An Anytime Algorithm for Computing Inconsistency Measurement

Yue Ma[1] Guilin Qi[2] Guohui Xiao[3,5] Pascal Hitzler[4] Zuoquan Lin[3]

[1]Laboratoire d'Informatique de Paris-Nord, Université Paris-Nord CNRS, France
[2]School of Computer Science and Engineering, Southeast University, Nanjing, China
[3]Department of Information Science, Peking University, China
[4]Kno.e.sis Center, Wright State University, Dayton, OH, USA
[5]Institut für Informationssysteme, Technische Universität Wien, Austria

Motivation

- Consistent KBs serve as useful knowledge resources v.s. inconsistent KBs imply any conclusion (meaningless!)
- For handling inconsistent KBs:
  - paraconsistent reasoning (1960s)
  - knowledge diagnose and repair (1980s)
  - Which approach should we take?

  $\rightsquigarrow$ inconsistency measurement: a guidance to choice different approaches (2000s)
- How about the computational aspects of inconsistent measurement?

Introductive Example

- $K = \{p, \neg q, r\} \rightsquigarrow$ consistent
- $K' = \{p, \neg q, r, \neg p \vee q\} \rightsquigarrow$ inconsistent
- $K'' = \{p, \neg p, q, \neg q\} \rightsquigarrow$ inconsistent

The inconsistency degrees (ID):

$$ID(K) = 0, ID(K') = \frac{1}{3}, ID(K'') = 1$$

Related Work and Our Contribution

Related work:

- Defining (various) inconsistency degrees:

  (1) syntax-based; (2) *semantics-based*
- Algorithms
  - for restricted KBs: [GrantHunter08] only deals with KBs in the form $Q_1x_1, ..., Q_nx_n. \bigwedge_i (P_i(t_1, ..., t_{m_i}) \wedge \neg P_i(t_1, ..., t_{m_i}))$, ;
  - with high complexity: [MaQiHLin2007] with exponential times of invoking a SAT solver

Our work:

- To show that computing IDs is intractable generally but can be approximated polynomially

Inconsistency Degree by 4-valued Semantics

The set of truth values
$\{t, f, BOTH, NONE\}$

A 4-model I:
$Var(K) \rightarrow \{t, f, BOTH, NONE\}$

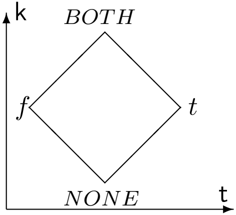

Figure: FOUR

- $Conflict(I, K) = \{p \mid p \in Var(K), p^I = BOTH\}$,
- $PreferModel(K) = \{I \mid \forall I' \in \mathcal{M}_4(K), |Conflict(I, K)| \leq |Conflict(I', K)|\}$.
- $ID(K) = \frac{|Conflict(I,K)|}{|Var(K)|}$, where I is a preferred model.

⇝ $K' = \{p, \neg q, r, \neg p \vee q\}: \ ID(K') = \frac{1}{3}$

⇝ $I_1 : p^{I_1} = BOTH, q^{I_1} = f, r^{I_1} = t,$
$I_2 : p^{I_2} = f, q^{I_2} = BOTH, r^{I_2} = t$

Computational Complexities

Given a propositional knowledge base K and a number $d \in [0, 1]$:

- ID$_{\leq d}$ (resp. ID$_{<d}$): is $ID(K) \leq d$ (resp. $ID(K) < d$)?
- ID$_{\geq d}$ (resp. ID$_{>d}$): is $ID(K) \geq d$ (resp. $ID(K) > d$)?
- EXACT-ID: is $ID(K) = d$?
- ID: what is the value of $ID(K)$?

Theorem

- *ID$_{\leq d}$ and ID$_{<d}$ are* **NP**-*complete;*
- *ID$_{\geq d}$ and ID$_{>d}$ are* **coNP**-*complete;*
- *EXACT-ID is* **DP**-*complete;*
- *ID is* $\mathbf{\Theta_2^P}$-*complete.*

Formal Definitions of Approximating IDs

Definition (Bounding Values)

Lower bounding value x*:* $x \leq ID(K)$; *Upper bounding value* y*:* $ID(K) \leq y$.

Definition (Bounding Models)

Given a preferred model I*:*

$$\begin{array}{rl} \text{Lower bounding model } I' \text{ of } K: & |Conflict(I', K)| \leq |Conflict(I, K)| \\ \text{Upper bounding model } I'' \text{ of } K: & |Conflict(I'', K)| \geq |Conflict(I, K)| \\ & \text{and } I'' \in \mathcal{M}_4(K) \end{array}$$

Formal Definitions of Approximating IDs

Definition (Bounding Values)

Lower bounding value x: $x \leq ID(K)$; Upper bounding value y: $ID(K) \leq y$.

Definition (Bounding Models)

Given a preferred model I:

$$\begin{aligned} \textit{Lower bounding model } I' \textit{ of } K: \quad & |\textit{Conflict}(I', K)| \leq |\textit{Conflict}(I, K)| \\ \textit{Upper bounding model } I'' \textit{ of } K: \quad & |\textit{Conflict}(I'', K)| \geq |\textit{Conflict}(I, K)| \\ & \textit{and } I'' \in \mathcal{M}_4(K) \end{aligned}$$

Requirements on Algorithms for Approximating IDs

An anytime approximating algorithm for computing inconsistency degrees should be able to produce two sequences $r_1, ..., r_m$ and $r^1, ..., r^k$:

$$r_1 \leq ... \leq r_m \leq ID(K) \leq r^k \leq ... \leq r^1, \quad (1)$$

such that these two sequences have the following properties:

- Tractability: $\exists . f(|K|)$, $g(|K|)$ s.t. computing r_i and r^j both stay tractable if $i \leq f(|K|)$ and $j \leq g(|K|)$;
- Convergence: $|ID(K) - r_{i+1}| < |ID(K) - r_i|$, $|ID(K) - r^i| < |ID(K) - r^{i+1}|$;
- Meaning: each r_i (r^j) corresponds to a lower (an upper) bounding model, which indicates the sense of the two sequences.

Requirements on Algorithms for Approximating IDs

An anytime approximating algorithm for computing inconsistency degrees should be able to produce two sequences $r_1, ..., r_m$ and $r^1, ..., r^k$:

$$r_1 \leq ... \leq r_m \leq ID(K) \leq r^k \leq ... \leq r^1, \quad (1)$$

such that these two sequences have the following properties:

- Tractability: $\exists . f(|K|)$, $g(|K|)$ s.t. computing r_i and r^j both stay tractable if $i \leq f(|K|)$ and $j \leq g(|K|)$;
- Convergence: $|ID(K) - r_{i+1}| < |ID(K) - r_i|$, $|ID(K) - r^i| < |ID(K) - r^{i+1}|$;
- Meaning: each r_i (r^j) corresponds to a lower (an upper) bounding model, which indicates the sense of the two sequences.

Requirements on Algorithms for Approximating IDs

An anytime approximating algorithm for computing inconsistency degrees should be able to produce two sequences $r_1, ..., r_m$ and $r^1, ..., r^k$:

$$r_1 \leq ... \leq r_m \leq ID(K) \leq r^k \leq ... \leq r^1, \tag{1}$$

such that these two sequences have the following properties:

- Tractability: $\exists . f(|K|), g(|K|)$ s.t. computing r_i and r^j both stay tractable if $i \leq f(|K|)$ and $j \leq g(|K|)$;
- Convergence: $|ID(K) - r_{i+1}| < |ID(K) - r_i|$, $|ID(K) - r^i| < |ID(K) - r^{i+1}|$;
- Meaning: each r_i (r^j) corresponds to a lower (an upper) bounding model, which indicates the sense of the two sequences.

Approximations from Above and Below

For a given w $(1 \leq w \leq |Var(K)|)$:

Theorem (Approximation from Above)

If K is w-4 satisfiable, then $ID(K) \leq 1 - w/|Var(K)|$.

Theorem (Approximation from Below)

If K is w-4 unsatisfiable, then $ID(K) \geq 1 - (w-1)/|Var(K)|$.

Definition. K is w-4 satisfiable iff. there is a subset $S \subseteq Var(K)$ such that K is S-4 satisfiable, i.e., K has a 4-model in the form of

$$p^{\mathfrak{I}} \in \begin{cases} \{B\} & \text{if } p \in Var(K) \setminus S, \\ \{N, t, f\} & \text{if } p \in S. \end{cases}$$

Approximations from Above and Below

For a given w $(1 \leq w \leq |\mathit{Var}(K)|)$:

Theorem (Approximation from Above)

If K is w-4 satisfiable, then $\mathit{ID}(K) \leq 1 - w/|\mathit{Var}(K)|$.

Theorem (Approximation from Below)

If K is w-4 unsatisfiable, then $\mathit{ID}(K) \geq 1 - (w-1)/|\mathit{Var}(K)|$.

Definition. K is w-4 satisfiable iff. there is a subset $S \subseteq \mathit{Var}(K)$ such that K is S-4 satisfiable, i.e., K has a 4-model in the form of

$$p^{\mathfrak{I}} \in \begin{cases} \{B\} & \text{if } p \in \mathit{Var}(K) \setminus S, \\ \{N, t, f\} & \text{if } p \in S. \end{cases}$$

Tractability of the Approximations

Theorem (Complexity)

There exists an algorithm for deciding if K is S-4 unsatisfiable in $\mathcal{O}(|K||S| \cdot 2^{|S|})$ time for any given $S \subseteq Var(K)$.

$\rightsquigarrow$ S-4 unsatisfiability can be computed in P-time, if $|S| = \mathcal{O}(\log |K|)$.

Tractable Anytime Algorithm

Suppose r_i, r^j are defined as follows $(1 \leq w \leq |Var(K)|)$:

$$r^j = 1 - w/|Var(K)|, \text{ where } K \text{ is } w\text{-4 satisfiable;}$$

$$r_i = 1 - \frac{w-1}{|Var(K)|}, \text{ where } K \text{ is } w\text{-4 unsatisfiable.}$$

- If $w = \mathcal{O}(\log |K|)$, computing upper bounds can be done in P-time w.r.t $|K|$.
- If w is limited by a constant, computing lower bounds can be done in P-time w.r.t. $|K|$.
- $r_i(r^j)$ corresponds to inconsistency degrees of K w.r.t. its upper (lower) bounding models.

$\rightsquigarrow$ Meets all the requirements given previously for tractable anytime algorithms.

Tractable Anytime Algorithm

Tow main sources of complexity to compute approximating inconsistency degrees:

1. *the complexity of* w*-4 satisfiability* $\rightsquigarrow$ solved by previous results
2. *the complexity of search space* $\rightsquigarrow$ a truncation strategy to limit the search space by the monotonicity of S-4 unsatisfiability:

   For all S, if K is S-4 unsatisfiable, K is S'-4 unsatisfiable for all $S' \supset S$.

Primary Experiment

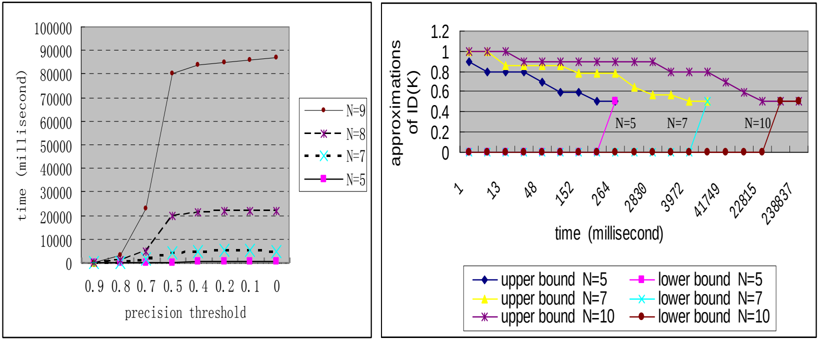

Figure: Evaluation results over KBs with $|K| = N^2 + 2N$ and $|Var(K)| = 2N$ for $N = 5, 7, 8, 9, 10$.

Conclusion and Outlook

- Conclusion
  1. Studied the problem complexity of ID (intractable, Θ_2^p-complete)
  2. Defined approximating inconsistency degrees
  3. Proposed a tractable anytime algorithm for computing approximating IDs
- Outlook
  1. To test the algorithm on more benchmark datasets
  2. To explore more optimization for the algorithm

Thanks for Your Attention!

Questions?