On Resolving Ambiguous Anaphoric Expressions in Imperative Discourse

Vasanth Sarathy & Matthias Scheutz

Human-Robot Interaction Laboratory
Tufts University
Medford, MA, USA

33rd AAAI Conference on Artificial Intelligence (2019)

Setting the Stage

antecedent

coreference

*"Pick up <u>the paper</u>. Give **<u>it</u>** to me."*

referent

anaphor

- Anaphors (e.g., pronouns) common in imperative discourse
- Goal: To resolve the real-world referent for "it" by selecting appropriate antecedent referring expression

Real-world is Complicated

Task-Oriented Dialogue Excerpt (Grosz 1978):

(1) E: First you have to remove the flywheel.
(2) A: How do I remove the flywheel?
(3) E: First, loosen the two allen head setscrews holding it to the shaft, then pull it off.
(4) A: OK.
(5) I can only find one screw. Where's the other one?
(6) E: On the hub of the flywheel.
(7) A: That's the one I found. Where's the other one?
(8) E: About ninety degrees around the hub from the first one.
(9) A: I don't understand. I can only find one. Oh wait, yes I think I was on the wrong wheel.

- Multiple choices
- Recency not enough
- Shifting salience
- Syntax unhelpful
- Semantically empty

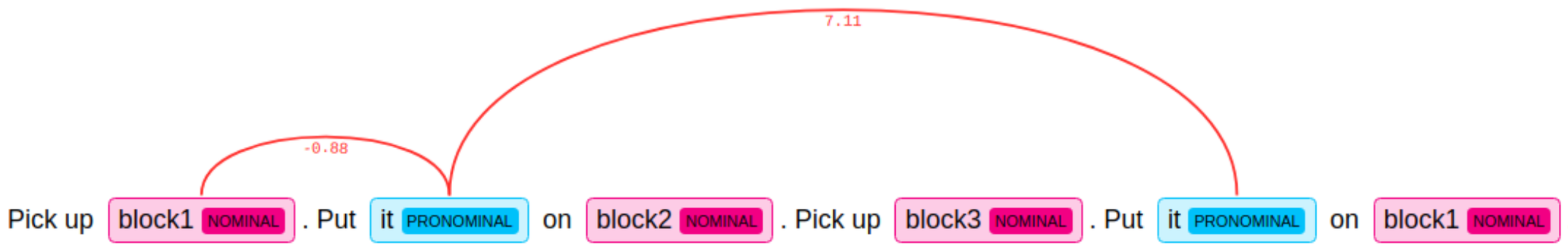
[1] Grosz, Barbara J. 1978. Focusing in Dialog. Proceedings of the 1978 Workshop on Theoretical Issues in Natural Processing, July 25-27,1978

Let's Simplify:
Enter Blocks World!

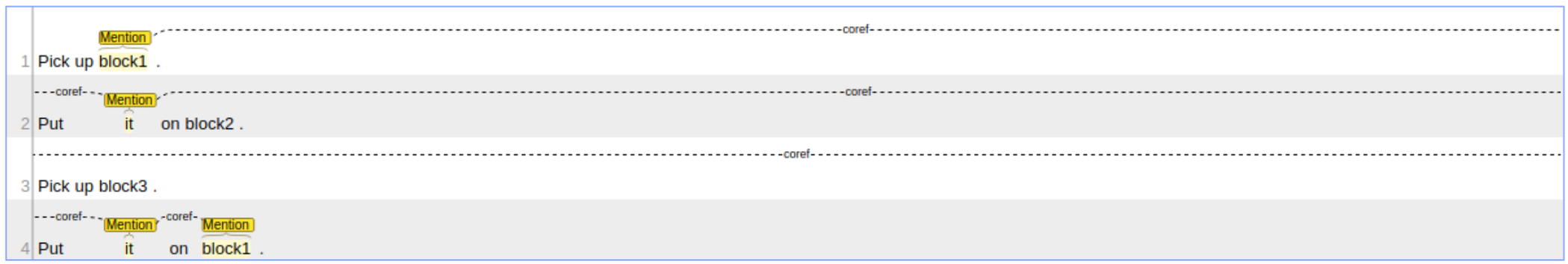
D1: "*Pick up block1. Put **it** on block2. Pick up block3. Put **it** on block1.*"

D2: "*Pick up block1. Slide block3 under **it**. Put **it** down.*"

Performance of State-of-the-Art Data-Driven Systems



Clark, K., and Manning, C. D., "Deep reinforcement learning for mention-ranking coreference models," EMNLP 2016.



Stanford CoreNLP 2018

- At best: distinction between choices statistically insignificant
- At worst: incorrect resolution

Key Contributions

- Characterize the general class of situated anaphora resolution problems
- Construct a proof-of-concept resolver using Answer Set Programming and Dempster-Shafer Theory
- Articulate domain-independent properties of the reasoners

Our Approach

- Imperative discourse consists of speech acts that require the listener to perform (or at least simulate) a sequence of actions
- Anaphora resolution is the task of associating each action with parameters in a way that “makes sense”
- “Makes sense” from the perspective of mutual knowledge : information that the speaker and listener both have (Clark and Marshall 2002)
  - Agent’s own capabilities
  - Expectations of its interactants
  - Normative expectations of the community

Clark, H. H., and Marshall, C. R. 2002. Definite reference and mutual knowledge. Psycholinguistics: critical concepts in psychology 414.

Reasoning Modes

- Goal: Select ***object candidate*** *that when paired with the specified* ***action*** *makes the <u>most</u> sense:*

- Three “starter” reasoning modes answering questions of:

1. Plausibility: ***Can*** the agent perform the <u>action</u> on (or with) an <u>object</u> candidate?

2. Speaker Intent: Is the speaker ***intending*** for the agent to perform the desired <u>action</u> on (or with) the <u>object</u> candidate?

3. Normativity: ***Should*** the agent perform the <u>action</u> on (or with) the <u>object</u> candidate?

A Quick Note on Reasoning

- Ours is a form of commonsense reasoning
- Different from traditional AI notions of commonsense reasoning used in pronoun disambiguation problems
  - Winograd Schemas: Commonsense reasoning about timeless facts (e.g., whether trees are larger than toothbrushes)
  - **Situated Anaphora Problems**: Requires very specific and fluid situational information as well as implicit normative knowledge and social reasoning.

Microtheories

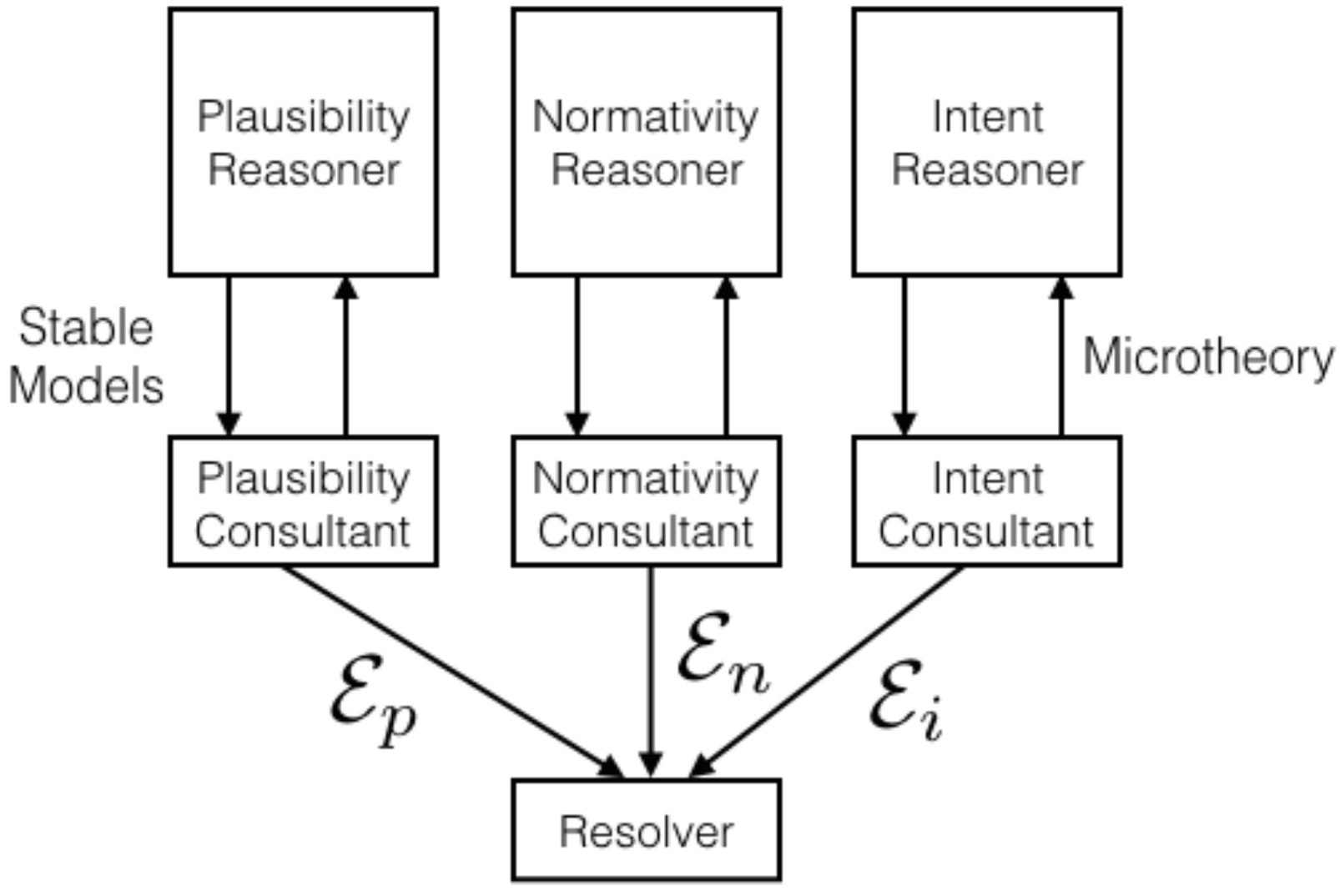
- Microtheories implemented as answer set programs contain relevant knowledge and reasoning capabilities.
    - Initially, a microtheory is incomplete and only contains rules about a domain that it handles (e.g., social norms).
    - Microtheories can then be filled in real-time when situational facts become available
    - Once the microtheory is filled, a reasoner can iterate over different object candidates
- Output from reasoners (uncertain evidence for various object candidates) are combined together

Implementation Choices

- **Why Answer Set Programming:** non-monotonic reasoning, choice rules, negation-as-failure and classical negation, cardinality constraints, incremental logic programming, and declarative specification

- **Why Dempster-Shafer Theory:** extends Bayesian theory, useful for set-valued random variables, no assumptions of priors needed, fuse evidence from heterogeneous sources

Consultant Architecture



Resolving with Uncertainty

- Each ambiguous pronoun has a set of object candidates (e.g., O = {block1, block2, block3})
- Reasoners provide DS-theoretic masses over these objects.
    - In DS-theory, unlike Bayesian theory, the masses of each object in O need not sum to 1. The sum of the masses of the power set 2^O must be 1.
- For each reasoner uncertainty (i.e., masses) comes from the proportion of stable models that reference the object candidate against all those that reference the action verb

Domain-Independent Aspects

- All micro theories have a common structure (Generate-Define-Test)
- For imperatives, the crucial relationship is between the *action verb* and the *object (or pronoun).*
- Each reasoner is defined by the specific relation between the action verb, *a*, and object variable, *O*
  - Plausibility reasoner: $occ(a(O),t)$
  - Normativity reasoner: $has(a,\ \mathrm{permissible},\ O)$
  - Speaker Intent reasoner: $has(a,\ \mathrm{speaker\_intends},\ O)$
- Generally, facts and definitions in the micro theories have general forms specified by $is(X,Y)$ and $has(X,Z,Y)$ syntax.

Proof-of-Concept

- Microtheories generalized over the following examples:

1. *“Pick up the knife. Cut the tomato. Put it down.”*

2. [washing dishes/cooking]
   “Pick up the knife. Cut the tomato. Pass it to me”

3. [bowl contains food]
   “Pick up the knife. Cut the tomato. Put it in the bowl”

See paper for ASP code and implementation details

Future Work

- Integration onto the NL pipeline of a robotic architecture, e.g., with Williams (2016)
    - Abstract object representation in an ASP program allows for multimodal information integration.
    - Formalizing the syntax and semantics of a high-level microtheory language
    - Exploring cases when additional reasoning modes are needed
    - Integration with data-driven systems
        - Learning Microtheories: Using data-driven approaches to learn the domain-independent rules in the micro theories

Williams, T., and Scheutz, M. 2016. A framework for resolving open-world referential expressions in distributed heterogeneous knowledge bases. In Proceedings of the 30th AAAI Conference on Artificial Intelligence.

Takeaways

- Anaphora (and reference) resolution can require reasoning about situational and embodied knowledge

- We consider the case of disambiguating the pronoun “it” in imperative discourse

- Resolving “it” requires reasoning about the plausibility, normatively and speaker intent associated with the action verb and objects in the discourse context

- We propose a consultant framework and proof-of-concept system for reasoning under uncertainty about object candidates

- We have only scratched the surface and argue for potentially fruitful research direction with practically and theoretically significant implications

Acknowledgments

- ONR MURI for funding this work
- Chitta Baral for reading and providing feedback on early drafts
- My colleagues at the HRI lab at Tufts for letting me talk about this endlessly - yes, it’s just that exciting!

Thank you! Questions?

Mission transcript for NASA Gemini V (1965)

00 01 41 10	CC	Gemini-5, this is Houston here. We still haven't received the Command Pilot blood pressure.
00 01 41 15	P	He was having a little trouble getting it in. He's got it in now and he's pumping it up.
00 01 41 21	CC	Okay, very good.