Grounding Language in an Intelligent Agent

James F. Allen
University of Rochester
and
Institute for Human and Machine Cognition (IHMC)

UpHill Battles Workshop, Nov. 2016

a long time ago, in a galaxy far, far away …

Natural Language Research was a central part of Artificial Intelligence

Language understanding was considered to be fully integrated with our knowledge and reasoning processes

The same reasoning processes that drive our perception, planning, learning and acting

The goal was to ground language in an intelligent agent

From Deep to Shallow

- focus turned to shallow language processing, driven solely by structural properties of language
  - remarkable progress was made of a wide range of applications
- deep understanding was viewed as too difficult
  - if you can’t build a corpus of the behavior, you can’t work on the problem

If you can't build a corpus you can't work on the problem…

- reduced the field to mostly studying problems that can be represented as simple sets of tags!
  - possibly structured (e.g., syntax)
- discounted traditional theory-driven engineering-based approaches to building systems
  - at least in academics, not in industry

Promising News

- Learning semantic parsers has become a hot area in the last decade
  - given language/representation or language/scene pairs, we can learn the semantic interpretation rules

BUT…

rules learned in one application/task/domain do not help in a new domain!!!!

Such work grounds language solely to a specific task and domain

Upper Bound of such techniques?

Maybe “dog-level” language understanding

- The Border Collie “Chaser” learned
  - well over a 1000 names of objects
  - dozens of different actions
    - including “imitate me”

A Few Requirements for grounding language in a reasoning agent

- a shared ontology: word and sentence meanings must map to concepts that are the stuff of reasoning
- a generic (domain general) capability to parse language to such a representation
- the integration of language usage with an agent's planning, intention recognition, reasoning and acting
  - language is just another thing an agent does

Why human-level understanding is difficult

A perfectly normal conversation at a grocery store ...

Customer: *black beans?*

clerk: *aisle 3.*

WITH A SUFFICIENT CORPUS, CURRENT TECHNIQUES CAN LEARN THAT "BLACK BEANS" PARSES TO DATABASE QUERY [LOCATION "BLACK BEANS" ?X]

BUT CHANGING THE DOMAIN TO A HOME ENVIRONMENT...

When arriving home ...

Spouse: *black beans?*

You: *Oh, sorry, I forget to get them.*

When cooking ...

Spouse: *black beans?*

You: *in the cupboard.*

When exploring nutrition options ...

Spouse: *black beans?*

You: *227 calories in a cup*

When cooking (adding black beans to a pot) ...

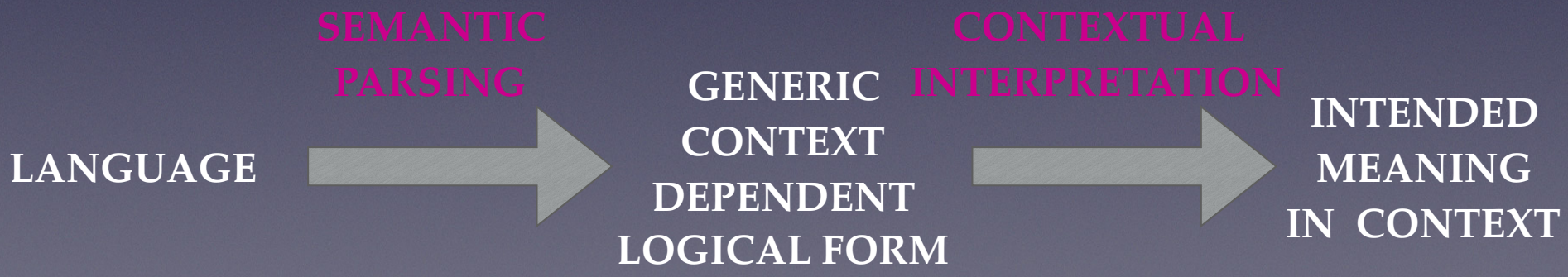
Spouse: *black beans?*

You: *don't you like them.*

DEEP UNDERSTANDING REQUIRES INTENTION RECOGNITION IN CONTEXT

Deep(er) understanding and intention recognition

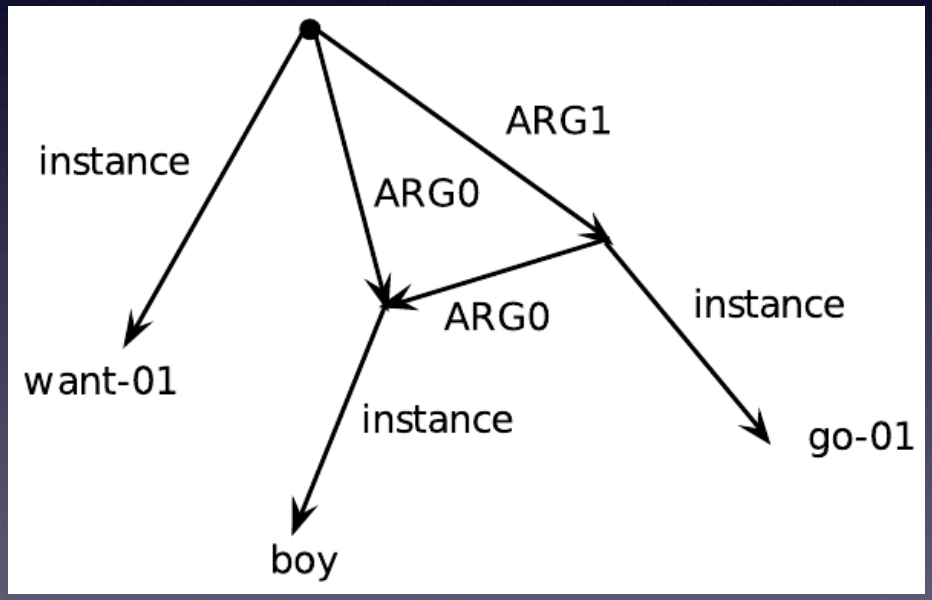
- the problem
  - language loses its one-to-one correspondence to intended meaning
  - intended meaning is intricately tied to the goals and intentions of the speaker/writer
- how to make progress on such a complex issue?



A PRACTICAL MIDDLE GROUND?

Promising news:
GENERIC semantic representation
AMR (Abstract meaning representation)

"THE BOY WANTS TO GO"



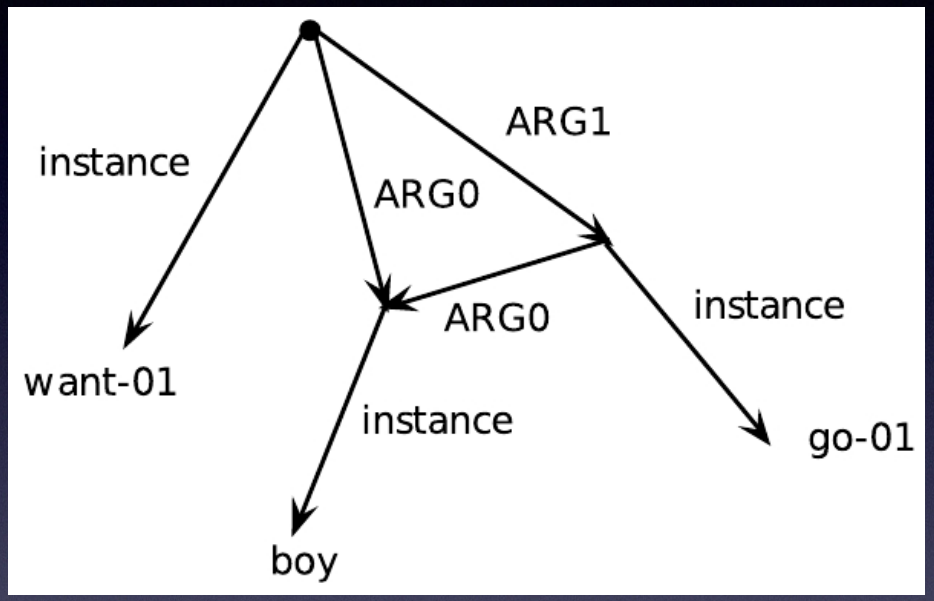
AMR

WHERE IT FALLS SHORT!

- senses are relative only to word, not an ontology
- most non-verbs are **not** sense tagged
- no quantifiers (can't distinguish "a boy", "the boy", "the boys", …)

Extending AMR

"THE BOY WANTS TO GO"



TRIPS LF GRAPH

ALLEN ET AL, 2008

TRIPS LF is a generic semantic representation

http://trips.ihmc.us/parser/LF%20Documentation.pdf

that is produced by a broad-coverage domain-general semantic parser

http://trips.ihmc.us/parser/cgi/step#

Grounding in an Agent:

Intention Recognition, Planning, Acting, …

The Black beans problem

- we have reasonable theories and algorithms (logical and probabilistic) of intention recognition and planning
- we have good core theories of how communication relates to agent's beliefs, desires and intentions
- but hindered by lack of commonsense knowledge
  - about the world
  - about everyday activities
  - about human problem solving behavior
- AND especially hindered by lack of effort on the problem!
  - no corpus => can't work on it!

Dialogue as Collaborative Problem Solving

Current Active Domains

Biocuration (BOB)

system and biologist develop causal models of cancer pathways)

Embodied Blocks World (CABOT

system and user collaborate setting goals, teaching each other, and building structures

World Modelers

System works with scientist to build world scale simulations (including climate, agriculture, economics, social unrest)

Medical coaching teens (ASMA)

System advises teens about managing their Asthma using text messaging

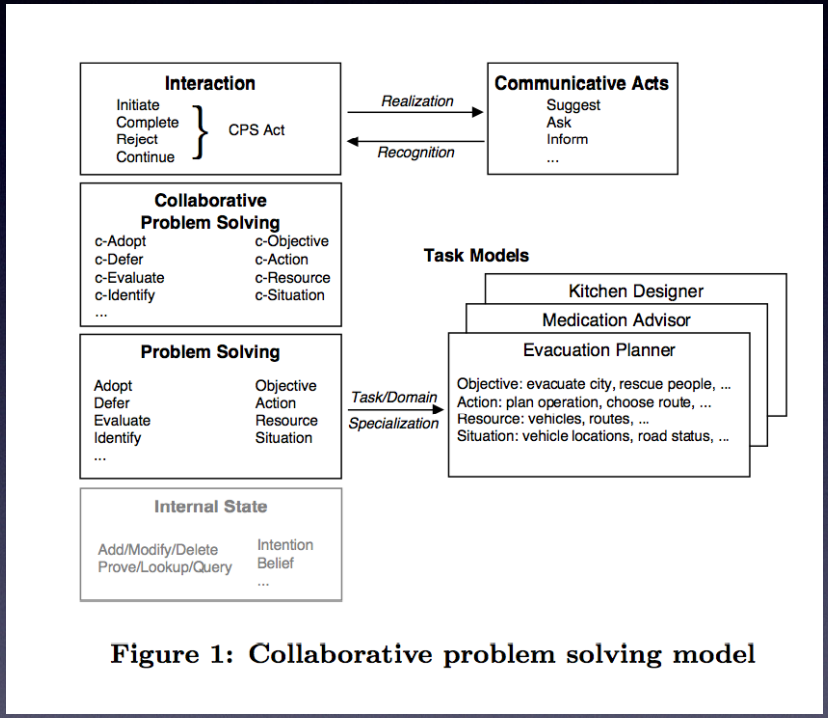


Figure 1: Collaborative problem solving model

Allen, J., et al. (2002). A Problem Solving Model for Collaborative Agents. AAMAS

Empirical Techniques and Evaluation

If Physicists only worked on problems they could evaluate we wouldn't have the theory of relativity (or most of the rest of theoretical physics, for that matter)

and its not just a matter of waiting until the evaluation techniques become available, in fact research problems should drive the development of new evaluation technologies

Recap: My two key problems

- Broad-coverage domain-general semantic parsing
- Integration of language into intelligent agent architectures

References

BROAD-COVERAGE SEMANTIC PARSING

Allen, J., et al. (2008). Deep Semantic Analysis of Text. Symposium on Semantics in Systems for Text Processing (STEP), Venice, Italy.

Dzikovska, M., et al. (2008). "Linking Semantic and Knowledge Representation in a Multi-Domain Dialogue System." Logic and Computation 18(3): 405-430.

Allen, J. (2009). Word senses, semantic roles and entailment. 5th Conference on Generative Approaches to the Lexicon. Pisa, Italy.

Allen, J. F. (2014). Learning a Lexicon for Broad-coverage Semantic Parsing. ACL Workshop on Semantic Parsing. Baltimore, MD.

TRIPS

Logical Form Documentation:

Lexicon and Ontology Browser: www.cs.rochester.edu/research/trips/lexicon/browse-ont-lex.html

Online Parsing/Web Services: trips.ihmc.us/parser/cgi/cabot

http://trips.ihmc.us/parser/cgi/bob

http://trips.ihmc.us/parser/cgi/step

LANGUAGE AND INTELLIGENT AGENTS

Allen, J. F. and C. R. Perrault (1980). "Analyzing Intention in Utterances." Artificial Intelligence 15(3).

Ferguson, G. and J. Allen (1998). TRIPS: An Integrated Intelligent Problem-Solving Assistant. National Conference on Artificial Intelligence (AAAI), Madison, WI, MIT Press.

Allen, J., et al. (2002). A Problem Solving Model for Collaborative Agents. 1st Int'l. Joint Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS-02), Bologna, Italy, ACM Press.

Allen, J. F. and G. Ferguson (2002). Human-Machine Collaborative Planning. 3rd Int'l. NASA Workshop on Planning and Scheduling for Space, Houston, TX.

Allen, J. F., et al. (2007). PLOW: A collaborative task learning agent. Named Best Paper, National Conference on Artificial Intelligence (AAAI), Vancouver, BC

Perera, I. and J. F. Allen (2013). SALL-E: A Situated Agent for Language Learning. 27th Conference on Artificial Intelligence (AAAI-13), Seattle, WA.

Rhee, H., et al. (2014). "Mobile phone-based asthma self-management aid for adolescents (mASMAA): a feasibility study." Patient Preference and Adherence 2014:8.

Postscript

- Don’t underestimate dogs!
  - in addition to grounding language in world
  - they do intention recognition
  - they do “zero-shot” learning using reasoning strategies