Inconsistency Measurement based on Variables in Minimal Unsatisfiable Subsets

Guohui Xiao   Yue Ma

Institute of Informatics, Vienna University of Technology
Theoretical Computer Science, Dresden University of Technology

ECAI 2012 — August 30, 2012

Overview

Outline

Background

- Consistent KBs are useful, but inconsistent KBs imply any conclusion (meaningless!)
- Inconsistency measurement:

  from "is inconsistent" to "how inconsistent"
- Ideas and approaches:
  - based on different views of *atomicity of inconsistency*
  - Semantics based approaches
  - Syntax based approaches
  - Semantics - syntax combined approaches (this paper)

Outline

Inconsistency Measurement by Multi-valued Semantics

Inconsistency Measurement by Multi-valued Semantics

- Multi-Valued Semantics
  - 4-valued, 3-valued, LP_m, Quasi-Classical, ...
  - $I : Var(K) \rightarrow \{t, f, Both, None\}$

Inconsistency Measurement by Multi-valued Semantics

- Multi-Valued Semantics
  - 4-valued, 3-valued, LP_m, Quasi-Classical, ...
  - $I : Var(K) \rightarrow \{t, f, Both, None\}$
- ID of K respect to I under i-semantics ($i = 3, 4, LP_m, Q$)

$$ID_i(K, I) = \frac{|\{p \mid p^I = B, p \in Var(K)\}|}{|Var(K)|}, \text{ if } I \models_i K$$

Inconsistency Measurement by Multi-valued Semantics

- Multi-Valued Semantics
  - 4-valued, 3-valued, LP_m, Quasi-Classical, ...
  - $I : Var(K) \to \{t, f, Both, None\}$
- ID of K respect to I under i-semantics ($i = 3, 4, LP_m, Q$)

$$ID_i(K, I) = \frac{|\{p \mid p^I = B, p \in Var(K)\}|}{|Var(K)|}, \text{ if } I \models_i K$$

- ID of K under under i-semantics ($i = 3, 4, LP_m, Q$)

$$ID_i(K) = \min_{I \models_i K} ID_i(K, I)$$

Inconsistency Degree under 4-valued Semantics

Truth values: $\{t, f, B, N\}$

4-model I:

$K \to \{t, B\}$

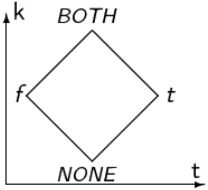


Figure : Four-Valued Logic

- $ID_4(K, I) = \frac{|\{p | p^I = B, p \in Var(K)\}|}{|Var(K)|}$

  $ID_4(K) = min_{I \models_4 K} ID_4(K),$

$\leadsto$ $K = \{p, \neg q, \neg p \vee q, r \vee s\}$

$\leadsto$ $I_1 : p^{I_1} = B, q^{I_1} = f, r^{I_1} = t, s^{I_1} = t,$

$I_2 : p^{I_2} = B, q^{I_2} = B, r^{I_2} = t, s^{I_2} = t$

$I_3 : p^{I_3} = B, q^{I_3} = B, r^{I_3} = t, s^{I_3} = N$

$\leadsto$ $ID_4(K, I_1) = \frac{1}{4}$, $ID_4(K, I_2) = \frac{2}{4}$

$ID_4(K, I_3) = \frac{2}{4}$

$ID_4(K) = \frac{1}{4}$

Inconsistency Degree under Quasi-Classical Semantics

Quasi-Classical (Q) interpretation:

- 4-valued interpretation
- Resolution laws are satisfied

  $I \models_Q \alpha \vee \beta,$

  $I \models_Q \neg\beta \vee \gamma$

  $\Rightarrow I \models_Q \alpha \vee \gamma$

- $ID_Q(K, I) = \frac{|\{p | p^I = B, p \in Var(K)\}|}{|Var(K)|}$

  $ID_Q(K) = min_{I \models_Q K} ID_Q(K),$

$\rightsquigarrow$ $K = \{p, \neg q, \neg p \vee q, r \vee s\}$

$\rightsquigarrow$ ~~$I_1 : p^{I_1} = B, q^{I_1} = f, r^{I_1} = t, s^{I_1} = t$~~

$I_2 : p^{I_2} = B, q^{I_2} = B, r^{I_2} = t, s^{I_2} = t$

$I_3 : p^{I_3} = B, q^{I_3} = B, r^{I_3} = t, s^{I_3} = N$

$\rightsquigarrow$ ~~$ID_Q(K, I_1) = \frac{1}{4}$~~, $ID_Q(K, I_2) = \frac{2}{4}$

$ID_Q(K, I_3) = \frac{2}{4}$

$ID_Q(K) = \frac{2}{4}$

Remark: $ID_4(K) = ID_3(K) = ID_{LPm}(K) \leq ID_Q(K)$ [Xiao et al., 2010]

MUS and MCS

Definition

A subset $U \subseteq K$ is an Minimal Unsatisfiable Subset (MUS), if

- U is unsatisfiable and
- $\forall C_i \in U, U \setminus \{C_i\}$ is satisfiable.

Definition

A subset $M \subseteq K$ is an Minimal Correction Subset (MCS), if

- $K \setminus M$ is satisfiable and
- $\forall C_i \in M, K \setminus (M \setminus \{C_i\})$ is unsatisfiable.

Example

Let $K = \{p, \neg p, p \vee q, \neg q, \neg p \vee r\}$. Then
$MUSes(K) = \{\{p, \neg p\}, \{\neg p, p \vee q, \neg q\}\}$ and
$MCSes(K) = \{\{\neg p\}, \{p, p \vee q\}, \{p, \neg q\}\}$.

Inconsistency Measurement by MUSes and MCSes

[Hunter and Konieczny, 2008]

The MI inconsistency measure is defined as the numbers of minimal inconsistent sets of K: $I_{MI}(K) = |MUSes(K)|$.
(minimal inconsistent sets = minimal unsatisfiable subsets)

Example

Let $K = \{p, \ \neg p, \ p \vee q, \ \neg q, \ \neg p \vee r\}$.

- $MUSes(K) = \{\{p, \neg p\}, \ \{\neg p, p \vee q, \ \neg q\}\}$
- $I_{MI}(K) = 2$

- Note that $I_{MI}(K)$ can be exponentially large

Why another Inconsistency Measurement?

- Combination of Semantics and Syntax based IDs
  - Shapley inconsistency measures [Hunter and Konieczny, 2006]: distribution of $ID_{\{4,Q,\ldots\}}$ among different formulas
  - Ours:
    combination of semantics and syntax based IDs in the KB level
- Expected properties:
  - Easier to compute than I_{MI}:
    - ★ I_{MI} tends to be difficult to compute or approximate because of exponentially many MUSes
  - More intuitive:
    - ★ For $K = \{a \wedge \neg a\}$ and $K' = \{a \wedge \neg a \wedge b \wedge \neg b\}$, we have $I_{MI}(K) = I_{MI}(K') = 1$, which is unintuitive
    - ★ Later we see ID_4 tends to be "small",
      while ID_Q tends to be "large"

Outline

Inconsistency Measurement by Variables in MUSes

Definition

For a given set of variables S and a given knowledge base K such that $Var(K) \subseteq S$, its MUS-variable based inconsistency degree, written $ID_{MUS}(K)$, is defined as:

$$ID_{MUS}(K) = \frac{|Var(MUSes(K))|}{|S|} .$$

Example

Let $K = \{p, \neg p, p \vee q, \neg q, \neg p \vee r\}$ and $S = Var(K) = \{p, q, r\}$, $MUSes(K) = \{\{p, \neg p\}, \{\neg p, p \vee q, \neg q\}\}$. Then $ID_{MUS}(K) = 2/3$.

Example

For $K = \{a \wedge \neg a\}$ and $K' = \{a \wedge \neg a \wedge b \wedge \neg b\}$, let $S = Var(K) \cup Var(K') = \{a, b\}$. Then we have $MUSes(K) = \{\{a \wedge \neg a\}\}$ and $MUSes(K') = \{\{a \wedge \neg a \wedge b \wedge \neg b\}\}$, $ID_{MUS}(K) = 1/2$ and

Inconsistency Measurement by Variables in MCSes

Similarly to $ID_{MUS}(K)$, we can define another inconsistency degree through MCS as follows:

Definition

For a given set of variables S and a given knowledge base K such that $Var(K) \subseteq S$, its MCS-variable based inconsistency degree, written $ID_{MCS}(K)$, is defined as follows:

$$ID_{MCS}(K) = \frac{|Var(MCSes(K))|}{|S|}.$$

Example

Let $K = \{p, \neg p, p \vee q, \neg q, \neg p \vee r\}$ and $S = Var(K)$, $MCSes(K) = \{\{\neg p\}, \{p, p \vee q\}, \{p, \neg q\}\}$, then $ID_{MCS}(K) = 2/3$.

$ID_{MUS} = ID_{MCS}$

- $MUSes(K)$ and $MCSes(K)$ are hitting sets dual of each other [Liffiton and Sakallah, 2008]

$\Rightarrow \bigcup MUSes(K) = \bigcup MCSes(K)$

$\Rightarrow Var(\bigcup MUSes(K)) = Var(\bigcup MCSes(K))$

$\Rightarrow ID_{MUS}(K) = ID_{MCS}(K)$

In the rest of the talk, the discussion is only about $ID_{MUS}(K)$,

ID_4 and ID_{MUS}

Corollary

Let U be an MUS, then $ID_4(U) = 1/|Var(U)|$.

The following theorem shows that $ID_4(K)$ can be determined by the cardinality minimal hitting sets of $MUSes(K)$.

Theorem

For a given KB K,

$$ID_4(K) = \frac{min_H\{|H| \mid \forall U \in MUSes(K), Var(U) \cap H \neq \emptyset\}}{|Var(K)|}.$$

Corollary

$ID_{MUS}(K) \geq ID_4(K)$.

ID_Q and ID_{MUS}

Lemma

Let U be an MUS, then U has only one Q-model which assigns B to all of its variables. Hence $ID_Q(U) = 1$.

Proposition

Let K be a KB and $\mathcal{I} \in PM_Q(K)$, then $Conflict(\mathcal{I}, K) \supseteq Var(MUSes(K))$.

Corollary

Let K be a KB, then $ID_Q(K) \geq ID_{MUS}(K)$.

Outline

Complexity Results

- $ID\text{-}MUS_{\geq k}$: Given a CNF KB, and a number k, deciding $ID_{MUS}(K) \geq k$.
- $ID\text{-}MUS$: Functional complexity of computing ID_{MUS}

Problem	Complexity
$ID\text{-}MUS_{\geq k}$	Σ_2^p-complete
$ID\text{-}MUS_{\leq k}$	Π_2^p-complete
$ID\text{-}MUS_{=k}$	D_2^p-complete
$ID\text{-}MUS$	$\mathrm{FP}^{\Sigma_2^p[\log]}$

Table : Complexity Results

- All the results are in the second layer of polynomial hierarchy
- Recall that ID_4 and ID_Q are in first layer

Outline

Anytime Algorithm

- Using MCS finder to find $MCSes(K)$
- Update ID_{MUS} by newly found MCS

Algorithm: Anytime Algorithm for $ID_{MUS}(K)$;
Input: K: KB as a set of clauses
Output: $ID_{MUS}(K)$

```
B ← {}                                   // variable set
N ← |Var(K)|
foreach M ∈ MCSes(K)                     // call MCS finder
do
    B ← B ∪ Var(M)                       // update B
    id ← |B|/N                           // new idmus lower bound
    print 'id_mus(K) ⩾ ', id
end
print 'id_mus(K) = ', id
return id
```

Prototype Implementation

- prototype implementation, called CAMUS_IDMUS
- by adapting the source code of CAMUS_MCS 1.02[1].

[1] http://www.eecs.umich.edu/~liffiton/camus/

Experiments

Table : Evaluation of CAMUS_IDMUS on DC Benchmark

Instance	#V	#C	#M	#4	#Q	#VM	T
C168_FW_SZ_41	1,698	5,387	>30,104	1	211	> 124	600.00
C168_FW_SZ_66	1,698	5,401	>16,068	1	182	> 69	600.00
C168_FW_SZ_75	1,698	5,422	>37,317	1	198	> 116	600.00
C168_FW_SZ_107	1,698	6,599	>51,597	1	189	> 92	600.00
C168_FW_SZ_128	1,698	5,425	>25,397	1	211	> 66	600.00
C168_FW_UT_2463	1,909	7,489	>109,271	1	436	> 168	600.00
C168_FW_UT_2468	1,909	7,487	>54,845	1	436	> 138	600.00
C168_FW_UT_2469	1,909	7,500	>56,166	1	436	> 150	600.00
C168_FW_UT_714	1,909	7,487	>84,287	1	436	> 92	600.00
C168_FW_UT_851	1,909	7,491	30	1	436	11	0.35
C168_FW_UT_852	1,909	7,489	30	1	436	11	0.35
C168_FW_UT_854	1,909	7,486	30	1	436	11	0.35
C168_FW_UT_855	1,909	7,485	30	1	436	11	0.35
C170_FR_SZ_58	1,659	5,001	177	1	157	54	0.46
C170_FR_SZ_92	1,659	5,082	131	1	163	46	0.10
C170_FR_SZ_95	1,659	4,955	175	1	23	23	0.20
C170_FR_SZ_96	1,659	4,955	1,605	1	125	43	0.36

Anytime Property of CAMUS_IDMUS

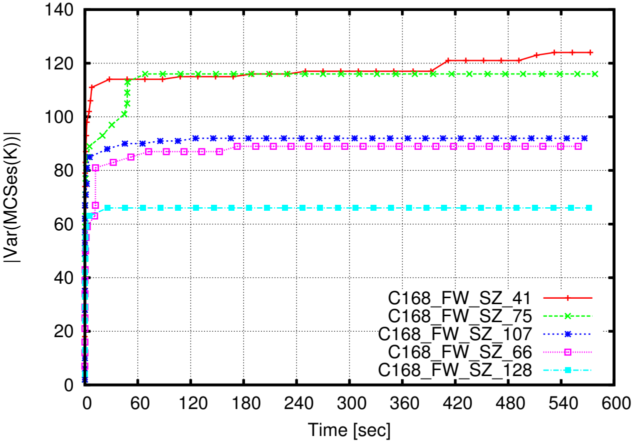


Figure : Anytime Property of CAMUS_IDMUS

Outline

Summary

- ID_{MUS}: inconsistency measurement by counting variables in *MUSes*
- $ID_4 \leq ID_{MUS} = ID_{MCS} \leq ID_Q$
- Complexity of ID_{MUS} is intractable: second layer of polynomial hierarchy
- The anytime algorithm and experiments show feasibility

- As a by-product, the relationship between *MUSes*, 4-models, Q-models are also interesting:

  informally, variables in $MUSes(K)$ are in between of the minimal 4-models and Q-models

Future Work

- Different inconsistency measurements have different views on inconsistency, we should combine them
- More efficient algorithm and implementations are needed

References

Grégoire, É., Mazure, B., and Piette, C. (2007).
Boosting a complete technique to find MSS and MUS thanks to a local search oracle.
In Veloso, M. M., editor, *IJCAI*, pages 2300–2305.

Hunter, A. and Konieczny, S. (2006).
Shapley inconsistency values.
In *Proc. of KR'06*, pages 249–259.

Hunter, A. and Konieczny, S. (2008).
Measuring inconsistency through minimal inconsistent sets.
In *Proc. of KR'08*, pages 358–366.

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

Xiao, G., Lin, Z., Ma, Y., and Qi, G. (2010).
Computing inconsistency measurements under multi-valued semantics by partial max-SAT solvers.
In *Proc. of KR'10*, pages 340–349.

Thanks!

MUS/MCS Finders

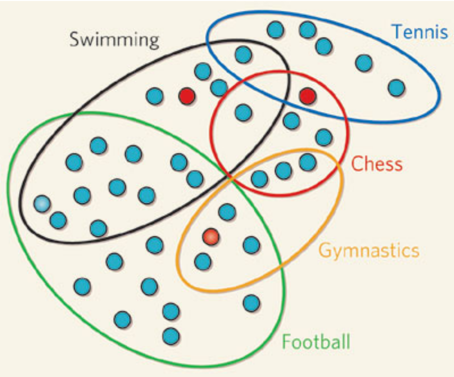
The state-of-the-art MCS/MUS finders are highly optimized Some of them are

- CAMUS (open sourced) [Liffiton and Sakallah, 2008],
- HYCAM [Grégoire et al., 2007].

Common steps in MUSes finders:

1. Computing MCSes with an incremental SAT solver
2. Using Hitting sets algorithm to find MUSes

Hitting Set



http://www.nature.com/nature/journal/v451/n7179/fig_tab/451639a_F1.html

- H is a hitting set of a set of sets Ω if $\forall S \in \Omega, H \cap S \neq \emptyset$.
- A hitting set H is irreducible if there is no other hitting set H', s.t. $H' \subsetneq H$.
- Remark: Hitting set problem in NP-complete

MUS/MCS Duality

Theorem [Liffiton and Sakallah, 2008]

Given an inconsistent knowledge base K:

- A subset M of K is an MCS of K iff M is an irreducible hitting set of $MUSes(K)$;
- A subset U of K is an MUS of K iff U is an irreducible hitting set of $MCSes(K)$.

Example

Let $K = \{p, \ \neg p, \ p \vee q, \ \neg q, \ \neg p \vee r\}$.

- $MUSes(K) = \{\{p, \neg p\}, \ \{\neg p, p \vee q, \ \neg q\}\}$
- $MCSes(K) = \{\{\neg p\}, \ \{p, p \vee q\}, \ \{p, \neg q\}\}$.

Clearly, $MUSes(K)$ and $MCSes(K)$ are hitting set duals of each other.