A Framework for Inconsistency-tolerant Reasoning with Sets of Models

Yehia Hatab, **Kai Sauerwald**, **Matthias Thimm**
Artificial Intelligence Group, University of Hagen, Germany
{yehia.hatab, kai.sauerwald, matthias.thimm}@fernuni-hagen.de

Abstract

We propose a framework for reasoning from inconsistent knowledge bases using minimal *hitting sets*, i. e., sets of interpretations such that each formula of the knowledge base is satisfied by at least one those interpretations. By additionally considering preference orders over minimal hitting sets, we can define a wide variety of non-monotonic inference relations. We consider concrete preference orders based on set inclusion, cardinality, the number of conflicting atoms within the hitting set, and using the Hamming distance between pairs of interpretations. We compare the resulting inference relations, characterize their logical properties, and position them relative to classical inference from maximal consistent subsets. Finally, we show that inference based on minimal conflicting atoms coincides with reasoning in Priest's 3-valued logic.

1 Introduction

A core objective in knowledge representation and reasoning (KRR) is to support sound and meaningful inference from knowledge bases (Harmelen, Lifschitz, and Porter 2007). Inconsistencies are an unavoidable reality in large-scale, heterogeneous, or evolving knowledge bases. Whether arising from data integration, imperfect information, or conflicting sources, such inconsistencies pose a significant challenge for KRR systems. Considerable research has addressed methods for identifying and resolving them, including approaches from inconsistency measurement (Thimm 2018), belief revision (Fermé and Wassermann 2018), or ontology engineering (Kalyanpur et al. 2006; Baader et al. 2018). However, resolving inconsistencies may be difficult or undesirable—either due to the cost of modification or the need to preserve original information. In such cases, reasoning directly under inconsistency becomes essential (Rescher and Manor 1970; Benferhat, Dubois, and Prade 1997; Lang and Marquis 2010).

Traditional approaches to inconsistency-tolerant reasoning often rely on syntactic repairs, such as extracting maximal consistent subsets or applying belief revision operators. While these methods have proven effective in many contexts, they frequently ignore the semantic structure of the underlying models and treat all formulas uniformly, irrespective of their role in the inconsistency. Moreover, existing semantic approaches typically lack mechanisms for flexible inference or fine-grained preference among models.

In this paper, we propose a new framework for inconsistency-tolerant reasoning based on sets of models—specifically, minimal hitting sets of interpretations guided by preference relations. This semantic perspective enables reasoning directly over interpretations, allowing for a wide spectrum of inference relations that range from cautious to permissive, depending on how preferred hitting sets are selected. This establishes a unified perspective that not only captures existing approaches as special cases but also supports the principled design of new inference strategies.

Our contribution lays the semantic foundation for a general framework of inconsistency-tolerant reasoning with sets of models. Rather than committing to a single inference relation, the framework is parameterized by a preference relation over hitting sets. We formally identify the basic properties that preference relations must satisfy to ensure well-behaved inference.

Our main contributions are as follows:

- We introduce a general framework for non-monotonic inference from inconsistent knowledge bases using minimal hitting sets of interpretations.
- We instantiate the framework using preference orders based on set inclusion, cardinality, conflicting atoms, and Hamming distance.
- We relate our framework to established approaches, including inference from maximal consistent subsets and reasoning in 3-valued logic.
- We define semantic conditions on preference relations that guarantee classical behavior for consistent knowledge and support principled non-monotonic inference under inconsistency.

Detailed proofs of all propositions and theorems are provided in the full paper.[1]

The remainder of this paper is structured as follows. Section 2 provides the necessary background on propositional logic and classical inference. In Section 3, we introduce our general framework for reasoning with minimal hitting sets. Section 4 presents various instantiations of preference orders and defines the resulting inference relations. Section 5

[1] https://zenodo.org/records/16538664

analyzes their relationships to inference from maximal consistent subsets and 3-valued logic. Section 6 shows the satisfaction of the non-monotonic properties of these inference modes. We conclude in Section 7 with a summary and directions for future research.

2 Background

Let At be some fixed propositional signature, i. e., a (possibly infinite) set of propositional variables (also called atoms), and let $\mathscr{L}(\mathsf{At})$ be the corresponding propositional language constructed using the usual connectives $\wedge$ (*conjunction*), $\vee$ (*disjunction*), $\rightarrow$ (*implication*), $\leftrightarrow$ (*bi-implication*) and $\neg$ (*negation*). Furthermore, $\top$ denotes an arbitrary tautology and $\bot$ denotes an arbitrary contradiction.

If ϕ is a formula, we write $\mathsf{At}(\phi)$ to denote the set of propositions appearing in ϕ. Similarly, for a set $\Phi = \{\phi_1, \ldots, \phi_n\}$, we write $\mathsf{At}(\Phi)$ for the propositions appearing in Φ. Furthermore, let $\bigwedge \Phi = \phi_1 \wedge \ldots \wedge \phi_n$ and $\neg\Phi = \{\neg\phi \mid \phi \in \Phi\}$.

Semantics to a propositional language are given by *interpretations* where an *interpretation* ω on At is a function $\omega : \mathsf{At} \rightarrow \{\mathsf{true}, \mathsf{false}\}$. Let $\Omega(\mathsf{At})$ 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 $\omega(a) = \mathsf{true}$. The relation $\models$ is extended to formulas in the usual way. For a set of formulas $\Phi \subseteq \mathscr{L}(\mathsf{At})$ we write $\omega \models \Phi$ if $\omega \models \phi$ for every $\phi \in \Phi$. 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 propositional variables that are assigned true by ω and $a_{n+1}, \ldots, a_m \in \mathsf{At}$ are those variables 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 $a\bar{b}c$. As usual, we use $\models$ to denote classical entailment. That is, $\models \subseteq \mathscr{P}(\mathscr{L}(\mathsf{At})) \times \mathscr{L}(\mathsf{At})$, which is a binary relation, where $\Phi \models \phi$ if and only if every interpretation that satisfies Φ must also satisfy ϕ. In the following, let Φ, Φ_1, Φ_2 be sets of formulas. Define the set of models $\mathsf{Mod}(\Phi) = \{\omega \in \Omega(\mathsf{At}) \mid \omega \models \Phi\}$. We write $\Phi_1 \models \Phi_2$ if $\mathsf{Mod}(\Phi_1) \subseteq \mathsf{Mod}(\Phi_2)$. We say that Φ_1, Φ_2 are *equivalent*, denoted by $\Phi_1 \equiv \Phi_2$, if $\mathsf{Mod}(\Phi_1) = \mathsf{Mod}(\Phi_2)$. If $\mathsf{Mod}(\Phi) = \emptyset$ holds, we say that Φ is *inconsistent*.

In this work, we are only considering knowledge bases that are non-empty and do not contain formulas that contradict themselves.

Definition 1. A knowledge base K is a non-empty finite set of formulas $K \subseteq \mathscr{L}(\mathsf{At})$, such that $K \neq \emptyset$ and, every formula $\phi \in K$ is consistent. Let $\mathbb{K}$ be the set of all knowledge bases.

Note that while every formula in the knowledge base K is consistent, the knowledge base itself can be inconsistent, for example $K = \{p, q, \neg p \vee \neg q\}$.

An important concept for reasoning under inconsistency is reasoning based on maximal consistent sets.

Definition 2. Let K be a knowledge base. A set $M \subseteq K$ is a *maximal consistent subset* of K if $M \not\models \bot$ and for all $M' \supsetneq M$ holds $M' \models \bot$. With $\mathsf{MCS}(K)$ we denote the set of all maximally consistent subsets of K.

Maximal consistent subsets provide a foundation for defining inference relations in the presence of inconsistency. We now recall two such classical inference relations introduced by Rescher and Manor 1970

Definition 3 (Rescher and Manor 1970)**.** Let K be a knowledge base.

- A formula ϕ is said to be an *inevitable consequence* of K, shortly $K {\mid\!\sim}^{\mathrm{mc}}_{i} \phi$, if $M \models \phi$ for all $M \in \mathsf{MCS}(K)$.
- A formula ϕ is said to be an *weak consequence* of K, shortly $K {\mid\!\sim}^{\mathrm{mc}}_{w} \phi$, if $M \models \phi$ for some $M \in \mathsf{MCS}(K)$.

The following example clarifies Definition 2 and shows inference is possible from maximally consistent subsets.

Example 1. Consider the knowledge base $K = \{p, q, r \wedge \neg p\}$. First, observe that K is inconsistent. Note that $\mathsf{MCS}(K) = \{\{p, q\}, \{q, r \wedge \neg p\}\}$ are the maximally consistent subsets of K. Hence, inevitable inference are some inferences are $K {\mid\!\sim}^{\mathrm{mc}}_{i} q$, and some weak consequence are $K {\mid\!\sim}^{\mathrm{mc}}_{w} p$, $K {\mid\!\sim}^{\mathrm{mc}}_{w} q$, ${\mid\!\sim}^{\mathrm{mc}}_{w} \neg p$, $K {\mid\!\sim}^{\mathrm{mc}}_{w} r$:

While inference based on maximal consistent subsets is a classical approach to reasoning under inconsistency, it does not take into account the semantic roles of individual formulas. For example, consider a formula such as r that is not directly involved in any contradiction. Despite this, it may be inferred with the same inference operator as p or $\neg p$ which are explicitly in conflict. This uniform treatment can obscure the distinction between formulas that merely coexist with inconsistency and those that contribute to it.

We now turn to Priest's 3-valued logic (Priest 1979), which extends classical truth values with a third option: both. This value represents cases where a formula is considered simultaneously true and false, capturing contradictions directly within the semantic evaluation.

Definition 4. A 3-valued interpretation υ on At is a function $\upsilon : \mathsf{At} \rightarrow \{true, false, both\}$.

The additional truth value *both* refers to the conflicting state when an atom is both *true* and *false* at the same time. The function υ is extended to formulas as shown in Table 1.

Definition 5. For a 3-valued interpretation υ and a formula ϕ, we write $\upsilon \models^3 \phi$ iff $\upsilon(\phi) \in \{true, both\}$. Let $\mathsf{Mod}^3(K)$ be the set of all 3-valued models of the knowledge base K.

The relation $\models^3$is extended naturally to sets of formulas: a 3-valued interpretation satisfies a set if it satisfies each individual formula in the set. For clarity, we use the symbol υ for 3-valued interpretations and the symbol ω for classical two-valued interpretations.

A 3-valued interpretation that assigns many atoms to *both* is not informative , as it often fails to yield a definitive truth value for many formulas. An illustrative extreme case is the interpretation υ_0 such that $\upsilon_0(a) = both$ for all $a \in \mathsf{At}$, which is a model of every formula. Thus, we define a notion of minimal models for a knowledge base that minimizes the set of atoms assigned *both*.

$\upsilon(\phi)$	$\upsilon(\psi)$	$\upsilon(\phi \wedge \psi)$	$\upsilon(\phi \vee \psi)$	$\upsilon(\neg\phi)$
false	*false*	*false*	*false*	*true*
false	*both*	*false*	*both*	*true*
false	*true*	*false*	*true*	*true*
both	*false*	*false*	*both*	*both*
both	*both*	*both*	*both*	*both*
both	*true*	*both*	*true*	*both*
true	*false*	*false*	*true*	*false*
true	*both*	*both*	*true*	*false*
true	*true*	*true*	*true*	*false*

Table 1: 3-valued interpretation truth tables.

Definition 6. A model υ of a knowledge base K is a minimal model of K if there is no other $\upsilon' \models^3 K$ with $(\upsilon')^{-1}(both) \subsetneq (\upsilon)^{-1}(both)$. Let $MinMod^3(K)$ denote the set of minimal models of K.

Based on Definition 6, we define an inference relation that considers only such minimal models.

Definition 7. Define the 3-valued inference relation $\mid\!\sim^3$ via

$$K \mid\!\sim^3 \phi \text{ iff } \upsilon \models^3 \phi \text{ for every } \upsilon \in MinMod^3(K).$$

Having reviewed both classical and 3-valued approaches to reasoning under inconsistency, we now turn to our hitting set-based framework, which builds on sets of interpretations to support a flexible and semantically grounded inference mechanism.

3 A General framework for Reasoning with Hitting Sets

In this section, we introduce a framework for inconsistency-tolerant reasoning built upon the concept of *hitting sets* of interpretations. This approach allows us to define inference relations that operate directly on sets of models, providing a more fine-grained and adaptable foundation for drawing conclusions from potentially conflicting information. We consider the following definition of a hitting set (Thimm 2016; Thimm 2014).

Definition 8. Let K be a knowledge base. A set $H \subseteq \Omega(\mathsf{At})$ is a *hitting set* of K, if for all $\phi \in K$ there is $\omega \in H$ with $\omega \models \phi$. Let $\mathscr{H}(K)$ be the set of all hitting sets of K.[2]

Hitting sets were originally introduced as a tool for measuring inconsistency in knowledge bases (Thimm 2016; Thimm 2014).

Consider the following example.

Example 2. Given the knowledge base $K_1 = \{p, \neg p \wedge q\}$. It has the following hitting sets:

$$\mathscr{H}(K_1) = \{\{pq, \bar{p}q\}, \{p\bar{q}, \bar{p}q\}, \{pq, \bar{p}\bar{q}, \bar{p}q\}, \{pq, p\bar{q}, \bar{p}q\}, \{p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}, \{pq, \bar{p}\bar{q}, \bar{p}q, p\bar{q}\}\}$$

[2]Note that in a more general sense, a set H is a *hitting set* of a set of sets S if $M \cap T \neq \emptyset$ for all $T \in S$. In our context, S is a set of formulas and each formula is represented by its set of models. So a hitting set H of a knowledge base K is the same as a hitting set of $\{\mathsf{Mod}(\phi) \mid \phi \in K\}$.

Although any set of interpretations that satisfies every formula in the knowledge base qualifies as a hitting set, many such sets may contain redundant or unnecessary elements. For instance, in Example 2, the set $pq, \bar{p}q$ is sufficient to satisfy the knowledge base, but larger sets are still technically valid hitting sets. To refine this notion for inference, we introduce preference relations over hitting sets to isolate those that are minimal in a well-defined sense.

We generalize the notion of minimality by considering minimality with respect to an arbitrary relation $\prec$ over hitting sets.

Definition 9. Given a knowledge base K. A hitting set $H \in \mathscr{H}(K)$ is *minimal* wrt. $\prec \subseteq 2^{\Omega} \times 2^{\Omega}$ if there is no $H' \in \mathscr{H}(K)$ with $H' \prec H$. Let $\mathscr{H}_{\prec}(K)$ denote the set of all minimal hitting sets of K wrt. $\prec$.

The idea of minimality depends on the specific preference relation $\prec$ applied. This flexibility allows us to tailor inference to different intuitions about what constitutes a "better" or "more informative" set of interpretations. In the next definition, we formalize four distinct modes of inference by quantifying over both preferred hitting sets and the interpretations they contain.

Definition 10. For $\prec \subseteq 2^{\Omega} \times 2^{\Omega}$, define inference relations $\mid\!\sim^{\text{ns}}$, $\mid\!\sim^{\text{nc}}$, $\mid\!\sim^{\text{ps}}$ and $\mid\!\sim^{\text{pc}}$ via

$$\begin{aligned}
K \mid\!\sim^{\text{ns}}_{\prec} \phi &\quad \text{if } \forall H \in \mathscr{H}_{\prec}(K) : \forall \omega \in H : \omega \models \phi \\
K \mid\!\sim^{\text{nc}}_{\prec} \phi &\quad \text{if } \forall H \in \mathscr{H}_{\prec}(K) : \exists \omega \in H : \omega \models \phi \\
K \mid\!\sim^{\text{ps}}_{\prec} \phi &\quad \text{if } \exists H \in \mathscr{H}_{\prec}(K) : \forall \omega \in H : \omega \models \phi \\
K \mid\!\sim^{\text{pc}}_{\prec} \phi &\quad \text{if } \exists H \in \mathscr{H}_{\prec}(K) : \exists \omega \in H : \omega \models \phi
\end{aligned}$$

for every knowledge base K and formula ϕ.

- If $K \mid\!\sim^{\text{ns}}_{\prec} \phi$ holds, we say that ϕ is a *necessary skeptical inference* of K (wrt. $\prec$).
- If $K \mid\!\sim^{\text{nc}}_{\prec} \phi$ holds, we say that ϕ is a *necessary credulous inference* of K (wrt. $\prec$).
- If $K \mid\!\sim^{\text{ps}}_{\prec} \phi$ holds, we say that ϕ is a *possible skeptical inference* of K (wrt. $\prec$).
- If $K \mid\!\sim^{\text{pc}}_{\prec} \phi$ holds, we say that ϕ is a *possible credulous inference* of K (wrt. $\prec$).

These four inference relations offer different perspectives on reasoning from sets of models. Necessary inferences (both skeptical and credulous) require that a formula holds across all minimal hitting sets, differing on whether every interpretation (skeptical) or some interpretation (credulous) within the hitting set must satisfy the formula. Possible inferences, in contrast, require only the existence of a minimal hitting set where the formula holds universally (skeptical) or partially (credulous). This distinction allows flexible reasoning patterns, depending on how cautious or permissive one wishes to be when accepting conclusions from possibly inconsistent knowledge bases.

Not every preference relation yields meaningful inference. In particular, we want to ensure that minimal hitting sets exist and that the framework aligns with classical logic when the knowledge base is consistent. To that end, we identify two critical properties for preference relations—well-foundedness and compatibility with classical consistency—and define proper relations as those that satisfy both. These ensure that our inference is well-behaved in both consistent and inconsistent cases.

Definition 11. Let $\prec \subseteq 2^\Omega \times 2^\Omega$ be a relation. We say that $S \subseteq 2^\Omega$ is *well-founded by* T if $T \in S$ and there is no $T' \in S$ such that $T' \prec T$. We say that $\prec$ is *well-founded for* S if S is well-founded by some $T \in S$. If $\prec$ is well-founded for all $S \subseteq 2^\Omega$, then we say that $\prec$ is *well-founded*.

In order to ensure that inference from inconsistent knowledge bases behaves in accordance with classical logic, it is essential to impose additional constraints on preference relations. One such constraint is *compatibility with classical consistency*, which requires that, in the absence of inconsistency, the framework reproduce the standard semantics of classical entailment. Intuitively, this means that when the knowledge base is consistent, every model of the knowledge base should be represented by a singleton minimal hitting set. This condition provides a bridge between non-monotonic reasoning in the presence of inconsistency and classical reasoning when no inconsistency is present.

Definition 12. Let $\prec \subseteq 2^\Omega \times 2^\Omega$ be a relation. We say $\prec$ is *compatible with classical consistency* if for every consistent knowledge base K the following holds:

- every $H \in \mathscr{H}_\prec(K)$ is a singleton, and
- $\bigcup_{H \in \mathscr{H}_\prec(K)} H = \mathsf{Mod}(K)$.

We define a preference relation $\prec \subseteq 2^\Omega \times 2^\Omega$ as proper if it is well-founded and compatible with classical consistency.

Definition 13. A relation $\prec \subseteq 2^\Omega \times 2^\Omega$ is called *proper* if it is well-founded and compatible with classical consistency.

The next proposition shows that for a well-founded relation, the set of minimal hitting sets is non-empty for every knowledge base.

Proposition 1. *Let $\prec \subseteq 2^\Omega \times 2^\Omega$ be a relation and let K be a knowledge base. The following statements hold:*

- *If $\prec$ is* well-founded*, then $\mathscr{H}_\prec(K)$ is non-empty.*
- *Each $H \in \mathscr{H}_\prec(K)$ is non-empty.*

When the preference relation is proper, we can ensure that our inference relations agree with classical entailment whenever the knowledge base is consistent. The next result confirms that both necessary skeptical and necessary credulous inference collapse to classical inference in such cases.

Corollary 1. *Let $\prec \subseteq 2^\Omega \times 2^\Omega$ be a relation, let K be a consistent knowledge base, and let ϕ be a formula. If $\prec$ is compatible with classical consistency, then $K \mid\!\sim^{ns}_{\prec} \phi$ iff $K \mid\!\sim^{nc}_{\prec} \phi$ iff $K \models \phi$.*

For consistent knowledge bases, the possible inference relations align with classical reasoning in a weaker manner. More precisely, they ensure that any inferred formula remains consistent with the original knowledge base.

Corollary 2. *Let $\prec \subseteq 2^\Omega \times 2^\Omega$ be a relation, K be a consistent knowledge base, and ϕ a formula. If $\prec$ is compatible with classical consistency, then $K \mid\!\sim^{ps}_{\prec} \phi$ iff $K \mid\!\sim^{pc}_{\prec} \phi$ iff $K \cup \{\phi\} \not\models \perp$.*

Together, these results establish proper preference relations as a foundation for well-behaved inference. They guarantee the existence of minimal hitting sets and ensure that, in the absence of inconsistency, the framework faithfully reproduces classical reasoning.

In the following, we instantiate different minimality criteria, highlighting their ability in prioritizing relevant interpretations and hitting sets for effective inference.

4 Instantiation of Minimality in Inference Relations

In the following subsections, we will instantiate the general framework from Definition 10 with 4 different approaches for $\prec$ and analyse the properties of the resulting inference relations.

4.1 Inference from Minimal Set Inclusion Hitting Sets

As a starting point, we consider minimizing hitting sets with respect to set inclusion.[3]

Definition 14. Define $\prec_{se} \subseteq 2^\Omega \times 2^\Omega$ via $H \prec_{se} H'$ iff $H \subsetneq H'$, for all $H, H' \subseteq 2^\Omega$.

We illustrate in the following example how set-inclusion minimality can lead to questionable inferences.

Example 3. The knowledge base $K_2 = \{p, q\}$ will have the minimal hitting sets $\mathscr{H}_{\prec_{se}}(K_2) = \{\{pq\}, \{p\bar{q}, \bar{p}q\}\}$ Hence, we obtain the inferences $K_2 \mid\!\sim^{\mathrm{pc}} \neg p$, $K_2 \mid\!\sim^{\mathrm{pc}} \neg q$.

Although intuitive and widely adopted, set-inclusion minimality does not preserve desirable logical behavior. In particular, it fails to ensure compatibility with classical inference, even when the knowledge base is consistent. Consequently, the relation $\prec_{se}$ is not proper.

Proposition 2. *The relation $\prec_{se}$ does not satisfy compatibility with classical consistency.*

This example demonstrates the failure of compatibility with classical consistency under set-inclusion minimality. As a result, the guarantees provided by Corollaries 1, 2 no longer apply. Notably, the inference mechanism permits the derivation of formulas such as $\neg p$ which, when added back to the knowledge base, would reintroduce inconsistency. Because of this, we will not further consider inference relations based on $\prec_{se}$.

4.2 Inference from Cardinality Minimal Hitting Sets

Next, we consider cardinality-based minimality, where preference is given to hitting sets containing the smallest number of interpretations. This approach aims to capture a form of parsimony in model selection. From this point forward, all

[3]The subscript "se" stands for set inclusion.

minimality criteria we consider will be compatible with classical consistency.

Definition 15. Define $\prec_\# \subseteq 2^\Omega \times 2^\Omega$ via $H \prec_\# H'$ iff $|H| < |H'|$, for all $H, H' \subseteq 2^\Omega$.

For convenience, we use $\mathscr{H}_\#(K)$ as an abbreviation for $\mathscr{H}_{\prec_\#}(K)$. Similarly, we abbreviate the corresponding relations $\mid\!\sim^{\text{ns}}_{\prec_\#}, \mid\!\sim^{\text{nc}}_{\prec_\#}, \mid\!\sim^{\text{ps}}_{\prec_\#}$, and $\mid\!\sim^{\text{pc}}_{\prec_\#}$ as $\mid\!\sim^{\text{ns}}_{\#}, \mid\!\sim^{\text{nc}}_{\#}, \mid\!\sim^{\text{ps}}_{\#}$, and $\mid\!\sim^{\text{pc}}_{\#}$, respectively.

To ensure that cardinality-based inference is well-defined and logically aligned with classical reasoning, we verify that this preference relation satisfies the conditions for being proper.

Proposition 3. *The relation* $\prec_\#$ *is* proper.

This implies that all minimal hitting sets under cardinality have the same size, which reinforces the regularity of the selection criterion.

Observation 1. *Let K be a consistent knowledge base. For all $H, H' \in \mathscr{H}_\#(K)$, it holds $|H| = |H'|$.*

Although cardinality-based minimality respects classical consistency, it does not control the internal diversity of interpretations within a hitting set. As a result, the inferred conclusions may still be counterintuitive or misleading, as illustrated next.

Example 4. Consider the knowledge base $K_3 = \{p, \neg p \wedge r, q\}$. It contains the following cardinality minimal hitting sets:

$$H_1 = \{pqr, \bar{p}qr\}, H_2 = \{pqr, \bar{p}\bar{q}r\}, H_3 = \{pq\bar{r}, \bar{p}qr\},$$
$$H_4 = \{pq\bar{r}, \bar{p}\bar{q}r\}, H_5 = \{p\bar{q}\bar{r}, \bar{p}qr\}, \text{ and } H_6 = \{p\bar{q}r, \bar{p}qr\}$$
$$\mathscr{H}_\#(K_3) = \{H_1, H_2, H_3, H_4, H_5, H_6\}$$

Now consider the formulas $p \vee r$, $\neg p$, $q \wedge r$, $\neg q$ and $p \wedge r$. One can observe the following:

1. $K_3 \mid\!\sim^{\text{ns}}_{\#} p \vee r$, since all interpretation in each minimal hitting set in $\mathscr{H}_\#(K_3)$ satisfy $p \vee r$
2. $K_3 \mid\!\sim^{\text{nc}}_{\#} \neg p$, since in every minimal hitting set in $\mathscr{H}_\#(K_3)$ there is an ω that models $\neg p$.
3. $K_3 \mid\!\sim^{\text{ps}}_{\#} q \wedge r$, since both interpretations of H_1 model $q \wedge r$.
4. $K_3 \mid\!\sim^{\text{pc}}_{\#} \neg q$, since there exist an interpretation in a hitting set that model $\neg q$, like the first interpretation in H_5.
5. $K_3 \not\mid\!\sim^{\text{ns}}_{\#} \neg p$, because there is some interpretation (like $\bar{p}qr$ in H_1d) that does not model $\neg p$
6. $K_3 \not\mid\!\sim^{\text{nc}}_{\#} q \wedge r$, because there is some hitting set H_4 where all interpretations does not model $q \wedge r$
7. $K_3 \not\mid\!\sim^{\text{ps}}_{\#} \neg p$ because for every minimal hitting set there will always be some interpretation that satisfies p since it exists as formula in the knowledge base.
8. $K_3 \not\mid\!\sim^{\text{pc}}_{\#} p \wedge r$ since no interpretation satisfies $p \wedge r$

Figure 1 summarizes the entailment of $\mid\!\sim^{\text{ns}}_{\#}, \mid\!\sim^{\text{nc}}_{\#}, \mid\!\sim^{\text{ps}}_{\#}, \mid\!\sim^{\text{pc}}_{\#}$.

Note that the inference $K \mid\!\sim^{\text{pc}}_{\#} \neg q$ is obtained. This outcome is arguably problematic because q is explicitly contained in the knowledge base, yet the cardinality inference framework allows deriving its negation. Furthermore, the formula q is not entailed under necessary skeptical inference, despite being unrelated to the inconsistency.

Formula	$K_3 \mid\!\sim^{\text{ns}}_{\#}$	$K_3 \mid\!\sim^{\text{nc}}_{\#}$	$K_3 \mid\!\sim^{\text{ps}}_{\#}$	$K_3 \mid\!\sim^{\text{pc}}_{\#}$
$p \vee r$	✓	✓	✓	✓
$\neg p$	✗	✓	✗	✓
$q \wedge r$	✗	✗	✓	✓
$\neg q$	✗	✗	✗	✓
$p \wedge r$	✗	✗	✗	✗

Figure 1: Entailment of $\mid\!\sim^{\text{ns}}_{\#}$, $\mid\!\sim^{\text{nc}}_{\#}$, $\mid\!\sim^{\text{ps}}_{\#}$, and $\mid\!\sim^{\text{pc}}_{\#}$ for the knowledge base $K_3 = \{p, \neg p \wedge r, q\}$ and the formulas from Example 4.

4.3 Inference from Minimal Conflicting Atoms Hitting Sets

Cardinality-based minimality minimizes the number of interpretations but not the differences between them. As a result, interpretations within a minimal hitting set may disagree on many atoms, which can lead to problematic inferences. To mitigate this, we introduce a new minimality criterion that penalizes such internal disagreement by minimizing the number of conflicting truth assignments across interpretations.

Definition 16. Let $H \subseteq \Omega$ be a set of interpretations. Define the *set of conflicting atoms* $CA(H)$ of H as

$$CA(H) = \{a \mid \exists \omega, \omega' \in H \text{ such that } \omega(a) \neq \omega'(a)\}$$

In other words $a \in CA(H)$ iff there are two interpretations in H that differ on their truth values of a. Using $CA(H)$ leads to a preference relation that ranks hitting sets by the set of atoms on which interpretations disagree.

Definition 17. Define[4] $\prec_{ca} \subseteq 2^\Omega \times 2^\Omega$ via $H \prec_{ca} H'$ iff $CA(H) \subsetneq CA(H')$, for any $H, H' \in 2^\Omega$.

To illustrate the difference between cardinality and conflicting-atoms minimality, we revisit the earlier knowledge base K_3 from Example 4. This comparison highlights how conflicting-atoms minimization avoids some of the unintuitive conclusions permitted by the cardinality criterion.

Example 5. In comparison to minimal cardinality hitting sets, the variation between interpretations in hitting sets decreases drastically. The minimal conflicting atoms hitting sets for K_3 are:

$$\mathscr{H}_{ca}(K_3) = \{\{pqr, \bar{p}qr\}\}$$

and therefore $\neg q$ is not entailed by any of the inference relations, while q is inferred by all of them.

As with previous criteria, we check whether the conflicting-atoms preference relation satisfies the requirements of being proper — namely, whether it guarantees minimal hitting sets and aligns with classical inference in the consistent case.

Proposition 4. *The relation* $\prec_{ca}$ *is* proper.

The property established above guarantees that for consistent knowledge bases, minimal hitting sets under conflicting atoms minimality criterion are singletons and therefore behave consistently with classical inference.

[4]The subscript "ca" stands for *conflicting atoms*.

4.4 Inference from Minimal Hamming Distance Hitting sets

We now consider another way to measure variation among interpretations in a hitting set, the Hamming distance (Also known as Dalal distance). Unlike $\prec_{ca}$, which considers the set of atoms where interpretations differ, Hamming distance quantifies the degree of such differences by summing pairwise discrepancies.

Definition 18. Define the Hamming distance $d : \Omega \times \Omega \to \mathbb{N}$ as

$$d(\omega, \omega') = |\{a \in \mathsf{At} | \omega(a) \neq \omega'(a)\}|$$

for all $\omega, \omega' \in \Omega(\mathsf{At})$. For a set $X \subseteq \Omega$ we write

$$D(X) = \sum_{\omega_1, \omega_2 \in X} d(\omega_1, \omega_2)$$

We extend the Hamming distance to sets of interpretations by summing all pairwise distances. This aggregate measure reflects the overall dissimilarity within a hitting set and underpins a preference for sets with greater internal coherence.

Definition 19. Define[5] $\prec_d \subseteq 2^\Omega \times 2^\Omega$ via $H \prec_d H'$ iff $D(H) < D(H')$, for any $H, H' \subseteq 2^\Omega$.

This ordering selects hitting sets with minimal total pairwise disagreement across interpretations. Unlike $\prec_{ca}$, the Hamming distance takes into account the degree of disagreement, not just its presence. As the next example shows, the two criteria are incomparable, i. e., neither strictly subsumes the other.

Example 6. Let $K = \{p \wedge q \wedge r, p \to q \wedge \neg r, \neg q \wedge \neg r\}$. Consider the following hitting sets:

$$H_1 = \{pqr, \bar{p}\bar{q}\bar{r}\} \qquad H_2 = \{pqr, p\bar{q}\bar{r}, pq\bar{r}\}$$
$$D(H_1) = 6 \qquad D(H_2) = 8$$
$$CA(H_1) = \{p, q, r\} \qquad CA(H_2) = \{q, r\}$$

Notice in Example 6 that $H_1 \in \mathscr{H}_d(K)$ and $H_1 \notin \mathscr{H}_{ca}(K)$. Hence, $\mathscr{H}_d(K) \not\subseteq \mathscr{H}_{ca}(K)$. Similarly for the other direction, $H_2 \in \mathscr{H}_{ca}(K)$ and $H_2 \notin \mathscr{H}_d(K)$

As with earlier criteria, we confirm that the Hamming distance ordering is proper.

Proposition 5. *The relation $\prec_d$ is* proper.

To conclude, Hamming-distance minimality provides another perspective on disagreements between interpretations by quantifying the total number of differing atomic assignments. Like the cardinality and conflicting-atoms approaches, it satisfies the basic properties of well-foundedness and compatibility with classical consistency. This ensures that the resulting inference relations are proper and maintain classical behavior for consistent knowledge bases, while enabling a refined form of inference under inconsistency.

[5]The subscript "d" stands for *Dalal*.

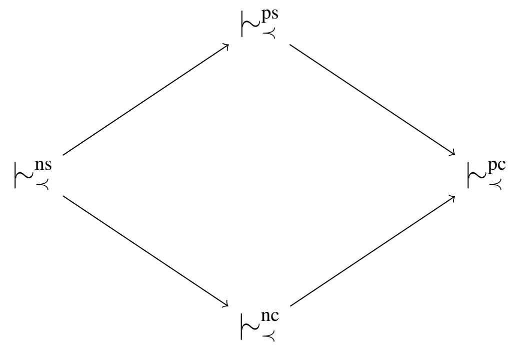

Figure 2: Relationship among $\mid\!\sim^{ns}_{\prec}$, $\mid\!\sim^{nc}_{\prec}$, $\mid\!\sim^{ps}_{\prec}$, $\mid\!\sim^{pc}_{\prec}$ for an arbitrary $\prec$, given that $\prec$ is well-founded. Arrow directions indicate the antecedent to consequent relation.

5 Relationships Among Inference Relations

In this section, we analyze how the four hitting-based inference relations—$\mid\!\sim^{ns}_{\prec}$, $\mid\!\sim^{nc}_{\prec}$, $\mid\!\sim^{ps}_{\prec}$, and $\mid\!\sim^{pc}_{\prec}$—relate to each other and to existing approaches, including inference from maximal consistent subsets and Priest's 3-valued logic. Understanding these relationships helps position our framework within the broader landscape of nonmonotonic reasoning and highlights the inferential strengths and limitations of each mode.

5.1 Relationships Among Hitting Sets Inference Relations

We begin by examining the relationships among the four inference types within our framework. Specifically, we identify which inference relations are strictly stronger than others and where no such containment holds. This comparative analysis clarifies how each inference mode varies in strength and independence.

We start with establishing general inclusion relationships among the four inference relations, abstracting away from any specific choice of ordering $\prec$.

Proposition 6. *Let K be a knowledge base, and let $\prec \subseteq 2^\Omega \times 2^\Omega$ be a well-founded relation. The following relationships among $\mid\!\sim^{ns}_{\prec}$, $\mid\!\sim^{nc}_{\prec}$, $\mid\!\sim^{ps}_{\prec}$, and $\mid\!\sim^{pc}_{\prec}$ hold:*

1. *If $K \mid\!\sim^{ns}_{\prec} \phi$, then $K \mid\!\sim^{nc}_{\prec} \phi$*
2. *If $K \mid\!\sim^{ns}_{\prec} \phi$, then $K \mid\!\sim^{ps}_{\prec} \phi$*
3. *If $K \mid\!\sim^{nc}_{\prec} \phi$, then $K \mid\!\sim^{pc}_{\prec} \phi$*
4. *If $K \mid\!\sim^{ps}_{\prec} \phi$, then $K \mid\!\sim^{pc}_{\prec} \phi$*

The next lemma captures the duality between necessary and possible inference, showing how negation interrelates the two modes.

Lemma 1. *Let K be a knowledge base, ϕ a formula, and $\prec \subseteq 2^\Omega \times 2^\Omega$. Then*

- *$K \mid\!\sim^{ps}_{\prec} \phi$ iff $K \not\mid\!\sim^{nc}_{\prec} \neg\phi$.*
- *$K \mid\!\sim^{pc}_{\prec} \phi$ iff $K \not\mid\!\sim^{ns}_{\prec} \neg\phi$.*

We now present an example that illustrates how different inference relations can yield distinct conclusions, emphasizing their incomparability.

Example 7. Let $K = \{p \wedge q \wedge r \wedge s \wedge t, \neg(p \vee q \vee r \vee s \vee t) \vee (\neg p \wedge q \wedge r \wedge s \wedge t), \neg(p \vee q \vee r \vee s \vee t) \vee (p \wedge \neg q \wedge r \wedge s \wedge t)\}$ be a knowledge base. It will have the following hitting sets:

$$H_1 = \{pqrst, \bar{p}qrst, p\bar{q}rst\} \qquad H_2 = \{pqrst, \bar{p}\bar{q}\bar{r}\bar{s}\bar{t}\}$$
$$\mathscr{H}_{ca}(K) = \mathscr{H}_d(K) = \{H_1\} \qquad \mathscr{H}_{\#}(K) = \{H_2\}$$

We say that two inference relations $\mid\!\sim_x$ and $\mid\!\sim_y$ are incomparable if there are K, K' and ϕ, ϕ' such that

$$K \mid\!\sim_x \phi \quad K \not\mid\!\sim_y \phi \quad K' \not\mid\!\sim_x \phi' \quad K' \mid\!\sim_y \phi'$$

The following proposition shows that inference relations considered in this paper, which are not covered by Proposition 6 are incomparable. That is, for distinct minimality relations, neither inference relation uniformly entails the other.

Proposition 7. *For all $I \in \{ns, nc, ps, pc\}$, and all $(x,y) \in \{(\#, ca), (\#, d), (d, ca)\}$, the relations $\mid\!\sim_x^I$ and $\mid\!\sim_y^I$ are incomparable.*

5.2 Relationship with Maximal Consistent Subsets Inference

We now compare our framework to inference based on maximal consistent subsets ($\mid\!\sim_i^{\mathrm{mc}}$ and $\mid\!\sim_w^{\mathrm{mc}}$), as introduced by Rescher and Manor (Rescher and Manor 1970). In particular, we focus on how cardinality-minimal hitting sets relate to the maximally consistent subsets of a knowledge base. The intuition is that every maximally consistent subset corresponds to at least one interpretation in a cardinality-minimal hitting set. This correspondence is made precise in the following result.

Proposition 8. *Let K be a knowledge base. Then*

$$\forall M \in MCS(K) : \exists H \in \mathscr{H}_{\#}(K) : \exists \omega \in H : \omega \models M$$

Thus, for each maximally consistent subset of the knowledge base, some interpretation in a cardinality-minimal hitting set satisfies it. This leads to the following theorem relating the respective inference notions.

Theorem 1. *It holds:*

1. $\mid\!\sim_{\#}^{ns} \subsetneq \mid\!\sim_i^{mc}$
2. $\mid\!\sim_w^{mc} \subsetneq \mid\!\sim_{\#}^{pc}$

This result situates our inference framework between the classical extremes. Necessary skeptical inference based on cardinality is more selective than Rescher and Manor's inevitable consequence, while possible credulous inference is more permissive than their weak consequence.

5.3 Relationship with 3-valued Inference

In this section, we consider a variant of 3-valued inference that selects models which minimize the use of the indeterminate truth value (both), as showed in Definition 6. We then examine how this semantics corresponds to inference in our framework based on hitting sets that minimize conflicting atoms.

To connect our framework with 3-valued reasoning, we define a transformation from a hitting set of classical interpretations to a single 3-valued interpretation. This construction captures the agreement across interpretations directly, while encoding any disagreement as the value *both*.

Definition 20. Let K be a knowledge base and $H \in \mathscr{H}(K)$ a hitting set of K. The *H-induced 3-valued interpretation* υ_H is the 3-valued interpretation υ_H defined as

$$\upsilon_H(a) = \begin{cases} true & \text{if } \forall \omega \in H : \omega(a) = \mathsf{true} \\ false & \text{if } \forall \omega \in H : \omega(a) = \mathsf{false} \\ both & \text{otherwise} \end{cases}$$

for all $a \in \mathsf{At}$.

By definition, any atom assigned *both* in the resulting 3-valued interpretation corresponds exactly to an atom with conflicting truth values across the hitting set.

Observation 2. *For any knowledge base K, and an H-induced 3-valued interpretation υ_H from some $H \in \mathscr{H}(K)$: it holds $CA(H) = \upsilon_H^{-1}(both)$.*

Different 3-valued interpretations may provide varying levels of information, depending on how many atoms they assign determinately (as *true* or *false*) versus ambiguously (as *both*). To formalize this, we introduce an information ordering that compares interpretations based on their determinacy.

Definition 21. Let υ, υ' be a 3-valued interpretation. Define the *information order* $\sqsubseteq$ via $\upsilon' \sqsubseteq \upsilon$ iff $\upsilon'(true)^{-1} \subseteq \upsilon(true)^{-1}$ and $\upsilon'(false)^{-1} \subseteq \upsilon(false)^{-1}$

Note that classical (two-valued) interpretations are maximal elements in this ordering, since they assign a definite truth value to every atom. Conversely, the interpretation mapping every atom to *both* is the least informative element.

Proposition 9. *Let ϕ be a formula, and υ, υ' be 3-valued interpretations. For every $\upsilon' \sqsubseteq \upsilon$ it holds that*

1. *If $\upsilon(\phi) = true$, then $\upsilon'(\phi) \in \{true, both\}$*
2. *If $\upsilon(\phi) = both$, then $\upsilon'(\phi) = both$*
3. *If $\upsilon(\phi) = false$, then $\upsilon'(\phi) \in \{false, both\}$*

The next result lifts the Proposition 9 from individual formulas to entire knowledge bases, showing that every H-induced 3-valued interpretation satisfies the original knowledge base.

Proposition 10. *Let K be a knowledge base. Then for all $H \in \mathscr{H}(K)$, $\upsilon_H \models^3 K$.*

Conversely, to complete the correspondence, we show the opposite direction of having a hitting set that can satisfy every formula in knowledge base K given that there exists a 3-valued interpretation $\upsilon \models^3 K$. We define next we introduce a construction that maps a 3-valued interpretation to a set of two-valued interpretations. This mapping reflects the idea that a 3-valued interpretation can be viewed as an abstraction over multiple classical interpretations. Specifically, we define the notion of υ-induced hitting sets, which extract from a 3-valued interpretation those two-valued interpretations that agree with it on all determinate (non-'*both*') assignments.

Definition 22. Let υ be a 3-valued interpretation. Define the *υ-induced hitting set H_υ* via

$$H_\upsilon = \{\omega \in \Omega(\mathsf{At}) \mid \upsilon \sqsubseteq \omega\}$$

3-valued induced hitting set H_υ corresponding to a 3-valued interpretation υ preserves all deterministic truth assignments while introducing variation only for atoms assigned *both*. Specifically, for every atom a where $\upsilon(a) = both$, the hitting set must include at least two interpretations $\omega_1, \omega_2 \in H_\upsilon$ such that $\omega_1(a) = \mathsf{true}$ and $\omega_2(a) = \mathsf{false}$. As a result, the following is observed:

Observation 3. *For any knowledge base K, and an υ-*induced hitting set H_υ *from some $\upsilon \models K$. It holds that $CA(H) = \upsilon_H^{-1}(both)$.*

Note that the relation $\models^3$ is invariant under certain classical equivalences, in particular De Morgan's law and distributivity rules.

Example 8. Given an interpretation $\upsilon \models^3 \neg(p \wedge q)$. Applying De Morgan's law to $\neg(p \wedge q)$ yields $\neg p \vee \neg q$. It is the case $\upsilon \models^3 \neg(p \wedge q)$ iff $\upsilon \models^3 \neg p \vee \neg q$.

Similarly, $\models^3$ is preserved under transformations such as associativity and distributivity of $\vee$ and $\wedge$, as well as the rewriting of implications into disjunctions. This invariance justifies assuming, without loss of generality, that all formulas are given in conjunctive normal form (CNF).

To ensure that the subsequent results apply uniformly, we restrict attention to knowledge bases in CNF. The following proposition shows that this restriction does not affect the generality of our results.

Proposition 11. *Let υ be a 3-valued interpretation and let ϕ be a formula. It holds that $\upsilon \models^3 \phi$ iff $\upsilon \models^3 \phi'$, where ϕ' is an equivalent conjunctive normal form of ϕ.*

As a first step toward establishing the equivalence between conflicting-atoms-based inference and 3-valued semantics, we consider the case of disjunctive formulas. The following lemma shows that if a 3-valued interpretation satisfies such a formula, then there exists a corresponding 2-valued interpretation—within the υ-induced hitting set—that also satisfies the formula. This forms the basis for lifting satisfaction from 3-valued models to their classical refinements.

Lemma 2. *Let ϕ be a clause and υ a 3-valued interpretation with $\upsilon \models^3 \phi$. Then there is $\omega \in H_\upsilon$ with $\omega \models \phi$.*

We now extend the previous result to full knowledge bases. The proposition shows that every 3-valued model of a knowledge base induces a hitting set of classical interpretations that together satisfy all formulas in the base.

Proposition 12. *Let K be a knowledge base, where $K = \{C_1, \dots, C_n\}$ and $C_1, \dots, C_n$ are clauses, and υ a 3-valued interpretation for K. If $\upsilon \models^3 K$ then $H_\upsilon \in \mathscr{H}(K)$.*

We now extend the previous result to full knowledge bases. The proposition shows that every 3-valued model of a knowledge base induces a hitting set of classical interpretations that together satisfy all formulas in the base.

Proposition 13. *Let K be a knowledge base, where $K = \{C_1, \dots, C_n\}$ and $C_1, \dots, C_n$ are clauses, and $\upsilon \in MinMod^3(K)$, then $H_\upsilon \in \mathscr{H}_{ca}(K)$.*

We now establish the converse: every minimal hitting set under the conflicting-atoms preference induces a minimal 3-valued model. This bidirectional correspondence fully aligns the two approaches.

Proposition 14. *Let K be a knowledge base where $K = \{C_1, \dots, C_n\}$ and $C_1, \dots, C_n$ are clauses. For each $H \in \mathscr{H}_{ca}(K)$, then $\upsilon_H \in MinMod^3(K)$.*

In the following, we finally complete the equivalence between minimal conflicting-atoms hitting sets coincides with inference in Priest's 3-valued logic, provided the knowledge base is in conjunctive normal form. This result connects our semantic hitting set-based approach to a well-established non-classical logic for reasoning under inconsistency.

Theorem 2. *Let K be a knowledge base such that $K = \{C_1, \dots, C_n\}$ and $C_1, \dots, C_n$ are clauses. $\Phi {\mid\!\sim}^3 \phi$ iff $\Phi {\mid\!\sim}^{nc}_{ca} \phi$*

This correspondence demonstrates that the conflicting-atoms minimality criterion not only aligns with intuitive notions of disagreement in interpretations but also captures the semantics of an established many-valued logic. By bridging our hitting set-based framework with Priest's 3-valued logic, we provide a principled foundation for reasoning under inconsistency that combines the clarity of classical models with the flexibility of non-classical valuation.

6 KLM Properties for Inference with Minimal Conflicting Atoms

In this section, we examine how our hitting set-based inference relations behave when knowledge bases are extended. This is crucial for assessing whether the framework satisfies desirable nonmonotonic reasoning principles, such as those captured by the KLM postulates. We begin by analyzing how minimal hitting sets evolve when a consistent formula is added to a knowledge base.

Proposition 15. *Let K be a knowledge base, let ϕ be a formula, and let $\prec \subseteq 2^\Omega \times 2^\Omega$ be a relation. If there is a hitting set $H \in \mathscr{H}_\prec(K)$ with $\omega \in H$ such that $\omega \models \phi$. Then, there is no $H' \in \mathscr{H}(K \cup \{\phi\})$ with $H' \prec H$, and $H \in \mathscr{H}_\prec(K \cup \{\phi\})$.*

To capture this behavior more generally, we introduce a property of preference relations called regularity. Intuitively, regularity ensures that if a formula is already satisfied within a minimal hitting set, then adding that formula to the knowledge base should not generate new, strictly preferred hitting sets.

Definition 23. Given a knowledge base K, and a consistent formula ϕ. A relation $\prec \subseteq 2^\Omega \times 2^\Omega$ is called *regular* if there exists some $H \in \mathscr{H}_\prec(K)$ with $\omega \in H$ such that $\omega \models \phi$, then $\mathscr{H}_\prec(K \cup \{\phi\}) \subseteq \mathscr{H}_\prec(K)$ holds.

Regularity requires a form of structural persistence in the preference relation: if a formula ψ holds in some minimal hitting set of a knowledge base K, then this hitting set must remain among the minimal ones after the formula is added. In other words, the existence of such a hitting set

$H \in \mathscr{H}_{\prec}(K)$ ensures that all hitting sets $\mathscr{H}_{\prec}(K \cup \{\psi\})$ must satisfy the same preference constraints.

We now verify that the main preference relations considered in this paper—cardinality, conflicting atoms, and Hamming distance—all satisfy the regularity condition.

Proposition 16. *The relations $\prec_{\#}$, $\prec_{ca}$, and $\prec_{d}$ are* regular.

We recall that regularity ensures a form of persistence: if a formula holds in some minimal hitting set of a knowledge base, then the inclusion of that formula in the knowledge base cannot introduce new strictly preferred hitting sets. The following propositions formalize this behavior.

Proposition 17. *Let K be a knowledge base, let ϕ be a consistent formula, and let $\prec \subseteq 2^{\Omega} \times 2^{\Omega}$ be a regular relation. If a minimal hitting set $H \in \mathscr{H}_{\prec}(K)$ contains a model of ϕ, then $\mathscr{H}_{\prec}(K \cup \{\phi\}) \subseteq \mathscr{H}_{\prec}(K)$.*

Next, we define the notion of hitting set equivalence, which characterizes when two knowledge bases yield the same collection of minimal hitting sets under a given preference relation.

Definition 24. Let K, K' be knowledge bases and $\prec \subseteq 2^{\Omega} \times 2^{\Omega}$. We say the knowledge bases K and K' are hitting set equivalent, or $K \equiv^{H}_{\prec} K'$ if $\mathscr{H}_{\prec}(K) = \mathscr{H}_{\prec}(K')$

With the structural properties of preference relations in place, we now assess how our inference framework aligns with well-known nonmonotonic reasoning principles — mainly System P, as introduced by Kraus, Lehmann, and Magidor (Kraus, Lehmann, and Magidor 1990). The system encompasses several desirable principles for nonmonotonic reasoning, which we briefly recall below.

The following properties are mostly those of System P, but note that knowledge bases may not contain inconsistent formulas (yet the knowledge bases may be inconsistent at all):

$\{\phi\} \mid\!\sim \phi$. (Ref)

If $\phi \models \psi$ and $K \mid\!\sim \phi$, then $K \mid\!\sim \psi$. (RW)

If $K \equiv^{H} K'$ and $K \mid\!\sim \phi$, then $K' \mid\!\sim \phi$. (LLE)

If ψ, $K \cup \{\psi\} \mid\!\sim \phi$ and $K \mid\!\sim \psi$, then $K \mid\!\sim \phi$. (Cut)

If ψ, $K \mid\!\sim \psi$ and $K \mid\!\sim \phi$, then $K \cup \{\psi\} \mid\!\sim \phi$. (CM)

If $\phi \mid\!\sim \chi$ and $\psi \mid\!\sim \chi$, then $\{\phi \vee \psi\} \mid\!\sim \chi$. (Or)

If $K \mid\!\sim \phi$ and $K \not\mid\!\sim \psi$, then $K \cup \{\psi\} \mid\!\sim \phi$. (RM)

In the area of non-monotonic logic, System P is known as the conservative core of non-monotonic reasoning, as it captures preferential entailment (Kraus, Lehmann, and Magidor 1990); respectively, all properties of System P, except for the property Or, are known due to Gabbay as basic properties of non-monotonic inference relations (Gabbay 1984).

The following theorem summarizes which of the KLM postulates are satisfied by our inference relations, assuming the underlying preference relation is regular and proper. The result shows that the framework adheres to nearly all key principles of nonmonotonic reasoning. In particular, necessary inference relations satisfy all properties of System P, including cautious monotony, while possible inference relations fail this one condition, reflecting their more permissive nature.

	$\mid\!\sim^{\text{ns}}_{\prec}$	$\mid\!\sim^{\text{nc}}_{\prec}$	$\mid\!\sim^{\text{ps}}_{\prec}$	$\mid\!\sim^{\text{pc}}_{\prec}$
Ref	✓	✓	✓	✓
RW	✓	✓	✓	✓
LLE	✓	✓	✓	✓
Cut	✓	✓	✓	✓
CM	✓	✓	✗	✗
Or	✓	✓	✓	✓
RM	✓	✓	✓	✓

Figure 3: Overview of the satisfaction (✓), respectively violation (✗), of inference properties by $\mid\!\sim^{\text{ns}}_{\prec}$, $\mid\!\sim^{\text{nc}}_{\prec}$, $\mid\!\sim^{\text{ps}}_{\prec}$, and $\mid\!\sim^{\text{pc}}_{\prec}$. For any regular proper relation $\prec \subseteq 2^{\Omega} \times 2^{\Omega}$

Theorem 3. *Let $\prec \subseteq 2^{\Omega} \times 2^{\Omega}$ be a proper regular relation.*

- *$\mid\!\sim^{ns}_{\prec}$, $\mid\!\sim^{nc}_{\prec}$, $\mid\!\sim^{ps}_{\prec}$ and $\mid\!\sim^{pc}_{\prec}$ satisfy Ref, RW, LLE, Cut, Or, and RM.*
- *$\mid\!\sim^{ns}_{\prec}$, $\mid\!\sim^{nc}_{\prec}$ satisfy CM*
- *$\mid\!\sim^{ps}_{\prec}$, $\mid\!\sim^{pc}_{\prec}$ violate CM*

This analysis confirms that our framework supports robust nonmonotonic reasoning when instantiated with regular proper preference relations. The necessary inference modes, in particular, satisfy all of the core KLM postulates, including cautious monotony, thereby aligning with the foundational principles of preferential entailment. While the possible inference modes trade some logical discipline for flexibility, their behavior remains well-characterized within the framework. Overall, these results underscore the expressive power and formal soundness of hitting set-based inference under appropriately constrained preference relations.

7 Concluding Remarks

We have introduced a general framework for inconsistency-tolerant reasoning grounded in the notion of hitting sets of classical interpretations. By defining inference relations based on minimal hitting sets under a range of preference criteria, our approach generalizes and extends classical MCS-based reasoning to a semantic, model-oriented setting. We characterized four distinct inference modes and systematically analyzed their logical properties. Our comparisons with maximal consistent subset inference, Priest's 3-valued logic, and the KLM postulates demonstrate both the expressive flexibility and the formal rigor of the framework. These results provide a foundation for principled reasoning in inconsistent environments. Some related work on 3-valued hitting set is of Coste-Marquis and Marquis 2008, where three-valued models are utilized to derive consistent two-valued models by systematically forgetting inconsistent information. Future work should examine the computational complexity of these inference relations, explore their deeper connections to other semantic nonmonotonic formalisms, and investigate applications in belief revision (Alchourron, Gärdenfors, and Makinson 1985), formal argumentation (Gabbay, Giacomin, and Simari 2024; Atkinson et al. 2017), and logics that integrate both syntactic and semantic criteria for minimality.

Acknowledgements

The research reported here was partially supported by the Deutsche Forschungsgemeinschaft (grant 424710479, "Predictive and Interactive Management of Potential Inconsistencies in Business Rules", MIB).

References

Alchourron, C.; Gärdenfors, P.; and Makinson, D. 1985. On the logic of theory change: Partial meet contraction and revision functions. *J. Symb. Log.* 50:510–530.

Atkinson, K.; Baroni, P.; Giacomin, M.; Hunter, A.; Prakken, H.; Reed, C.; Simari, G. R.; Thimm, M.; and Villata, S. 2017. Towards artificial argumentation. *AI Magazine* 38(3):25–36.

Baader, F.; Kriegel, F.; Nuradiansyah, A.; and Peñaloza, R. 2018. Making repairs in description logics more gentle. In Thielscher, M.; Toni, F.; and Wolter, F., eds., *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona, 30 October - 2 November 2018*, 319–328. AAAI Press.

Benferhat, S.; Dubois, D.; and Prade, H. 1997. Some syntactic approaches to the handling of inconsistent knowledge bases: A comparative study part 1: The flat case. *Studia Logica* 58.

Coste-Marquis, S., and Marquis, P. 2008. Recovering consistency by forgetting inconsistency. In Hölldobler, S.; Lutz, C.; and Wansing, H., eds., *Logics in Artificial Intelligence*, 113–125. Berlin, Heidelberg: Springer Berlin Heidelberg.

Fermé, E., and Wassermann, R. 2018. On the logic of theory change: iteration of expansion. *J. Braz. Comp. Soc.* 24(1):8:1–8:9.

Gabbay, D.; Giacomin, M.; and Simari, G. 2024. *Handbook of Formal Argumentation. Volume 3*. College Publications.

Gabbay, D. M. 1984. Theoretical foundations for non-monotonic reasoning in expert systems. In Apt, K. R., ed., *Logics and Models of Concurrent Systems - Conference proceedings, Colle-sur-Loup (near Nice), France, 8-19 October 1984*, volume 13 of *NATO ASI Series*, 439–457. Springer.

Harmelen, F.; Lifschitz, V.; and Porter, B. 2007. The handbook of knowledge representation. *Elsevier Science San Diego, USA* 1034.

Kalyanpur, A.; Parsia, B.; Sirin, E.; and Grau, B. C. 2006. Repairing unsatisfiable concepts in OWL ontologies. In Sure, Y., and Domingue, J., eds., *The Semantic Web: Research and Applications, 3rd European Semantic Web Conference, ESWC 2006, Budva, Montenegro, June 11-14, 2006, Proceedings*, volume 4011 of *Lecture Notes in Computer Science*, 170–184. Springer.

Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artif. Intell.* 44(1-2):167–207.

Lang, J., and Marquis, P. 2010. Reasoning under inconsistency: A forgetting-based approach. *Artificial Intelligence* 174(12):799–823.

Priest, G. 1979. The logic of paradox. *Journal of Philosophical Logic* 8(1):219–241.

Rescher, N., and Manor, R. 1970. On inference from inconsistent premisses. *Theory and Decision* 1:179–217.

Thimm, M. 2014. Towards large-scale inconsistency measurement. In *Proceedings of the 37th German Conference on Artificial Intelligence (KI'14)*, 195–206.

Thimm, M. 2016. Stream-based inconsistency measurement. *International Journal of Approximate Reasoning* 68:68–87.

Thimm, M. 2018. On the evaluation of inconsistency measures. In Grant, J., and Martinez, M. V., eds., *Measuring Inconsistency in Information*, volume 73 of *Studies in Logic*. College Publications.