Knowledge Representation Formalisms for Hybrid Intelligence

Loan Ho, Victor de Boer, Stefan Schlobach, Myrthe Tielman and Birna van Riemsdijk

Problem Description

Explanation is a prominent topic in AI. However, the explanations provided by existing approaches often lack critical information, in particular when the data comes with preferences. In this project,

1. we aim to **represent the inconsistent knowledge bases (KBs).**
2. we will investigate how to **explain query answering** in the context of **inconsistent KBs**, particularly when the date comes with preferences.

In the context of HI scenarios, inconsistencies in knowledge bases (KBs) can also occur for a variety of reasons. These include shifting preferences, user's motivation and or external conditions (for example, available resources and environment can vary over time) (Loan Ho et.al. 2022). We will investigate how to **represent the HI scenarios** and **provide explanations** during query answering.

Argumentation for knowledge base inconsistencies in hybrid intelligence scenarios, KR4HI workshop 2022, Loan Ho , Victor de Boer , M. Birna van Riemsdijk , Stefan Schlobach and Myrthe L. Tielman.

Challenges

- How to explain query answers to users in the case of inconsistency and preferences?
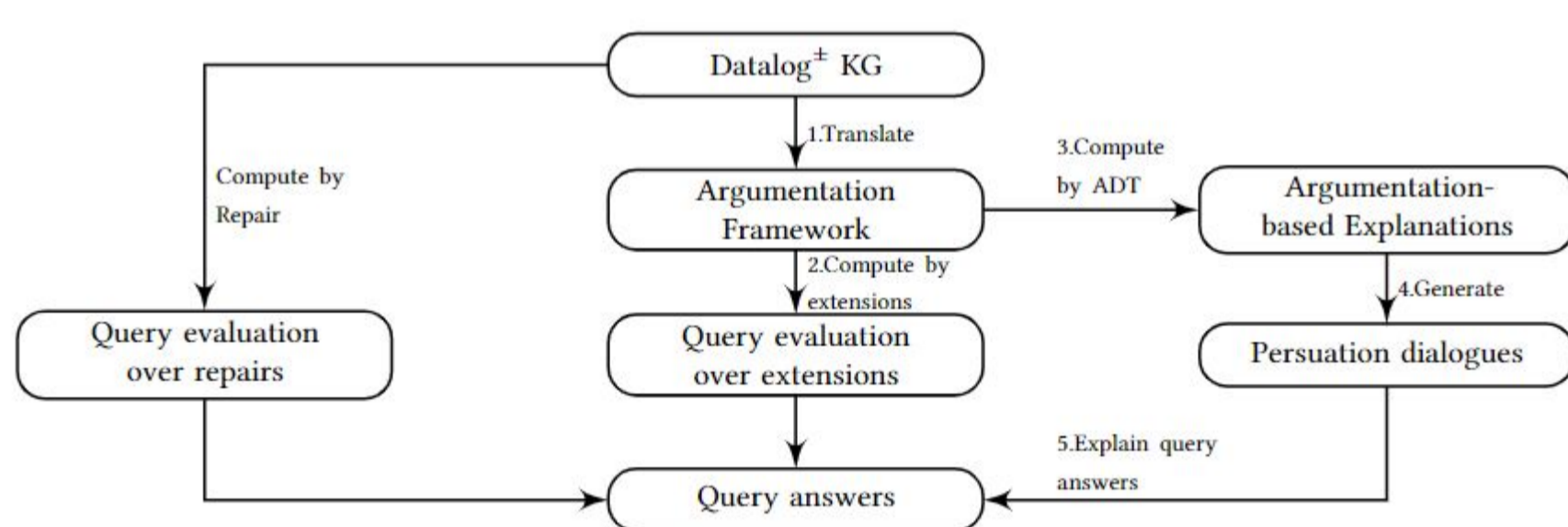- Where can preferences be elicited from?

Figure 1: Conceptual Framework

Methodology

We proposed a general conceptual framework for explaining query answers in KB, in which:

- Step 1: A faithful translation of prioritized Datalog$^\pm$ KB into a *novel logic-based argumentation framework with preferences (LAFP).*

- Step 2: By leveraging the link between repairs extensions of our LAFP we show that this translation preserves entailment (the answers are given).

- Step 3: We use argumentation to explain answers for w.r.t. different semantics. We propose a notion of explanation containing both causes and sets of conflicts. To compute such argumentation-based explanation, we use admissible dispute trees.

- Step 4: Based on these explanations, *persuasion dialogues* are constructed

Outcomes
- A general conceptual framework for explaining query answers in KB via Persuasion Dialogues
- An implementation to show the applicability of this framework to users

Future Works
- From a human-computer interaction perspective, we are planning to perform experiments with our new explanations in real-data applications with real users

- Qualitatively evaluate our explanation by human evaluation

- Analyze the computational complexity of computing the argument-based explanations both empirically and theoretically.
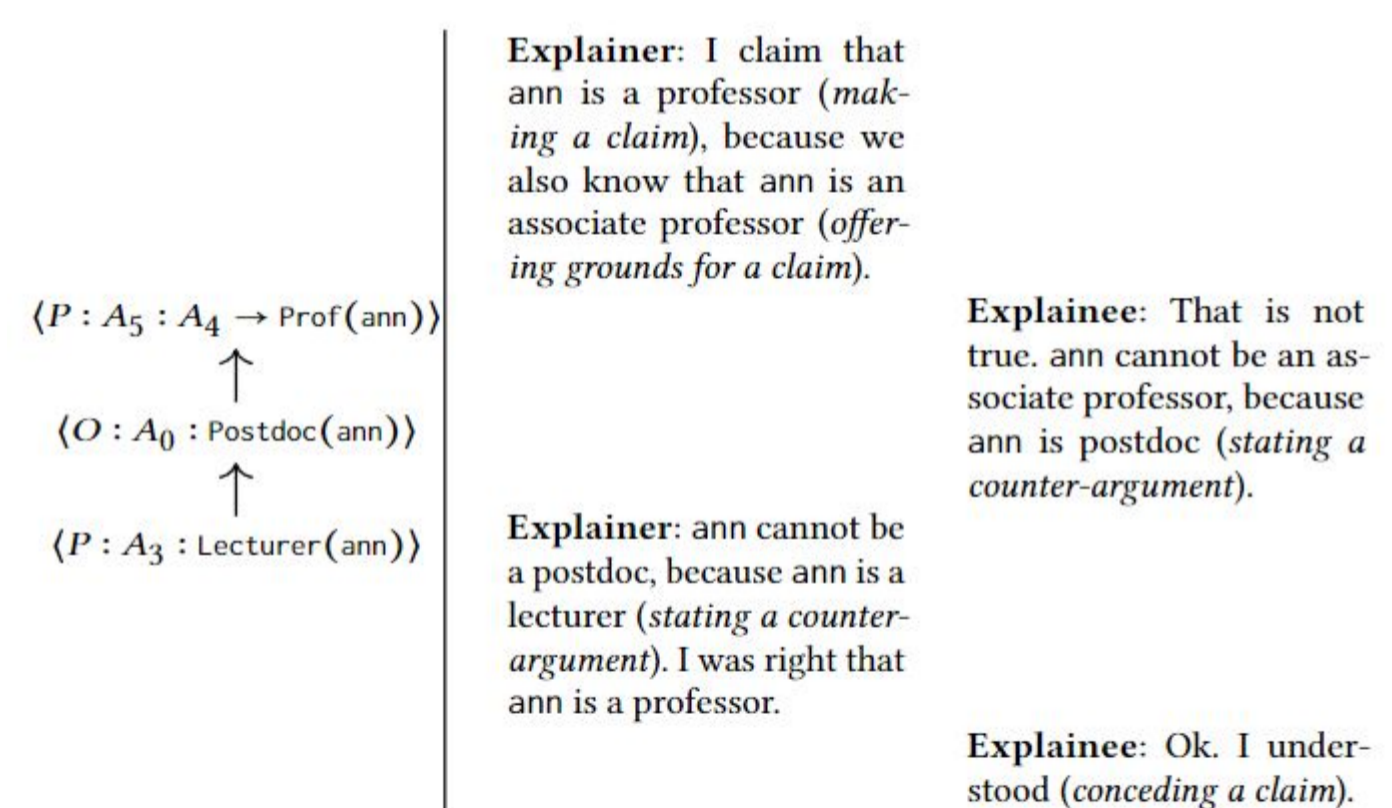

Figure 5: A persuasion dialogue created from $\mathcal{T}_{\mathcal{E}_2}$

Source code: https://github.com/LoanHo88/LAFP-framework.git