

Agent Requirements for Effective and Efficient Task-Oriented Dialog

Shiwali Mohan

Palo Alto Research Center

James Kirk, Aaron Mininger, John E. Laird

University of Michigan

Task-oriented dialog

Expert: *First you have to remove flywheel.*

Novice: *How do I remove the flywheel?*

Expert: *First, loosen the two allen setscrews holding it to the shaft, then pull it off.*

Novice: *OK. I can only find one screw. Where is the other one?*

....

Expert: *Use the whelpuller. Do you know how to use it?*

Novice: *No.*

Expert: *Do you know how it looks like?*

Grosz and Sidner (1986)

- Referential
 - objects, relationships, events, collections
- Multi-modal
 - language with gestures, expressions etc.
- Mixed-initiative
- Multi-purpose
 - collaboration, learning, teaching, ...

Language for robots

Tellex et al. 2011



Commands from the corpus

- Go to the first crate on the left and pick it up.
 - Pick up the pallet of boxes in the middle and place them on the trailer to the left.
 - Go forward and drop the pallets to the right of the first set of tires.
 - Pick up the tire pallet off the truck and set it down
-

Matuszek et al. 2012

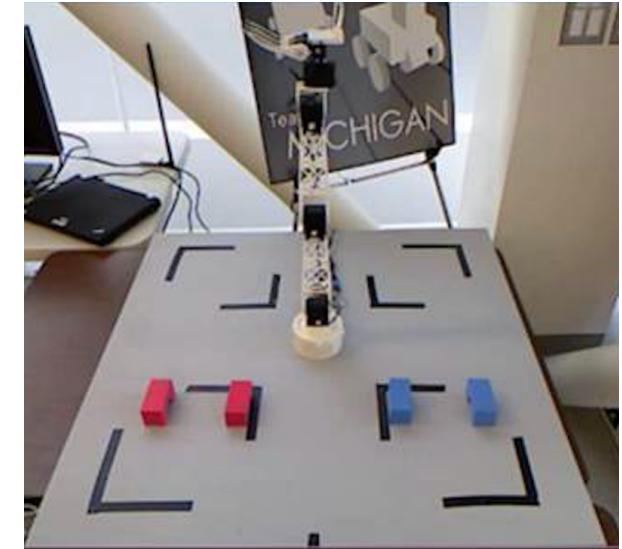


Mericli et al. 2014



Designing for human-robot dialog

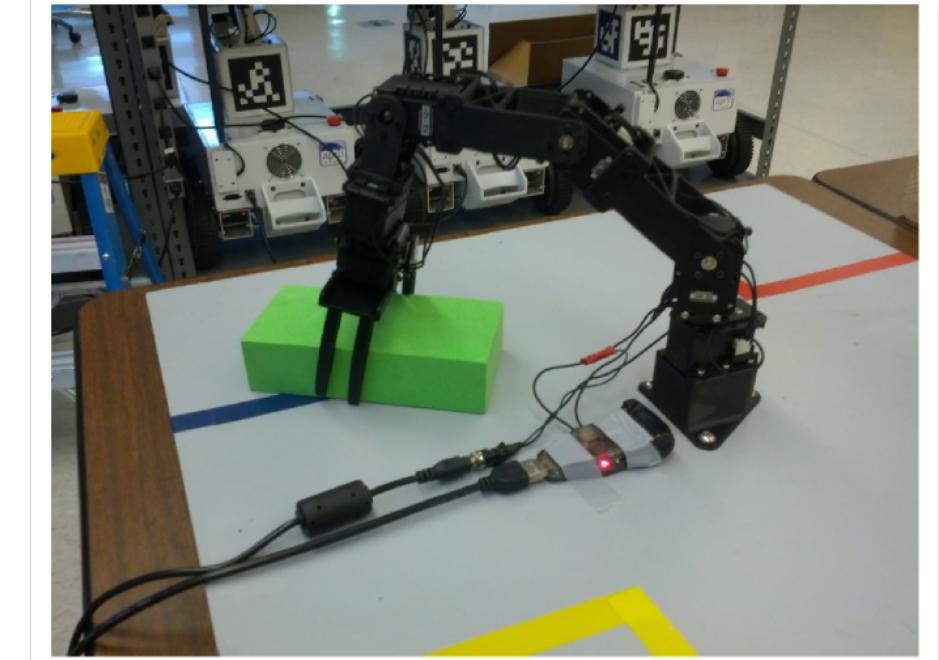
- How do we go from language to dialog?
 - situated, flexible, mixed-initiative, ...
- Rosie @ UoM
 - for interactive task learning (ITL)
 - Soar cognitive architecture (*Laird 2012*)
 - tasks (fetch, deliver, move, store)
 - games (tic-tac-toe, tower-of-hanoi)
- AAAI-FSS 2015: identify a set of 9 system level requirements
 - may be incomplete, inconsistent, ...



Rosie: learning to play tower-of-hanoi

1. Encode a common reference scheme

- Task dialog is referential
 - *that green block*
- Agent must connect language to sensory inputs and control structures
- Rosie learns interactively
Mohan et al. ACS 2012

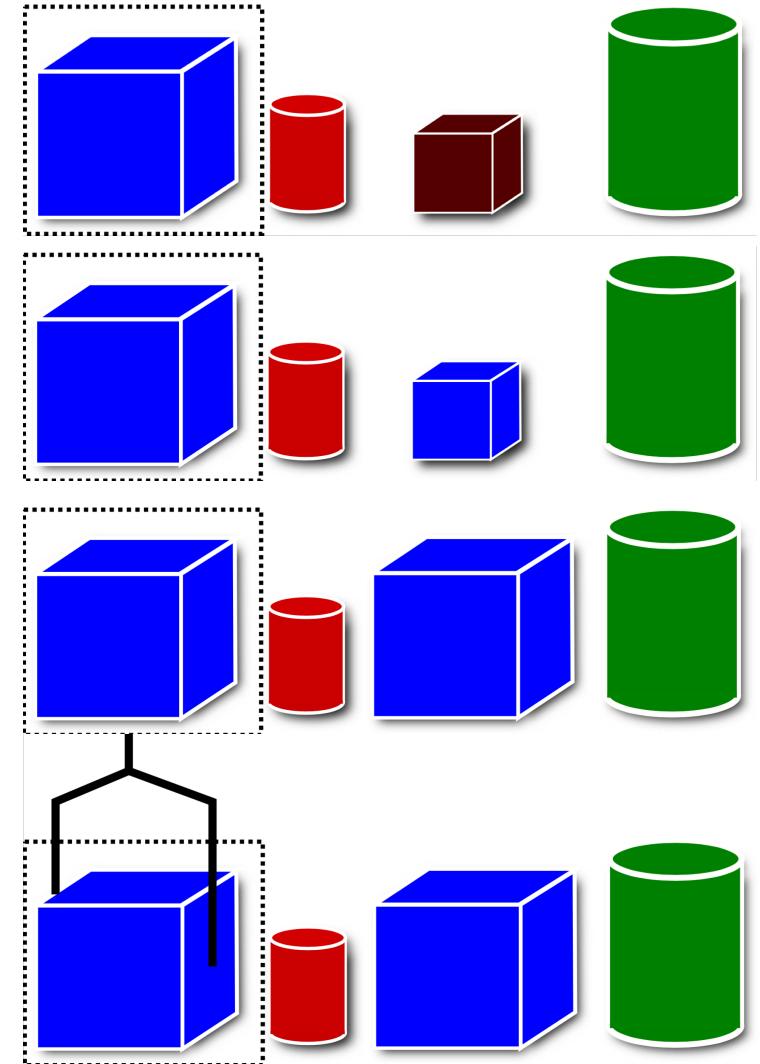


2. Encode information-exchange protocols

- Different sentential structure suggests a different interpretation/actions
 - *Where is the milk?*
 - *Put the book over there.*
 - Agent must connect structure with interpretation and response.
 - Rosie has a simple referential grammar
 - noun-phrase -> objects
 - prepositions -> spatial relationships
 - adjectives -> object state
 - Imperative sentence -> action
- Mohan, Mininger, & Laird ACS 2014*

3. Incorporate non-linguistic context

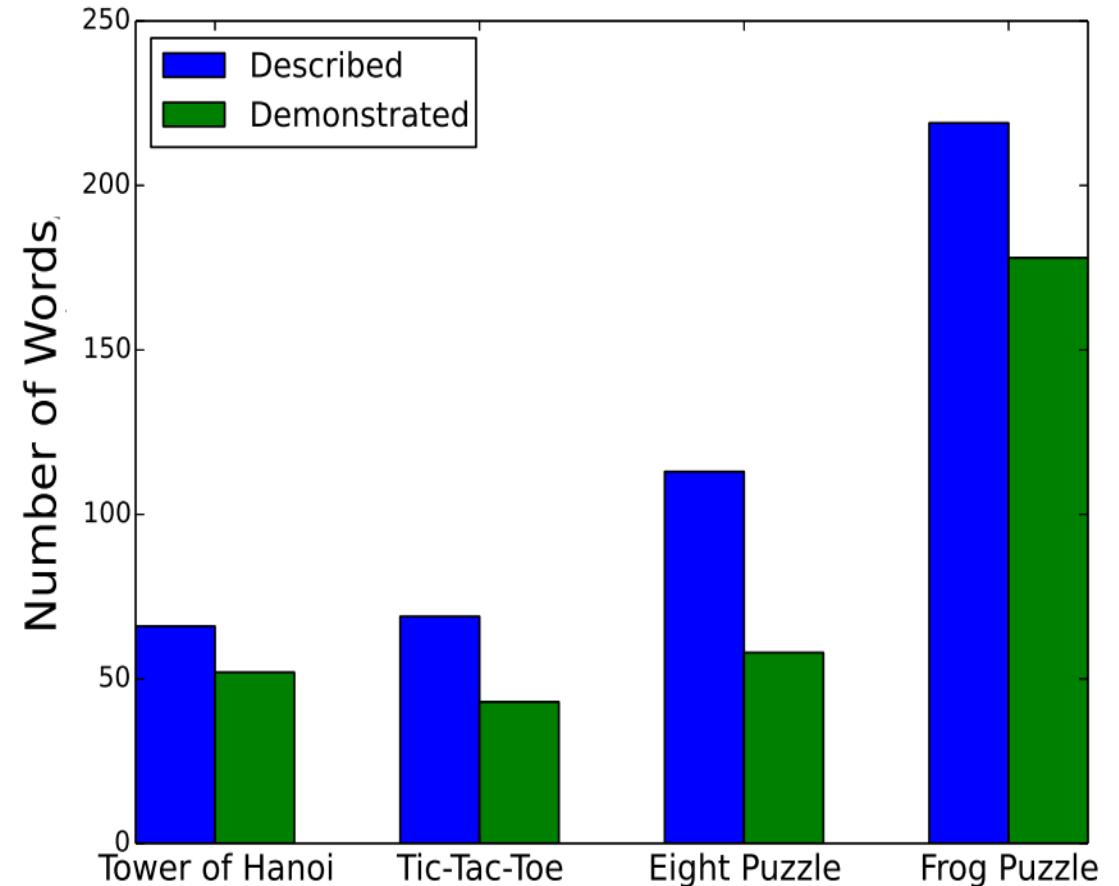
- Language is contextual, ambiguous, incomplete
 - *Move the blue block to the pantry.*
 - *Take out the trash.*
 - *What is red? A color.*
 - Agent must use other sources of information, knowledge
 - Rosie uses a variety of information
 - scene, task structure, dialog history etc.
- Mohan, Mininger, and Laird ACS 2014*



4. Incorporate information from multiple interaction modalities

- Dialog is multi-modal
 - demonstration, gestures may be employed for efficiency
- Agent must integrate information online, real-time
- Rosie can learn from instructions + demonstrations

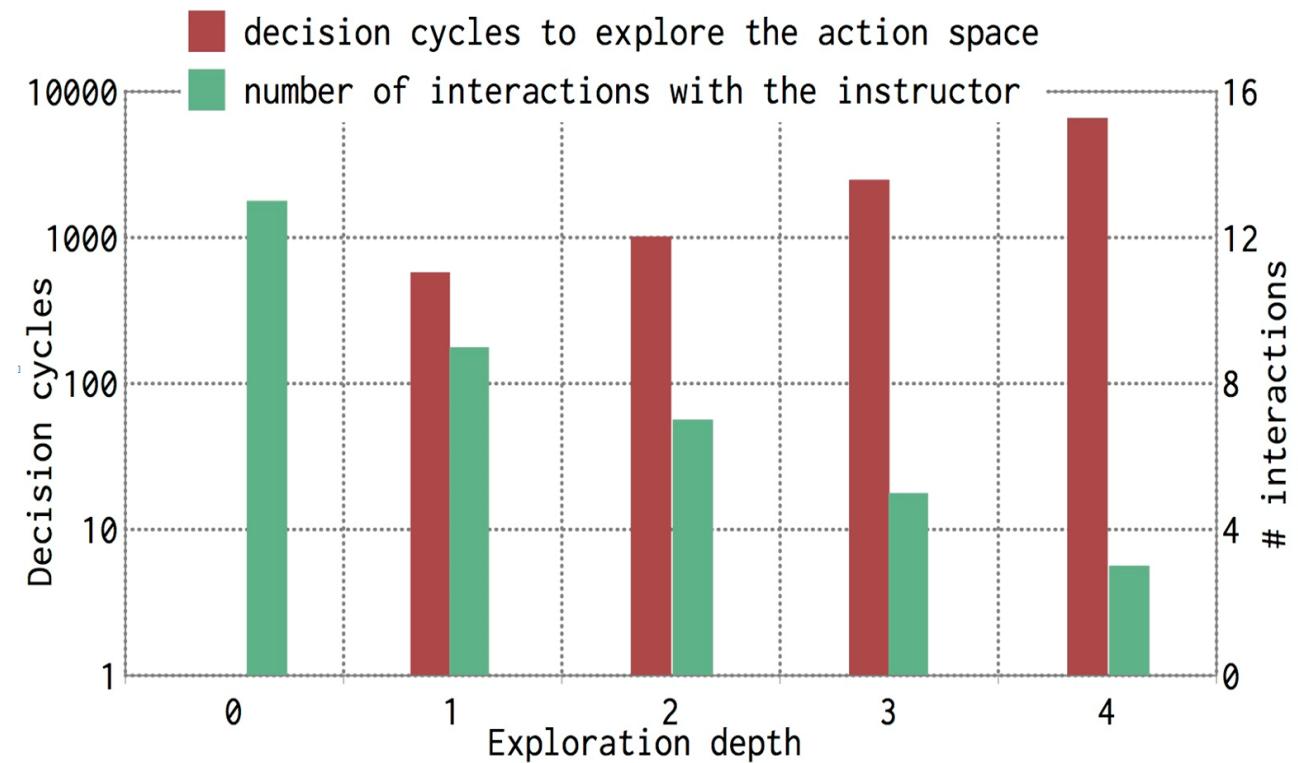
Kirk, Mininger, & Laird 2016 (in submission)



5. Apply reasoning strategies online

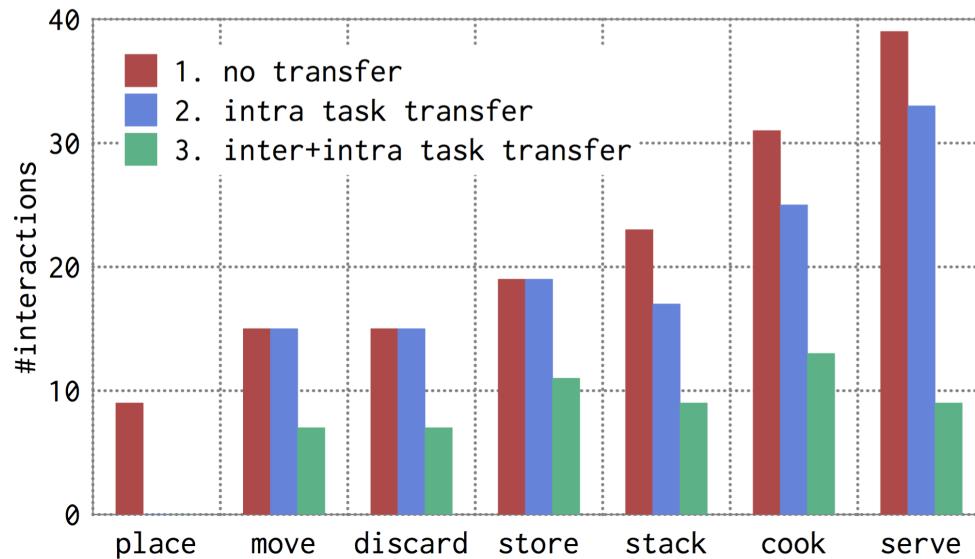
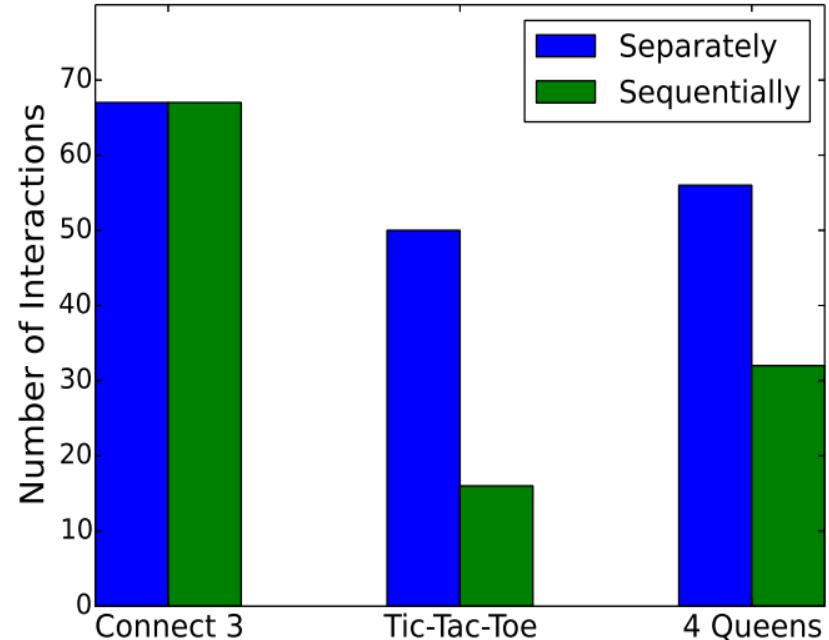
- Dialog is mixed-initiative
 - both participants bear responsibility in information exchange
- Agent must actively reason about dialog
- Rosie applies search to attempt task execution during instruction

Mohan and Laird AAAI 2014



6. Integrate prior knowledge

- Some task knowledge is often shared between different tasks
 - knowledge often structured hierarchically
 - Long-lived agents must support transfer of knowledge to remain efficient at learning
 - Rosie can transfer many kinds and forms of knowledge
 - actions, goals, visual properties, defined concepts
 - structure, tasks, policy
- Kirk and Laird ACS 2014,
Mohan and Laird AAAI 2014*



Other requirements

7. Implement a model of collaborator's knowledge, intention, and goals.
8. Communicate own knowledge, intentions, and goals.
9. Support informative failure states.
 - problem-spaces, impasses in Soar

Expert: Move the green block to the pantry

Rosie: What is green?

Conclusions and Future Work

- AI methods can be leveraged to generate desired behaviors
 - relational, hierarchical representation
 - search-based methods
 - architecture: Soar
- Rosie: a case study in flexible human-robot dialog
 - some good properties, lacks in others
- How do we evaluate if we are making progress?
 - user-studies
 - comparative studies
- What are some other requirements?

Rosie: Learning mobile tasks