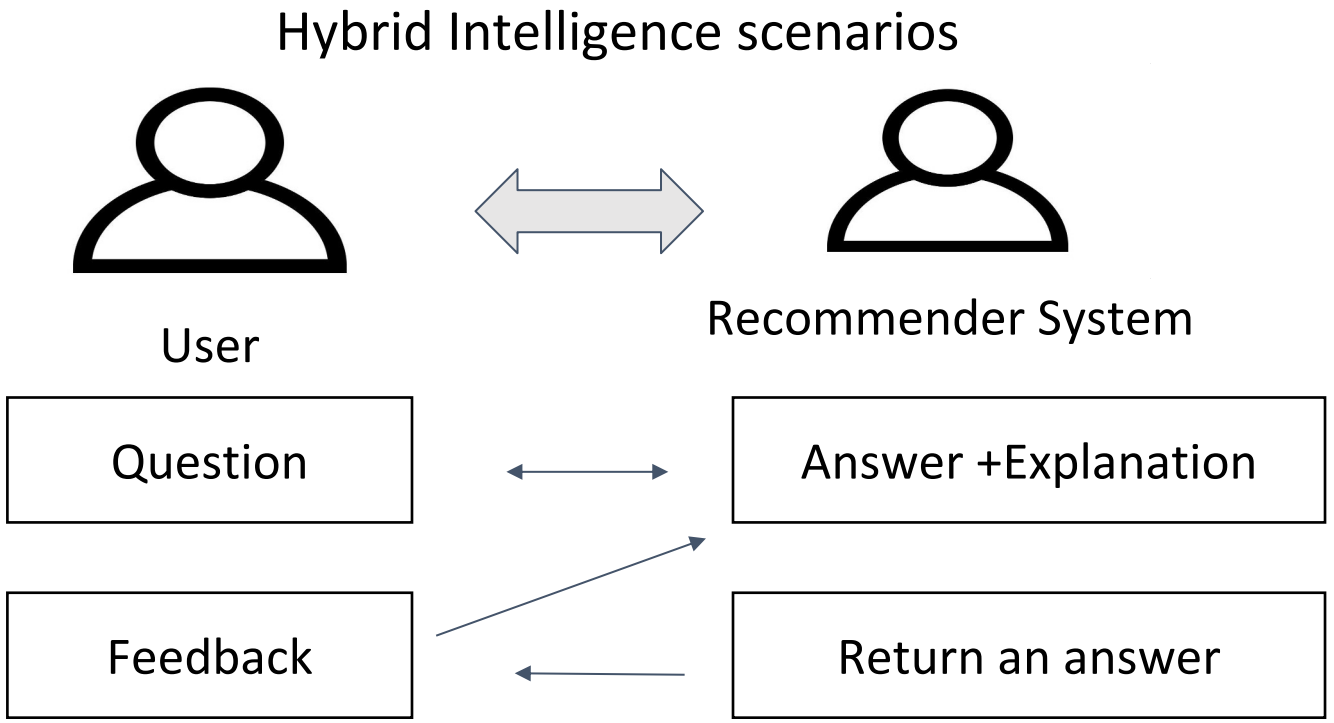
Abstract Argumentation for Hybrid Intelligence Scenarios

Loan Ho

Knowledge Representation & Reasoning Group

Hybrid Intelligence (HI)

According to the paper Akata et.al., *"Hybrid Intelligence (HI) systems* which combine *human and artificial intelligence* and attempt to *integrate human and machines* rather than use AI to *replace human intelligence*."



Knowledge Representation and Reasoning (KR) in AI

Hybrid Intelligence scenarios

Solution

Representation

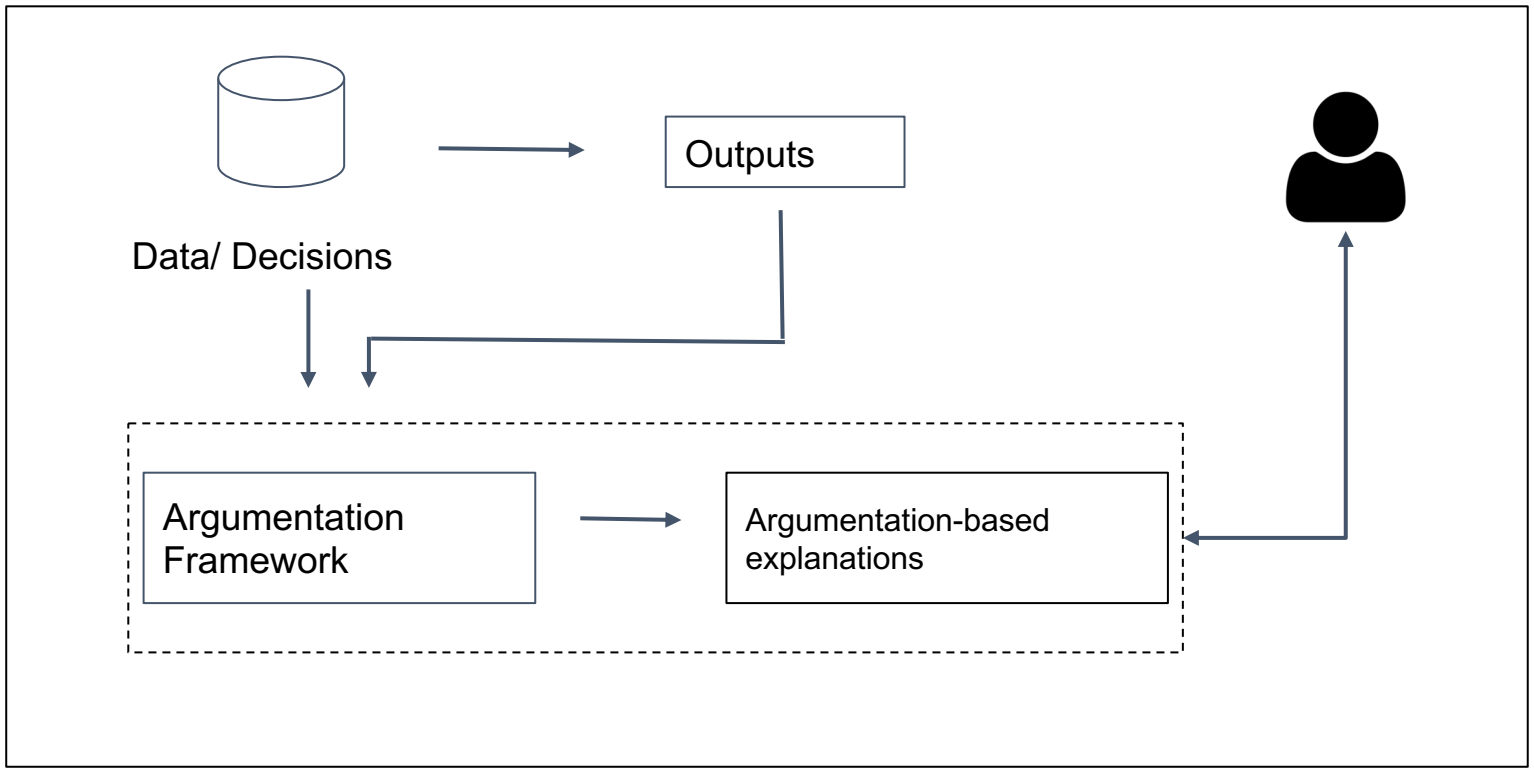
?

Automated Reasoning output

Research Questions

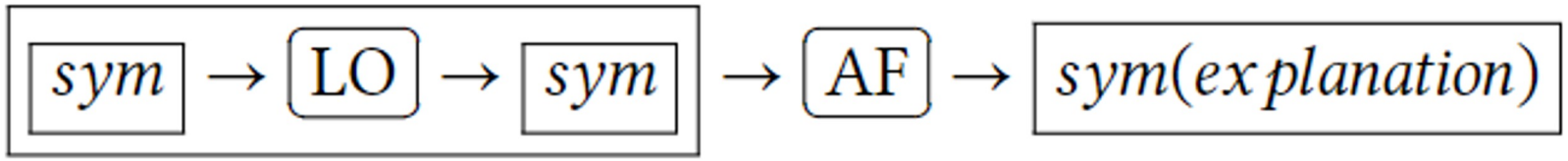
- **RQ1.** What are the capabilities of Argumentation in representing and reasoning knowledge of HI scenarios in the presence of inconsistencies?
- **RQ2.** How can Argumentation enable Explainability in HI scenarios?

XAI by computational argumentation

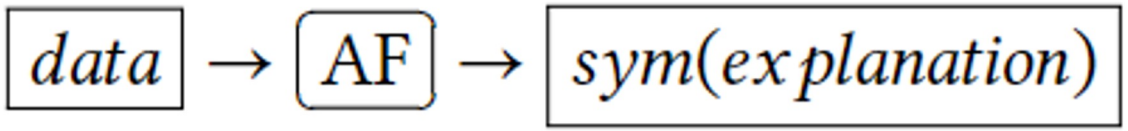


Translation of the HI scenarios into AF

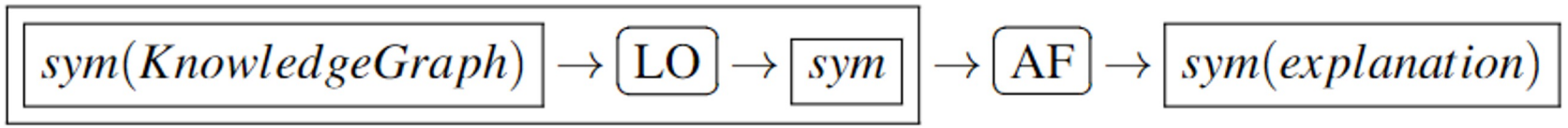
- Mediate translation method



- Immediate translation method



Example - Project 09



***Scenario**: The agent schedules the meeting at 10am. Unfortunately, the manager got sick, and he will not be able to join the meeting. He cancels the meeting at 10am.*

=> User wants to know why the meeting cancelled at 10am?

Consider $\mathscr{K}_1 = \{\mathscr{R}_1, \mathscr{C}_1, \mathscr{F}_1\}$, in which:

$\mathscr{R}_1 = \{R_1 : manager(x) \rightarrow bookMeeting(x, y, t_1),$

$R_2 : manager(x) \wedge gotSick(x, sick) \rightarrow cancelMeeting(x, y, t_2)\};$

$\mathscr{C}_1 = \{C : \forall x,y\ bookMeeting(x, y, t_1) \wedge cancelMeeting(x, y, t_2) \wedge t_1 = t_2 \rightarrow \bot\};$

$\mathscr{F}_1 = \{f_1 : manager(Tim), f_2 : bookMeeting(Tim, meetingA, 10am),$

$f_3 : gotSick(Tim, sick), f_4 : cancelMeeting(Tim, meetingA, 10am)\}.$

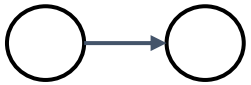
Abstract Argumentation Framework - Computational Argumentation

Argumentation Framework (abstraction of debate)

Argumentation Semantics (evaluation of debate)

Properties (goodness of semantics)

A1: Tim books a meeting at 10am



A2: Tim doesnot books a meeting at 10am

e.g. **preferred** semantics

(set of "winning"arguments)

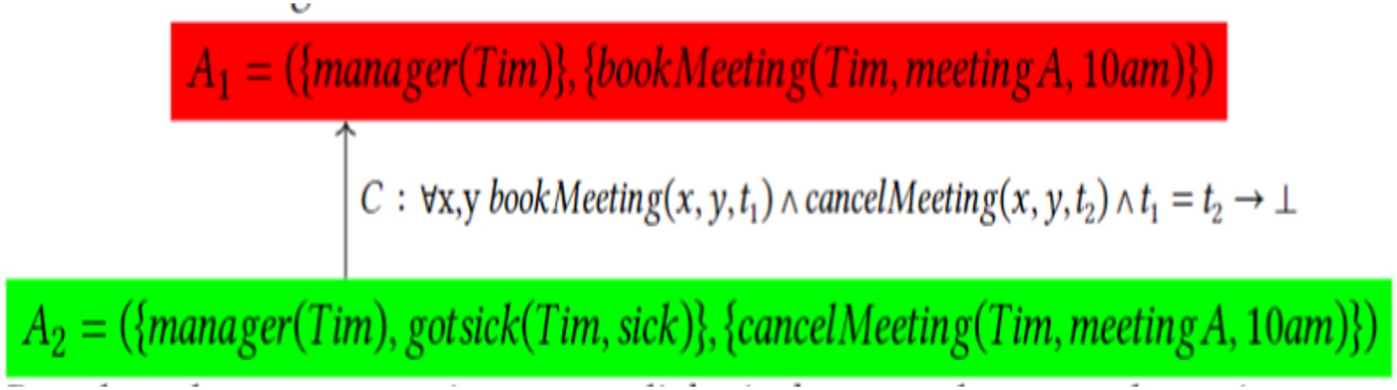
Is the meeting booked at 10am? **NO**

$$A_1 = (\{manager(Tim)\}, \{bookMeeting(Tim, meetingA, 10am)\})$$

$$C : \forall x,y\ bookMeeting(x,y,t_1) \wedge cancelMeeting(x,y,t_2) \wedge t_1 = t_2 \rightarrow \bot$$

$$A_2 = (\{manager(Tim), gotsick(Tim, sick)\}, \{cancelMeeting(Tim, meetingA, 10am)\})$$

Argumentation-based Explanation



User: Why not *bookMeeting(Tim,meetingA,10am)*?

System: Because *cancelMeeting(Tim,meetingA,10am)*

User: I understand there is a reason why meeting A is not booked at 10am by Tim.

Survey Research

- Research Methodology
- Participants
- Materials and procedure

Survey Research - Research Methodology

- **S1:** We conduct a survey for HI project members and the survey were analyzed by using qualitative data analysis methods.
  - Determine participants who are PhD candidates working on HI project members.
  - Design survey questions.
  - Conduct the survey by asking the participants for information through a questionnaire, which is online.
- **S2:** We investigate how Argumentation can assist in representing and reasoning inconsistent KBs of the scenarios and how Argumentation can support the vision of explainable AI.
  - Translate KBs of HI scenarios to AF.
  - Describe how Argumentation enables Explainability according to what they explain (i.e. providing explanations through *Decision-Making*, *Justification* of an opinion, and *Dialogues*).

Survey Research

- **Participants**
  - Conduct a survey among 26 sub-projects of the HI.
  - Five of the participants did not respond to our survey, which resulted in a final number of 21 contributing participants.

- **Materials and procedure**
  - Conduct a survey by asking the participants for information through a questionnaire (in online).
  - Conducted interviews (both online and face to face) focused on the projects that most clearly deal with inconsistencies.

Result survey

Project	Name of the project	Use cases/ Scenarios	Conflicting information	Using logic formalism for knowledge representation	Translation of the HI scenarios into AF	Type of problem	Reason for no (unknown) mapping to AF
2	Aligning learning and reasoning systems for responsible HI	Study a method for evaluating the decision-making of machine learning systems in fictional and non-fictional legal domains	-	-	-	-	no conflicting information
3	Mining texts for perspectives for human-machine deliberation	Deliberation platform to increase the scale of the discussions from relatively small to crowd-scale	Yes	-	Yes	Explanation through dialogue	
5	Computational deliberation	Self driving cars with communication, robot navigation with communication, multi-agent games with communication, social dilemmas with communication	Yes	-	Unknown	Explanation through dialogue	using synthetic simple numeric data or image data
6	Computational theory of mind for collaborative privacy	Human-agent and agent-agent negotiation/cooperation	Yes	-	Unknown	Explanation through dialogue	no having formal representations for knowledge (data)
8	Hybrid approaches to sequential decision making for explainable behaviour	Digital Assistants (Conversational Search, Conversational Recommender Systems)	Yes	Yes (Knowledge Graph)	Yes	Decision making	
9	AutoAI for dynamic data	Assistant Agents that can assist employees in calendar scheduling within a company	Yes	Yes	Yes	Decision making	
10	Continual Learning and Deep Generative Modeling for Adaptive Systems	Train a generative model on a steam of data without revisiting past data	-	-	-	-	no conflicting information
11	A Design Pattern Language for HI Teams	The pattern language is used in designing and thinking about HI projects, so that it leads the designer in reaching their goals	-	-	-	-	no conflicting information
12	Identity, reference and perspective in personal communication models	Consider an application in nursery homes to help people having memory problems with daily activities (e.g., taking medicine at a certain time)	-	-	-	-	no conflicting information
13	Monitoring and constraining adaptive systems	N/A - The work focuses on fundamental research into integrating interpretable KRR with learning in the context of adaptive systems	-	-	-	-	no using data
14	Interactive Machine Reasoning for Responsible HI	Behaviour support agents in the healthy lifestyle domain	Yes	Yes	Yes	Explanation through dialogue	

Result survey

Project	Name of the project	Use cases/ Scenarios	Conflicting information	Using logic formalism for knowledge representation	Translation of the HI scenarios into AF	Type of problem	Reason for no (unknown) mapping to AF
16	Context Modeling for Dialog Systems	Collaborative grounded dialog between humans and machines	-	-	-	-	no conflicting information
19	Explaining data-driven decisions with legal, ethical or social impact to end users	Develop explainable AI techniques (possibly on the basis of argumentation theory)	Yes	-	Yes	Decision making	
20	Enabling Co-regulation for long-term engaging semi-structured conversations	A patient interview robot (A robot that gathers PROM data from patients through spoken interaction)	Yes	-	Yes	Explanation through dialogue	
22	Including ethical and legal considerations in the HI design process	Design, development, and deployment of AI systems in public policy settings	Yes	-	Unknown	Justification	no having formal representations for knowledge (data)
23	Causal RL and Reasoning	Simulated and real games/robot environments such as Atari or locomotion tasks	Yes	-	Yes	Decision making	Using propositional logic to express prior knowledge
24	Explainable De-biasing in Learning from Interactions	Ranking systems, Recommender Systems where a ranked list of items is shown to the user	Yes	-	Unknown	Decision making	Data is queries (in text-form), documents (in text-form) and some kind of relevance signal, either click logs or relevant
26	Knowledge Representation Formalisms for Hybrid Intelligence	Model knowledge of human and artificial agents in different HI scenarios	Yes	Yes	Yes	Justification	
27	Functional Priors for Fast Bayesian Adaptation under Human Guidance	Consider robust Bayesian machine learning workflows in real world scenarios which typically involve humans	-	-	-	-	no conflicting information
30	Common sense reasoning for embodied agents	Use commonly available knowledge graphs to enhance the everyday task execution of robots	Yes	Yes (Knowledge Graph- OWL)	Yes	Decision making	
32	The role of trust in people in cooperative human-robot interaction	Social dialogue between a human and an embodied agent in multimodal environments	Yes	Yes (Knowledge Graph)	Yes	Justification	

Summary Result

- Clarify 14 out of 21 HI projects having scenarios with inconsistent information, and the reason of inconsistencies.
- For 10 out of the 14 projects, we analyzed how to apply Argumentation to model the specific representation knowledge.
- Categorize 14 projects based on the type of problems that Argumentation can address in their use-cases.
- We did not analyse the remaining projects since conflicting information is not available in their scenarios or the projects currently do not using data or knowledge.

Limitations

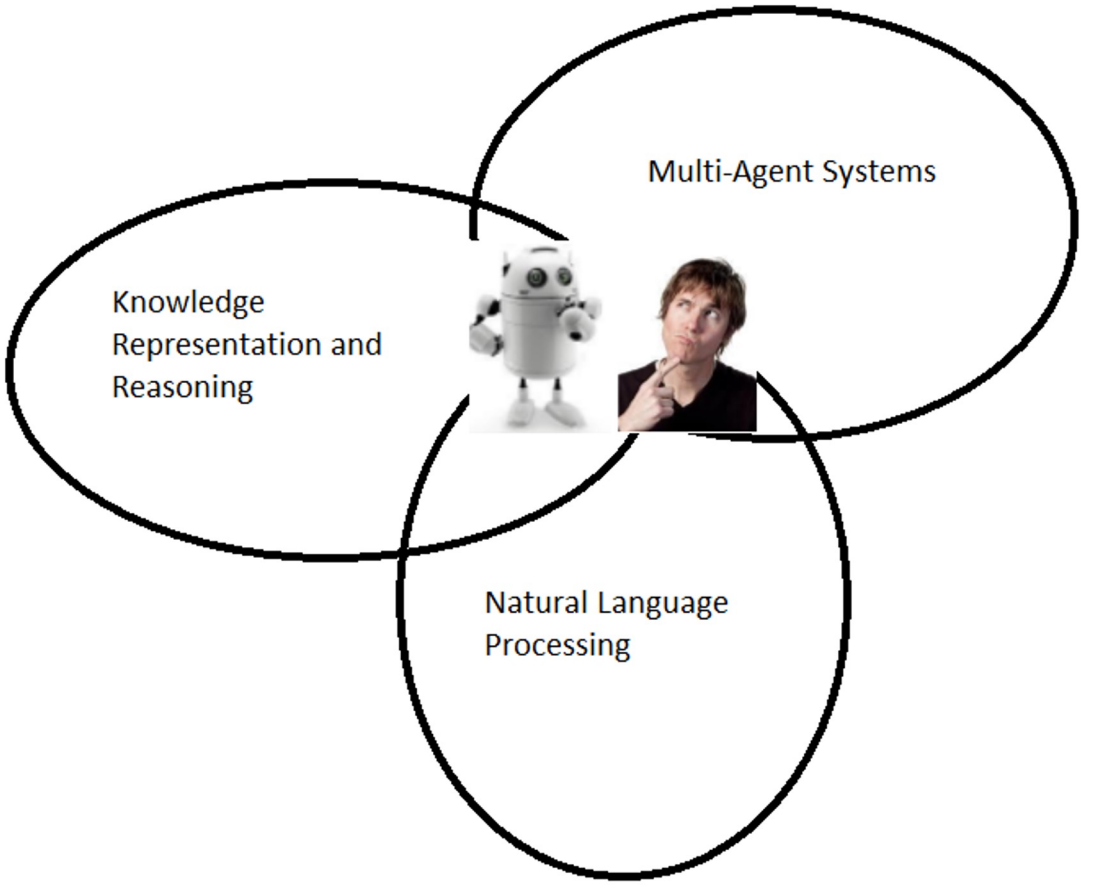
- We chose to focus on projects of the HI Centre.
- Data/ knowledge from these dialogues expressed in natural language or synthetic simple numeric data or documents => still *challenge.*
- Various projects having massive data in real-world application => the use of argumentation based explanation is still a challenge.

Conclusions

- We outline potential HI scenarios in different application domain.
- We demonstrate *the capabilities of Argumentation in representation and reasoning inconsistent KB* of HI scenarios.
- We show *how Argumentation can enable Explainability* in the HI systems, for solving various types of problems in *decision-making*, *justification* of an opinion, and *dialogues*.

Future Work

- Materialize human–machine dialogue from human text dialogues in the HI scenarios.
- Causality could be achieved by reasoning over each step that led to a decision and explain why alternatives were left out => combine Argumentation and causality for this purpose.



Thank you for your attention!

Feel free to ask questions