Algorithmic Approaches for Inconsistency Measurement

Isabelle Kuhlmann

Artificial Intelligence Group,
University of Hagen, Germany

Cape-KR
February 11, 2025

Motivation

In Artificial Intelligence, we cannot avoid the occurrence of conflicting (*inconsistent*) information.

Examples

- Different expert opinions or assessments
- Noisy/distorted sensor data
- Conflicting results from machine learning (e. g., rule mining)

Inconsistencies can occur in virtually any area of application.

Motivation

Inconsistency Measurement offers an analytical perspective

- Quantitative assessment of the level of inconsistency
  - We assign a numerical value which indicates the inconsistency level
- Allows for comparison of different formalizations
- Can assist automatic reasoning mechanisms
- May help to *identify*, and therefore later *remove* conflicts

Application Examples

- Analysis of inconsistencies in news reports (Hunter, 2006)
- Support of collaborative software requirements specifications (Martinez et al., 2004)
- Monitoring and maintenance of quality in database settings (Bertossi, 2018)
- Handling of inconsistencies in business processes (Corea et al., 2021, 2022)

There is clearly a need for practical working solutions!

Overview

Overview

Intuition

An inconsistency measure assigns a value to a (propositional) knowledge base.

- The larger the value, the more severe the inconsistency
- Consistent knowledge bases have the value 0

Definition

Let $\mathbb{K}$ be the set of all (propositional) knowledge bases.
An *inconsistency measure* $\mathcal{I}$ is a function $\mathcal{I} : \mathbb{K} \to \mathbb{R}^{\infty}_{\geq 0}$ that satisfies $\mathcal{I}(\mathcal{K}) = 0$ iff $\mathcal{K}$ is consistent, for all $\mathcal{K} \in \mathbb{K}$.

- Different definitions of a “conflict”
- Different formalisms
  - *Examples:* propositional logic, linear temporal logic, data bases, ...

Example

$$\mathcal{K} = \{cloudy \wedge rainy,\ \neg cloudy,\ rainy \vee sunny\}$$

Different perspectives on inconsistency:

- Which atoms need to be removed in order to make $\mathcal{K}$ consistent?
  $\mathcal{K} = \{\boxed{cloudy} \wedge rainy,\ \neg \boxed{cloudy},\ rainy \vee sunny\}$
- Which atom *occurrences* need to be removed in order to make $\mathcal{K}$ consistent?
  $\mathcal{K} = \{\boxed{cloudy} \wedge rainy,\ \neg cloudy,\ rainy \vee sunny\}$ or

  $\mathcal{K} = \{cloudy \wedge rainy,\ \neg \boxed{cloudy},\ rainy \vee sunny\}$
- Which formulas comprise a *minimal unsatisfiable subset* (MUS)?
  $\mathcal{K} = \{\boxed{cloudy \wedge rainy,\ \neg cloudy},\ rainy \vee sunny\}$

Example: Contension Inconsistency Measure

The contension inconsistency measure is based on Priest's *three-valued logic*:

- In addition to true (t) and false (f), this logic includes a third value which indicates *paradoxical*, or *both true and false* (b)
- A *three-valued interpretation* ω^3 is a function that assigns one of the three truth values to each atom in a given knowledge base:

$$\omega^3 : \mathrm{At}(\mathcal{K}) \to \{t, f, b\}$$

- A three-valued *model* is an interpretation where each formula $\phi \in \mathcal{K}$ is assigned either t or b.

x	y	$x \wedge y$	$x \vee y$
t	t	t	t
t	b	b	t
t	f	f	t
b	t	b	t
b	b	b	b
b	f	f	b
f	t	f	t
f	b	f	b
f	f	f	f

x	$\neg x$
t	f
b	b
f	t

Example: Contension Inconsistency Measure

Definition

The *set of models* w.r.t. $\mathcal{K}$ is defined as

$$\text{Models}(\mathcal{K}) = \{\omega^3 \mid \forall \phi \in \mathcal{K}, \omega^3(\phi) = t \text{ or } \omega^3(\phi) = b\}.$$

We can divide the domain of an interpretation ω^3 into two sets:

- One contains those atoms that are assigned a classical truth value (t, f)
- One contains those atoms that are assigned b

Definition

$$\text{Conflictbase}(\omega^3) = \{x \in \text{At}(\mathcal{K}) \mid \omega^3(x) = b\}$$

Intuition

The contension inconsistency measure $\mathcal{I}_c$ describes the minimum number of atoms in $\mathcal{K}$ that need to be assigned truth value b in order to make $\mathcal{K}$ consistent.

Definition

$$\mathcal{I}_c(\mathcal{K}) = \min\{|\mathrm{Conflictbase}(\omega^3)| \mid \omega^3 \in \mathrm{Models}(\mathcal{K})\}.$$

Example: $\mathcal{K}_1 = \{x \wedge y, \neg x, y \vee z\}$

- Let ω_1^3 be an interpretation with $\omega_1^3(y) = \omega_1^3(z) = t$ and $\omega_1^3(x) = b$
- ω_1^3 is a model of $\mathcal{K}_1$
- $\mathrm{Conflictbase}(\omega_1^3) = \{x\}$
- $\mathcal{I}_c(\mathcal{K}_1) = |\mathrm{Conflictbase}(\omega_1^3)| = |\{x\}| = 1$

Forgetting-Based Inconsistency Measure (Intuition)

We look for the minimal number of atom *occurrences* that need to be "forgotten" (i. e., replaced by either $\top$ or $\bot$) in order to render the given knowledge base consistent.

Hitting Set Inconsistency Measure (Intuition)

We search for the cardinality-minimal set of interpretations s. t. each formula is satisfied by at least one of those interpretations, subtracted by 1.

Distance-Based Inconsistency Measure (Intuition)

We aim to find an interpretation with an "optimal" distance to the models of the formulas in the given knowledge base.

Problematic Inconsistency Measure (Intuition)

We count how many formulas are included in at least one minimal unsatisfiable subset (MUS).

MUS-Variable-Based Inconsistency Measure (Intuition)

We calculate the fraction of the signature that is involved in at least one MUS.

Observation

- There exists a plethora of different inconsistency measures in the literature
- Only few works consider the topic of inconsistency measurement from an *algorithmic* perspective

There are different "dimensions" to explore:

- Different application domains/formalisms, e.g.:
  - Different logics
  - Databases
  - Argumentation
- Different complexity classes (Thimm and Wallner, 2019)
- Different problem solving paradigms
  - Depending on application domain and complexity class

Complexity-wise, inconsistency measurement is hard in general (Thimm and Wallner, 2019)

- The decision problems corresponding to the "easiest" measures are on the first level of the polynomial hierarchy
  - Contension, forgetting-based, hitting set, distance-based
- Some measures are also higher up in the polynomial hierarchy
  - The problematic and the MUS-variable-based measure are on the second level

Overview

We considered a total of 8 inconsistency measures for propositional logic

- 6 are on the first level of the polynomial hierarchy
- 2 are on the second level

In addition, we considered 4 inconsistency measures for *linear temporal logic on fixed traces* ($\mathrm{LTL}_{\mathrm{ff}}$) (Kuhlmann et al., 2023a; Corea et al., 2024)

- 2 are on the first level of the polynomial hierarchy
- 2 are on the second level

Algorithmic approaches:

- Answer set programming (ASP)
- Boolean satisfiability (SAT) solving
- Maximum satisfiability (MaxSAT) solving
- SAT-based Counterexample-guided abstraction refinement (CEGAR)

Overview

	First Level	Second Level
PL	ASP, SAT, MaxSAT	ASP, CEGAR
$\mathrm{LTL}_{\mathrm{ff}}$	ASP, MaxSAT	ASP, CEGAR

Experimental Evaluation

Objective

- Compare the ASP approaches to other approaches in terms of runtime

First-level measures:

- Naive (brute-force) methods[1] (Kuhlmann and Thimm, 2020, 2021)
- Iterative SAT approach (Kuhlmann et al., 2022)
- MaxSAT (Niskanen et al., 2023)

Second-level measures:

- Iterative MUS/MCS Enumeration (Kuhlmann et al., 2023b)
- Counterexample-guided abstraction refinement (CEGAR) (Kuhlmann et al., 2023b)

[1] http://tweetyproject.org/

There is no dedicated benchmark data set for inconsistency measurement.

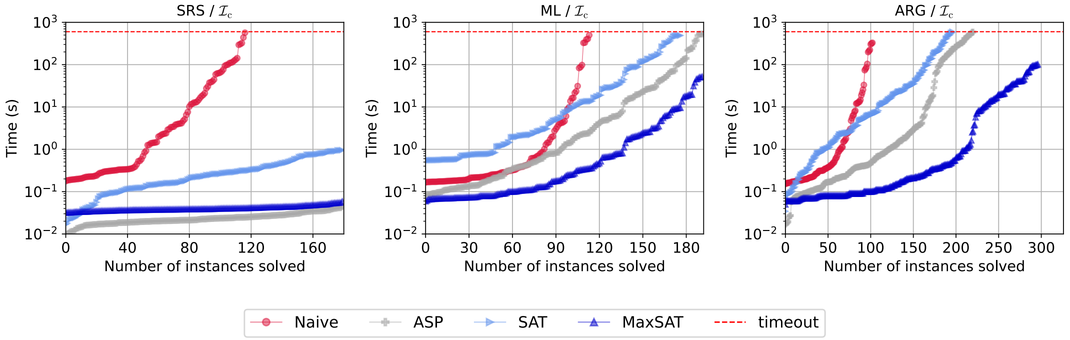
- We can either synthesize completely new data or "translate" data sets from other fields of application.

Data Sets

- **SRS** data set: Synthetic dataset created using the *SyntacticRandomSampler*[2]
  - 180 knowledge bases
  - Smallest instances: signature size 3; 5–15 formulas
  - Largest instances: signature size 30; 50–100 formulas
- **ML** data set: 192 KBs learnt from machine learning data (*Animals with Attributes*)
- **ARG** data set: 326 KBs extracted from benchmark data of the International Competition on Computational Models of Argumentation 2019

[2] `http://tweetyproject.org/api/1.14/net/sf/tweety/logics/pl/util/SyntacticRandomSampler.html`

Runtime results regarding the *contension* inconsistency measure (cactus plots):



Timeout: 600 s

Observations

- As expected, the naive approach cannot compete with the others
- The ASP approach outperforms the iterative SAT approach
- Overall, the MaxSAT approach outperforms all other approaches
- This pattern also (broadly) applies to the other 5 measures on the first level of the polynomial hierarchy

Overview

Conclusion

We proposed algorithmic approaches for computing inconsistency measures

- 8 measures for propositional logic + 4 for $\mathrm{LTL}_{\mathrm{ff}}$
  - 3 approaches for each propositional first-level measure
  - 2 approaches for each $\mathrm{LTL}_{\mathrm{ff}}$ first-level measure
  - 2 approaches for each second-level measure
- 3 data sets for propositional logic + 3 for $\mathrm{LTL}_{\mathrm{ff}}$

Future Work?

- Consider other inconsistency measures and/or other algorithmic approaches
- Approximation approaches and preprocessing
- Improve existing approaches
  - Evaluate different optimization strategies for ASP, try heuristics, etc.

Thank you for your attention!

Bibliography

Bertossi, L. (2018). Measuring and computing database inconsistency via repairs. In *Proceedings of the 12th International Conference on Scalable Uncertainty Management (SUM'18)*.

Corea, C., Grant, J., and Thimm, M. (2022). Measuring inconsistency in declarative process specifications. In *International Conference on Business Process Management*, pages 289–306. Springer.

Corea, C., Kuhlmann, I., Thimm, M., and Grant, J. (2024). Paraconsistent reasoning for inconsistency measurement in declarative process specifications. *Information Systems*, 122:102347.

Corea, C., Thimm, M., and Delfmann, P. (2021). Measuring inconsistency over sequences of business rule cases. In Bienvenu, M., Lakemeyer, G., and Erdem, E., editors, *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning (KR'21)*, pages 656–660. IJCAI Organization.

Hunter, A. (2006). How to act on inconsistent news: Ignore, resolve, or reject. *Data & Knowledge Engineering*, 57(3):221–239.

Kuhlmann, I., Corea, C., and Grant, J. (2023a). An ASP-based framework for solving problems related to declarative process specifications. In *Proceedings of the 21st International Workshop on Non-Monotonic Reasoning (NMR)*, pages 129–132.

Kuhlmann, I., Gessler, A., Laszlo, V., and Thimm, M. (2022). A comparison of ASP-based and SAT-based algorithms for the contension inconsistency measure. In *15th International Conference on Scalable Uncertainty Management*, pages 139–153. Springer.

Kuhlmann, I., Niskanen, A., and Järvisalo, M. (2023b). Computing MUS-based inconsistency measures. In *Proceedings of the 18th European Conference on Logics in Artificial Intelligence*, Lecture Notes in Computer Science, pages 745–755. Springer.

Kuhlmann, I. and Thimm, M. (2020). An algorithm for the contension inconsistency measure using reductions to answer set programming. In *14th International Conference on Scalable Uncertainty Management*, pages 289–296. Springer.

Kuhlmann, I. and Thimm, M. (2021). Algorithms for inconsistency measurement using answer set programming. In *Proceedings of the 19th International Workshop on Non-Monotonic Reasoning (NMR)*, pages 159–168.

Martinez, A. B. B., Arias, J. J. P., and Vilas, A. F. (2004). On measuring levels of inconsistency in multi-perspective requirements specifications. In *Proceedings of the 1st Conference on the Principles of Software Engineering (PRISE'04)*.

Niskanen, A., Kuhlmann, I., Thimm, M., and Järvisalo, M. (2023). MaxSAT-based inconsistency measurement. In *Proceedings of the 26th European Conference on Artificial Intelligence*, pages 1779–1786. IOS Press.

Thimm, M. and Wallner, J. P. (2019). On the complexity of inconsistency measurement. *Artificial Intelligence*, 275:411–456.