

Authors

Yurun GU
SAMOVAR,
Telecom SudParis,
IP Paris

Keywords

Database, DC Violation,
Query Answering,
Inconsistency Measure,
Why-Provenance,
Answer Explainability

Contributions

1. Three novel tuple measures and their rational properties,
2. Framework of quantifying query answer inconsistency and their properties,
3. Computational complexity analysis and experimental comparisons show that:

- IM_{pim} is the best tuple inconsistency measure;
- $ICQA_{shap}^{IM}$ is the best answer inconsistency measure.

Definitions

- Tuple measure IM : the inconsistency contribution of $t \in I$ w.r.t. a constraint set Δ
- Answer measure $ICQA$: quantified inconsistency of a query answer by systematically aggregating the inconsistencies of its contributing tuples

Partners

BACKGROUND

- A database(DB) violates integrity rules - constraints
- Denial constraint (DC) forbids co-occurrence patterns
- Measuring inconsistency is useful for handling DC violations

EXAMPLE

$I \{ \begin{array}{l} t1: Age(Alice, 20) \quad t2: Age(Alice, 21) \quad q(x):-Age(x,21) \\ t3: Age(Bob, 21) \quad t4: Age(Bob, 22) \quad q(l)={Alice, Bob} \end{array} \right.$
DC: no two different ages for one person
Minimal inconsistent subset (MIS): $\{t1, t2\}, \{t3, t4\}$

How to quantify inconsistency of tuples and query answers? And which measures are better?

TUPLE MEASURE PROPERTIES

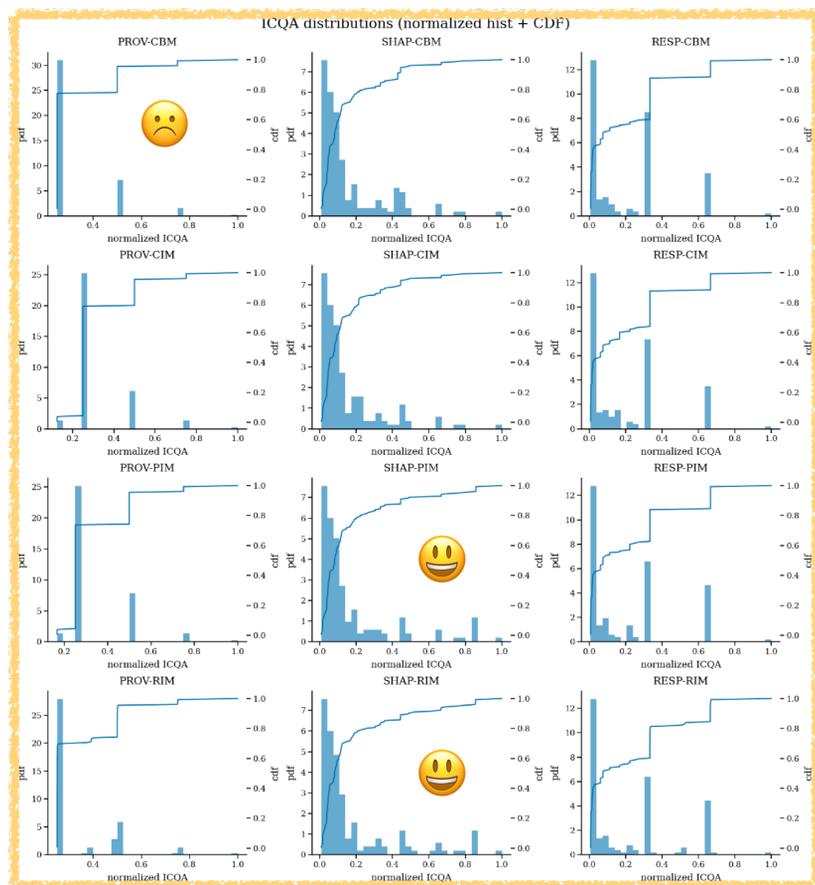
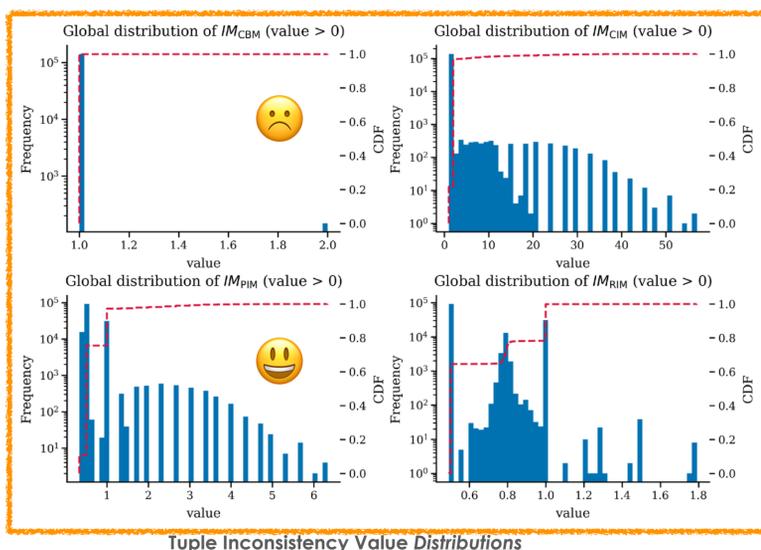
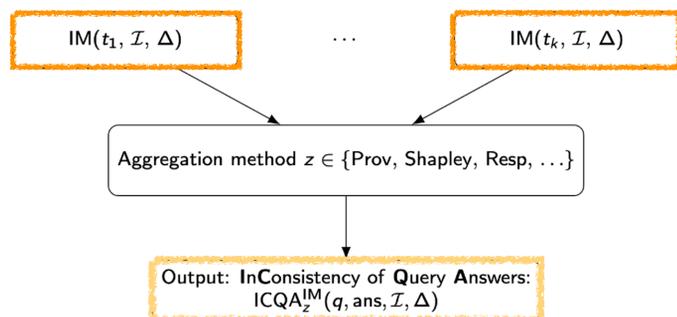
Free tuple independence (FI), Positivity (PO), Centrality Conflict (CC), Data Monotonicity (DM)

	IM_{PIM}	IM_{CIM}	IM_{RIM}	IM_{CBM}	IM_{shap}^{MIS}	IM_P	IM_{dr}	$IM_{MIS\#}$
FI	✓	✓	✓	✓	⊗	✓	✓	✓
PO	✓	✓	✓	✓	⊗	⊗	⊗	⊗
CC	✓	⊗	✓	⊗	✓	⊗	✓	✓
DM	✓	✓	⊗	✓	✓	✓	⊗	✓

ICQA MEASURE PROPERTIES

Blamelessness, Additivity, Monotonicity

- $ICQA_z^{IM}$ satisfies **Blamelessness** for any z and IM .
- $ICQA_{Prov}^{IM}$ satisfies **Additivity** for any IM .
- $ICQA_{Prov}^{IM}$ satisfies **Monotonicity** if IM satisfies Data Monotonicity.



Experimental evaluation on TPC-H dataset

CONCLUSION

		Inconsistency Measuring Methods						
		CBM	Shapley	KB	Provenance	Responsibility	Condition	
Answer Explaining Methods	Provenance	Green, Karvounarakis, and Tannen (2007)	Issa, Bonifati, and Toumani (2020)	Livshits and Kimelfeld (2022)	Raddaoui, Straßer, and Jabbour (2024)	PIM	RIM	CIM
	Shapley	Bienvenu, Figueira, and Lafourcade (2025); Livshits et al. (2021a); Reshef, Kimelfeld, and Livshits (2019)	Issa, Bonifati, and Toumani (2020)	$\{P, FP\#^P, FPRAS, ?\}$	$\{P, coNP-C\}$	P	NP-hard	P
	Responsibility	Meliou et al. (2010)	Issa, Bonifati, and Toumani (2020)	$ICQA_{shap}^{CBM}$	$ICQA_{shap}^{KB}$	$ICQA_{Prov}^{PIM}$	$ICQA_{Prov}^{RIM}$	$ICQA_{Prov}^{CIM}$
			$\{P, NP-hard\}$	$\{P, NP-hard\}$	$\{P, coNP-C\}$	$\{P, coNP-C\}$	NP-hard	NP-hard

Table: Existing works of tuple measure, answer explaining method and our works of tuple measure and answer inconsistency