

A Game-Theoretic Perspective on Inconsistency Handling

Yakoub Salhi

Univ. Artois, CNRS, CRIL, F-62300 Lens, France
salhi@cril.fr

Abstract

This paper introduces a game-theoretic framework for restoring consistency in propositional bases. The process is modeled as an interactive dialogue between two agents: a Proponent, who seeks to isolate a unique, consistent subset by posing strategic questions, and an Opponent, who aims to obstruct that goal through adversarial responses. We show that this framework provides a foundation for quantifying the effort involved in restoring consistency, revealing a connection between this effort and entropy in information theory. Focusing on the case where consistency is achieved by isolating a single maximal consistent subset, we establish links between the structure and number of such subsets and the existence of winning strategies. Finally, we demonstrate how the quantified restoration effort can serve as a basis for measuring inconsistency.

1 Introduction

Inconsistencies in knowledge-based applications stem from multiple factors, such as conflicting information sources, noisy data, context dependency, and dynamic changes. These inconsistencies cause important challenges for reasoning processes in artificial intelligence, especially in the field of knowledge representation and reasoning. To address these challenges, a variety of concepts, approaches, and tools have been developed to analyze and handle inconsistency, including argumentation frameworks (e.g. [Besnard and Hunter, 2008]), inconsistency measures (e.g. [Hunter and Konieczny, 2010; Thimm, 2018]), and non-explosive inference relations (e.g. [Rescher and Manor, 1970; Priest, 1991]).

A common approach to dealing with inconsistencies is to modify the initial knowledge base to obtain a consistent one that can be used with classical reasoning; a process we call consistency restoration. This modification can involve forgetting variables [Lin and Reiter, 1994; Lang and Marquis, 2002], ignoring pieces of knowledge [Rescher and Manor, 1970; Benferhat *et al.*, 1995], or applying belief revision and updating operators [Alchourrón *et al.*, 1985; Gärdenfors, 1992; Fermé and Hansson, 2011].

In this work, we model the process of restoring consistency in propositional logic as a dialogue between two play-

ers: the Proponent and the Opponent. The Proponent’s goal is to achieve consistency by identifying a unique, consistent propositional base. To do this, she asks questions to gather necessary information. Conversely, the Opponent, who controls the responses to these questions, aims to prevent her from achieving her goal.

Our game-theoretic framework is inspired by two key sources in the literature. The first is the work of Lorenzen [Lorenzen, 1960] on dialogue games in logic, which provides a semantics grounded in interactive dialogues (see also [Blass, 1992; Pitts and Dybjer, 1997; Fermüller and Ciabattoni, 2003]). In many formulations of game semantics, a game involves two players: the Proponent, whose goal is to prove statements, and the Opponent, who challenges the application of proof rules.

The second source of inspiration is the question-based interpretation of entropy in information theory [Shannon, 1948a; Shannon, 1948b]. This interpretation links the number of questions needed to solve a search problem with the amount of information. For instance, the entropy of a system reflects the minimal number of binary (yes/no) questions required to determine an outcome [van der Lubbe and Hoeve, 1997; Cover and Thomas, 2001].

Dealing with inconsistency through our game-theoretic approach brings several advantages. By considering the problem as a game, we can analyze and develop optimal strategies for agents to resolve conflicts, such as decisions about which pieces of information to question and how to allocate limited resources. Furthermore, introducing the Opponent adds an adversarial dimension that captures uncertainty. This allows us to study how agents can overcome this obstacle to restore consistency.

Our main contributions are as follows. We first introduce a general game-theoretic framework in which consistency restoration is modeled as a series of questions posed by the Proponent and answered by the Opponent. We then define a measure that quantifies the expected cost associated with a Proponent’s strategy for achieving consistency. This provides a quantitative perspective to the process of consistency restoration. Next, we focus on a specific scenario where restoring consistency involves identifying a single maximal consistent subset of the base. In this context, we establish results that link winning strategies to the number and structure of these subsets. Finally, we show how the quantified restora-

tion effort can serve as the basis for defining an inconsistency measure, a function that assesses the severity of conflicts in propositional bases. We study the properties of this measure and examine several related computational aspects.

2 Preliminaries

Let us start by defining the language of classical propositional logic. We begin with a constant \perp , denoting *false*, and a countably infinite set of propositional variables, denoted by Prop . The set of well-formed formulas (wffs), denoted WFF , is the smallest set that satisfies:

- \perp is a wff.
- Every propositional variable $p \in \text{Prop}$ is a wff.
- If ϕ and ψ are wffs, then $(\neg\phi)$ and $(\phi \wedge \psi)$ are wffs.

The additional connectives \vee , \rightarrow , and \leftrightarrow are defined, as usual, using \wedge and \neg .

We use p, q, r (with possible subscripts) to denote propositional variables, and we use Greek letters such as ϕ, ψ, χ to denote formulas. For any syntactic entity X (e.g., a formula or a set of formulas), the set of propositional variables occurring in X is denoted by $\text{var}(X)$.

Given a set of formulas Q , we use \hat{Q} to denote the set $Q \cup \{\neg\phi : \phi \in Q\}$.

A *propositional base* (PB) is a finite subset of WFF .

A *Boolean interpretation*, or simply *interpretation*, is a function ω that assigns a truth value in $\{0, 1\}$ to each formula in WFF such that: $\omega(\neg\phi) = 1 - \omega(\phi)$ and $\omega(\phi \wedge \psi) = \omega(\phi) \times \omega(\psi)$. An interpretation ω is called a *model* of a formula ϕ , written $\omega \models \phi$, if and only if $\omega(\phi) = 1$.

A formula is *consistent* if it has at least one model; otherwise, it is *inconsistent*.

Given a PB K and a formula ϕ , we say that K *entails* ϕ , written $K \vdash \phi$, if, for every interpretation ω , whenever $\omega \models \bigwedge K$ (the conjunction of all formulas in K), it follows that $\omega \models \phi$. Note that $\omega \models \bigwedge \emptyset$ holds for every interpretation ω . When $K = \{\psi\}$ contains a single formula, we may write $\psi \vdash \phi$ instead of $\{\psi\} \vdash \phi$. Two formulas ϕ and ψ are said to be *logically equivalent*, written $\phi \equiv \psi$, if both $\phi \vdash \psi$ and $\psi \vdash \phi$ hold.

Using the compactness theorem, we obtain the following properties: a countable set of formulas S is consistent iff for every finite subset S' of S , $\bigwedge S'$ is consistent; $S \vdash \phi$ if and only if there exists a finite subset S' of S such that $S' \vdash \phi$.

To indicate that a set of formulas S is consistent (resp. inconsistent), we sometimes write $S \not\vdash \perp$ (resp. $S \vdash \perp$).

A *minimal inconsistent subset* (MIS) of a PB K is a subset $M \subseteq K$ that is inconsistent, and such that for every $\phi \in M$, the set $M \setminus \{\phi\}$ is consistent.

A *maximal consistent subset* (MCS) of K is a subset $M \subseteq K$ that is consistent, and such that for every $\phi \in K \setminus M$, the set $M \cup \{\phi\}$ is inconsistent. We use $\text{MCS}(K)$ to denote the set of all MCSes of K .

A formula in a PB K is called *free* if it does not belong to any MIS of K . We denote the set of all free formulas in K by $\text{free}(K)$. Conversely, a formula that is not free is referred to as *problematic*. In particular, the inconsistent formulas in a

PB are always problematic. We use $\text{Inc}(K)$ to denote the set of inconsistent formulas in K .

We define the equivalence relation \sim_b on formulas as follows: $\phi \sim_b \psi$ iff $\phi \equiv \psi$ or $\phi \equiv \neg\psi$.

We use \mathbb{R}_∞^+ to denote the set of positive real numbers equipped with the usual order \leq and extended by an additional greatest element ∞ .

A *binary tree* is a rooted, ordered tree where each node can have at most two children, referred to as the *left child* and *right child*. A binary tree is considered *full* if every node has either 0 or 2 children.

The set of leaf nodes of a tree \mathcal{T} is denoted $\text{Lv}(\mathcal{T})$.

The *depth* of a node v is defined as the length (i.e., the number of edges) in the path from the root of the tree to v . We denote the depth of the node v in the binary tree \mathcal{T} as $d(v, \mathcal{T})$. Further, we use $d(\mathcal{T})$ to denote the depth of \mathcal{T} .

3 Consistency-Restoration Game

3.1 Informal Description

In this work, we represent the process of restoring consistency in a PB as a dialogue between two players: the Proponent and the Opponent. The game is as follows: the Proponent aims to restore consistency by identifying a unique, consistent PB; she does this by asking questions to gather information. The Opponent controls the responses to the Proponent's questions and tries to prevent the Proponent from achieving her goal.

The Proponent should choose questions that will most effectively eliminate uncertainties. Additionally, with a limited effort budget, the Proponent needs to prioritize the most informative questions. The Opponent's consistent responses shape the Proponent's deductions. Indeed, we consider that the Opponent cannot contradict herself.

To illustrate our framework, consider a medical researcher (Proponent) investigating a patient's condition. The researcher's PB initially contains conflicting information about the patient's symptoms, possible diseases, and test results. This inconsistency arises due to contradictory patient data. The researcher conducts medical tests (questions) to resolve these conflicts. The Opponent provides consistent test results grounded in the true state of the patient's condition. We interpret the test answers as originating from the Opponent to emphasize that these responses may present difficulties for the Proponent.

However, what happens if the Proponent cannot access the results of all these tests due to constraints such as a limited laboratory budget? In our framework, we refer to this constraint as the effort budget, where each question is assigned a specific weight representing its cost or effort required.

In addition, what are the potential consistent PBs associated with the responses provided by the Opponent? To address this question, we introduce the concept of a possibility function, denoted as Λ . This function captures the range of possible outcomes and their corresponding consistent PBs derived from the Opponent's answers.

A winning strategy for the Proponent is a systematic plan that outlines how to handle every possible response from the Opponent, ensuring the achievement of a single consistent PB without exceeding the allocated effort budget.

3.2 Formal Description

Game Input. We define a *game input* (GI) as a tuple (K, Q, E, b) , where: K is the considered PB, Q is a countable set of possible questions, each being a wff, $E : Q \rightarrow \mathbb{R}_\infty^+$ is a weight function assigning a number in \mathbb{R}_∞^+ to each question in Q ($E(q)$ corresponds to the effort required to answer q), and $b \in \mathbb{R}_\infty^+$ is the *effort budget*: an upper bound on the total weight of the questions that the Proponent can ask.

Game. A *consistency-restoration game*, or simply a game, associated with the GI (K, Q, E, b) involves two players: the Proponent (P) and the Opponent (O). The player P tries to identify a unique consistent PB by asking questions from Q , while O tries to prevent P from succeeding. At each step: (1) P selects a formula ϕ from Q , and (2) O responds by choosing either ϕ or its negation $\neg\phi$.

More precisely, a game \mathcal{G} of (K, Q, E, b) is a possibly infinite alternating sequence of moves of the form: $P \rightarrow \phi_1 \rightarrow O \rightarrow \psi_1 \rightarrow P \rightarrow \phi_2 \rightarrow O \rightarrow \psi_2 \rightarrow \dots \rightarrow P \rightarrow \phi_k \rightarrow O \rightarrow \psi_k \rightarrow \dots$, where for each index i , $\phi_i \in Q$ and $\psi_i \in \{\phi_i, \neg\phi_i\}$.

In the sequel, we use the following notational conventions for a game \mathcal{G} :

- $\alpha(\mathcal{G})$ denotes the set of O's answers ψ_i ,
- $\kappa(\mathcal{G})$ denotes the set of P's questions ϕ_i ,
- $\Xi(\mathcal{G})$ denotes the total effort $\sum_i E(\phi_i)$.

We impose the following condition on all games:

- *Opponent's Consistency (OC)*: For any game \mathcal{G} , the set of responses from O is consistent; that is, $\alpha(\mathcal{G})$ is consistent.

Possibility Function. A player's success in the game is determined by a mechanism called the Q -possibility function Λ . This function associates each PB K and each consistent set of formulas $A \subseteq \widehat{Q}$ with a set of consistent PBs. In essence, $\Lambda(K, A)$ identifies all possible consistent PBs that P can have and that are compatible with O's responses A . This function is crucial because it helps P determines which beliefs she can have without conflicts in light of O's responses. It guides her toward strategies that lead to a successful outcome.

We require the following rationality postulate on Q -possibility functions:

- *Consistency With Responses (CWR)*: For every PB K and every consistent set of responses $A \subseteq \widehat{Q}$, $S \cup A$ is consistent for any $S \in \Lambda(K, A)$.
- *Identity on Consistent Propositional Bases (ICPB)*: For every consistent PB K , $\Lambda(K, \emptyset) = \{K\}$.
- *Dominance Between Answers (DBA)*: For every PB K , every pair of consistent sets $A, A' \subseteq \widehat{Q}$, if $\bigwedge A \vdash \bigwedge A'$, then $\Lambda(K, A) \subseteq \Lambda(K, A')$.

CWR ensures that O's answers do not contradict any consistent subset that P considers possible. ICPB says that if K itself is already consistent, asking no questions must leave K unchanged as the unique possibility. DBA states that having more conclusive answers ($\bigwedge A \vdash \bigwedge A'$) cannot enlarge the set of possibilities. In other words, stronger sets of answers reduce or maintain, but never expand, the set of possible consistent bases.

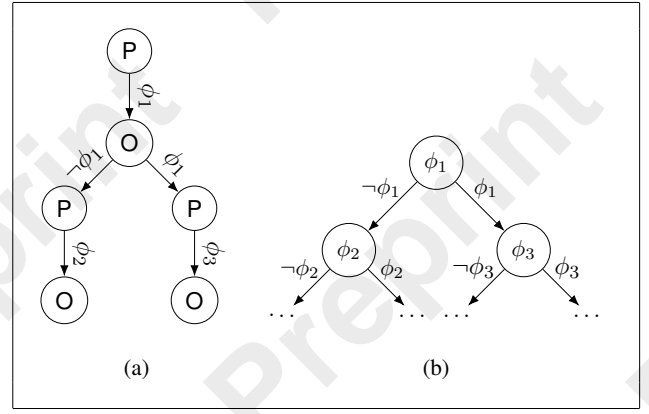


Figure 1: A P-Strategy of a Consistency-Restoration Game

Proponent Λ -Wins. The player P Λ -wins the game \mathcal{G} of (K, Q, E, b) if the game is finite, adheres to the effort budget, and enables P to reduce the possible consistent PBs that are compatible with O's responses to exactly one, according to the Q -possibility function Λ . Formally, P Λ -wins \mathcal{G} when \mathcal{G} is finite, $\Xi(\mathcal{G}) \leq b$ and $|\Lambda(K, \alpha(\mathcal{G}))| = 1$.

Example 1. Consider the PB $K = \{p, q, \neg p \vee \neg q, r\}$ and the set of questions $Q = K$. The effort function E is simply defined by $E(\phi) = 1$ for every $\phi \in Q$, and the effort budget is $b = 2$. The following sequence is a game for (K, Q, E, b) :

$$P \rightarrow \neg p \vee \neg q \rightarrow O \rightarrow \neg p \vee \neg q \rightarrow P \rightarrow p \rightarrow O \rightarrow \neg p$$

If we define a Q -possibility function Λ for any set of O's responses by $\Lambda(K, A) = \{M \in \text{MCS}(K) : M \cup A \not\vdash \perp\}$, we find that the only MCS of K that is consistent with the responses $\{\neg p \vee \neg q, \neg p\}$ is $\{q, \neg p \vee \neg q, r\}$. This means that P Λ -wins this game.

3.3 Winning Strategies

To develop a strategy for player P, we must consider how P reacts to each possible response from O. More precisely, for each question ϕ that P selects, we need to consider both scenarios: when O affirms ϕ and when O negates ϕ . This explains why we represent P's strategy as a full binary tree.

Definition 1 (P-Strategy). A P-strategy of a tuple (K, Q, E, b) is a pair $\sigma = \langle \mathcal{T}, \Phi \rangle$, where:

- \mathcal{T} is a full binary tree in which every branch is finite;
- Φ is a function that assigns to each internal node of \mathcal{T} a formula from Q ;
- for each complete branch (i.e., path from the root to a leaf) (v_1, \dots, v_k) in \mathcal{T} , it holds that $\sum_{i=1}^{k-1} E(\Phi(v_i)) \leq b$.

We then define a context function that captures the O's responses accumulated along each branch of the tree.

Definition 2 (Context Function \mathcal{C}_σ). Given a P-strategy $\sigma = \langle \mathcal{T}, \Phi \rangle$, the function $\mathcal{C}_\sigma(v)$ is defined recursively for each node v in \mathcal{T} as follows:

- If v is the root node, then $\mathcal{C}_\sigma(v) = \emptyset$;

- If v is the left child of a node v' , then $C_\sigma(v) = C_\sigma(v') \cup \{\neg\Phi(v')\}$;
- If v is the right child of a node v' , then $C_\sigma(v) = C_\sigma(v') \cup \{\Phi(v')\}$.

Definition 3 (Λ -Winning P-Strategy). Given a Q -possibility function Λ , a Λ -winning P-strategy is a P-strategy σ where for any leaf node l , $|\Lambda(K, C_\sigma(l))| = 1$ or $C_\sigma(l)$ is inconsistent. A Λ -winning P-strategy σ is said to be pure if, for any leaf node l , $|\Lambda(K, C_\sigma(l))| = 1$.

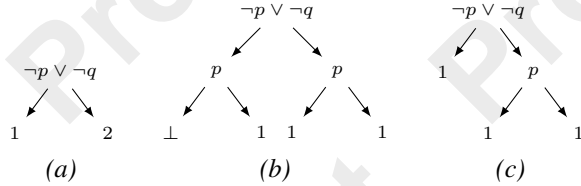
Note that a Λ -winning P-strategy cannot consist of a single node: if no questions are needed, the corresponding strategy tree must be empty.

Given a P-strategy σ , we use $\text{questions}(\sigma)$ to denote the set of questions occurring in σ .

In Figure 1, Subfigure (b) on the right illustrates a tree representing a P-strategy, while Subfigure (a) on the left displays the corresponding player moves.

To visually represent a P-strategy $\sigma = \langle \mathcal{T}, \Phi \rangle$, we depict each internal node of the tree \mathcal{T} with its associated formula; each leaf nodes l is labeled with the number $|\Lambda(K, C_\sigma(l))|$ or \perp , depending on whether $C_\sigma(l)$ is consistent or not.

Example 2. Consider again the GI $g = (K, Q, E, b)$ and the Q -possibility function Λ described in Example 1. In what follows, some P-strategies of g :



The trees (b) and (c) correspond to Λ -winning P-strategies, with tree (c) being the unique pure strategy.

Whether the tree in a P-strategy necessarily contains a finite number of nodes is a natural consideration. This result is, in fact, ensured by König's Lemma: every finitely branching infinite tree contains at least one infinite branch [Rival, 2012].

The following theorem establishes that reasoning about the effort required for consistency restoration can be limited to pure Λ -winning P-strategies.

Theorem 1. For every GI g and every Q -possibility function Λ , if g admits a Λ -winning P-strategy $\sigma = (\mathcal{T}, \Phi)$, then g also admits a pure Λ -winning P-strategy $\sigma' = (\mathcal{T}', \Phi')$, such that there exists an injective function f from $\text{Lv}(\mathcal{T}')$ to $\text{Lv}(\mathcal{T})$ satisfying the following condition: for every $l \in \text{Lv}(\mathcal{T}')$, $C_{\sigma'}(l) \subseteq C_\sigma(f(l))$.

The next proposition shows that in a pure Λ -winning P-strategy, internal nodes on the same branch have distinct, non-equivalent labels (w.r.t. \sim_b), and each label is both consistent and not a tautology.

Proposition 1. Let $g = (K, Q, E, b)$ be a game input, Λ a Q -possibility function, and $\sigma = (\mathcal{T}, \Phi)$ a pure Λ -winning P-strategy for g . Then, the following properties hold:

1. For any two distinct internal nodes v and v' on the same branch in \mathcal{T} , we have $\Phi(v) \not\sim_b \Phi(v')$.
2. For every internal node v in \mathcal{T} , $\Phi(v)$ is both consistent and not a tautology.

In the sequel, Λ -PWS will refer to pure Λ -winning P-strategy.

4 Quantifying Consistency-Restoration Effort

In this section, we show how to quantify the consistency restoration effort associated with a given strategy by interpreting it as the proponent's expected cost to secure a win. We focus on GIs with an unbounded effort budget, i.e., $b = \infty$.

Consider a GI $g = (K, Q, E, \infty)$ and a Λ -PWS $\sigma = (\mathcal{T}, \Phi)$ for g . For each node v in the strategy tree \mathcal{T} , we define $E(v) \in \mathbb{R}_\infty^+$ inductively as follows:

- If v is the only node in \mathcal{T} , then $E(v) = 0$.
- If v is the root node of \mathcal{T} , then $E(v) = E(\Phi(v))$.
- Otherwise, if v is a child of v' , then $E(v) = E(\Phi(v)) + E(v')$.

In particular, for any leaf node l in \mathcal{T} , the value $E(l)$ corresponds to $\Xi(\mathcal{G})$, where \mathcal{G} is the game defined by the complete path from the root to l . Intuitively, $E(l)$ is the total effort expended along the branch leading to l .

To compute the expected effort of the strategy σ , assume each question is equally likely to receive an affirmative or a negative response from O . Thus, each binary choice leads to two equally probable outcomes, and the probability of following any particular branch of depth $d(l, \mathcal{T})$ is $1/2^{d(l, \mathcal{T})}$.

Thus, the expected effort value of σ is then given by:

$$\text{EV}(\sigma) = \sum_{l \in \text{Lv}(\mathcal{T})} \frac{1}{2^{d(l, \mathcal{T})}} E(l).$$

This expected effort $\text{EV}(\sigma)$ can be interpreted as the average cost to ensure a P's victory, taking into account all possible response patterns from O .

Quantifying the effort required to restore consistency offers several practical benefits. For example, it enables a direct comparison between different sets of questions. This helps to identify and select those that are most effective in achieving consistency with minimal cost. It also facilitates resource management: the effort budget can be adjusted based on the expected effort value. Furthermore, this approach provides a foundation for defining inconsistency measures, as we will illustrate later.

It is noteworthy that the effort value (EV) can be linked to the concept of *entropy* in information theory [Shannon, 1948a; Shannon, 1948b]. Recall that, if p_1, p_2, \dots, p_n are probabilities such that $\sum_{i=1}^n p_i = 1$, the entropy $H(p_1, p_2, \dots, p_n)$ is defined as:

$$H(p_1, p_2, \dots, p_n) = \sum_{i=1}^n p_i \log \left(\frac{1}{p_i} \right),$$

where the base of the logarithm is chosen based on the application; here, we adopt base 2.

Now, if E is defined such that $E(\phi) = 1$ for each question ϕ , we can express the effort value as:

$$\text{EV}(\sigma) = H \left(\frac{1}{2^{d(l_1, \mathcal{T})}}, \dots, \frac{1}{2^{d(l_n, \mathcal{T})}} \right),$$

where $\text{Lv}(\mathcal{T}) = \{l_1, \dots, l_n\}$.

The connection between EV and entropy arises from viewing each question as analogous to a bit of information. This

perspective aligns the notion of effort required for consistency restoration with the quantification of information: each question reduces uncertainty about the consistent knowledge base that \mathbf{P} should consider. Notably, this view resonates with existing interpretations of entropy in the literature, where entropy is seen as the minimal number of binary questions needed to identify an outcome (e.g., see [van der Lubbe and Hoeve, 1997; Cover and Thomas, 2001]).

5 Identifying the Appropriate MCS

In this section, we further simplify our setting by imposing two restrictions:

1. We restrict \mathbf{P} 's questions to the elements initially present in the given possibility base K , i.e., $Q = K$.
2. We adopt the uniform effort function E_1 , defined as $E_1(\phi) = 1$ for every $\phi \in K$. Assigning a uniform cost of 1 means that the effort in a game corresponds to the number of asked questions.

Considering these restrictions, we use (K, b) to denote a GI instead of (K, K, E_1, b) .

Additionally, we focus on a specific K -possibility function Λ_{mc} defined as:

$$\Lambda_{mc}(K, A) = \{M \in \text{MCS}(K) : M \cup A \not\models \perp\}.$$

It is straightforward to see that Λ_{mc} satisfies the properties CWR, ICPB and DBA.

By focusing on these simplified conditions, we achieve a baseline scenario that is conceptually clear. Since all questions come from K and have the same cost, we remove extraneous variables that could complicate the analysis. Moreover, using the K -possibility function Λ_{mc} links our game-theoretic framework to well-known inconsistency handling approaches based on MCSes [Rescher and Manor, 1970; Brewka, 1989; Konieczny *et al.*, 2019]. Several results developed in this section will serve for the analysis of our proposed inconsistency measure.

Theorem 2. *For every Λ_{mc} -PWS $\sigma = (\mathcal{T}, \Phi)$ of a GI (K, b) , the number of leaf nodes of \mathcal{T} is equal to $|\text{MCS}(K)|$.*

Proof. We first show that for every $M \in \text{MCS}(K)$, there exists a leaf node l in \mathcal{T} such that M is the unique MCS satisfying $\mathcal{C}_\sigma(l) \cup M \not\models \perp$. This implies that the number of leaf nodes in \mathcal{T} is at least $|\text{MCS}(K)|$.

Let M be an MCS of K and ω a model of M . Define $R = \{\phi \in \widehat{K} : \omega \models \phi\}$. Since ω is a model of both M and R , we have $M \cup R \not\models \perp$. Furthermore, we know that there exists a leaf node l such that $\mathcal{C}_\sigma(l) \subseteq R$. Because $|S| = 1$ with $S = \{M' \in \text{MCS}(K) : M' \cup \mathcal{C}_\sigma(l) \not\models \perp\}$, it follows that $S = \{M\}$.

Next, assume, for the sake of contradiction, that there exists an MCS M of K and two distinct leaf nodes l and l' such that both $\mathcal{C}_\sigma(l) \cup M \not\models \perp$ and $\mathcal{C}_\sigma(l') \cup M \not\models \perp$. By definition, there must exist a formula $\phi \in K$ such that $\phi \in \mathcal{C}_\sigma(l) \cup \mathcal{C}_\sigma(l')$ and $\neg\phi \in \mathcal{C}_\sigma(l) \cup \mathcal{C}_\sigma(l')$. Consequently, we have $\phi \in M$ and $\phi \notin M$, leading to a contradiction. \square

Since \mathcal{T} is a full binary tree in every Λ_{mc} -PWS $\sigma = (\mathcal{T}, \Phi)$, we derive the next corollary.

Corollary 1. *For every Λ_{mc} -PWS $\sigma = (\mathcal{T}, \Phi)$ of a GI (K, b) , the following properties hold:*

1. *The number of nodes of \mathcal{T} is equal to $2 \times |\text{MCS}(K)| - 1$.*
2. *$d(\mathcal{T}) \geq \lceil \log_2(|\text{MCS}(K)|) \rceil$.*

Using the foregoing corollary, we obtain the following result.

Theorem 3. *For every GI $g = (K, b)$ with $b < \lceil \log_2(|\text{MCS}(K)|) \rceil$, g does not admit any Λ_{mc} -PWS.*

The next proposition states that in any Λ_{mc} -PWS, no question asked during the strategy corresponds to a free formula.

Proposition 2. *Let $g = (K, b)$ be a GI. For every Λ_{mc} -PWS $\sigma = (\mathcal{T}, \Phi)$ of g , it holds that $\text{questions}(\sigma) \cap \text{free}(K) = \emptyset$.*

Proof. Let $\sigma = (\mathcal{T}, \Phi)$ be a Λ_{mc} -PWS of g . Assume, for the sake of contradiction, that there exists an internal node v in \mathcal{T} such that $\Phi(v) = \phi$ is free in K . Knowing that ϕ is free, $\phi \in M$ holds for every $M \in \text{MCS}(K)$. Thus, no MCS M satisfies $M \cup \{\neg\phi\} \not\models \perp$. However, as v is an internal node associated with ϕ , there exists a leaf l such that $\neg\phi \in \mathcal{C}_\sigma(l)$. Consequently, $|\{M' \in \text{MCS}(K) : M' \cup \mathcal{C}_\sigma(l) \not\models \perp\}| = 0$ holds. This contradicts the fact that σ is pure. \square

Now, we present a theorem that characterizes types of game inputs that always admit a Λ_{mc} -PWS.

Theorem 4. *For every GI $g = (K, b)$ with $b \geq |(K \setminus S)/\sim_b|$ such that $S = \text{free}(K) \cup \text{Inc}(K)$, g admits a pure Λ_{mc} -winning P -strategy.*

In particular, Λ_{mc} -PWS always exists when there is no budget limit, i.e., $b = \infty$.

The theorem below describes the common structure of Λ_{mc} -PWSES for MISes.

Theorem 5. *Let K be an MIS. Then, (K, b) admits a pure Λ_{mc} -PWS $\sigma = (\mathcal{T}, \Phi)$ iff $b \geq |K| - 1$ and σ satisfies the following properties:*

1. *The left child of every internal node in \mathcal{T} is a leaf node.*
2. *The depth of \mathcal{T} is $|K| - 1$.*
3. *For every two distinct internal nodes v and v' in \mathcal{T} , $\Phi(v) \neq \Phi(v')$.*

6 Inconsistency Measurement

Inconsistency measures are designed to quantify the amount of contradiction present in a knowledge base. Widely studied in the literature, these measures often use postulate-based frameworks to capture different aspects of inconsistency (e.g., see [Hunter and Konieczny, 2010; Ammoura *et al.*, 2017; Thimm, 2018; Bona *et al.*, 2019; Besnard and Grant, 2020; Corea *et al.*, 2021]).

Definition 4 (Inconsistency Measure). *An inconsistency measure is a mapping \mathcal{I} from PBs to \mathbb{R}_∞^+ that satisfies the following property for every PB K :*

CONSISTENCY: $\mathcal{I}(K) = 0$ if and only if K is consistent.

Beyond CONSISTENCY, the literature proposes various other postulates (e.g., [Besnard, 2014; Thimm, 2018]) to further characterize inconsistency measures. Among these, we focus on:

- **FREE FORMULA:** If $\phi \in \text{Free}(K)$, then $\mathcal{I}(K \setminus \{\phi\}) = \mathcal{I}(K)$.
- **MONOTONICITY:** If $K \subseteq K'$, then $\mathcal{I}(K) \leq \mathcal{I}(K')$.
- **EQUIVALENCE:** If $\phi \equiv \psi$, then $\mathcal{I}(K \cup \{\phi\}) = \mathcal{I}(K \cup \{\psi\})$.
- **EQUAL CONFLICT:** If M and M' are MISes of K with $|M| = |M'|$, then $\mathcal{I}(M) = \mathcal{I}(M')$.
- **ATTENUATION:** If M and M' are MISes such that $|M| > |M'|$, then $\mathcal{I}(M) < \mathcal{I}(M')$.

6.1 A Game-based Measure

Building on the restriction described in Section 5, we define the inconsistency measure \mathcal{I}_{mc}^w as follows:

$$\mathcal{I}_{mc}^w(K) = \min\{\text{EV}(\sigma) \mid \sigma \in \text{PWinS}(g, \Lambda_{mc})\} + |\text{Inc}(K)|$$

where $g = (K, \infty)$ and $\text{PWinS}(g, \Lambda_{mc})$ denotes the set of all Λ_{mc} -PWS for g .

This function mainly represents the minimal expected effort required to restore consistency. Additionally, assigning a cost of 1 to each inconsistent formula aligns with the approach used in the MCS-based measure introduced in [Grant and Hunter, 2011].

By mainly observing that every inconsistent PB either contains at least one inconsistent formula or has at least two MCSs, we obtain the following result.

Proposition 3. *The function \mathcal{I}_{mc}^w is an inconsistency measure.*

In the proposition below, we focus on the specific case where the PB is an MIS. This result primarily follows from the structure of Λ_{mc} -PWSes described in Theorem 5.

Proposition 4. *If K is an MIS, then the following holds:*

$$\mathcal{I}_{mc}^w(K) = 2 - \frac{1}{2^{n-2}}$$

where $n = |K|$.

In addition to satisfying CONSISTENCY, the measure \mathcal{I}_{mc}^w inherits various properties from our game framework:

- **FREE FORMULA and EQUIVALENCE:** It follow from the fact that no winning strategy needs to query free formulas (Proposition 2).
- **EQUIVALENCE:** It follow from the fact that logically equivalent formulas can be substituted without changing the set of consistent subsets.
- **EQUAL CONFLICT:** If two MISes M and M' have the same cardinality, $\mathcal{I}_{mc}^w(M) = \mathcal{I}_{mc}^w(M')$ (Proposition 4). This reflects the equal cost of resolving conflicts of the same size.

Note that MONOTONICITY does not hold in general because adding formulas to K changes both the base itself and the question set, which can reduce the cost of restoring consistency.

For example, let $K = \{p, \neg p, q, \neg q\}$. Any Λ_{mc} -PWS for (K, ∞) must ask at least two questions in each branch, leading to $\mathcal{I}_{mc}^w(K) = 2$. However, if we extend K by adding

$S = \{p \wedge q, p \wedge \neg q, \neg p \wedge q, \neg p \wedge \neg q\}$, then \mathcal{I}_{mc}^w can actually decrease. Indeed, a strategy that exclusively queries these newly added formulas achieves: $\mathcal{I}_{mc}^w(K \cup S) = \frac{1}{2} + \frac{2}{4} + 2 \times \frac{3}{8} = 1.75$. Thus, extending K has reduced the minimal expected effort from 2 to 1.75, showing the failure of MONOTONICITY.

Moreover, by applying Proposition 4, we observe that \mathcal{I}_{mc}^w exhibits a property opposite to ATTENUATION: if M and M' are two MISes such that $|M| > |M'|$, then $\mathcal{I}_{mc}^w(M) > \mathcal{I}_{mc}^w(M')$. Intuitively, resolving larger MISes requires identifying the correct MCS among a greater number of conflicting formulas, which increases the expected effort.

6.2 Computational Aspects

To compute \mathcal{I}_{mc}^w , the restoration effort value of a strategy can be determined from the shape of its associated tree and the number of inconsistent formulas. In what follows, for a given full binary tree \mathcal{T} , we use $\text{EV}(\mathcal{T})$ to denote the value $\sum_{l \in \text{Lv}(\mathcal{T})} \frac{d(l, \mathcal{T})}{2^{d(l, \mathcal{T})}}$.

Now, we identify the tree structure that minimizes the effort value.

Theorem 6. *Let \mathcal{T} be a finite full binary tree such that each internal node has a child that is a leaf node. Then, for every tree \mathcal{T}' with $|\text{Lv}(\mathcal{T}')| = |\text{Lv}(\mathcal{T})|$, it holds that $\text{EV}(\mathcal{T}') \geq \text{EV}(\mathcal{T})$.*

Proof (Sketch). The proof proceeds by induction on $k = |\text{Lv}(\mathcal{T})|$:

Base Case ($k = 1$): Straightforward (there is only one possible binary tree shape that consists of a single node).

Inductive Step: Assume the property holds for all $k \leq m$. Consider $k = m + 1$. Let \mathcal{T}' be a full binary tree such that $|\text{Lv}(\mathcal{T}')| = |\text{Lv}(\mathcal{T})|$. Define \mathcal{T}_r as the tree obtained by removing the two leaves of depth m from \mathcal{T} . Similarly, define \mathcal{T}'_r as the tree obtained from \mathcal{T}' by removing any two leaves having the same parent.

By the induction hypothesis, $\text{EV}(\mathcal{T}'_r) \geq \text{EV}(\mathcal{T}_r)$. Now, note that:

$$\text{EV}(\mathcal{T}) = \text{EV}(\mathcal{T}_r) + \frac{1}{2^{m-1}}, \quad \text{EV}(\mathcal{T}') = \text{EV}(\mathcal{T}'_r) + \frac{1}{2^{d-1}},$$

where d is the depth of the removed leaves in \mathcal{T}' . Since $d \leq m$ and $\text{EV}(\mathcal{T}'_r) \geq \text{EV}(\mathcal{T}_r)$, it follows that $\text{EV}(\mathcal{T}') \geq \text{EV}(\mathcal{T})$. \square

The following corollary is mainly a consequence of the foregoing theorem and Theorem 2.

Corollary 2. *Let K be a PB. Then, the following holds:*

$$\mathcal{I}_{mc}^w(K) \geq 2 - \frac{1}{2^{n-2}} + |\text{Inc}(K)|$$

where $n = |\text{MCS}(K)|$.

The proposition below outlines a scenario where the previously established bound is achieved.

Proposition 5. *Let K be a PB such that for every $M \in \text{MCS}(K)$, $M \setminus (\bigcup(\text{MCS}(K) \setminus M)) \neq \emptyset$. Then, the following holds:*

$$\mathcal{I}_{mc}^w(K) = 2 - \frac{1}{2^{n-2}} + |\text{Inc}(K)|$$

where $n = |\text{MCS}(K)|$.

Algorithm 1 Greedy Algorithm for Approximating \mathcal{I}_{mc}^w

Require: A PB K and the set of its MCSes $S = \text{MCS}(K)$

- 1: **function** BUILDTREE(S)
- 2: **if** $|S| = 1$ **then**
- 3: **return** a leaf node identifying the single MCS in S
- 4: **end if**
- 5: Let $\phi = \arg \min_{\phi \in \bigcup S \setminus \bigcap S} G(\phi, S)$
- 6: Recursively build left and right subtrees:
 - $T_l \leftarrow \text{BUILDTREE}(S_{-\phi})$
 - $T_r \leftarrow \text{BUILDTREE}(S_\phi)$
- 7: **return** the tree (ϕ, T_l, T_r)
- 8: **end function**
- 9: BUILDTREE(S)

Proof. To prove the result, we construct a Λ_{mc} -PWS $\sigma = (\mathcal{T}, \Phi)$ such that each internal node has a child that is a leaf node.

Let $\text{MCS}(K) = \{M_1, \dots, M_n\}$. For each M_i , pick $\phi_i \in M_i \setminus (\bigcup (\text{MCS}(K') \setminus \{M_i\}))$. We define \mathcal{T} as a full binary tree of depth $n - 1$ where each right child of an internal node is a leaf. Then, we define Φ as follows: for every internal node v of depth k , $\Phi(v) = \phi_{k+1}$.

We can easily check $\sigma = (\mathcal{T}, \Phi)$ is a Λ_{mc} -PWS. In particular, for every internal node v , knowing that there exists exactly one MCS containing $\Phi(v)$ explains why its right child is a leaf. \square

Since we use the effort function E_1 and due to the connection between the consistency restoration effort and entropy, the following relationship holds for every Λ_{mc} -PWS $\sigma = (\mathcal{T}, \Phi)$:

$$\text{EV}(\sigma) = H\left(\frac{1}{2^{d(l_1, \mathcal{T})}}, \dots, \frac{1}{2^{d(l_n, \mathcal{T})}}\right),$$

where $\text{Lv}(\mathcal{T}) = \{l_1, \dots, l_n\}$.

Recall that the maximum value of $H(p_1, p_2, \dots, p_n)$ is achieved when $p_1 = p_2 = \dots = p_n$ (e.g., see [Cover and Thomas, 2001]). Consequently, the maximum value of EV occurs when $d(l, \mathcal{T}) = \log_2(n)$ for each leaf node $l \in \text{Lv}(\mathcal{T})$. Thus, the tree shape that maximizes the effort value corresponds to a balanced tree. Based on this observation, we derive the following proposition:

Proposition 6. *Let K be a PB. Then, the inequality $\mathcal{I}_{mc}^w(K) \leq \log_2(|\text{MCS}(K)|) + |\text{Inc}(K)|$ holds.*

Finding the optimal strategy to compute \mathcal{I}_{mc}^w is a challenging problem. By leveraging the connection between this problem, decision trees, and information theory, we propose a greedy approach to approximate the inconsistency value. This approach is inspired by the decision tree learning algorithm ID3 [Quinlan, 1986; Mitchell, 1997]. Intuitively, each MCS can be viewed as a distinct class, while the formulas of the propositional base serve as the features.

To define our algorithm, we first introduce the concept of *information gain*, which quantifies the reduction in entropy when the set of MCSes is split using a given formula ϕ . The

entropy associated with a set of MCSes S is defined as:

$$H(S) = |S| \left(\frac{1}{|S|} \log_2(|S|) \right) = \log_2(|S|),$$

where we assume that all elements of S are equally probable. Note that $H(S)$ is only defined for non-empty S .

Given a formula ϕ and a set of MCSes S , the *information gain* $G(\phi, S)$ is defined as:

$$G(\phi, S) = H(S) - \left(\frac{|S_\phi|}{|S|} H(S_\phi) + \frac{|S_{-\phi}|}{|S|} H(S_{-\phi}) \right),$$

where $S_\phi = \{M \in S : \phi \in M\}$ and $S_{-\phi} = \{M \in S : \phi \notin M\}$.

Consider the function $f(i) = (n-i) \log_2(n-i) + i \log_2(i)$ ($1 < i < n-1$) which is strictly concave and attains its global minimum at $i = \frac{n}{2}$. Thus, maximizing $G(\phi, S)$ rewards balanced splits (exactly the goal of the classical ID3 algorithm). In our setting balanced partitions are not desirable; we instead seek highly unbalanced splits. Consequently, we choose to minimize $G(\phi, S)$ when selecting the next formula.

The greedy approach used in Algorithm 1 can be summarized as follows:

- At each step, select ϕ that minimizes the G .
- Partition $S = \text{MCS}(K)$ into two subsets, S_ϕ and $S_{-\phi}$.
- Recursively repeat this process for the subsets S_ϕ and $S_{-\phi}$ until each leaf of the tree corresponds to a single MCS.

Observe that, since the selected formula ϕ at each step belongs to $\bigcup S \setminus \bigcap S$, it holds that $S_{-\phi} \neq \emptyset$ and $S_\phi \neq \emptyset$.

In the next proposition, we identify a scenario in which Algorithm 1 allows us to compute the exact value of \mathcal{I}_{mc}^w .

Proposition 7. *Suppose (K, ∞) admits a Λ_{mc} -PWS σ whose tree \mathcal{T} ensures that each internal node has at least one child that is a leaf. Then Algorithm 1 computes one of the optimal Λ_{mc} -PWSes.*

A notable drawback of the greedy algorithm is its reliance on having all MCSes, which is generally recognized as a highly challenging task.

7 Conclusion and Perspectives

We developed a game-theoretic framework to restore consistency in propositional bases. It represents the process as a strategic interaction between a Proponent, who aims to isolate a unique consistent base, and an Opponent, who controls the responses. By assigning a cost to each question, the framework reflects the concept of limited resources. This facilitates the definition of measures that evaluate the expected effort required to restore consistency. Drawing on the quantification of consistency restoration effort, we showed how to derive a new inconsistency measure.

This framework provides numerous opportunities for future work. First, analyzing the complexity of identifying optimal strategies and quantifying restoration effort represents a critical direction for future research. Second, the foundational concepts could be adapted to more expressive logics, such as description logics and modal logics. Third, exploring variants that incorporate non-binary questions presents interesting possibilities for further study.

References

- [Alchourrón *et al.*, 1985] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.
- [Ammoura *et al.*, 2017] Meriem Ammoura, Yakoub Salhi, Brahim Oukacha, and Badran Raddaoui. On an MCS-based inconsistency measure. *International Journal of Approximate Reasoning*, 80:443–459, 2017.
- [Benferhat *et al.*, 1995] Salem Benferhat, Didier Dubois, and Henri Prade. How to Infer from Inconsistent Beliefs without Revising? In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, IJCAI 95. Montréal Québec, Canada, 2 Volumes*, pages 1449–1457. Morgan Kaufmann, 1995.
- [Besnard and Grant, 2020] Philippe Besnard and John Grant. Relative Inconsistency Measures. *Artificial Intelligence*, 280:103231, 2020.
- [Besnard and Hunter, 2008] Philippe Besnard and Anthony Hunter. *Elements of Argumentation*. MIT Press, 2008.
- [Besnard, 2014] Philippe Besnard. Revisiting Postulates for Inconsistency Measures. In Eduardo Fermé and João Leite, editors, *Logics in Artificial Intelligence - 14th European Conference, JELIA 2014, Funchal, Madeira, Portugal*, volume 8761 of *Lecture Notes in Computer Science*, pages 383–396. Springer, 2014.
- [Blass, 1992] Andreas Blass. A Game Semantics for Linear Logic. *Annals of Pure and Applied Logic*, 56(1-3):183–220, 1992.
- [Bona *et al.*, 2019] Glauber De Bona, John Grant, Anthony Hunter, and Sébastien Konieczny. Classifying Inconsistency Measures Using Graphs. *Journal of Artificial Intelligence Research*, 66:937–987, 2019.
- [Brewka, 1989] Gerhard Brewka. Preferred Subtheories: An Extended Logical Framework for Default Reasoning. In N. S. Sridharan, editor, *Proceedings of the 11th International Joint Conference on Artificial Intelligence, IJCAI 89. Detroit, MI, USA*, pages 1043–1048. Morgan Kaufmann, 1989.
- [Corea *et al.*, 2021] Carl Corea, Matthias Thimm, and Patrick Delfmann. Measuring Inconsistency over Sequences of Business Rule Cases. In Meghyn Bienvenu, Gerhard Lakemeyer, and Esra Erdem, editors, *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning, KR 2021, Online event*, pages 655–660, 2021.
- [Cover and Thomas, 2001] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2001.
- [Fermé and Hansson, 2011] Eduardo L. Fermé and Sven Ove Hansson. AGM 25 Years - Twenty-Five Years of Research in Belief Change. *Journal of Philosophical Logic*, 40(2):295–331, 2011.
- [Fermüller and Ciabattoni, 2003] Christian G. Fermüller and Agata Ciabattoni. From Intuitionistic Logic to Gödel-Dummett Logic via Parallel Dialogue Games. In *33rd IEEE International Symposium on Multiple-Valued Logic (ISMVL 2003), Tokyo, Japan*, pages 188–196. IEEE Computer Society, 2003.
- [Gärdenfors, 1992] Peter Gärdenfors, editor. *Belief Revision*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1992.
- [Grant and Hunter, 2011] John Grant and Anthony Hunter. Measuring Consistency Gain and Information Loss in Stepwise Inconsistency Resolution. In Weiru Liu, editor, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 11th European Conference, ECSQARU 2011, Belfast, UK*, volume 6717 of *Lecture Notes in Computer Science*, pages 362–373. Springer, 2011.
- [Hunter and Konieczny, 2010] Anthony Hunter and Sébastien Konieczny. On the measure of conflicts: Shapley Inconsistency Values. *Artificial Intelligence*, 174:1007–1026, 2010.
- [Konieczny *et al.*, 2019] Sébastien Konieczny, Pierre Marquis, and Srdjan Vesic. Rational Inference Relations from Maximal Consistent Subsets Selection. In Sarit Kraus, editor, *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019. Macao, China*, pages 1749–1755. ijcai.org, 2019.
- [Lang and Marquis, 2002] Jérôme Lang and Pierre Marquis. Resolving Inconsistencies by Variable Forgetting. In Dieter Fensel, Fausto Giunchiglia, Deborah L. McGuinness, and Mary-Anne Williams, editors, *Proceedings of the Eight International Conference on Principles and Knowledge Representation and Reasoning, KR’02. Toulouse, France*, pages 239–250. Morgan Kaufmann, 2002.
- [Lin and Reiter, 1994] Fangzhen Lin and Raymond Reiter. Forget It! In *Proceedings of the AAAI Fall Symposium on Relevance*, pages 154–159, 1994.
- [Lorenzen, 1960] Paul Lorenzen. Logik und Agon. *Atti Del XII Congresso Internazionale di Filosofia*, 4:187–194, 1960.
- [Mitchell, 1997] Tom M. Mitchell. *Machine learning, International Edition*. McGraw-Hill Series in Computer Science. McGraw-Hill, 1997.
- [Pitts and Dybjer, 1997] Andrew M. Pitts and P.Editors Dybjer, editors. *Semantics of Interaction: an Introduction to Game Semantics*, pages 1–32. Publications of the Newton Institute. Cambridge University Press, 1997.
- [Priest, 1991] G. Priest. Minimally inconsistent LP. *Studia Logica*, 50:321–331, 1991.
- [Quinlan, 1986] J. Ross Quinlan. Induction of Decision Trees. *Machine Learning*, 1(1):81–106, 1986.
- [Rescher and Manor, 1970] Nicholas Rescher and Ruth Manor. On Inference from Inconsistent Premises. *Theory and Decision*, 1:179–217, 1970.
- [Rival, 2012] I. Rival. *Graphs and Order: The Role of Graphs in the Theory of Ordered Sets and Its Applications*. Nato Science Series C. Springer Netherlands, 2012.

- [Shannon, 1948a] Claude E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, 1948.
- [Shannon, 1948b] Claude E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(4):623–656, 1948.
- [Thimm, 2018] Matthias Thimm. On the Evaluation of Inconsistency Measures. In John Grant and Maria Vanina Martinez, editors, *Measuring Inconsistency in Information*, volume 73 of *Studies in Logic*. College Publications, February 2018.
- [van der Lubbe and Hoeve, 1997] J.C.A. van der Lubbe and H.J. Hoeve. *Information Theory*. Cambridge University Press, 1997.