Understanding Social Interactions Using Incremental Abductive Inference

Benjamin Meadows
Pat Langley
Miranda Emery

Department of Computer Science
The University of Auckland
Private Bag 92019
Auckland 1142 NZ

Thanks to Paul Bello, Will Bridewell, and Alfredo Gabaldon for discussions that aided this research, which was partly funded by ONR Grant No. N00014-10-1-0487.

Social Understanding

Humans understand many social interactions with little effort; we easily generate hypotheses about:

- Other agents' beliefs and goals;
- Their beliefs and goals about others' mental states;
- Their awareness / ignorance of the true situation; and
- Even their intentions to deceive third parties.

Such abilities are a distinctive feature of human intelligence and thus a natural target for cognitive systems research.

Some Related Paradigms

The task of social understanding is related to a number of other research paradigms, including:

- Activity recognition (e.g., Aggarwal & Ryoo, 2011)
- Plan recognition (e.g., Goldman, Geib, & Miller, 1999)
- Behavior explanation (e.g., Malle, 1999)
- Collaborative planning (Rao, Georgeoff, & Sonenberg, 1992)
- Story understanding (e.g., Wilensky, 1978; Mueller, 2002)

Each differs in important ways, but we will incorporate a number of their ideas into our work.

Social Understanding in Fables

Aesop-like fables present an interesting variation on the task of social understanding:

> *The Lion and the Sheep.* A lion is too old to hunt animals for prey. The lion announces he is sick. The sheep, believing he is harmless, follows social convention and visits the lion's caves to pay respects to the ill. The lion kills and devours him.

Such stories are usually brief, focus on goal-directed behavior, and center on high-level social interaction / communication.

Explanations of these fables revolve around agents' beliefs and goals about other agents' beliefs and goals.

Theoretical Tenets

We propose four theoretical claims about the operation of social understanding; we maintain that it:

- Involves inference about the participating agents' *mental states* (beliefs / goals about activities and environment);
- Involves the *abductive* generation of *explanations* through the introduction of default assumptions;
- Operates in an *incremental* fashion to process observations that arrive sequentially; and
- Proceeds in a *data-driven* manner because understanding arises from observations about agents' activities.

These assumptions place constraints on our computational account of this important process.

The UMBRA System

This suggests that we use UMBRA, an abductive inference system developed previously that:

- Accepts observations and adds them to working memory
- Incrementally extends an explanation by:
  - Finding rules with antecedents that unify with memory elements
  - Tentatively completing each rule instance's missing antecedents
  - Selecting the rule instance R with best evaluation score
  - Adding R's inferred elements to memory as default assumptions
- Continues until no further observations arrive

This data-driven strategy aims to produce a *coherent explanation* in terms of available knowledge.

UMBRA is similar in spirit to AbRA (Bridewell & Langley, 2011).

Previous Results with UMBRA

In previous work, we have run UMBRA on plan understanding tasks that involve single agents.

- We provided the system with hierarchical task networks and observations of people's actions.
- On the Monroe corpus, a commonly used testbed, UMBRA's precision and recall were similar to those for other systems.

These results encouraged us to extend the software to handle tasks that require social understanding.

Extension 1: Timing and Constraints

To support social understanding, we have extended UMBRA's representation to incorporate:

- Start and end times for each belief and goal:
  - belief(lion, prey(sheep), 6:00, s1)
  - goal(lion, healthy(lion), 12:00, 12:30)
- Constraints on timing and equality:
  - constraint(fox, between(s2, s4, 8:00, s5), 5:35, 6:00)
  - constraint(lion, nequal(sheep, s3), 5:00, s2)

Constraints are first-class structures in both working and long-term memory, at the same level as beliefs and goals.

Extension 2: Embedded Structures

The extended UMBRA also represents agents' mental states, some of which involve embedded structures:

- belief(fox, has(crow, grapes, 09:30, s1), 09:31, s2)
- goal(crow, acquire_edible_food(crow, s3, s4))
- belief(snake,
  belief(lion, at_location(lion, river, 09:00, s5), 09:02, s6), 09:02, s7)
- belief(snake,
  goal(fox, trade(crow, fox, grapes, grain, 09:40, s8), 09:30, s9), 09:30, s10)
- goal(lion, belief(sheep, sick(lion, 09:00, 24:00), 09:45, s12), 09:00, s13)

Embedded structures appear in working memory and social rules, but *not typically in domain-level knowledge*.

Extension 3: Inference Processes

These representational changes also required some extensions to UMBRA's inference mechanisms:

- Introduction of start times for inferences based on current cycle;
- Adding timing and equality constraints to working memory as inferences when rules fire;
- Using constraints to eliminate rule applications that would create inconsistent default assumptions; and
- Reasoning over embedded beliefs and goals using rules with non-embedded structures.

We did not alter the basic abduction mechanism to operate over social knowledge, despite the latter's abstract character.

Empirical Claims About UMBRA

We make three claims about our extensions to UMBRA to let it support social understanding:

- The system generates appropriate explanations and inferences for fables from partial information;
- The ability to apply knowledge at different levels of embedding is critical to this functionality; and
- High-level knowledge about social interactions is also essential to generating reasonable explanations.

We have designed and carried out experiments designed to test these claims.

A Testbed for Social Understanding

We devised eight fables that require social understanding at different levels of complexity:

- *Nested understanding:* The observing agent interprets another agent's mental states and/or plan based on observed behavior.
- *Deeply nested understanding:* The observing agent infers another agent's inferences about a third agent's mental states.
- *Inferring mistakes:* The observing agent infers that another agent has mistaken beliefs, the reasons for them, and the true account.
- *Reasoning about opportunism:* The observing agent understands how another agent has capitalized upon another's false beliefs.
- *Reasoning about deception:* The observing agent infers that another agent engenders false beliefs in a third agent to achieve some goal.

We have used these scenarios to test UMBRA's ability to construct social explanations.

A Testbed for Social Understanding

We also created relevant knowledge for these eight scenarios that includes:

- About 60 distinct skills / operators
  - alternative decompositions
  - many with overlapping conditions
  - only ten percent used in any 'correct' fable explanation
  - about 500 domain-level conditions, excluding constraints
- About 100 distinct domain-level predicates

Domain knowledge typically describes physical situations and activities at a single level of embedding.

Social knowledge uses multiple levels of embedding to support reasoning about others' mental states.

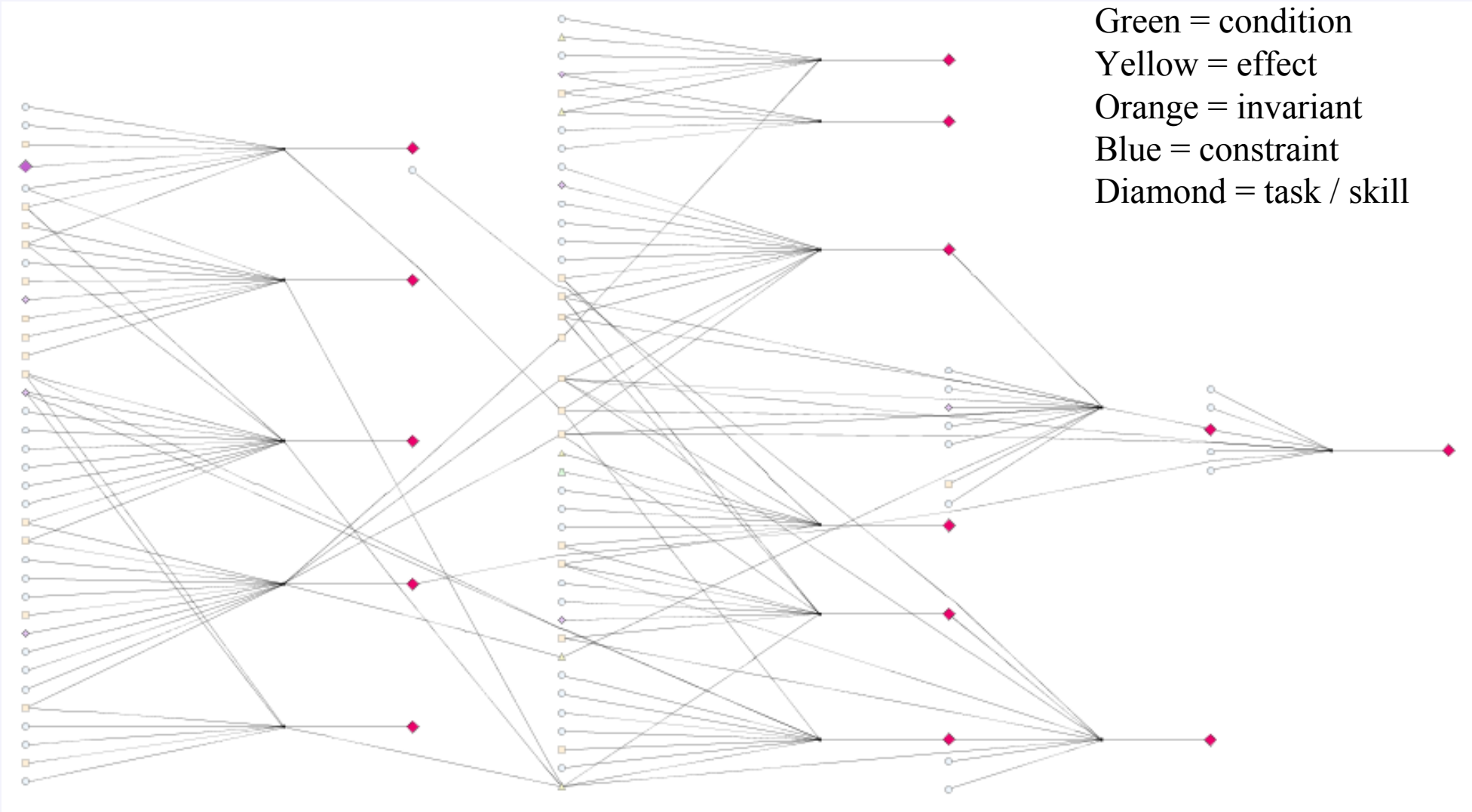
Social Predicates

UMBRA's social knowledge includes 13 some predicates that describe personal interactions:

- announce_genuine, announce_wrong, announce_false
- interpret_as_real, interpret_as_real_agent, interpret_as_real_attributed
- interpret_as_image, interpret_as_image_attributed
- become_jealous
- judge_not_a_threat
- pretend_attribute
- suggest_trade_good_faith, suggest_trade_bad_faith

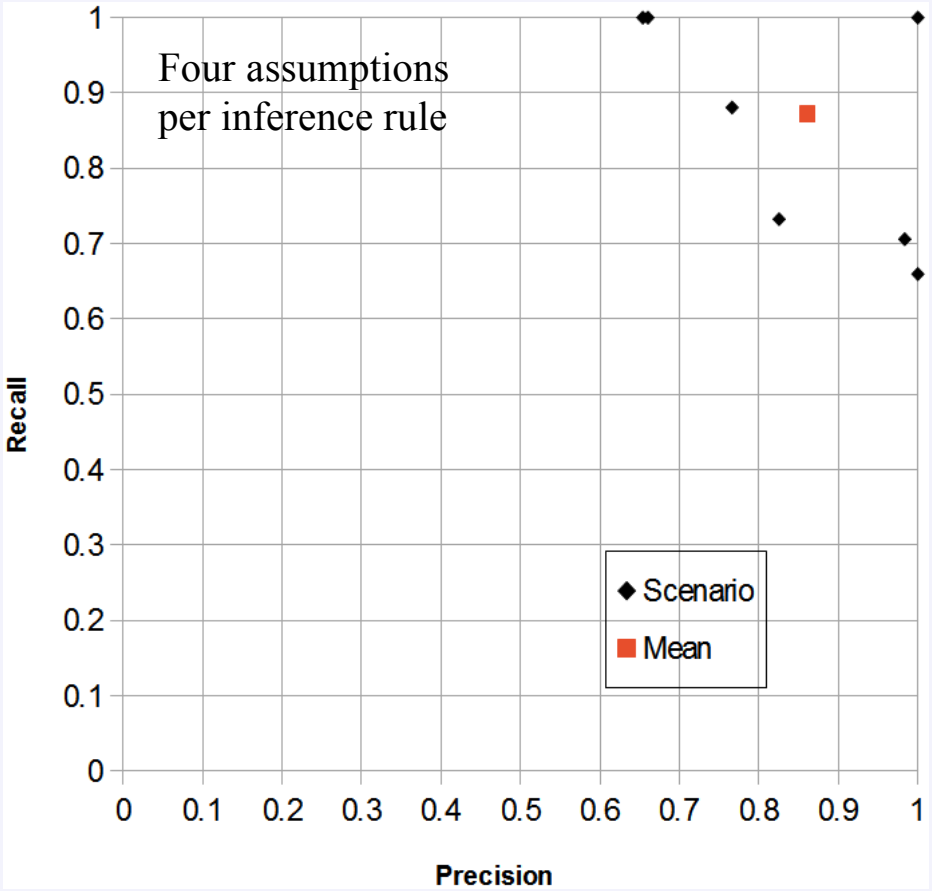
Each of these refers to activities that alter the mental states of participating agents.

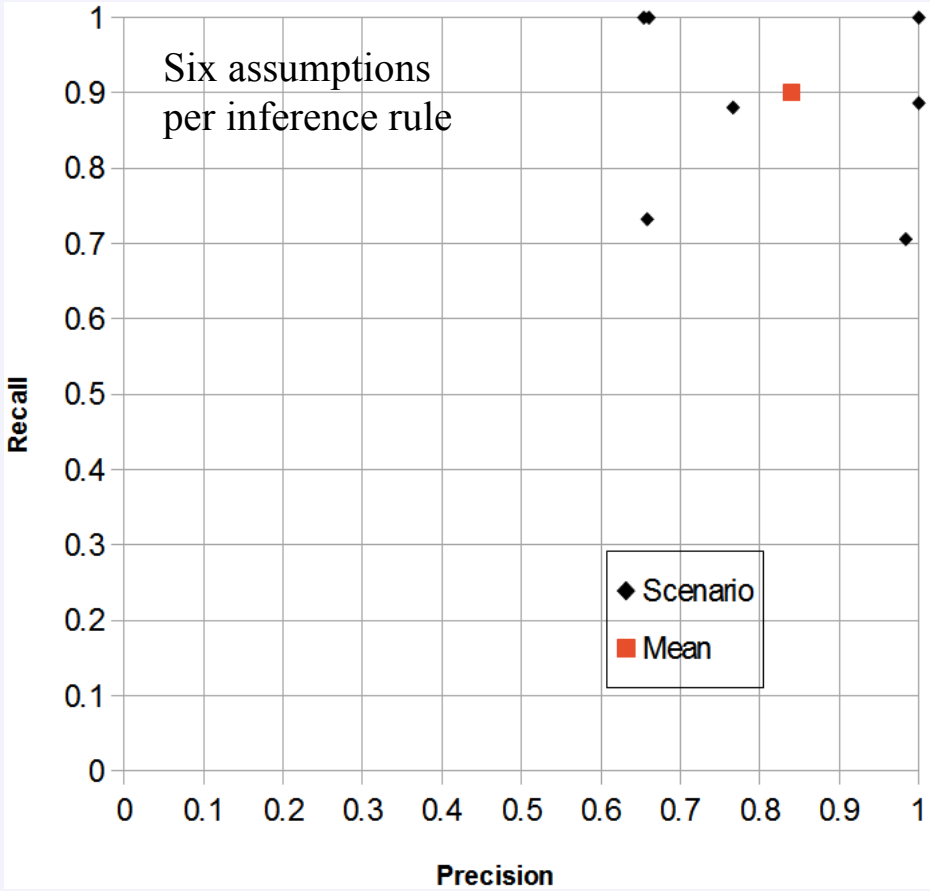
Structure of a Fable Explanation



Basic Results on Fable Understanding

The extended UMBRA draws correct inferences with high precision and recall given less than 40 percent of the target explanations.

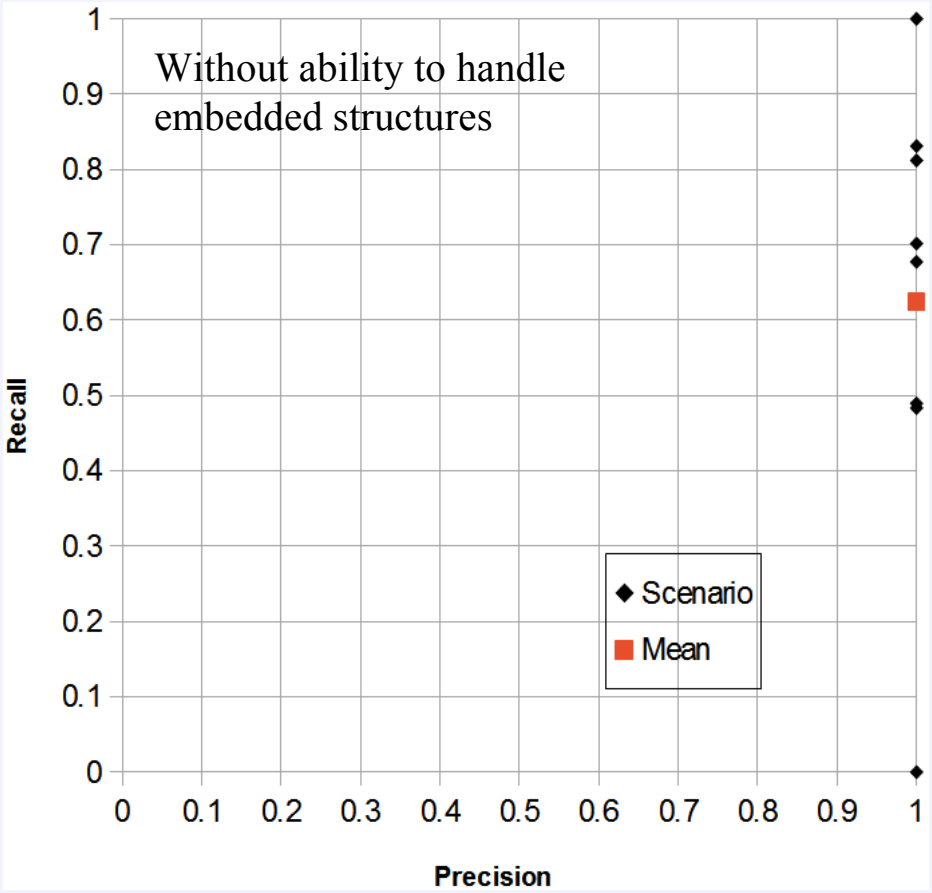


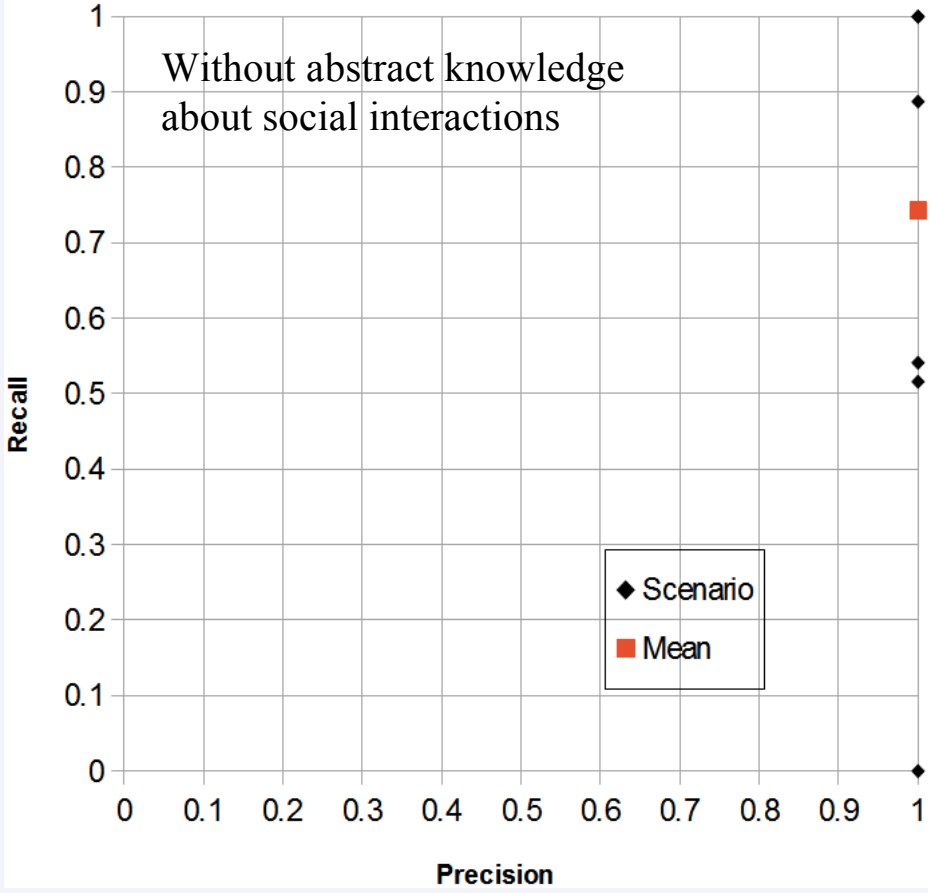


Changes to the system's parameters have little effect on these scores.

Results from Lesion Studies

We also ran UMBRA with its ability to handle embedded structures and its social knowledge removed.





Even when given all terminal literals, recall was still reduced greatly.

Related Research

Our approach relies centrally on three assumptions that have been explored in previous research:

- *Social cognition relies on representing and reasoning about models of other agents' mental states.*
  - Fahlman (2011), Bello (2012), Bridewell and Isaac (2011)
- *Plan understanding involves a process of incremental abduction that constructs an explanation of observed inputs.*
  - Ng and Mooney (1990), Bridewell and Langley (2011)
- *Social understanding depends on general knowledge about social interactions and their effects on mental states.*
  - Wilensky (1978), Winston (2012)

Our work incorporates ideas from these earlier traditions, but it combines them in novel ways to support social understanding.

Concluding Remarks

We have extended UMBRA, which constructs explanations with an incremental form of abductive inference, to:

- Represent other agents' mental states as embedded structures
- Encode information about timing and constraints
- Store domain-independent knowledge about social interactions
- Reason over this content to understand Aesop-like fables

Experiments suggest that our approach can create plausible and coherent social explanations from partial information.

In future work, we plan to extend UMBRA to revise assumptions when needed and to learn new social structures.

End of Presentation