective*. Cambridge University Press, 1994. (Cited on pages 3, 8, 16, 20, 22, 23, 51, 159, and 163.)

J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Springer, 1988. (Cited on pages 3, 8, 49, and 51.)

J. Pearl. System Z: a natural ordering of defaults with tractable applications to default reasoning. In *Proceedings of the 3rd conference on Theoretical* *aspects of reasoning about knowledge (TARK'90)*, pages 121–135. Morgan Kaufmann Publishers, 1990. (Cited on page 155.)

Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1998. (Cited on pages 16 and 20.)

Sylwia Polberg and Dragan Doder. Probabilistic abstract dialectical frameworks. In *Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA'14)*, 2014. (Cited on page 5.)

James Popple. *A Pragmatic Legal Expert System*. Applied Legal Philosophy Series. Dartmouth Publishing Group, 1996. (Cited on page 1.)

Nico Potyka and Matthias Thimm. Consolidation of probabilistic knowledge bases by inconsistency minimization. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 729–734, August 2014. (Cited on pages 6 and 14.)

Nico Potyka and Matthias Thimm. Probabilistic reasoning with inconsistent beliefs using inconsistency measures. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, July 2015. (Cited on pages 6 and 14.)

Nico Potyka. Linear Programs for Measuring Inconsistency in Probabilistic Logics. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14)*, pages 568–577, 2014. (Cited on pages 6, 66, 98, and 155.)

Henry Prakken and Giovanni Sartor. *Computational Logic: Logic Programming and Beyond. Essays In Honour of Robert A. Kowalski, Part II*, volume 2048 of *LNCS*, pages 342–380. Springer-Verlag, 2002. (Cited on page 32.)

H. Prakken. A Study of Accrual of Arguments, with Applications to Evidential Reasoning. In *Proc. of the Tenth Int. Conf. on Artificial Intelligence and Law*, pages 85–94, 2005. (Cited on page 29.)

Henry Prakken. An Abstract Framework for Argumentation with Structured Arguments. *Argument and Computation*, 1:93–124, 2009. (Cited on page 5.)

Alessandro Previti and Joao Marques-Silva. Partial MUS Enumeration. In *Proceedings of the 27th Conference on Artificial Intelligence (AAAI'13)*, pages 818–825, 2013. (Cited on page 88.)

G. Priest. Logic of Paradox. *Journal of Philosophical Logic*, 8:219–241, 1979. (Cited on pages 66, 69, 70, 97, and 109.)

Guilin Qi and Anthony Hunter. Measuring Incoherence in Description Logic-based Ontologies. In *Proceedings of the 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference (ISWC'07/ASWC'07)*, pages 381–394, Berlin, Heidelberg, 2007. Springer International Publishing. (Cited on page 66.)

Iyad Rahwan and Kate Larson. Mechanism design for abstract argumentation. In *Proceedings of AAMAS 2008*, 2008. (Cited on page 44.)

Iyad Rahwan and Kate Larson. Argumentation and Game Theory. In Iyad Rahwan and Guillermo R. Simari, editor, *Argumentation in Artificial Intelligence*, pages 321–339. Springer-Verlag, 2009. (Cited on page 5.)

Iyad Rahwan and Guillermo R. Simari, editors. *Argumentation in Artificial Intelligence*. Springer-Verlag, 2009. (Cited on pages 4 and 15.)

Iyad Rahwan, Kate Larson, and Fernando A. Tohmé. A characterisation of strategy-proofness for grounded argumentation semantics. In *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence*, pages 251–256, July 2009. (Cited on page 44.)

R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980. (Cited on pages 3 and 48.)

M. Richardson and P. Domingos. Markov logic networks. *Machine Learning*, 62(1–2):107–136, 2006. (Cited on pages 10, 47, 49, 50, 51, 155, and 159.)

Tjitze Rienstra, Matthias Thimm, and Nir Oren. Opponent Models with Uncertainty for Strategic Argumentation. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, 2013. (Cited on pages 5, 152, 153, 155, 162, 163, 165, 166, and 167.)

Tjitze Rienstra. Towards a probabilistic dung-style argumentation system. In *Proceedings AT2012*, 2012. (Cited on page 5.)

W. Rödder. Conditional logic and the principle of entropy. *Artificial Intelligence*, 117:83–106, 2000. (Cited on pages 155 and 159.)

B. Roth, R. Riveret, A. Rotolo, and G. Governatori. Strategic Argumentation: A Game Theoretical Investigation. In *ICAIL '07: Proceedings of the 11th International Conference on Artificial Intelligence and Law*, pages 81–90, New York, NY, USA, 2007. ACM Press. (Cited on page 44.)

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education, second edition, 2003. (Cited on page 1.)

Eveline R. Sacramento, Marco A. Casanova, Karin K. Breitman, Antonio L. Furtado, José Antonio F. de Macĕdo, and Vănia M. P. Vidal. Dealing with Inconsistencies in Linked Data Mashups. In *Proceedings of the 16th International Database Engineering and Applications Sysmposium (IDEAS'12)*, pages 175–180, 2012. (Cited on pages 11 and 66.)

Ralf Schweimeier and Michael Schroeder. A parameterised hierarchy of argumentation semantics for extended logic programming and its application to the well-founded semantics. *Theory and Practice of Logic Programming*, 5(1–2):207–242, 2003. (Cited on pages 155 and 161.)

Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976. (Cited on page 3.)

Claude E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27:379–423, 1948. (Cited on page 6.)

Edward .H. Shortliffe and Bruce G. Buchanan. A Model of Inexact Reasoning in Medicine. *Mathematical Biosciences*, 23(3–4):351–379, 1975. (Cited on page 1.)

Amit Singhal. Introducing the Knowledge Graph: Things, Not Strings (May 16, 2012), 2012. Official Blog of Google, http://googleblog.blogspot.co.uk/2012/05/introducing-knowledge-graph-things-not.html. Retrieved June 26, 2015. (Cited on page 1.)

Martin Slota and Joao Leite. On Semantic Update Operators for Answer-Set Programs. In H. Coelho, R. Studer, and M. Wooldridge, editors, *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI'10)*, pages 957–962. IOS Press, 2012. (Cited on page 3.)

E. Sperner. Ein Satz über Untermengen einer endlichen Menge (in German). *Mathematische Zeitschrift*, 27(1):544–548, 1928. (Cited on pages 13, 98, 117, 131, 135, and 136.)

Wolfgang Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In W.L. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics*, volume 2, pages 105–134. Kluwer Academic Publishers, 1988. (Cited on page 158.)

Tamon Stephen and Timothy Yusun. Counting inequivalent monotone boolean functions. Technical report, arXiv:1209.4623, 2012. (Cited on pages 119, 120, 121, and 133.)

Frieder Stolzenburg, Alejandro Garcia, Carlos I. Chesnevar, and Guillermo R. Simari. Computing generalized specificity. *Journal of Non-Classical Logics*, 13(1):87–113, 2003. (Cited on pages 155 and 161.)

Matthias Thimm and Alejandro J. Garcia. Classification and Strategical Issues of Argumentation Games on Structured Argumentation Frameworks. In Wiebe van der Hoek, Gal A. Kaminka, Yves Lespérance, Michael Luck, and Sandip Sen, editors, *Proceedings of the Ninth International Joint Conference on Autonomous Agents and Multi-Agent Systems 2010 (AAMAS'10)*, May 2010. (Cited on pages 35, 44, 152, 153, 155, 161, 163, and 165.)

Matthias Thimm and Gabriele Kern-Isberner. A distributed argumentation framework using defeasible logic programming. In Philippe Besnard, Sylvie Doutre, and Anthony Hunter, editors, *Proceedings of the 2nd International Conference on Computational Models of Argument (COMMA'08)*, number 172 in Frontiers in Artificial Intelligence and Applications, pages 381–392. IOS Press, May 2008. (Cited on page 153.)

Matthias Thimm and Gabriele Kern-Isberner. On controversiality of arguments and stratified labelings. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*, September 2014. (Cited on page 14.)

Matthias Thimm. *Probabilistic Reasoning with Incomplete and Inconsistent Beliefs*, volume 331 of *Dissertations in Artificial Intelligence*. IOS Press, December 2011. (Cited on pages 152, 153, 159, 163, and 165.)

Matthias Thimm. A Probabilistic Semantics for Abstract Argumentation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, 2012. (Cited on pages 5, 43, 153, 155, and 161.)

Matthias Thimm. Dynamic preference aggregation under preference changes. In *Proceedings of the Fourth Workshop on Dynamics of Knowledge and Belief (DKB'13)*, September 2013. (Cited on pages 153, 155, and 163.)

Matthias Thimm. Inconsistency Measures for Probabilistic Logics. *Artificial Intelligence*, 197:1–24, April 2013. (Cited on pages 6, 14, 52, 61, 66, 68, 73, 90, 98, 133, 152, 153, 155, 159, 163, 164, and 165.)

Matthias Thimm. Coherence and Compatibility of Markov Logic Networks. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14)*, pages 891–896, August 2014. (Cited on pages 6 and 66.)

Matthias Thimm. Strategic Argumentation in Multi-Agent Systems. *Künstliche Intelligenz, Special Issue on Multi-Agent Decision Making*, 28(3):159–168, 2014. (Cited on pages 5 and 8.)

Matthias Thimm. Towards Large-scale Inconsistency Measurement. In *Proceedings of the 37th German Conference on Artificial Intelligence (KI'14)*, pages 195–206, September 2014. (Cited on pages 6, 67, and 75.)

Matthias Thimm. Tweety - A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation. In *Proceedings of the 14th Int. Conference on Principles of Knowledge Representation and Reasoning (KR'14)*, pages 528–537, 2014. (Cited on pages 83 and 89.)

Matthias Thimm. Stream-based inconsistency measurement. *International Journal of Approximate Reasoning*, 68:68–87, January 2016. (Cited on pages 98, 100, 107, 113, and 127.)

Alice Toniolo, Timothy J. Norman, and Katia P. Sycara. Argumentation schemes for collaborative planning. In David Kinny, Jane Yung jen Hsu, Guido Governatori, and Aditya K. Ghose, editors, *PRIMA*, volume 7047 of *Lecture Notes in Computer Science*, pages 323–335. Springer, 2011. (Cited on page 31.)

Miklos A. Vasarhelyi, Enrique Bonson, and Rani Hoitash, editors. *Artificial Intelligence in Accounting & Auditing: International Perspectives*. Rutgers Series in Accounting Information Systems. Markus Wiener Publishers, 2005. (Cited on page 1.)

Bart Verheij, Stefan Szeider, and Stefan Woltran, editors. *Computational Models of Argument - Proceedings of COMMA 2012, Vienna, Austria, September 10-12, 2012*, volume 245 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, 2012. (Cited on pages 176 and 181.)

Bart Verheij. Arguments and their strength: Revisiting pollock's anti-probabilistic starting points. In *Proceedings of the Fifth International Conference on Computational Models of Argumentation (COMMA'14)*, 2014. (Cited on page 5.)

Toby Walsh. Representing and reasoning with preferences. *AI Magazine*, 28(4):59–70, 2007. (Cited on pages 155 and 163.)

G. Weiss, editor. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, Cambridge, MA, USA, second edition, 2013. (Cited on page 155.)

Y. Wu and M. Caminada. A labelling-based justification status of arguments. *Studies in Logic*, 3(4):12–29, 2010. (Cited on pages 7, 15, 17, and 28.)

Guohui Xiao and Yue Ma. Inconsistency Measurement based on Variables in Minimal Unsatisfiable Subsets. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI'12)*, pages 864–869, 2012. (Cited on pages 6, 66, 68, 98, 100, and 108.)

Guohui Xiao, Zuoquan Lin, Yue Ma, and Guilin Qi. Computing Inconsistency Measurements under Multi-Valued Semantics by Partial Max-SAT Solvers. In *Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR'10)*, pages 340–349, 2010. (Cited on page 87.)

Liping Zhou, Houkuan Huang, Guilin Qi, Yue Ma, Zhisheng Huang, and Youli Qu. Measuring Inconsistency in DL-Lite Ontologies. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT'09)*, pages 349–356, Washington, DC, USA, 2009. IEEE Computer Society. (Cited on pages 6, 66, and 98.)