Inconsistency-Tolerant Reasoning with Classical Logic and Large Databases

Jui-Yi Kao Stanford University

Presenting on joint work with:

Timothy L. Hinrichs University of Chicago
Michael Genesereth Stanford University

Challenge 1: Inconsistencies

- Occasional errors and disagreements are unavoidable in real-world data.
  - Data acquisition error
  - Out-of-sync
  - Genuine disagreement: Julius Caesar birth year
  - Semantic disagreement: measuring GDP
  - Approximation – apparent contradictions

Jui-Yi Kao
Stanford University

Tolerate Inconsistency

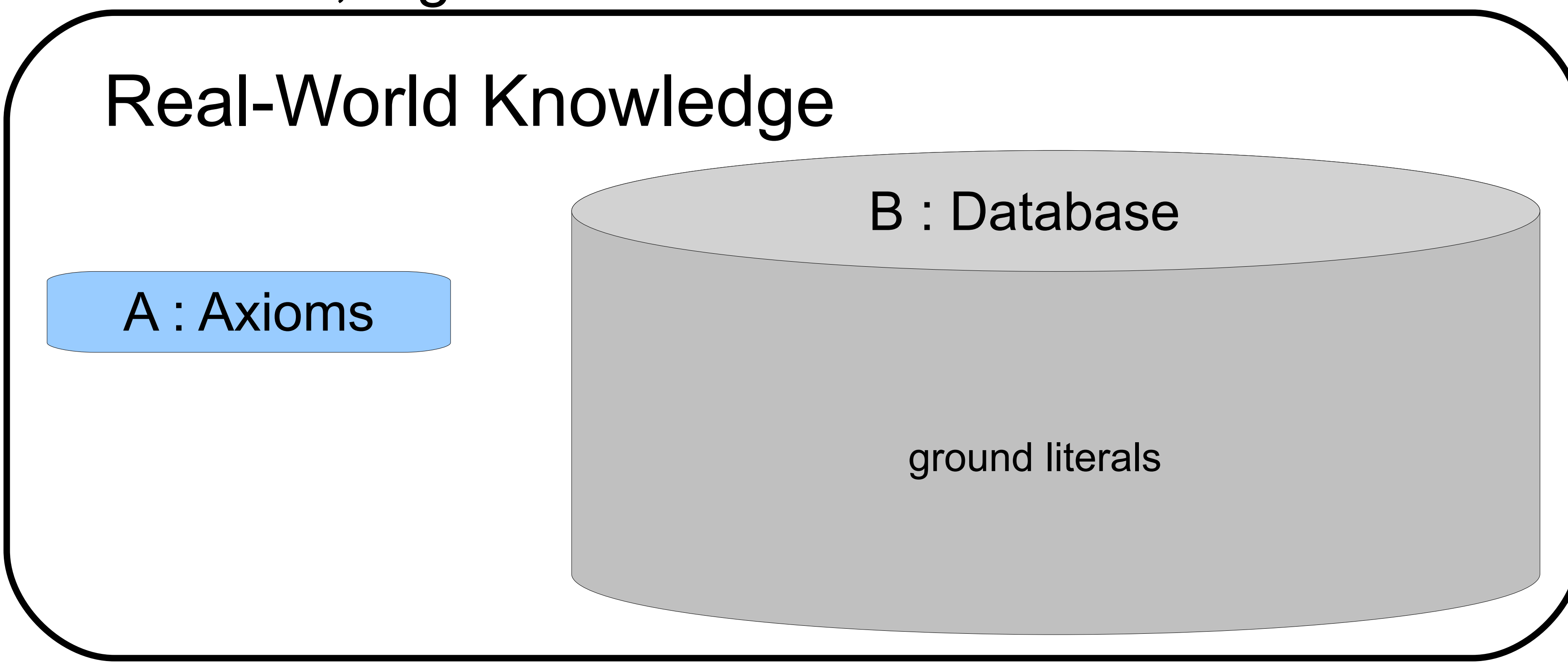
- Classical logic does not tolerate inconsistency
  - If $K \models \bot$ then $K \models \varphi$ for any sentence φ
- Many inconsistency-tolerant reasoning methods
  - Strict Existential Entailment

Challenge 2: Large Premise Set

- Vast amounts of data stored in relational databases
  - 10 Petabytes in Yahoo!'s Everest
- Most automated reasoning systems not designed to handle large premise sets

Real-World Knowledge

- Knowledge in the real world split naturally into
  - Data, represented in databases
  - Axioms, logical sentences



Presentation Outline

- Definition of Strict Existential Entailment
- Naïve method
- Our approach: compilation

Strict Existential Entailment

- Given a set of axioms A and a database B,
  - $A, B \models_E \mathbf{I}(\overline{\mathbf{a}})$

    $\Leftrightarrow$ a consistent portion B* of B classically entails $\mathbf{I}(\overline{\mathbf{a}})$

    – ie. $\exists B^* \sqsubset B \cdot A \cup B^* \not\models \bot$ and $A \cup B^* \models \mathbf{I}(\overline{\mathbf{a}})$

- Strict entailment for short

Jui-Yi Kao
Stanford University

Example

- Axioms A:
  - $p(X,Y) \vee \neg q(Z,a) \vee r(Z)$
  - $p(a,U) \vee \neg q(U,a)$
- Database B:
  - $\neg p(a,b)$
  - $q(a,a)$
  - $q(b,a)$

Jui-Yi Kao
Stanford University

Example

- Axioms A:
  - p(X,Y) ∨ ¬q(Z,a) ∨ r(Z)
  - p(a,U) ∨ ¬q(U,a)
- Database B:
  - ¬p(a,b) [1]
  - q(a,a)
  - q(b,a)

→

r(a)
¬q(b,a)

Jui-Yi Kao
Stanford University

Example

- Axioms A:
  - p(X,Y) ∨ ¬q(Z,a) ∨ r(Z)
  - p(a,U) ∨ ¬q(U,a)
- Database B:
  - ¬p(a,b)
  - q(a,a)
  - q(b,a)

2

r(a)
¬q(b,a)

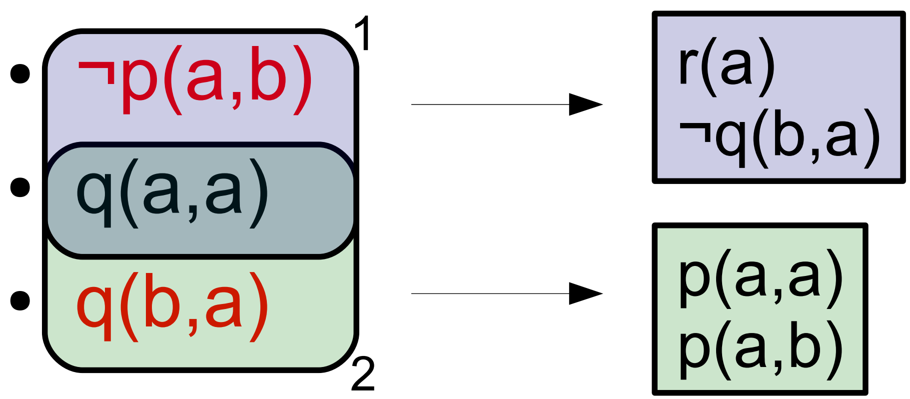
→

p(a,a)
p(a,b)

Jui-Yi Kao
Stanford University

Example

- Axioms A:
  - p(X,Y) ∨ ¬q(Z,a) ∨ r(Z)
  - p(a,U) ∨ ¬q(U,a)
- Database B:



- r(b) is excluded!

Jui-Yi Kao
Stanford University

Naïve Method

- Consider each consistent (maximal) subset of the data
- Find the the classically entailed conclusions for each subset
- There may be exponentially many consistent maximal subsets!

p(A,B)

A	a1	a1	a2	a2	...	an	an
B	b0	b1	b0	b1	...	b0	b1

Axiom: $p(X,Y) \wedge p(X,Z) \rightarrow Y = Z$

A relation of **2n** tuples has **2^n** consistent maximal portions!

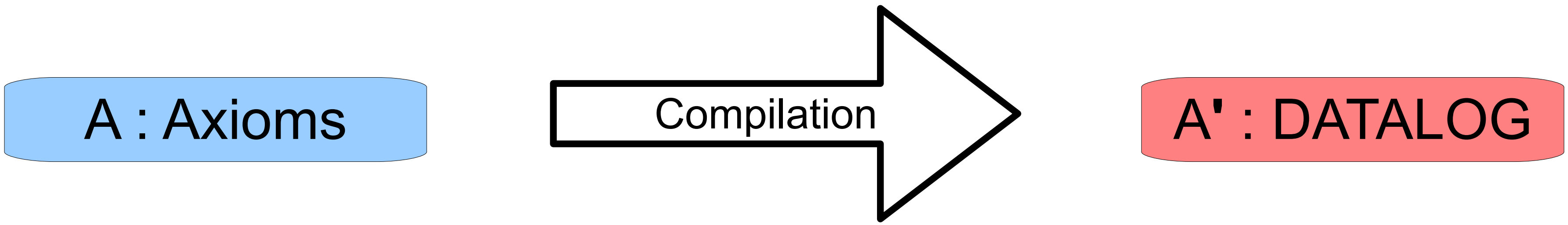
Jui-Yi Kao
Stanford University

Concentrate on the Axioms

- Axioms A:
  - p(X,Y) ∨ ¬q(Z,a) ∨ r(Z)
  - p(a,U) ∨ ¬q(U,a)
- Database B:
  - ¬p(a,b)
  - q(a,a)
  - q(b,a)
- Deduction:
  - `¬p(a,b)      q(b,a)      r(Z)`
  - `p(a,U)      ¬q(U,a)`

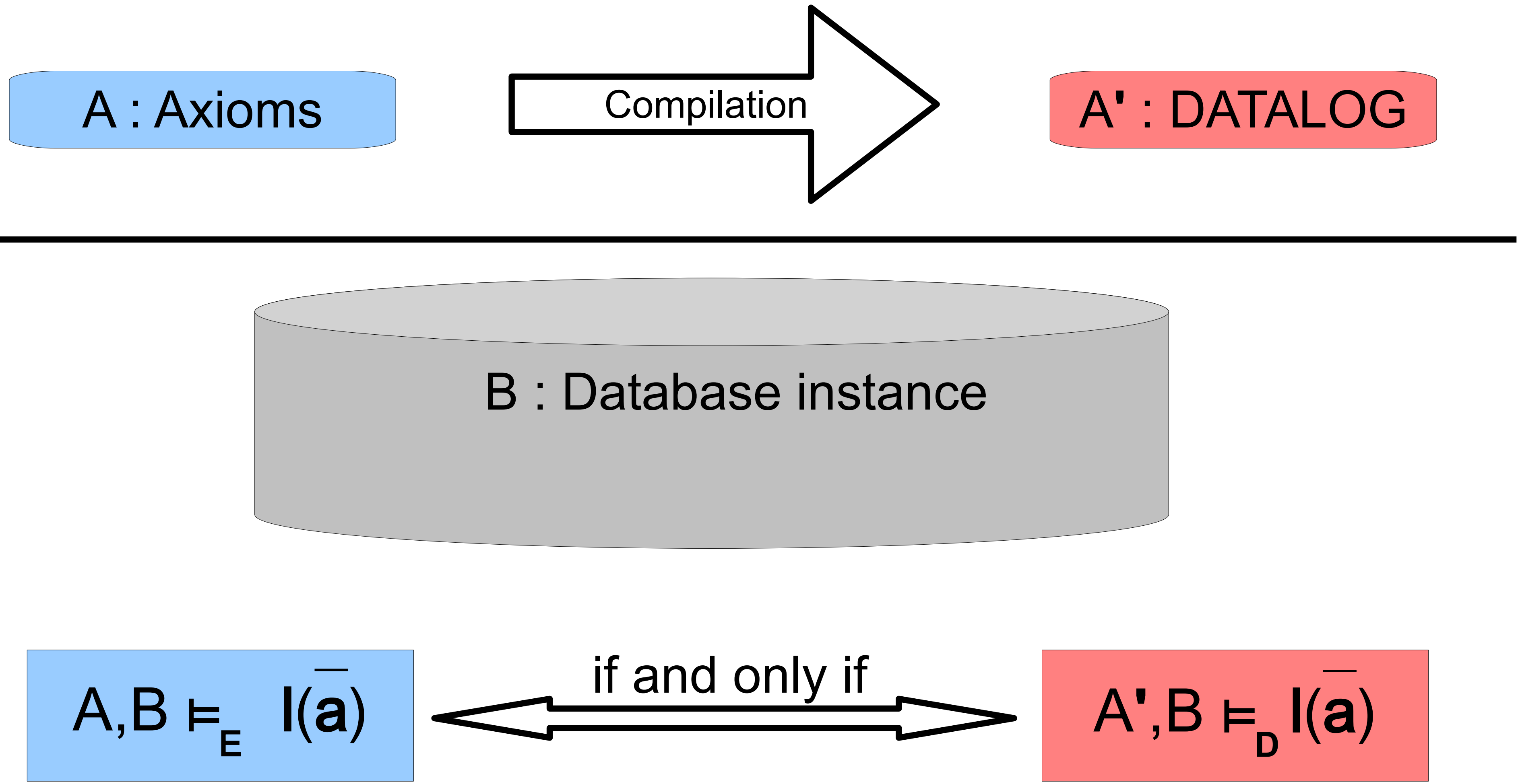
Jui-Yi Kao
Stanford University

Inconsistency-Tolerant Compilation Approach



Jui-Yi Kao
Stanford University

Inconsistency-Tolerant Compilation Approach



Jui-Yi Kao
Stanford University

Setting

- Axioms A: first-order logic with equality:
  - Function-free
  - Universal clause
- Relational database B
- Domain closure assumption
- Unique names assumption

Jui-Yi Kao
Stanford University

Compilation to DATALOG

p(X,Y) ¬q(Z,a) r(Z)
p(a,U) ¬q(U,a)

*See Algorithm 1 in paper

Jui-Yi Kao
Stanford University

Compilation to DATALOG

$p(X,Y) \quad \neg q(Z,a) \quad r(Z)$

$p(a,U) \quad \neg q(U,a)$

resolution

contrapositives

$$
\begin{aligned}
p(X,Y) &\leftarrow q(Z,a) \wedge \neg r(Z) \\
\neg q(Z,a) &\leftarrow p(X,Y) \wedge \neg r(Z) \\
r(Z) &\leftarrow \neg p(X,Y) \wedge q(Z,a) \\
\neg q(U,a) &\leftarrow \neg p(a,U) \\
p(a,U) &\leftarrow q(U,a)
\end{aligned}
$$

*See Algorithm 1 in paper

Jui-Yi Kao
Stanford University

Compilation to DATALOG

$$\begin{array}{lll} p(X,Y) & \neg q(Z,a) & r(Z) \\ p(a,U) & \neg q(U,a) & \end{array}$$

resolution

contrapositives

$$\begin{array}{rcllll} p(X,Y) & \leftarrow & q(Z,a) & \wedge & \neg r(Z) \\ \neg q(Z,a) & \leftarrow & p(X,Y) & \wedge & \neg r(Z) \\ r(Z) & \leftarrow & \neg p(X,Y) & \wedge & q(Z,a) \\ \neg q(U,a) & \leftarrow & \neg p(a,U) & & \\ p(a,U) & \leftarrow & q(U,a) & & \end{array}$$

DATALOG

$$\begin{array}{lcllll} p^{+}(X,Y) & :- & q(Z,a) & \wedge & \neg r(Z) \\ q^{-}(Z,a) & :- & p(X,Y) & \wedge & \neg r(Z) \\ r^{+}(Z) & :- & \neg p(X,Y) & \wedge & q(Z,a) \\ q^{-}(U,a) & :- & \neg p(a,U) & & \\ p^{+}(a,U) & :- & q(U,a) & & \end{array}$$

*See Algorithm 1 in paper

Jui-Yi Kao
Stanford University

Inconsistency

- Database B:
  - ¬p(a,b)
  - q(a,a)
  - q(b,a)

```
r+(Z) :- ¬p(X,Y) ∧ q(Z,a)
```

```
r+(a) :- ¬p(a,b) ∧ q(a,a)
```

```
r+(b) :- ¬p(a,b) ∧ q(b,a)
```

Jui-Yi Kao
Stanford University

Inconsistency

- Database B:
  - ¬p(a,b)
  - q(a,a)
  - q(b,a)

$r^{+}(Z)$:- ¬p(X,Y) ∧ q(Z,a)

$r^{+}(a)$:- ¬p(a,b) ∧ q(a,a)

~~r^{+}~~(b) :- ¬p(a,b) ∧ q(b,a)

Jui-Yi Kao
Stanford University

Augment for Inconsistency

- rule: `r⁺(Z) :- ¬p(X,Y) ∧ q(Z,a)`
- Negated rule body

  ¬b : p(X,Y) ∨ ¬q(Z,a)
- Axiom clause c : p(a,U) ∨ ¬q(U,V)

*See Algorithm 3 in paper

Jui-Yi Kao
Stanford University

Augment for Inconsistency

- rule: `r⁺(Z) :- ¬p(X,Y) ∧ q(Z,a)`
- Rule body

  $b: \neg p(X,Y) \vee q(Z,a)$
- Axiom clause $c: p(a,U) \vee \neg q(U,V)$
- $c, b\sigma \models \bot \Leftrightarrow [X = a \wedge Y = Z]\sigma$
- Augmented rule:

  `r⁺(Z) :- ¬p(X,Y) ∧ q(Z,a) ∧ ¬[X = a ∧ Y = Z]`

*See Algorithm 3 in paper

Jui-Yi Kao
Stanford University

Evaluate on Example Data

- Database B:
  - ¬p(a,b)
  - q(a,a)
  - q(b,a)

$r^+(Z)$:- ¬p(X,Y) ∧ q(Z,a) ∧ ¬[X = a ∧ Y = Z]

$r^+(a)$:- ¬p(a,b) ∧ q(a,a) ∧ ¬[a = a ∧ b = a]

$r^+(b)$:- ¬p(a,b) ∧ q(b,a) ∧ ¬[a = a ∧ b = b]

Jui-Yi Kao
Stanford University

Termination

- The compilation algorithm terminates when the input axioms A has a finite closure under resolution and factoring.

Jui-Yi Kao
Stanford University

Sound and Complete

- Theorem:
  - Assume:
    - Function-free universal axioms in FOL with =
    - Domain closure assumption
    - Unique names assumption
  - The compilation is sound and complete for strict existential entailment.

$$A,B \models_E I(\bar{a}) \quad \text{if and only if} \quad A',B \models_D I(\bar{a})$$

*See Theorems 1 and 2 in paper

Jui-Yi Kao
Stanford University

Features

- Compile independently of data
- flat Datalog$^\neg$ ⊏ RA ⊏ SQL
  - Polynomial data complexity
- Simple layer over existing DBMS
  - Custom code ignores data
  - Low cost of adoption
  - Leverage current state-of-the-art infrastructure
- Reuse on different/evolving data

Related Work

- Inconsistency tolerance based on classical logic
  - (Hunter 1998; Besnard & Hunter 2005; Konieczny, Lang & Marquis 2005; Huang, van Harmelen & ten Teije 2005; Zamansky & Avron 2006; Flouris et al. 2006; Subrahmanian & Amgoud 2007; Hunter and Konieczny 2008; Everaere, Konieczny, and Marquis 2008; Besnard and Hunter 2008)
- Knowledge compilation
  - (Darwiche & Marquis 2002; Selman & Kautz 1996; Nagy, Lukacsy & Szeredi 2006; Calvanese et al. 2008; Besnard & Hunter 2006; Hinrichs & Genesereth 2008; Cadoli & Mancini 2002; Nagy, Lukacsy & Szeredi 2006; Calvanese et al. 2008; Flouris et al. 2006; Huang, van Harmelen & ten Teije 2005; Gomez, Chesnevar & Simari 2008)

Future Work

- Existential quantification
- When resolution takes long:
  - compile into recursive Datalog or Prolog
- Give relationship between conclusions
  - rebuttal and undercutting

Inconsistency-Tolerant Reasoning with Classical Logic and Large Databases

Jui-Yi Kao Stanford University

Presenting on joint work with:

Timothy L. Hinrichs University of Chicago
Michael Genesereth Stanford University

Presented on July 9, 2009 at the Symposium on Abstraction, Reformulation, and Approximation (SARA 2009) Lake Arrowhead, CA, U.S.A.

T. L. Hinrichs, J.-Y. Kao, M. Genesereth. Inconsistency-Tolerant Reasoning with Classical Logic and Large Databases. SARA 2009