Deontic Reasoning Based On Inconsistency Measures

Ofer Arieli[1] , **Kees van Berkel**[2] , **Badran Raddaoui**[3] , **Christian Straßer**[4]
[1]School of Computer Science, Tel-Aviv Academic College, Israel
[2] Institute for Logic and Computation, TU Wien, Austria
[3]SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, France
[4]Institute for Philosophy II, Ruhr Universität Bochum, Germany

Abstract

Conflicts are inherent to normative systems. In this paper, we explore a novel approach to normative reasoning by quantifying the amount of conflicts within normative systems. We refine the idea from classical logic, according to which a formula is a consequence of a knowledge base in case its negation renders the knowledge base inconsistent. In our approach, whether a formula is a logical consequence depends, for instance, on its negation's marginal contribution to the inconsistency of the given knowledge base. Accordingly, various inconsistency measures and corresponding (nonmonotonic and paraconsistent) normative entailment relations are analyzed relative to a number of logical properties. To illustrate our approach, we adopt Input/Output logic, a renowned formalism in deontic logic, specifically designed for defeasible normative reasoning. As an application, the resulting entailment relations provide recommendations to agents for minimizing norm conflicts, and may be incorporated in a number of implementations (like the Tweety libraries and the LogiKey framework) by involving inconsistency measurements in normative reasoning.

1 Introduction

Normative systems fulfil an essential role in various aspects of life (e.g., ethics, law, business protocols, AI) and one of their central aims is to align agents' conduct in order to attain specific goals (Chopra, van der Torre, and Verhagen 2018). With the increasing demand for compliant autonomous systems, such alignment goals are also paramount to knowledge representation (KR) and AI (Gabriel 2020). By their very nature, however, normative systems are sensitive to conflicts (Nute 1997). Over the past decades, a range of mechanisms has been developed to effectively reason with normative systems in spite of the presence of conflicts, aiming at the re-establishment of consistency.

The presence of conflicts itself, however, provides unique insight into the normative code and factual context at hand. An immediate question, therefore, is what information can be inferred from the type and amount of conflicts within a normative code? For instance, one may assess the degree of conflicts of a normative system or the marginal contribution that each individual norm makes to render the system inconsistent. Surprisingly, this topic has not yet been addressed in deontic logic, the field of formal normative reasoning. In this paper, we address this problem by analyzing notions of *inconsistency measures* and their induced nonmonotonic consequence relations. To the best of our knowledge, this is the first thorough study of reasoning with quantified inconsistencies in normative settings.

Managing conflicting information poses significant challenges for its representation and the respective conclusion making (Martinez et al. 2013). To our benefit, measuring inconsistencies is, in fact, a well-accepted research topic in KR, offering methodologies for quantifying the extent of contradiction, thereby facilitating a deeper understanding of the principal sources of conflicts in knowledge bases (Hunter and Konieczny 2010; Grant and Martinez 2018), and more recently in databases (Livshits et al. 2021; Parisi and Grant 2023). However, the known inconsistency measures are not directly applicable to the context of normative systems for two reasons: (i) norms usually require different conflict handling mechanisms, and (ii) normative knowledge bases often employ richer languages (see (Gabbay et al. 2013)).

In order to address these challenges, we reformulate various inconsistency measures from the literature to apply to a prominent formalism for defeasible normative reasoning, namely, *Input/Output (I/O) logic* (Makinson and van der Torre 2001). Then, we employ these inconsistency measures to establish novel entailment relations that extend those of a well-established class of I/O logics. Intuitively, our entailment relations are based on the idea that a formula is a logical consequence if the addition of its negation to the normative knowledge base would increase its inconsistency. Both the inconsistency measures and the entailment relations are evaluated throughout the paper with respect to a series of rationality postulates and desired properties. Our study reveals that while the measure-based approach allows for a significant number of combinations of inconsistency measures and entailment relations, one has to carefully choose among them, as several combinations lead to non-intuitive inferences. Luckily, there are some combinations with appealing characteristics. In this paper, we consider in greater details one such family of entailment relations: those that are based on the inconsistency measure $\mathcal{I}_{\mathsf{MRN}}$, defined by *minimal correction norm sets*.

In a broader perspective, we note that the entailment relations introduced in this paper can be used to recommend conduct to agents. Recommendations, sometimes referred to as weak obligations (Lellmann, Gulisano, and Ciabattoni 2021), not only constitute instructions for achieving specific outcomes, they aim to minimize negative side-effects compared to the alternative action (Royce Sadler 1984). Side-effects are, in this context, norm conflicts and violations.

Our study has also strong links to other areas in knowledge representation and reasoning, such as nonmonotonic logics, belief revision, and causal reasoning. The underlying formalisms may be incorporated in applications involving inconsistency measurements and normative reasoning, e.g., the Tweety libraries for logical aspects of AI and KR (Thimm 2014), and the LogiKey framework (Benzmüller, Parent, and van der Torre 2020), in the context of which I/O logic and other deontic logics have been implemented in the generic automated theorem prover HOL/Isabelle.

The rest of the paper is organized as follows: In Section 2, we introduce normative systems and consider some basic notions concerning norm inconsistency. In Section 3, we employ these notions to extend various existing inconsistency measures to the normative setting and investigate the logical properties of these measures. In Section 4, we define several normative entailment relations based on these inconsistency measures, and we study their logical properties in Section 5. Related work and future research are reported on in Section 6. Due to space restrictions, we only provide some of the proofs and counterexamples.

2 Normative Systems

We start with some preliminaries concerning the basic normative systems and their consequence relations (Section 2.1), and important notions for evaluating (and reasoning with) inconsistencies in normative knowledge bases (Section 2.2).

2.1 Input/Output Logic

Our formalism is based on *Input/Output (I/O) logic* (Makinson and van der Torre 2000; 2001), a highly versatile general defeasible reasoning paradigm that, over the past decades, has seen a wide variety of applications in the field of knowledge representation and reasoning. Applications range from causal, epistemic, and legal reasoning (Bochman 2014; Ciabattoni, Parent, and Sartor 2021) to complexity and automated deduction results (Ciabattoni and Rozplokhas 2023), additionally showing strong connections to Reiter's default logic (Parent 2011). The main merit and focus point of I/O logic lies in its application to normative reasoning (Parent and van der Torre 2013).

Definition 1. Let $\mathcal{L}$ be a propositional language containing a countable set of propositional variables, and the propositional constants $\top$ and $\bot$, representing truth and falsity. A **normative system**, or **normative knowledge base**, is a triple $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$ consisting of sets of facts $\mathcal{F} \subseteq \mathcal{L}$, constraints $\mathcal{C} \subseteq \mathcal{L}$, and norms $\mathcal{N} \subseteq \{(\varphi, \psi) \mid \varphi, \psi \in \mathcal{L}\}$.

A norm $n = (\varphi, \psi) \in \mathcal{N}$ is read as "Given φ, it ought to be that ψ." We refer to φ (resp. ψ) by $\mathsf{body}(n)$ (resp. $\mathsf{head}(n)$). Various mechanisms have been developed for detaching obligations (propositional formulas) from norms and facts, and for resolving conflicts that arise during the detachment process. Constraints are formulas with which detached obligations must be classically consistent, and constitute a generalization of logical consistency checks on the output generated by norms.[1] For instance, the fact φ triggers the norm (φ, ψ) from which the obligation ψ is detached, whereas the constraint $\neg\psi$ blocks the application of (φ, ψ).

Example 1. The following normative system represents a central challenge for normative reasoning under conflicts, referred to as a *contrary-to-duty* scenario (Hilpinen and McNamara 2013). In this scenario, an agent needs to determine what to do given the fact that a norm is violated:

$$\mathcal{K}_{\mathcal{CTD}} = \langle \{\neg p\}, \{\neg p\}, \{(\top, p), (p, \neg a), (\neg p, a)\} \rangle.$$

Let p stand for "agent x keeps her promise," and a for "agent x apologizes." There are three norms involved: $(\top, p)$ is the default norm that "agent x ought to keep her promise," $(p, \neg a)$ expresses that "if x keeps the promise, she should not apologize," and $(\neg p, a)$ is the contrary-to-duty norm for "if x does not keep her promise, she ought to apologize." The facts are that the agent did not keep her promise: she violated the default norm $(\top, p)$. The constraint is set to $\neg p$ since agent x needs to know what she ought to do given the fact that she did not keep her promise. The desired detachable conclusion is that "agent x ought to apologize" (Hilpinen and McNamara 2013).

To illustrate our approach, we briefly recall the four central I/O detachment operations and refer to (Parent and van der Torre 2018) for an extensive survey of I/O systems.

Definition 2. (Makinson and van der Torre 2001) Let $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$ be a normative knowledge base and let $\mathsf{Cn}(\Delta) = \{\varphi \mid \Delta \vdash \varphi\}$ denote set closure under classical entailment $\vdash$. Detachment (det) from a set of norms $\mathcal{N}' \subseteq \mathcal{N}$ w.r.t. $\mathcal{F}$ is characterized by the following set:

$$\mathsf{det}(\mathcal{F}, \mathcal{N}') = \{\psi \mid \varphi \in \mathcal{F} \textit{ for some } (\varphi, \psi) \in \mathcal{N}'\}.$$

A basic output operation out_1 is then defined in terms of the logical closure of norm detachment to the logical closure of the facts, i.e.,

$$\mathsf{out}_1(\mathcal{F}, \mathcal{N}) = \mathsf{Cn}(\mathsf{det}(\mathsf{Cn}(\mathcal{F}), \mathcal{N})).$$

A stronger version that allows for disjunctive reasoning by cases with facts and norms is given by:

$$\mathsf{out}_2(\mathcal{F}, \mathcal{N}) = \bigcap \{\mathsf{Cn}(\mathsf{det}(\mathcal{V}, \mathcal{N})) \mid \mathcal{F} \subseteq \mathcal{V} \text{ and } \mathcal{V} \text{ is complete}\},$$

where a set $\mathcal{V} \subseteq \mathcal{L}$ is *complete* iff it is maximally consistent (over the whole language $\mathcal{L}$) or equal to the set of all the formulas of $\mathcal{L}$.[2]

The operation out_3 allows for the successive detachment (i.e., chaining) of norms. This operation can be defined iteratively by $\mathsf{out}_3(\mathcal{F}, \mathcal{N}) = \bigcup_{i \geq 1} \mathsf{out}_3^i(\mathcal{F}, \mathcal{N})$, where

[1] In I/O logic, the role of the constraints $\mathcal{C}$ is to block norms that may otherwise yield inconsistencies with $\mathcal{C}$, hence $\mathcal{N}$ and $\mathcal{C}$ can be jointly inconsistent. In the literature, for application contexts, $\mathcal{C}$ is intended to be consistent, but this is not a technical requirement.

[2] The formal role of constraints is discussed in Section 2.2.

$\mathsf{out}_3^1(\mathcal{F},\mathcal{N}) = \mathsf{out}_1(\mathcal{F},\mathcal{N})$, $\mathsf{out}_3^{n+1}(\mathcal{F},\mathcal{N}) = \mathsf{Cn}(\mathsf{det}(\mathcal{F} \cup \mathsf{out}_3^n(\mathcal{F},\mathcal{N})),\mathcal{N})$. A compact way of writing out_3 is:

$$\mathsf{out}_3(\mathcal{F},\mathcal{N}) = \bigcap\{\mathsf{Cn}(\mathsf{det}(\mathcal{A},\mathcal{N})) \mid \mathcal{F} \subseteq \mathcal{A} = \mathsf{Cn}(\mathcal{A}),\ \mathsf{det}(\mathcal{A},\mathcal{N}) \subseteq \mathsf{Cn}(\mathcal{A})\}.$$

Finally, out_4 combines out_2 with out_3:

$$\mathsf{out}_4(\mathcal{F},\mathcal{N}) = \bigcap\{\mathsf{Cn}(\mathsf{det}(\mathcal{V},\mathcal{N})) \mid \mathcal{F} \subseteq \mathcal{V},\ \mathsf{det}(\mathcal{A},\mathcal{N}) \subseteq \mathcal{V},\ \text{and } \mathcal{V} \text{ is complete}\}.$$

Example 2. Consider$\mathcal{N} = \{(\top, v), (v, u), (p, s), (q, s)\}$, $\mathcal{F} = \{p \vee q\}$ and $\mathcal{C} = \emptyset$. Then, $\mathsf{det}(\mathcal{F},\mathcal{N}) = \emptyset$ while $\mathsf{out}_1(\mathcal{F},\mathcal{N}) = \mathsf{Cn}(\{v\})$, since $\top \in \mathsf{Cn}(\mathcal{F})$. Also, $\mathsf{out}_2(\mathcal{F},\mathcal{N}) = \mathsf{Cn}(\{v, s\})$ and $\mathsf{out}_3(\mathcal{F},\mathcal{N}) = \mathsf{Cn}(\{v, u\})$. So, $s \in \mathsf{out}_2(\mathcal{F},\mathcal{N}) \setminus \mathsf{out}_3(\mathcal{F},\mathcal{N})$ while $u \in \mathsf{out}_3(\mathcal{F},\mathcal{N}) \setminus \mathsf{out}_2(\mathcal{F},\mathcal{N})$. Finally, $\mathsf{out}_4(\mathcal{F},\mathcal{N}) = \mathsf{Cn}(\{v, u, s\})$.

Our example illustrates a more general point: For every set $\mathcal{F}$ of facts and set $\mathcal{N}$ of norms, it holds that $\mathsf{det}(\mathcal{F},\mathcal{N}) \subseteq \mathsf{out}_1(\mathcal{F},\mathcal{N}) \subseteq \mathsf{out}_2(\mathcal{F},\mathcal{N}) \subseteq \mathsf{out}_4(\mathcal{F},\mathcal{N})$ and $\mathsf{det}(\mathcal{F},\mathcal{N}) \subseteq \mathsf{out}_1(\mathcal{F},\mathcal{N}) \subseteq \mathsf{out}_3(\mathcal{F},\mathcal{N}) \subseteq \mathsf{out}_4(\mathcal{F},\mathcal{N})$.

Example 3. Recall Example 1. Clearly, out_1 and out_2 yield, what is known as, a pragmatic oddity (Hilpinen and McNamara 2013): "to keep the promise" (p) and "apologize for not keeping it" (a). What is more, out_3 and out_4 yield inconsistent output due to the detachment of both a and $\neg a$. In other words, the latter operations perpetuate implicit inconsistencies in a given knowledge base. To overcome such oddities and inconsistencies, Makinson and van der Torre (2001) use constraints to control the output, identify maximal consistent subsets of norms and thereby re-establish consistency. We demonstrate this in the the next section.

2.2 Maximally Consistent, Minimally Inconsistent, and Minimal Correction Norm Sets

Traditionally, nonmonotonic I/O logics are investigated with respect to identifying *maximally consistent* norm sets. We recall this notion here. Moreover, we extend the I/O formalism with the concepts of *minimally inconsistent* and *minimal correction norm sets* (adopted from (Reiter 1987)).

Definition 3. Let $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$ be a normative system and $\mathsf{out} \in \{\mathsf{out}_1, \dots, \mathsf{out}_4\}$ a fixed output operation. We say that $\mathcal{N}' \subseteq \mathcal{N}$ is a:[3]

- **consistent norm set** (of $\mathcal{K}$ w.r.t. out), if $\mathsf{out}(\mathcal{F},\mathcal{N}') \cup \mathcal{C}$ is classically consistent (i.e., $\mathsf{out}(\mathcal{F},\mathcal{N}') \cup \mathcal{C} \nvdash \bot$). Else, $\mathcal{N}'$ is an inconsistent norm set.
- **maximally consistent norm set (MCN)** (of $\mathcal{K}$ w.r.t. out), if for each consistent norm set $\mathcal{N}'' \subseteq \mathcal{N}$ it holds that $\mathcal{N}' \not\subset \mathcal{N}''$.
- **minimally inconsistent norm set (MIN)** (of $\mathcal{K}$ w.r.t out), if $\mathsf{out}(\mathcal{F},\mathcal{N}') \cup \mathcal{C} \vdash \bot$ and for every $\mathcal{N}'' \subset \mathcal{N}'$ it holds that $\mathsf{out}(\mathcal{F},\mathcal{N}'') \cup \mathcal{C} \nvdash \bot$.
- **minimal correction norm set (MRN)** (of $\mathcal{K}$ w.r.t. out), if $\mathcal{N} \setminus \mathcal{N}'$ is a maximally consistent norm set of $\mathcal{K}$ w.r.t. out.

[3]Some logical relations between the sets defined below can be expressed in terms of hitting sets, see (Liffiton and Sakallah 2008).

The set of maximally consistent norm sets (resp. the set of minimally inconsistent norm sets, the set of minimal correction norm sets) of $\mathcal{N}$ over $\mathcal{K}$ and out_i ($1 \leq i \leq 4$) is denoted $\mathsf{MCN}_i(\mathcal{K})$ (resp. $\mathsf{MIN}_i(\mathcal{K})$, $\mathsf{MRN}_i(\mathcal{K})$). When the output operation is fixed or arbitrary we shall omit the subscript i.

Henceforth, we say that a knowledge base $\mathcal{K}$ is **consistent** whenever $\mathsf{MIN}(\mathcal{K}) = \emptyset$, and inconsistent otherwise.

Based on Definition 3, we consider the following notions:

- *Problematic norms* are those norms that are part of at least one minimally inconsistent norm set within $\mathcal{K}$. We denote by $\mathsf{prob}(\mathcal{K})$ the set of problematic norms in $\mathcal{K}$, i.e., $\mathsf{prob}(\mathcal{K}) = \{n \in \mathcal{N} \mid \mathcal{N} \in \mathsf{MIN}(\mathcal{K})\}$.
- A norm $n \in \mathcal{N}$ is a *self-contradictory norm* if $\{n\}$ is a minimally inconsistent norm set. We write $\mathsf{SelfCont}(\mathcal{K})$ to denote the set of self-contradictory norms in $\mathcal{K}$.

In Examples 4-8 below, norm sets are identified relative to an arbitrary out operation, unless explicitly stated otherwise.

Example 4. Reconsider the normative system $\mathcal{K}_{\mathcal{CTD}}$ of Example 1. We have one minimally inconsistent norm set, namely $\mathcal{N}' = \{(\top, p)\}$. Note here that the set is inconsistent for a generalized notion of consistency w.r.t. $\mathcal{C} = \{\neg p\}$. $\mathcal{N}'$ is also the unique minimal correction norm set of $\mathcal{K}_{\mathcal{CTD}}$ (since $\{(\neg p, a), (p, \neg a)\}$ is the unique MCN). Now, if $\mathcal{C} = \emptyset$, then under out_i ($i \in \{3, 4\}$), $\mathsf{MIN}(\mathcal{K}_{\mathcal{CTD}}) = \{\mathcal{N}\}$ whereas $\mathsf{MRN}(\mathcal{K}_{\mathcal{CTD}}) = \{\{(\top, p)\}, \{(p, \neg a)\}, \{(\neg p, a)\}\}$ (we discuss MCN in Example 9).

Example 5 (Double conflict)**.** Let $\mathcal{K}_5 = \langle \emptyset, \emptyset, \mathcal{N}_5 \rangle$ with $\mathcal{N}_5 = \{(\top, p), (\top, q), (\top, \neg p \wedge \neg q)\}$. Then, $\mathsf{MIN}(\mathcal{K}_5) = \{\{(\top, p), (\top, \neg p \wedge \neg q)\}, \{(\top, q), (\top, \neg p \wedge \neg q)\}\}$, and $\mathsf{MCN}(\mathcal{K}_5) = \mathsf{MRN}(\mathcal{K}_5) = \{\{(\top, p), (\top, q)\}, \{(\top, \neg p \wedge \neg q)\}\}$.

Example 6 (Binary conflict)**.** Consider $\mathcal{K}_6 = \langle \mathcal{F}_6, \mathcal{C}_6, \mathcal{N}_6 \rangle$ with $\mathcal{F}_6 = \mathcal{C}_6 = \emptyset$ and $\mathcal{N}_6 = \{(\top, p), (\top, \neg p)\}$. Then, $\mathcal{N}_6$ is the unique minimally inconsistent norm set. So, all norms in $\mathcal{K}_6$ are problematic. Moreover, $\mathsf{MCN}(\mathcal{K}_6) = \mathsf{MRN}(\mathcal{K}_6) = \{\{(\top, p)\}, \{(\top, \neg p)\}\}$. If we consider $\mathcal{K}_6' = \langle \mathcal{F}_6, \mathcal{C}_6', \mathcal{N}_6 \rangle$, with $\mathcal{C}_6' = \{p\}$, we have $\mathsf{MIN}(\mathcal{K}_6) = \mathsf{MRN}(\mathcal{K}_6) = \{\{(\top, \neg p)\}\}$. In this case, the norm $(\top, \neg p)$ is self-contradictory. Moreover, $\mathsf{MCN}(\mathcal{K}_6) = \{\{(\top, p)\}\}$.

Example 7 (Triple conflict)**.** Let $\mathcal{K}_7 = \langle \mathcal{F}_7, \mathcal{C}_7, \mathcal{N}_7 \rangle$ where $\mathcal{F}_7 = \mathcal{C}_7 = \emptyset$ and $\mathcal{N}_7 = \{(\top, p), (\top, q), (\top, \neg(p \wedge q))\}$. We have $\mathsf{MIN}(\mathcal{K}_7) = \{\mathcal{N}_7\}$, while $\mathsf{MCN}(\mathcal{K}_7) = \{\mathcal{N}_7 \setminus \{n\} \mid n \in \mathcal{N}_7\}$ and $\mathsf{MRN}(\mathcal{K}_7) = \{\{n\} \mid n \in \mathcal{N}_7\}$.

Example 8 (2 vs. 1)**.** Let $\mathcal{K}_8 = \langle \mathcal{F}_8, \mathcal{C}_8, \mathcal{N}_8 \rangle$ with $\mathcal{F}_8 = \{q_1, q_2, q_3\}$, $\mathcal{C}_8 = \emptyset$ and $\mathcal{N}_8 = \{(q_1, p), (q_2, p), (q_3, \neg p)\}$. Here, there are two minimally inconsistent norm sets, $\{(q_1, p), (q_3, \neg p)\}$ and $\{(q_2, p), (q_3, \neg p)\}$, and it holds that $\mathsf{MCN}(\mathcal{K}_8) = \mathsf{MRN}(\mathcal{K}_8) = \{\{(q_1, p), (q_2, p)\}, \{(q_3, \neg p)\}\}$.

In (Makinson and van der Torre 2001), inference relations of I/O logic are defined using maximal consistent sets.[4]

Definition 4. Let $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$ be a normative knowledge base:

[4]Other inferences based on I/O logic, such as the ones based on argumentative reasoning (Arieli, van Berkel, and Straßer 2024; van Berkel and Straßer 2022), are outside the scope of this paper.

- We say that φ **skeptically follows** from $\mathcal{K}$ (w.r.t. out), if $\varphi \in \mathsf{out}(\mathcal{F}, \mathcal{N}')$ for *all* $\mathcal{N}' \in \mathsf{MCN}(\mathcal{K})$.
- We say that φ **credulously follows** from $\mathcal{K}$ (w.r.t. out), if $\varphi \in \mathsf{out}(\mathcal{F}, \mathcal{N}')$ for *some* $\mathcal{N}' \in \mathsf{MCN}(\mathcal{K})$.

Example 9. Reconsider the knowledge base $\mathcal{K}_{\mathcal{CTD}}$ from Example 1. In this case, $\mathsf{out}_1(\mathcal{F}, \mathcal{N}) = \mathsf{Cn}(\{p, \neg a, a\})$ is inconsistent. Using the set of constraints $\mathcal{C} = \{\neg p\}$, $\mathcal{K}_{\mathcal{CTD}}$ yields one maximally consistent norm set $\mathcal{N}' = \{(\neg p, a), (p, \neg a)\}$. The norm $(\top, p)$ yields an inconsistency with the given constraint. The norm $(p, \neg a)$ is part of $\mathcal{N}'$ but is not triggered in the reasoning process. It follows that the single output set is $\mathsf{out}_1(\mathcal{F}, \mathcal{N}') = \mathsf{Cn}(\{a\})$, which means (both skeptically and credulously) that the agent ought to apologize. Notice that for the empty constraint set $\mathcal{C} = \emptyset$, we have three maximally consistent norm sets $\mathcal{N}' = \{(\top, p), (p, \neg a)\}$, $\mathcal{N}'' = \{(\top, p), (\neg p, a)\}$, and $\mathcal{N}''' = \{(\neg p, a), (p, \neg a)\}$. Using skeptical inference, then, we conclude $p \vee a$, but neither p nor a.

Inference relations (such as those in Example 9) treat all the elements of MCN uniformly. In the next section, we show that by utilizing inconsistency measures, we are able to introduce more nuances.

3 Inconsistency Measures for Normative Knowledge Bases

We are now in the position to introduce various methods for quantifying conflicts within normative knowledge bases, employing the notion of '*inconsistency measure*.' Such measures may enhance the understanding of conflicts occurring in normative systems.

Definition 5. Let $\mathbb{K}$ denote the class of all normative systems $\mathcal{K}$ from Definition 1. An **inconsistency measure** is a function $\mathcal{I}$ that maps a normative knowledge base $\mathcal{K}$ to a real value, that is $\mathcal{I} : \mathbb{K} \to \mathbb{R}_0^+ \cup \{\infty\}$, such that $\mathcal{I}(\mathcal{K}) = 0$ if $\mathcal{K}$ is consistent.

Intuitively, the higher an inconsistency measure, the more conflicting is a normative system.

We consider the following six syntactic inconsistency measures from the literature[5] (see, e.g., (Bona et al. 2019)) and extend them to the context of normative systems.
For a knowledge base $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$, we define:

- **Drastic measure:**
$$\mathcal{I}_d(\mathcal{K}) = \begin{cases} 0 & |\mathsf{MIN}(\mathcal{K})| = 0 \ \& \ \mathcal{C} \nvdash \bot \\ 1 & |\mathsf{MIN}(\mathcal{K})| \neq 0 \ \& \ \mathcal{C} \nvdash \bot \\ \infty & \mathcal{C} \vdash \bot \end{cases}$$
- $\mathsf{MIN}_{\#}$**-based measure:**
$$\mathcal{I}_{\#}(\mathcal{K}) = \begin{cases} |\mathsf{MIN}(\mathcal{K})| & \mathcal{C} \nvdash \bot \\ \infty & \mathcal{C} \vdash \bot \end{cases}$$
- $\mathsf{MIN}_{\cup}$**-based measure:**
$$\mathcal{I}_{\mathsf{prob}}(\mathcal{K}) = \begin{cases} |\bigcup \mathsf{MIN}(\mathcal{K})| & \mathcal{C} \nvdash \bot \\ \infty & \mathcal{C} \vdash \bot \end{cases}$$
- MIN_{Σ}**-based measure:**
$$\mathcal{I}_{\mathsf{MIN}}(\mathcal{K}) = \begin{cases} 0 & |\mathsf{MIN}(\mathcal{K})| = 0 \ \& \ \mathcal{C} \nvdash \bot \\ \sum_{M \in \mathsf{MIN}(\mathcal{K})} \frac{1}{|M|} & |\mathsf{MIN}(\mathcal{K})| \neq 0 \ \& \ \mathcal{C} \nvdash \bot \\ \infty & \mathcal{C} \vdash \bot \end{cases}$$
- MCN**-based measure:**
$$\mathcal{I}_{\mathsf{MCN}}(\mathcal{K}) = \begin{cases} (|\mathsf{MCN}(\mathcal{K})| + |\mathsf{SelfCont}(\mathcal{K})|) - 1 & \mathcal{C} \nvdash \bot \\ \infty & \mathcal{C} \vdash \bot \end{cases}$$
- MRN**-based measure:**
$$\mathcal{I}_{\mathsf{MRN}}(\mathcal{K}) = \begin{cases} \min_{M \in \mathsf{MRN}(\mathcal{K})} |M| & \mathcal{C} \nvdash \bot \\ \infty & \mathcal{C} \vdash \bot \end{cases}$$

We refer to, e.g., (Bona et al. 2019) for further discussions on inconsistency measures like the ones above.

The following table illustrates the different inconsistency measures considered above for the previous examples.

	$\mathcal{CTD}$	$\mathcal{K}_5$	$\mathcal{K}_6$	$\mathcal{K}_7$	$\mathcal{K}_8$
$\mathcal{I}_d$	1	1	1	1	1
$\mathcal{I}_{\#}$	1	2	1	1	2
$\mathcal{I}_{\mathsf{prob}}$	1	3	2	3	3
$\mathcal{I}_{\mathsf{MIN}}$	1	$1/2 \cdot 2$	$1/2$	$1/3$	$1/2 \cdot 2$
$\mathcal{I}_{\mathsf{MCN}}$	1	1	1	2	1
$\mathcal{I}_{\mathsf{MRN}}$	1	1	1	1	1

Various criteria have been studied for characterizing inconsistency measures and evaluating their plausibility (see, e.g., (Hunter and Konieczny 2010; Besnard 2014; Thimm 2017; Bona et al. 2019)). In what follows, we consider the properties of *monotonicity* and *dominance* (Hunter and Konieczny 2010), as well as several other new postulates that offer guidance for a better understanding of the measures by facilitating a comparison between them. These postulates are also used to ensure some desirable properties of the entailment relations induced by the inconsistency measures, defined in the subsequent sections.

In what follows, we write $\mathcal{K} \oplus \varphi^x$ for $x \in \{f, c, n\}$ to denote the addition of φ to $\mathcal{K}$ as a fact $(x = f)$, constraint $(x = c)$, or norm $(x = n)$. Similarly, we define $\ominus$ to denote the removal from $\mathcal{K}$ of facts or constraints or norms. Given a formula $\varphi \in \mathcal{L} \cup \{(\psi, \varphi) \mid \psi, \varphi \in \mathcal{L}\}$, we write $\mathcal{K} \oplus_x \varphi$ for $\mathcal{K} \oplus \varphi^x$ $(x \in \{f, c, n\})$. We also use the following notions:

- Δ is a *minimal truth set* iff $\vdash \bigvee \Delta$ while $\nvdash \bigvee \Delta'$ for all $\Delta' \subset \Delta$.
- Δ is a *minimal conflict* (also known as *minimally inconsistent set*) iff $\vdash \neg \bigwedge \Delta$ while $\nvdash \neg \bigwedge \Delta'$ for all $\Delta' \subset \Delta$.
- *Remainder sets* in the context of normative systems are defined as follows: $\langle \mathcal{F}, \mathcal{N}, \mathcal{C} \rangle \bot_x \varphi$ denotes for $x = c$ the set of all $\mathcal{K}' = \langle \mathcal{F}, \mathcal{N}, \mathcal{C}' \rangle$ for which $\mathcal{C}'$ is a $\subset$-maximal subset of $\mathcal{C}$ for which $\mathcal{C}' \nvdash \varphi$. The case where $x = f$ is similar.

In what follows, we suppose that $\oplus \in \{\oplus_f, \oplus_c\}$ is fixed.[6] We introduce the following properties:

1. $\oplus$**-Monotonicity:** $\mathcal{I}(\mathcal{K}) \leq \mathcal{I}(\mathcal{K} \oplus \varphi)$.
   The degree of inconsistency cannot decrease when facts resp. constraints are added to $\mathcal{K}$.

[5] Semantic inconsistency measures based on multi-valued logics (Grant and Hunter 2023) are left for future work.

[6] We will consider norm additions ($\oplus_n$) in future work.

2. $\oplus$**-Equivalence**: $\mathcal{I}(\mathcal{K} \oplus \varphi) = \mathcal{I}(\mathcal{K} \oplus \psi)$ if φ and ψ are logically equivalent (i.e., $\{\varphi\} \vdash \psi$ and $\{\psi\} \vdash \varphi$).
   Substituting a fact resp. constraint by a logically equivalent formula is irrelevant for the inconsistency of $\mathcal{K}$.
3. $\oplus$**-Weakening:** $\mathcal{I}(\mathcal{K} \oplus \varphi) \geq \mathcal{I}(\mathcal{K} \oplus \psi)$ if $\varphi \vdash \psi$.[7]
   Replacing a fact/constraint with one of its consequences does not render the knowledge base more inconsistent.
4. $\oplus$**-Bi-innocent choice**: $\mathcal{I}(\mathcal{K} \oplus \varphi) \leq \mathcal{I}(\mathcal{K})$ or $\mathcal{I}(\mathcal{K} \oplus \neg\varphi) \leq \mathcal{I}(\mathcal{K})$.
   A formula and its negation cannot both render a knowledge base more inconsistent (when added as facts resp. constraints).
5. $\oplus$**-Innocent choice**: where Δ is a minimal truth set, there is a $\varphi \in \Delta$ for which $\mathcal{I}(\mathcal{K} \oplus \varphi) \leq \mathcal{I}(\mathcal{K})$.
   The addition of at least one member of a minimal truth set does not render a knowledge base more inconsistent.
6. $\oplus$**-Subtractive bi-$\ominus$-innocent choice:** $\mathcal{I}(\mathcal{K} \oplus \varphi) \geq \mathcal{I}(\mathcal{K} \ominus \neg\varphi)$ or $\mathcal{I}(\mathcal{K} \oplus \neg\varphi) \geq \mathcal{I}(\mathcal{K} \ominus \varphi)$.
7. $\oplus$**-Subtractive bi-$\perp$-innocent choice:** $\mathcal{I}(\mathcal{K} \oplus \varphi) \geq \mathcal{I}(\mathcal{K}')$ for all $\mathcal{K}' \in \mathcal{K}\perp\neg\varphi$, or $\mathcal{I}(\mathcal{K} \oplus \neg\varphi) \geq \mathcal{I}(\mathcal{K}')$ for all $\mathcal{K}' \in \mathcal{K}\perp\varphi$.
   The last two properties mean that, for any formula φ, if removing it leads to an increase of inconsistency as compared to adding its negation, then it cannot be the case that removing its negation leads to an increase of inconsistency as compared to adding φ.
8. $\oplus$**-Subtractive $\ominus$-innocent choice**: if Δ is a minimal conflict, there is a $\varphi \in \Delta$ s.t. $\mathcal{I}(\mathcal{K} \oplus \varphi) \geq \mathcal{I}(\mathcal{K} \ominus \neg\varphi)$.
9. $\oplus$**-Subtractive $\perp$-innocent choice**: if Δ is a minimal conflict, there is a $\varphi \in \Delta$ s.t. $\mathcal{I}(\mathcal{K} \oplus \varphi) \geq \mathcal{I}(\mathcal{K}')$ for all $\mathcal{K}' \in \mathcal{K}\perp\neg\varphi$.
   The last two properties are a conservative extension of $\oplus$-subtractive bi-($\ominus$ resp. $\perp$)-innocent choice to minimal conflicts of arbitrary length.
10. $\oplus$**-Contrastive innocent choice**: where Δ is a minimal truth set, there is a $\varphi \in \Delta$ s.t. $\mathcal{I}(\mathcal{K} \oplus \varphi) \leq \mathcal{I}(\mathcal{K} \oplus \neg\varphi)$.
    For any minimal truth set, the addition of at least one of its members should not increase the inconsistency of a knowledge base as compared to adding its negation.
11. $\oplus$**-Bi-guilty choice**: $\mathcal{I}(\mathcal{K} \oplus \varphi) \geq \mathcal{I}(\mathcal{K})$ or $\mathcal{I}(\mathcal{K} \oplus \neg\varphi) \geq \mathcal{I}(\mathcal{K})$.
    It cannot happen that both the addition of a formula and the addition of its negation lead to a decrease of the inconsistency of a knowledge base.
12. $\oplus$**-Guilty choice**: where Δ is a minimal conflict, there is a $\varphi \in \Delta$ for which $\mathcal{I}(\mathcal{K} \oplus \varphi) \geq \mathcal{I}(\mathcal{K})$.
    It cannot happen that for every formula contained in a minimal conflict, the addition of it leads to a decrease of the inconsistency of a knowledge base.
13. $\oplus$**-Free independence**: $\mathcal{I}(\mathcal{K} \oplus \varphi) \leq \mathcal{I}(\mathcal{K}) < \mathcal{I}(\mathcal{K} \oplus \neg\varphi)$ if $\varphi \in \mathsf{out}(\mathcal{F}, \bigcap \mathsf{MCN}(\mathcal{K}))$.
    Incorporating a formula not involved in any normative conflict (a) cannot increase the inconsistency, but (b) adding its negation does.
14. $\oplus$**-Strong independence**: $\mathcal{I}(\mathcal{K} \oplus \varphi) \leq \mathcal{I}(\mathcal{K}) < \mathcal{I}(\mathcal{K} \oplus \neg\varphi)$ if $\varphi \in \bigcap_{\mathcal{N}' \in \mathsf{MCN}(\mathcal{K})} \mathsf{out}(\mathcal{F}, \mathcal{N}')$.
    Incorporating a formula not involved in a universal conclusion of normative conflicts (a) cannot increase the inconsistency, but (b) adding its negation does.
15. $\oplus$**-Consistency by cases**: If $\mathcal{I}(\mathcal{K} \oplus \varphi) > \mathcal{I}(\mathcal{K})$ and $\mathcal{I}(\mathcal{K} \oplus \psi) > \mathcal{I}(\mathcal{K})$, then $\mathcal{I}(\mathcal{K} \oplus (\varphi \vee \psi)) > \mathcal{I}(\mathcal{K})$.
    If φ and ψ increase the inconsistency of a knowledge base, then so will $\varphi \vee \psi$.
16. $\oplus$**-Upper bounding**: If $\mathcal{I}(\mathcal{K} \oplus \neg\varphi) > \mathcal{I}(\mathcal{K})$ and $\mathcal{I}(\mathcal{K} \oplus \psi) \leq \mathcal{I}(\mathcal{K})$ then $\mathcal{I}(\mathcal{K} \oplus \varphi \oplus \psi) \leq \mathcal{I}(\mathcal{K} \oplus \varphi)$.
    If $\neg\varphi$ increases the inconsistency in $\mathcal{K}$ while ψ does not, then ψ also does not increase the inconsistency of $\mathcal{K} \oplus \varphi$.

The following proposition shows some relations between the properties considered bove.

Proposition 1. *Let $\oplus \in \{\oplus_c, \oplus_f\}$. Below are some relations among the principles:*

- *($\oplus$-subtractive ($\ominus$ resp. $\perp$)-innocent choice) implies ($\oplus$-subtractive bi-($\ominus$ resp. $\perp$)-innocent choice).*
- *($\oplus$-innocent choice) and ($\oplus$-monotonicity) implies ($\oplus$-contrastive innocent choice).*
- *($\oplus$-innocent choice) implies ($\oplus$-bi-innocent choice)*
- *($\oplus$-strong independence) implies ($\oplus$-free independence).*
- *($\oplus$-monotonicity) implies ($\oplus$-subtractive ($\ominus$ resp. $\perp$)-innocent choice) and ($\oplus$-guilty choice).*
- *($\oplus$-guilty choice) implies ($\oplus$-bi-guilty choice).*

Due to limited space, we focus in the remainder of this section on the inconsistency measure $\mathcal{I}_{\mathsf{MRN}}$ for the reason that, compared to the other measures, it performs surprisingly well with respect to the list of specified principles.

Proposition 2. *Let* $\mathsf{out} \in \{\mathsf{out}_1, \ldots, \mathsf{out}_4\}$ *and* $\oplus \in \{\oplus_c, \oplus_f\}$. *Then,* $\mathcal{I}_{\mathsf{MRN}}$ *satisfies:*

- *$\oplus$-monotonicity, $\oplus$-equivalence, $\oplus$-weakening, $\oplus$-guilty choice and $\oplus$-bi-guilty choice for* out.
- *$\oplus_c$-innocent choice, $\oplus_c$-contrastive innocent choice, $\oplus_c$-subtractive ($\ominus_c$ resp. $\perp$)-innocent choice, $\oplus_c$-bi-innocent choice, $\oplus_c$-subtractive bi-innocent choice, $\oplus_c$-consistency by cases, $\oplus_c$-strong and free independence, and $\oplus_c$-upper bounding for* out.
- *$\oplus_f$-innocent choice, $\oplus_f$-contrastive innocent choice, $\oplus_f$-subtractive ($\ominus_f$ resp. $\perp$)-innocent choice, $\oplus_f$-bi-innocent choice, $\oplus_f$-subtractive bi-innocent choice, $\oplus_f$-consistency by cases, and $\oplus_f$-upper bounding for* out_2 *and* out_4.

Proof. The proofs of most of the properties are based on the following facts. Let $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$ be a normative system and $\mathcal{N}' \in \mathsf{MCN}(\mathcal{K} \oplus \varphi)$. Then,

1. there is an $\mathcal{N}'' \in \mathsf{MCN}(\mathcal{K})$ such that $\mathcal{N}' \subseteq \mathcal{N}''$.

[7] In (Hunter and Konieczny 2010), this criterion is known as 'dominance,' with the difference that φ is required to be consistent. Here, we consider that if φ is inconsistent, it is deemed to be at least as inconsistent as any other fact/constraint.

2. $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K}) \leq |\mathcal{N} \setminus \mathcal{N}'|$.
3. $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K}) = |\mathcal{N} \setminus \mathcal{N}'|$ iff $|\mathcal{N} \setminus \mathcal{N}'|$ is minimal among the members of $\mathsf{MCN}(\mathcal{K})$ iff $|\mathcal{N}'|$ is maximal among the members of $\mathsf{MCN}(\mathcal{K})$.

Fact (1) follows from Lindenbaum's lemma; Facts (2)–(3) follow from the definition of $\mathcal{I}_{\mathsf{MRN}}$. Based on this, for illustration, we demonstrate the proofs of two properties.

$\oplus$**-monotonicity:** Consider an $\mathcal{N}' \in \mathsf{MCN}(\mathcal{K}\oplus\varphi)$ such that $|\mathcal{N} \setminus \mathcal{N}'|$ is minimal. By (3), $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus \varphi) = |\mathcal{N} \setminus \mathcal{N}'|$. By (1), there is a $\mathcal{N}'' \in \mathsf{MCN}(\mathcal{K})$ for which $\mathcal{N}' \subseteq \mathcal{N}''$. Clearly, $|\mathcal{N} \setminus \mathcal{N}''| \leq |\mathcal{N} \setminus \mathcal{N}'|$. By (2), $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K}) \leq |\mathcal{N} \setminus \mathcal{N}''| \leq \mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus \varphi)$.

$\oplus$**-weakening:** Suppose that $\varphi \vdash \psi$ and let $\mathcal{N}' \in \mathsf{MCN}(\mathcal{K}\oplus \varphi)$ be such that $|\mathcal{N}'|$ is maximal among the elements in $\mathsf{MCN}(\mathcal{K} \oplus \varphi)$. By (3), $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus \varphi) = |\mathcal{N} \setminus \mathcal{N}'|$.

- Consider first $\oplus = \oplus_f$. Thus, $\mathsf{out}(\mathcal{F} \cup \{\varphi\}, \mathcal{N}') \cup \mathcal{C} \nvdash \bot$. Clearly, $\mathsf{out}(\mathcal{F} \cup \{\psi\}, \mathcal{N}') \cup \mathcal{C} \nvdash \bot$. So, there is a $\mathcal{N}''$ for which $\mathcal{N}' \subseteq \mathcal{N}''$ and $\mathcal{N}'' \in \mathsf{MCN}(\mathcal{K} \oplus_f \psi)$. By (2), $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus_f \psi) \leq |\mathcal{N} \setminus \mathcal{N}''| \leq |\mathcal{N} \setminus \mathcal{N}'| = \mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus_f \varphi)$.
- Consider now $\oplus = \oplus_c$. We have, $\mathsf{out}(\mathcal{F}, \mathcal{N}') \cup \mathcal{C} \cup \{\varphi\} \nvdash \bot$. Clearly also $\mathsf{out}(\mathcal{F}, \mathcal{N}') \cup \mathcal{C} \cup \{\psi\} \nvdash \bot$. So, there is a $\mathcal{N}''$ for which $\mathcal{N}' \subseteq \mathcal{N}''$ and $\mathcal{N}'' \in \mathsf{MCN}(\mathcal{K} \oplus_c \psi)$. By (2), $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus_c \psi) \leq |\mathcal{N} \setminus \mathcal{N}''| \leq |\mathcal{N} \setminus \mathcal{N}'| = \mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus_c \varphi)$. □

There are some principles that are not satisfied by $\mathcal{I}_{\mathsf{MRN}}$ in some contexts. To illustrate this, we consider one case:

Example 10. Free and Strong independence do not hold for $\mathcal{I}_{\mathsf{MRN}}$ with $\oplus_f$ and any $\mathsf{out} \in \{\mathsf{out}_1, \mathsf{out}_2, \mathsf{out}_3, \mathsf{out}_4\}$. Indeed, consider for instance a knowledge base $\mathcal{K}$ with $\mathcal{C} = \emptyset$, $\mathcal{F} = \{\neg p\}$ and $\mathcal{N} = \{(\neg p, p), (\bot, \neg p)\}$. Then, $p \in \bigcap_{\mathcal{N}' \in \mathsf{MCN}(\mathcal{K})} \mathsf{out}(\mathcal{F}, \mathcal{N}') = \mathsf{out}(\mathcal{F}, \bigcap_{\mathcal{N}' \in \mathsf{MCN}(\mathcal{K})} \mathcal{N}')$ since $\mathsf{MCN}(\mathcal{K}) = \{\mathcal{N}\}$. However, since $\mathsf{MCN}(\mathcal{K} \oplus_f p) = \{\{(\neg p, p)\}, \{(\bot, \neg p)\}\}$, we have that $\mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus_f p) = 1 > 0 = \mathcal{I}_{\mathsf{MRN}}(\mathcal{K}) = \mathcal{I}_{\mathsf{MRN}}(\mathcal{K} \oplus_f \neg p)$.

4 Normative Reasoning with Inconsistency Measures

We are now ready to define normative entailments based on the inconsistency measures considered in the previous section. As demonstrated in Example 3, a weakness of the output operators of I/O logic (Definition 2) is that their set of conclusions may be inconsistent or odd, thus not informative. To resolve this, Makinson and van der Torre (2001) introduced their notion of inference based on MCN. We recall skeptical entailment:

$$\mathcal{K} \mid\!\sim^{i}_{\mathsf{MCN}} \varphi \quad \text{iff} \quad \varphi \in \textstyle\bigcap_{\mathcal{N}' \in \mathsf{MCN}(\mathcal{K})} \mathsf{out}_i(\mathcal{F}, \mathcal{N}').$$

As before, φ is understood as an agent's obligation given $\mathcal{K}$.

In this section, we investigate alternative ways of defining nonmonotonic entailments, extending those in (Makinson and van der Torre 2001). We do so by utilizing the defined inconsistency measures. The basic idea is that bringing about φ is *recommended* in case $\neg\varphi$ would render $\mathcal{K}$ more inconsistent. The aim of such recommendations is to minimize violations and conflicts (Royce Sadler 1984) (cf. Section 1). Below, we also consider some alternatives for corresponding entailment relations, incorporating two ways of how recommendations can be generated. Namely, we say φ is *recommended* whenever:

1. adding φ as a fact (i.e., a hypothetical consequence of the agent's conduct), or
2. putting φ as a constraint on the agent's conduct

leads to less conflicts.

Definition 6. Let $\mathcal{I}$ be an inconsistency measure for a fixed out-function, and let $x \in \{f, c\}$. Let $\mathcal{K}$ be a normative knowledge base and φ a formula in $\mathcal{L}$.

- **Negative marginal contribution.**
  $$\mathcal{K} \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}} \varphi \quad \text{iff} \quad \mathcal{I}(\mathcal{K} \oplus_x \neg\varphi) > \mathcal{I}(\mathcal{K})$$
  In words, φ follows from $\mathcal{K}$ if adding $\neg\varphi$ to the knowledge base strictly increases the inconsistency of $\mathcal{K}$.
- **Positive marginal contribution.**
  $$\mathcal{K} \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc+}} \varphi \quad \text{iff} \quad \mathcal{I}(\mathcal{K} \oplus_x \varphi) < \mathcal{I}(\mathcal{K})$$
  In words, φ follows from $\mathcal{K}$ if adding φ to the knowledge base strictly decreases the inconsistency of $\mathcal{K}$.
- **Contrastive marginal contribution.**
  $$\mathcal{K} \mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}} \varphi \quad \text{iff} \quad \mathcal{I}(\mathcal{K} \oplus_x \neg\varphi) > \mathcal{I}(\mathcal{K} \oplus_x \varphi)$$
  In words, φ follows from $\mathcal{K}$ if adding $\neg\varphi$ to the knowledge base leads to more inconsistency than adding φ.
- **Subtractive marginal contribution.**
  $$\mathcal{K} \mid\!\sim^{\mathcal{I},x}_{\mathsf{smc}} \varphi \quad \text{iff} \quad \mathcal{I}(\mathcal{K} \oplus_x \varphi) < \mathcal{I}(\mathcal{K} \ominus_x \neg\varphi)$$
  In words, φ follows from $\mathcal{K}$ if adding φ to $\mathcal{K}$ strictly decreases the inconsistency compared to removing $\neg\varphi$ from $\mathcal{K}$.
- **Remainder marginal contribution.**
  $$\mathcal{K} \mid\!\sim^{\mathcal{I},x}_{\mathsf{rmc}} \varphi \quad \text{iff} \quad (\forall \mathcal{K}' \in \mathcal{K}\bot\varphi)\ \mathcal{I}(\mathcal{K} \oplus_x \varphi) < \mathcal{I}(\mathcal{K}')$$
  In words, φ follows from $\mathcal{K}$ if adding φ to $\mathcal{K}$ strictly decreases the inconsistency, compared to the result of removing φ from $\mathcal{K}$ in terms of the remainder sets.

Remark 1. Entailments based on inconsistency measures have also been defined on the basis of flat knowledge bases of the type $\langle \mathcal{F} \rangle$. E.g., in (Salhi 2021) an entailment similar to $\mid\!\sim^{\mathcal{I},f}_{\mathsf{mc-}}$ is defined by: $\mathcal{F} \mid\!\sim \varphi$ iff $\mathcal{I}(\mathcal{F} \cup \{\neg\varphi\}) > \mathcal{I}(\mathcal{F})$. Similar for the idea based on contrastive marginal contributions. See also (Liu, Besnard, and Doutre 2023) for a systematic study of these kinds of entailment relations.

Before studying their logical properties (Sect. 5), we illustrate these entailment relations by means of some examples.

Example 11. Consider a knowledge base $\mathcal{K}$ where $\mathcal{F} = \mathcal{C} = \emptyset$ and $\mathcal{N} = \{(\top, \neg p), (\top, p \wedge u), (\top, p \wedge v)\}$. Table 1 below indicates some inferences for $\mathsf{out} \in \{\mathsf{out}_1, \mathsf{out}_2, \mathsf{out}_3, \mathsf{out}_4\}$, $\varphi \in \{p, \neg p, \neg\neg p\}$ and $\star \in \{\mathsf{mc-}, \mathsf{mc+}, \mathsf{cmc}, \mathsf{rmc}\}$. Among others, this example demonstrates the productivity[8] of some of our entailments, as e.g., some of them allow to infer p (the obligation according to the majority of the norms in $\mathcal{N}$), while $\mathcal{K} \not\mid\!\sim_{\mathsf{MCN}} p$. Traditional approaches based on MCN would neither conclude p nor $\neg p$ in such dilemma scenarios. However, the normative system $\mathcal{K}$ recommends p because, even though there is a symmetric conflict between p and $\neg p$, the knowledge base contains strictly more normative reasons (i.e., norms) in favour of performing p than $\neg p$. In other words, by following the recommendation the agent minimizes the amount of violations: by performing p the agent violates one norm whereas by performing $\neg p$ she violates two. This approach, reducing the amount of violations, is also related to the idea of accrual of reasons (Gordon and Walton 2016; Hirose 2014).

	$\mathcal{I}_{\#}$	$\mathcal{I}_{\mathsf{prob}}$	$\mathcal{I}_{\mathsf{MRN}}$	$\mathcal{I}_{\mathsf{MCN}}$
$\mathcal{K}$	2	3	1	1
$\mathcal{K} \oplus_c p$	1	1	1	1
$\mathcal{K} \oplus_c \neg\neg p$	1	1	1	1
$\mathcal{K} \oplus_c \neg p$	2	2	2	2
$\mathcal{K} \oplus_f \varphi$	2	3	1	1
$\mathcal{K}' \in \mathcal{K} \bot_c \neg p$	2	3	1	1
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{mc}-} p$			✓	✓
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{mc}-} \neg p$				
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{mc}+} p$	✓	✓		
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{mc}+} \neg p$		✓		
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{cmc}} p$	✓	✓	✓	✓
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{cmc}} \neg p$				
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{rmc}} p$	✓	✓		
$\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{rmc}} \neg p$		✓		
$\mathcal{K} \mid\!\sim^{\mathcal{I},f}_{\star} p$				
$\mathcal{K} \mid\!\sim^{\mathcal{I},f}_{\star} \neg p$				

Table 1: Inferences for Example 11

The following example shows that the idea of subtractive marginal contribution is highly sensitive to the syntax of the removed formula, unlike remainder marginal contribution.

Example 12. Consider $\mathcal{K}$ with $\mathcal{N} = \{(\top, p), (\top, \neg p)\}$, $\mathcal{F} = \emptyset$ and $\mathcal{C} = \{\neg p\}$. Let $\mathcal{I} = \mathcal{I}_{\mathsf{prob}}$. Then, $\mathcal{I}(\mathcal{K}) = \mathcal{I}(\mathcal{K} \oplus_c \neg p) = \mathcal{I}(\mathcal{K} \ominus_c \neg\neg\neg p) = \mathcal{I}(\mathcal{K} \oplus_c \neg\neg\neg p) = 1$, while $\mathcal{I}(\mathcal{K} \ominus_c \neg p) = 2$. Thus, $\mathcal{K} \mid\!\sim^{\mathcal{I},c}_{\mathsf{smc}} \neg p$, although $\mathcal{K} \not\mid\!\sim^{\mathcal{I},c}_{\mathsf{smc}} \neg\neg\neg p$. Also, $\mathcal{K} \bot_c \neg p = \mathcal{K} \bot_c \neg\neg\neg p = \{\langle \mathcal{F}, \emptyset, \mathcal{N} \rangle\}$ and $\mathcal{I}(\langle \mathcal{F}, \emptyset, \mathcal{N} \rangle) = 2$, hence $\mathcal{K} \mid\!\sim^{\mathcal{I}_{\#},c}_{\mathsf{rmc}} \neg p$ and $\mathcal{K} \mid\!\sim^{\mathcal{I}_{\#},c}_{\mathsf{rmc}} \neg\neg\neg p$.

The entailments based on $\oplus_f$ have a rather peculiar behavior, as can be seen from the following examples.

[8] An entailment relation is *productive* if it is more liberal, i.e., it subsumes existing normative entailments.

Example 13. Let $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$ where $\mathcal{F} = \mathcal{C} = \emptyset$ and $\mathcal{N} = \{(\top, p)\}$. Then, for $\mathcal{I} \in \{\mathcal{I}_{\#}, \mathcal{I}_{\mathsf{prob}}, \mathcal{I}_{\mathsf{MRN}}, \mathcal{I}_{\mathsf{MCN}}\}$ and $\star \in \{\mathsf{mc}-, \mathsf{mc}+, \mathsf{cmc}, \mathsf{rmc}\}$, $\mathcal{I}(\mathcal{K} \oplus_f p) = \mathcal{I}(\mathcal{K}) = \mathcal{I}(\mathcal{K} \oplus_f \neg p) = \mathcal{I}(\mathcal{K} \ominus_f p) = \mathcal{I}(\mathcal{K} \ominus_f \neg p) = 0$. Thus, $\mathcal{K} \not\mid\!\sim^{\mathcal{I},f}_{\star} p$. This seems highly counter-intuitive.

Example 14. Let $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$ where $\mathcal{F} = \mathcal{C} = \emptyset$ and $\mathcal{N} = \{(\neg r, p), (\neg r, \neg p), (\top, u)\}$. We consider, for instance $\mathcal{I} = \mathcal{I}_{\mathsf{prob}}$. Then, $\mathcal{I}(\mathcal{K}) = \mathcal{I}(\mathcal{K} \oplus_f r) = 0$, while $\mathcal{I}(\mathcal{K} \oplus_f \neg r) = 2$. So, $\mathcal{K} \mid\!\sim^{\mathcal{I},f}_{\mathsf{mc}-} r$ and $\mathcal{K} \mid\!\sim^{\mathcal{I},f}_{\mathsf{cmc}} r$. Nevertheless, like in the previous example, $\mathcal{K} \not\mid\!\sim^{\mathcal{I},f}_{\mathsf{mc}-} u$ and $\mathcal{K} \not\mid\!\sim^{\mathcal{I},f}_{\mathsf{cmc}} u$.

Remark 2. Example 14 reveals that entailments based on $\oplus_f$ recommend bringing about states that do not trigger conflicts and that such entailments are not productive.

5 Logical Properties

The variety of entailment relations introduced in Definition 6, as well as the various possibilities to combine each one of these entailments with inconsistency measures such as the ones considered in Section 3, create great degrees of freedom in choosing the intended normative inference system. In this section, we study some of the logical properties of these inference relations, to gain a clearer understanding of their strengths and weaknesses. Then, we consider in some more details particular inference relations that appear to satisfy several desired properties.

First, we adapt to our setting some known properties for evaluating entailments for defeasible reasoning, in the spirit of the general patterns for nonmonotonic reasoning, presented by Kraus, Lehman, and Magidor (1990).

Definition 7. Given a knowledge base $\mathcal{K}$ and the consequence relation $\vdash$ of classical logic, we identify the following list of $\mid\!\sim$-properties:[9]

- **Direct consistency (dcon):** $\mathcal{K} \not\mid\!\sim \neg\varphi$, if $\mathcal{K} \mid\!\sim \varphi$.
- **Consistency (con):** $\{\varphi \mid \mathcal{K} \mid\!\sim \varphi\} \nvdash \bot$.
- **Introduction of conjunction (and):** $\mathcal{K} \mid\!\sim \varphi \wedge \psi$, if $\mathcal{K} \mid\!\sim \varphi$ and $\mathcal{K} \mid\!\sim \psi$.
- **Logical equivalence (le):** $\mathcal{K} \mid\!\sim \varphi$ iff $\mathcal{K} \mid\!\sim \psi$ for $\vdash \varphi \leftrightarrow \psi$.
- **Right weakening (we):** $\mathcal{K} \mid\!\sim \psi$ implies $\mathcal{K} \mid\!\sim \varphi$, if $\psi \vdash \varphi$.
- **Free formulas (free):** $\mathcal{K} \mid\!\sim \varphi$ if $\varphi \in \mathsf{out}(\mathcal{F}, \bigcap \mathsf{MCN}(\mathcal{K}))$.
- **Strong inference (strong):** $\mathcal{K} \mid\!\sim \varphi$, if $\varphi \in \mathsf{out}(\mathcal{F}, \mathcal{N}')$ for all $\mathcal{N}' \in \mathsf{MCN}(\mathcal{K})$.
- **Cautious monotonicity (cm):** $\mathcal{K} \mid\!\sim \psi$, if $\mathcal{K} \mid\!\sim \varphi$ and $\mathcal{K} \oplus \varphi \mid\!\sim \psi$.
- **Cautious cut (ct):** $\mathcal{K} \mid\!\sim \psi$, if $\mathcal{K} \mid\!\sim \varphi$ and $\mathcal{K} \oplus \varphi \mid\!\sim \psi$.
- **Cautious reflexivity (cref):** $\mathcal{K} \oplus \varphi \mid\!\sim \varphi$ if $\varphi \nvdash \bot$.[10]

Theorem 1. *Table 2 depicts some relations between the properties of the inconsistency measures and those of the entailment relations based on them.*

[9] Each property quantifies over all $\mathcal{K} \in \mathbb{K}$ and formulas φ, ψ.

[10] Although (cref) is a well-known property, we note that it is not desirable for normative entailments and, indeed, I/O logics are designed to not satisfy (cref) (Makinson and van der Torre 2001). Intuitively, if something is factually (and so consistently) the case, this does not necessarily mean that it ought to be the case.

Property	Entailment(s)	Condition(s) on $\mathcal{I}$
(dcon)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$	none
(dcon)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-bi-innocent choice
(dcon)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc+}}$	$\oplus_x$-bi-guilty choice
(dcon)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{smc}}$	$\oplus_x$-bi-subtractive $\ominus$-innocent choice
(dcon)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{rmc}}$	$\oplus$-bi-subtractive $\perp$-innocent choice
(con)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$	$\oplus_x$-contrastive innocent choice
(con)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-innocent choice
(con)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc+}}$	$\oplus_x$-guilty choice
(con)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{smc}}$	$\oplus_x$-subtractive $\ominus$-innocent choice
(con)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{rmc}}$	$\oplus_x$-subtractive $\perp$-innocent choice
(free)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$, $\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-free independence
(strong)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$, $\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-strong independence
(cm)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$, $\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-monotonicity, $\oplus_x$-bi-innocent choice
(ct)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-upper bounding
(ct)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$	$\oplus_x$-bi-innocent choice, $\oplus_x$-upper bounding
(le)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$, $\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	none
(le)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc+}}$, $\mid\!\sim^{\mathcal{I},x}_{\mathsf{rmc}}$	none
(we)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$, $\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-weakening
(we)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc+}}$	$\oplus_x$-weakening
(and)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$	$\oplus_x$-consistency by cases
(and)	$\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$	$\oplus_x$-consistency by cases, $\oplus_x$-monotonicity $\oplus_x$-bi innocent choice

Table 2: Conditions on the inconsistency measures for satisfying logical properties of the corresponding entailment, for $x \in \{c, f\}$.

Proof. Below, we demonstrate some cases. Let $x \in \{c, f\}$.

(dcon) and $\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$: The claim directly follows by the definition of the entailment relation.

(con) and $\mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}$: Assume (con) fails. So, for some $\mathcal{K}$ there are $\varphi_1, \ldots, \varphi_n$ such that $\mathcal{K} \mid\!\sim \varphi_i$ for all $i = 1, \ldots, n$ and $\varphi_1, \ldots, \varphi_n \vdash \perp$. Without loss of generality, suppose that $\{\varphi_1, \ldots, \varphi_n\}$ is a minimal conflict. Then, $\{\neg\varphi_1, \ldots, \neg\varphi_n\}$ is a minimal truth set. Also, $\mathcal{I}(\mathcal{K} \oplus_x \neg\varphi_i) > \mathcal{I}(\mathcal{K})$. This is a violation of ($\oplus_x$-innocent choice).

(free) and $\mid\!\sim \in \{\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}\}$: Consider a formula $\varphi \in \mathsf{out}(\mathcal{F}, \bigcap \mathsf{MCN}(\mathcal{K}))$ and suppose that (free independence) holds. Then, $\mathcal{I}(\mathcal{K} \oplus_x \varphi) \leq \mathcal{I}(\mathcal{K}) < \mathcal{I}(\mathcal{K} \oplus_x \neg\varphi)$. This immediately implies that $\mathcal{K} \mid\!\sim \varphi$. The converse is similar.

(strong) and $\mid\!\sim \in \{\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}\}$: Similar to the previous case.

(le) and $\mid\!\sim \in \{\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc+}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{rmc}}\}$: These claims follow from the following fact. Let $\mathcal{K} = \langle \mathcal{F}, \mathcal{C}, \mathcal{N} \rangle$, let φ and ψ be logically equivalent. Then, for $i \in \{1, \ldots, 4\}$, (i) $\mathsf{out}_i(\mathcal{N}, \mathcal{F} \cup \{\varphi\}) = \mathsf{out}_i(\mathcal{N}, \mathcal{F} \cup \{\psi\})$, (ii) if $\Xi \in \{\mathsf{MIN}_i, \mathsf{MRN}_i, \mathsf{MCN}_i\}$ and $x \in \{c, f\}$, then $\Xi(\mathcal{K} \oplus_x \varphi) = \Xi(\mathcal{K} \oplus_x \psi)$, (iii) $\mathcal{K}\perp\varphi = \mathcal{K}\perp\psi$. □

We turn now to some specific cases. The following examples provide some negative results (due to reasons of space we do not provide an exhaustive list of counter-examples).

Example 15. Let $\mathcal{K} = \langle \emptyset, \emptyset, \emptyset \rangle$. Then, $\mathcal{K} \oplus_x p \not\mid\!\sim p$ for each $\mid\!\sim \in \{\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc+}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{smc}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{rmc}}\}$ and $x \in \{c, f\}$.

Example 16. Example 11 illustrates a failure of (dcon) for $\mid\!\sim^{\mathcal{I}_{\mathsf{prob}},c}_{\mathsf{mc+}}$ and $\mid\!\sim^{\mathcal{I}_{\mathsf{prob}},x}_{\mathsf{smc}}$. Example 12 shows a failure of (le) for $\mid\!\sim^{\mathcal{I}_{\#},c}_{\mathsf{rmc}}$. Example 13 illustrates a failure of (free) for $\oplus_f$-based entailments.

In contrast to the previous examples, we note that careful choices of the inconsistency measures and merging operators *do lead* to robust entailment relations. Namely, by Proposition 2 one can show that $\mathcal{I}_{\mathsf{MRN}}$-based entailments are such relations. The following theorem vindicates this.

Theorem 2. *Let $\mid\!\sim \in \{\mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}, \mid\!\sim^{\mathcal{I},x}_{\mathsf{mc-}}\}$, where $\mathcal{I} = \mathcal{I}_{\mathsf{MRN}}$, $x \in \{c, f\}$, and $\mathsf{out} \in \{\mathsf{out}_1, \ldots, \mathsf{out}_4\}$. Then, $\mid\!\sim$ satisfies*

- *(we) and (le);*
- *(free) and (strong) for $x = c$;*
- *(dcon) for $\mid\!\sim = \mid\!\sim^{\mathcal{I},x}_{\mathsf{cmc}}$;*
- *(dcon), (con), (and), (cm) and (ct) for $x = c$;*
- *(dcon), (con), (and), (cm) and (ct) for $x = f$ and $\mathsf{out} \in \{\mathsf{out}_2, \mathsf{out}_4\}$.*

Proof. Follows from Proposition 2 and Theorem 1. □

6 Related Work, Conclusion, and Outlook

Related work. Applications of inconsistency measures are diverse, ranging from network intrusion detection (McAreavey et al. 2011), to conflict management in ontologies (Ma et al. 2007), rule-based expert systems in internal medicine (Picado-Muiño 2011), as well as reasoning with temporal and spatial information (Condotta, Raddaoui, and Salhi 2016), and software requirements engineering (Mu, Liu, and Jin 2012). There is a considerable amount of work on inconsistency measures (see, e.g., the surveys in (Grant and Martinez 2018; Thimm and Wallner 2019)) and how they can be used for drawing conclusions in the context of propositional logics (Jabbour, Ma, and Raddaoui 2014; Mu, Wang, and Wen 2014; Jabbour et al. 2016; Salhi 2021; Liu, Besnard, and Doutre 2023), probabilistic conditional logic (Thimm 2013), non-monotonic logic (Ulbricht, Thimm, and Brewka 2018), relational databases (Livshits et al. 2021; Parisi and Grant 2023), and business rule bases (Corea and Thimm 2020). In the context of a normative setting, the properties of entailment relations that are induced by inconsistency measures are studied in this paper relative to the well-known postulates of defeasible inference in (Kraus, Lehmann, and Magidor 1990).

A number of formalisms have been proposed for reasoning about inconsistencies in the light of normative conflicts (e.g., (Beirlaen, Straßer, and Meheus 2013)), including some that are based on I/O logics. Most of these approaches are based on the identification of maximal consistent subsets of norms, e.g., extended with priority orderings (Liao et al. 2018; Parent 2011). Interestingly, (Parent 2011) shows a strong connection between I/O logic and Reiter's default logic. Furthermore, I/O logic knows extensions with permissive norms which are, e.g., derivable through absence of norms to the contrary (Olszewski, Parent, and van der Torre 2021). Recently, scholars became increasingly interested in characterizations of I/O reasoning, and normative reasoning in general, using formal argumentation (Peirera et al. 2017; Liao et al. 2018; Straßer and Arieli 2019; van Berkel and Straßer 2022; Arieli, van Berkel, and Straßer 2024). Nevertheless, these approaches, whether through traditional MCN approaches or instantiated argumentation frameworks, do not allow for the quantification of norm conflicts.

Conclusion. To the best of our knowledge, this is the first study on the integration of inconsistency measures in the context of normative reasoning. Such a combination provides nonmonotonic inference systems that are adequate to the specific characteristics of normative systems, combining factual contexts, norms, and constraints. In order to do so, we enhanced the well-established I/O formalism with various inconsistency measures. Accordingly, we provided an in-depth analysis of these measures and the resulting entailment relations, the latter are interpreted to yield recommendations to agents. Figure 1 provides a schematic illustration of our conceptual contribution (encased by a dashed line) to the I/O paradigm and defeasible reasoning in general.

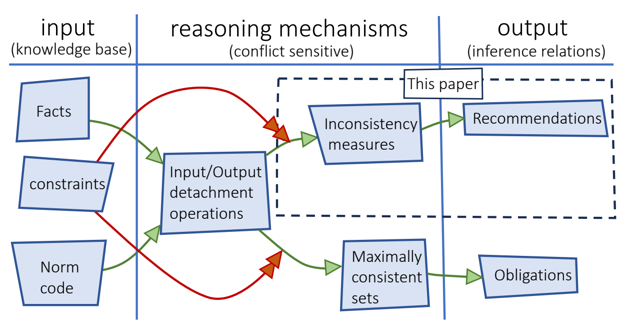

Figure 1: A schematic description of the framework.

This paper utilizes inconsistency measurement (a) to provide a postulate-driven study of normative reasoning, and (b) to introduce a novel approach to nonmonotonic entailment. Accordingly, this work is relevant not only to normative reasoning per-se, but has links to other disciplines of knowledge representation and reasoning, such as formalisms for consistency tolerance by inconsistency measurements, belief revision and causal reasoning, and nonmonotonic logic in general.

Like related works on inconsistency measures for deductive systems (see, e.g., (Liu, Besnard, and Doutre 2023)), the conclusions regarding the quality of the inferences obtained in our case are not unequivocal. As we have seen, combinations of specific measures and entailment relations have several advantages as well as shortcomings over other combinations, or in comparison to other accepted normative inference systems that are not based on inconsistency measures. Since this work does not pretend to be an exhaustive review of all the accepted measures and all the resulting relations induced by them, we focused on various properties of the inconsistency measures, in particular $\mathcal{I}_{\mathsf{MRN}}$, and their effect on commonly studied properties of the resulting (normative) entailment relations. We highlight three takeaways:

1. Revision by constraints ($\oplus_c$) is in general superior to revision by facts ($\oplus_f$);
2. Many inconsistency measures from the literature do not perform well as a basis for normative reasoning;
3. Nevertheless, some entailment relations induced by the inconsistency measure $\mathcal{I}_{\mathsf{MRN}}$ offer an attractive approach to normative reasoning. Besides satisfying central properties of defeasible inference, they also allow for a more fine-grained analysis of a normative system than standard approaches based on maximal consistent subsets.

Outlook. Our results pave the way for various future research directions such as investigating local inconsistency measures which identify in how many minimally inconsistent norm sets a norm plays a part. Such local measures may prove useful for normative system revision procedures, e.g., by identifying norms that are the most problematic in generating conflicts. In this respect, we will study the properties of inconsistency measures through adding and retracting norms to a normative system.

We also aim to investigate computational complexity, which we conjecture is the same as the complexity of reasoning of maximal consistent sets, since the former is based (via a linear transformation) on the latter.

Another future work avenue would be to differentiate and study inconsistency measures for different types of conflicts, such as contrary-to-duty and specificity cases. In relation to this, we also intend to include the norm synthesis problem: identifying the set of norms that prevents conflicting states (Shoham and Tennenholtz 1995).

Last, while this paper adopted the I/O formalism at its base, a natural topic for further exploration is to see what other defeasible normative reasoning mechanisms may benefit from reasoning with various inconsistency measures and entailment relations; e.g., defeasible deontic logics (Nute 1997; Olivieri et al. 2024).

Acknowledgements

We thank the reviewers for their helpful comments. The first author was supported by the Israel Science Foundation, grant No. 550/19. The second author and the fourth author were supported by the Deutsche Forschungsgesellschaft (DFG) under the grant LoDeX: 511915728. The third author was supported by the Agence Nationale de la Recherche (ANR) under grant EXPIDA: ANR-22-CE23-0017.

References

Arieli, O.; van Berkel, K.; and Straßer, C. 2024. Defeasible normative reasoning: A proof-theoretic integration of logical argumentation. In *AAAI*, 10450–10458.

Beirlaen, M.; Straßer, C.; and Meheus, J. 2013. An inconsistency-adaptive deontic logic for normative conflicts. *J. Philos. Log.* 42(2):285–315.

Benzmüller, C.; Parent, X.; and van der Torre, L. W. N. 2020. Designing normative theories for ethical and legal reasoning: Logikey framework, methodology, and tool support. *Artif. Intell.* 287:103348.

van Berkel, K., and Straßer, C. 2022. Reasoning with and about norms in logical argumentation. In *COMMA*, 332 – 343.

Besnard, P. 2014. Revisiting postulates for inconsistency measures. In *JELIA*, 383–396.

Bochman, A. 2014. Dynamic causal calculus. In *KR*, 188–197.

Bona, G. D.; Grant, J.; Hunter, A.; and Konieczny, S. 2019. Classifying inconsistency measures using graphs. *J. Artif. Intell. Res.* 66:937–987.

Chopra, A.; van der Torre, L.; and Verhagen, H. 2018. *Handbook of Normative Multiagent Systems*. College Publications.

Ciabattoni, A., and Rozplokhas, D. 2023. Streamlining input/output logics with sequent calculi. In *KR*, 146–155.

Ciabattoni, A.; Parent, X.; and Sartor, G. 2021. A kelsenian deontic logic. In *Frontiers in Artificial Intelligence and Applications, Legal Knowledge and Information Systems*, volume 346. IOS Press. 141–150.

Condotta, J.; Raddaoui, B.; and Salhi, Y. 2016. Quantifying conflicts for spatial and temporal information. In *KR*, 443–452.

Corea, C., and Thimm, M. 2020. Towards inconsistency measurement in business rule bases. In *ECAI*, volume 325, 704–711.

Gabbay, D.; Horty, J. F.; Parent, X.; van der Meyden, R.; and van der Torre, L. 2013. *Handbook of Deontic Logic and Normative Systems, Volume 1*. United Kingdom: College Publications.

Gabriel, I. 2020. Artificial intelligence, values, and alignment. *Minds and machines* 30(3):411–437.

Gordon, T. F., and Walton, D. 2016. Formalizing balancing arguments. In *COMMA*, 327–338.

Grant, J., and Hunter, A. 2023. Semantic inconsistency measures using 3-valued logics. *Int. J. Approx. Reason.* 156:38–60.

Grant, J., and Martinez, M. V. 2018. *Measuring Inconsistency in Information*. College Publications.

Hilpinen, R., and McNamara, P. 2013. Deontic logic: A historical survey and introduction. In Gabbay, D.; Horty, J. F.; Parent, X.; van der Meyden, R.; and van der Torre, L., eds., *Handbook of Deontic Logic and Normative Systems, Volume 1*. College Publications. 3–136.

Hirose, I. 2014. *Moral Aggregation*. Oxfods University Press.

Hunter, A., and Konieczny, S. 2010. On the measure of conflicts: Shapley inconsistency values. *Artif. Intell.* 174(14):1007–1026.

Jabbour, S.; Ma, Y.; Raddaoui, B.; Sais, L.; and Salhi, Y. 2016. A MIS partition based framework for measuring inconsistency. In *KR*.

Jabbour, S.; Ma, Y.; and Raddaoui, B. 2014. Inconsistency measurement thanks to MUS decomposition. In *AAMAS*.

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

Lellmann, B.; Gulisano, F.; and Ciabattoni, A. 2021. Mīmāṃsā deontic reasoning using specificity: a proof theoretic approach. *Artificial Intelligence and Law* 29(3):351–394.

Liao, B.; Oren, N.; van der Torre, L.; and Villata, S. 2018. Prioritized norms in formal argumentation. *Journal of Logic and Computation* 29(2):215–240.

Liffiton, M. H., and Sakallah, K. A. 2008. Algorithms for computing minimal unsatisfiable subsets of constraints. *J. Autom. Reason.* 40(1):1–33.

Liu, X.; Besnard, P.; and Doutre, S. 2023. Paraconsistent inference relations induced from inconsistency measures. *Int. J. Approx. Reason.* 152:183–197.

Livshits, E.; Kochirgan, R.; Tsur, S.; Ilyas, I. F.; Kimelfeld, B.; and Roy, S. 2021. Properties of inconsistency measures for databases. In *SIGMOD*, 1182–1194.

Ma, Y.; Qi, G.; Hitzler, P.; and Lin, Z. 2007. Measuring inconsistency for description logics based on paraconsistent semantics. In *ECSQARU*, 30–41.

Makinson, D., and van der Torre, L. 2000. Input/Output logics. *Journal of Philosophical Logic* 383–408.

Makinson, D., and van der Torre, L. 2001. Constraints for Input/Output logics. *Journal of Philosophical Logic* 30(2):155–185.

Martinez, M. V.; Molinaro, C.; Subrahmanian, V. S.; and Amgoud, L. 2013. *A General Framework for Reasoning On Inconsistency*. Springer Briefs in Computer Science. Springer.

McAreavey, K.; Liu, W.; Miller, P.; and Mu, K. 2011. Measuring inconsistency in a network intrusion detection rule set based on snort. *Int. J. Semantic Comput.* 5(3):281–322.

Mu, K.; Liu, W.; and Jin, Z. 2012. A blame-based approach to generating proposals for handling inconsistency in software requirements. *Int. J. Knowl. Syst. Sci.* 3(1):1–17.

Mu, K.; Wang, K.; and Wen, L. 2014. Approaches to measuring inconsistency for stratified knowledge bases. *Int. J. Approx. Reason.* 55(2):529–556.

Nute, D. 1997. *Defeasible deontic logic*, volume 263. Springer Science & Business Media.

Olivieri, F.; Governatori, G.; Cristani, M.; Rotolo, A.; and Sattar, A. 2024. Deontic meta-rules. *Journal of Logic and Computation* 34(2):261–314.

Olszewski, M.; Parent, X.; and van der Torre, L. 2021. Input/output logic with a consistency check-the case of permission. In *DEON*, 358–375. College publications.

Parent, X., and van der Torre, L. 2013. Input/Output logic. In Gabbay, D.; Horty, J. F.; Parent, X.; van der Meyden, R.;

and van der Torre, L., eds., *Handbook of Deontic Logic and Normative Systems, Volume 1*. College Publications. 499–544.

Parent, X., and van der Torre, L. 2018. Input/output logics with a consistency check. In *DEON*. College Publications.

Parent, X. 2011. Moral particularism in the light of deontic logic. *Artificial Intelligence and Law* 19(2):75–98.

Parisi, F., and Grant, J. 2023. On measuring inconsistency in definite and indefinite databases with denial constraints. *Artif. Intell.* 318:103884.

Peirera, C. d. C.; Tettamanzi, A. G.; Villata, S.; Liao, B.; Malerba, A.; Rotolo, A.; and van der Torre, L. 2017. Handling norms in multi-agent system by means of formal argumentation. *IfCoLog Journal of Logics and Their Applications* 4(9):1–35.

Picado-Muiño, D. 2011. Measuring and repairing inconsistency in probabilistic knowledge bases. *International Journal of Approximate Reasoning* 52(6):828–840.

Reiter, R. 1987. A theory of diagnosis from first principles. *Artif. Intell.* 32(1):57–95.

Royce Sadler, D. 1984. Evaluation and the logic of recommendations. *Evaluation Review* 8(2):261–268.

Salhi, Y. 2021. Inconsistency measurement for paraconsistent inference. In *IJCAI*, 2033–2039.

Shoham, Y., and Tennenholtz, M. 1995. On social laws for artificial agent societies: Off-line design. *Artif. Intell.* 73(1-2):231–252.

Straßer, C., and Arieli, O. 2019. Normative reasoning by sequent-based argumentation. *J. Log. Comput.* 29(3):387–415.

Thimm, M., and Wallner, J. P. 2019. On the complexity of inconsistency measurement. *Artif. Intell.* 275:411–456.

Thimm, M. 2013. Inconsistency measures for probabilistic logics. *Artif. Intell.* 197:1–24.

Thimm, M. 2014. Tweety: A comprehensive collection of java libraries for logical aspects of artificial intelligence and knowledge representation. In *KR*.

Thimm, M. 2017. On the compliance of rationality postulates for inconsistency measures: A more or less complete picture. *Künstliche Intell.* 31(1):31–39.

Ulbricht, M.; Thimm, M.; and Brewka, G. 2018. Measuring strong inconsistency. In *AAAI*, 1989–1996.