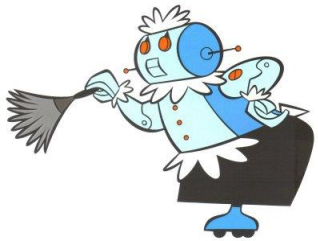
Extending Interactive Task Learning with One-Shot Goal Demonstrations

James Kirk, Aaron Mininger, John Laird

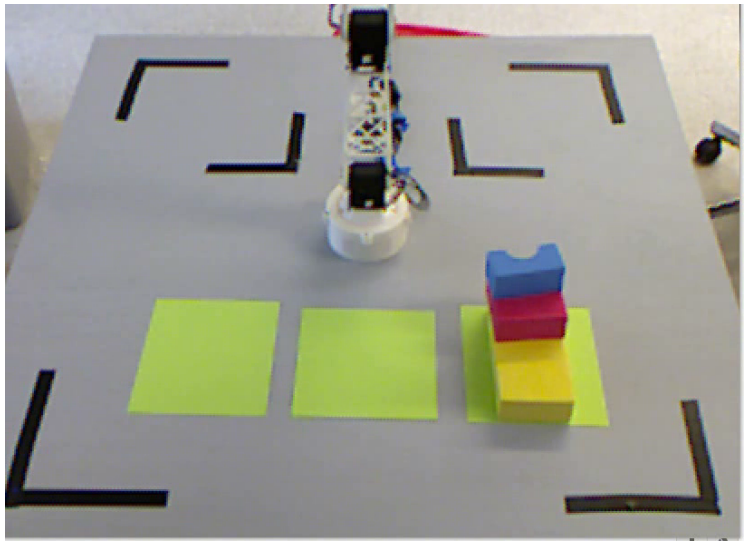
Pending Review for IROS 2015

Soar Workshop 2015

Rosie an ITL Agent



- Tabletop robot
  - Robotic arm for manipulation
  - Kinect sensor for vision
  - Speech (Google) and recognition (CMU sphinx)
- Learns through situated interactive instruction using limited natural language
- Learns concepts about
  - Spatial prepositions (*on, right of, near*)
  - Object attributes (*red, rectangle*)
  - Actions (*move, store*)
  - Games (*tic-tac-toe, tower of hanoi)*

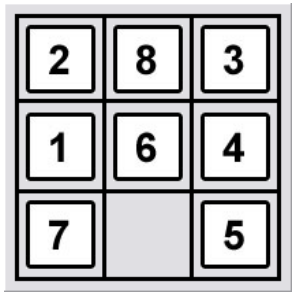


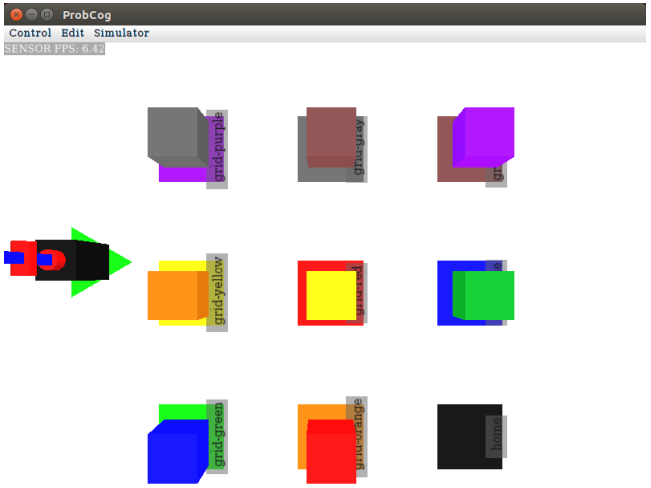
Two Approaches to Goal Demonstrations

- Games/Puzzles
  - Originally: Goal conditions are described
  - Now: Allow for goals states to be demonstrated instead
- Actions/Tasks
  - Originally: The goal for a specific action example is described
  - Now: Omit goal description and use information from example policy to estimate goal state
- Different approaches to refining feature selection in goal state

Instructional Game Learning

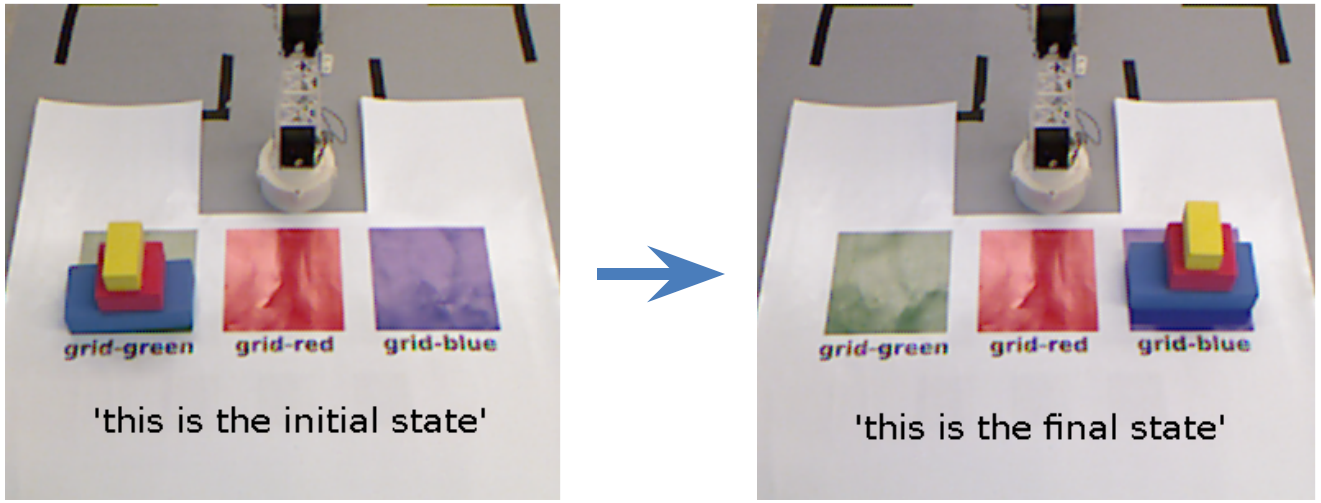
- Rosie prompts the instructor to define the conditions for each action, failure condition, and goal
- Example: Eight puzzle goal (Rosie, **Instructor**)
  - **The goal is eight-puzzle-matched.**
  - Describe objects and conditions for the goal.
  - **A red block is on a red location.**
  - **A blue block is on a blue location.**
  - **An orange block is on a orange location.**
  - **A green block is on a green location.**
  - **A yellow block is on a yellow location.**
  - **A purple block is on a purple location.**
  - **A brown block is on a brown location.**
  - **A gray block is on a gray location.**



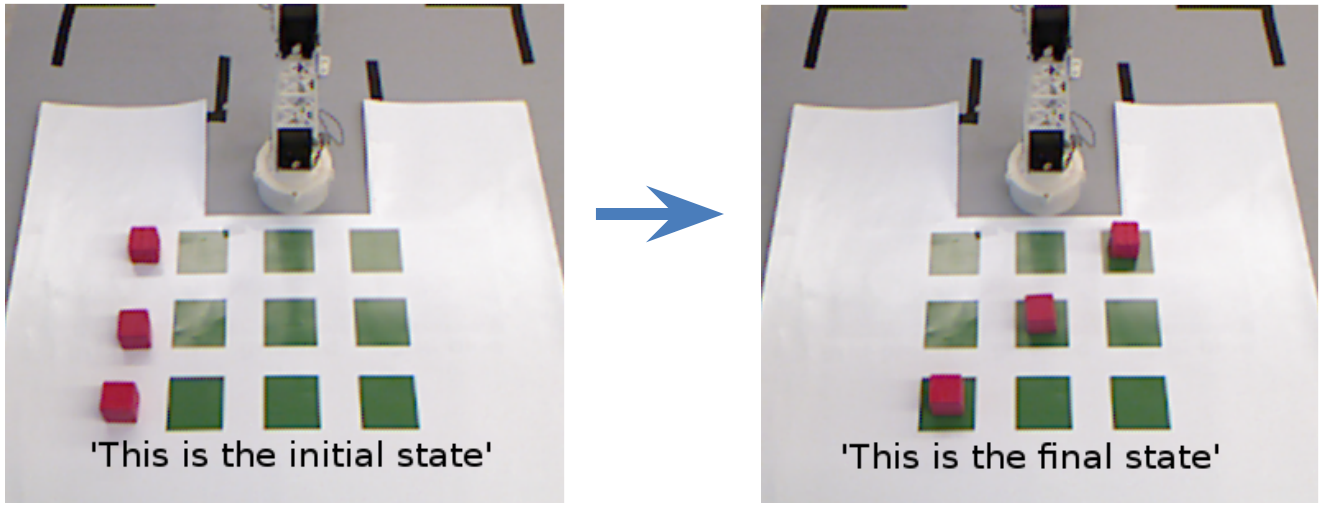


Game Goal Demonstration

Tower of Hanoi



Tic-Tac-Toe



Instructional Action Learning

- Rosie is taught actions through descriptions of the goal state during a specific grounded example
- Rosie can generalize to a more general action policy through EBL
- Example: teaching store (Rosie responses omitted)
  - **Store the red block**
  - **The goal is the red block is in the pantry and the pantry is closed**
  - **Open the pantry**
  - **Pick up the red block**
  - **Put the red block in the pantry**
  - **Close the pantry**
  - **You are done**
- Instead of acquiring a sequence of actions, Rosie also can perform a search to find described goal
- Mohan, S. and Laird, J. 2014. Learning Goal-Oriented Hierarchical Tasks from Situated Interactive Instruction. *AAAI Conference on Artificial Intelligence*, Quebec City, Canada.

Algorithm: State difference

- State representation consists of a set of Object *O* and Predicates *P*
  - Unary predicates describe conditions on specific objects, such red(o_1)
  - Binary, Trinary predicates describe conditions between objects, such as on(o_1,o_2)
- State difference calculates the new predicates in the final state and associated objects to create goal state estimate
- Additionally we add any predicates exclusively over the new set of objects *O*

Objects

O_1-O_3

L_1-L_9

Predicates

red(O_1)

red(O_2)

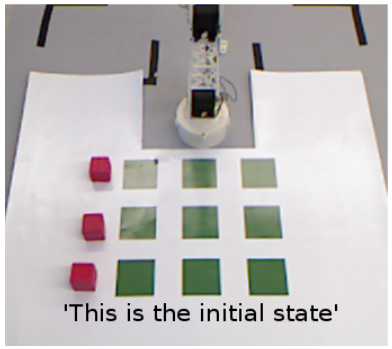
red(O_3)

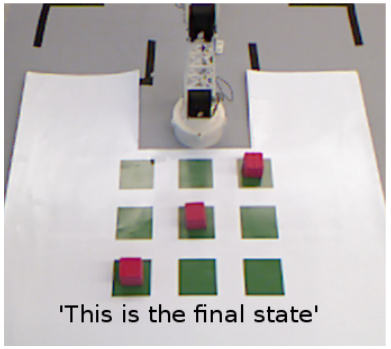
linear(L_1,L_2,L_3)

linear(L_1,L_5,L_9)

linear(L_1,L_4,L_7)

. . .





Objects

O_1, O_2, O_3

L_1, L_5, L_9

Predicates

on(O_1,L_1)

on(O_2,L_5)

on(O_3,L_9)

red(O_1)

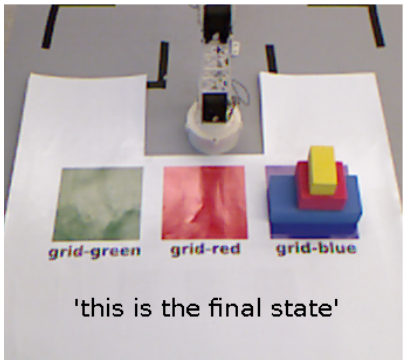
red(O_2)

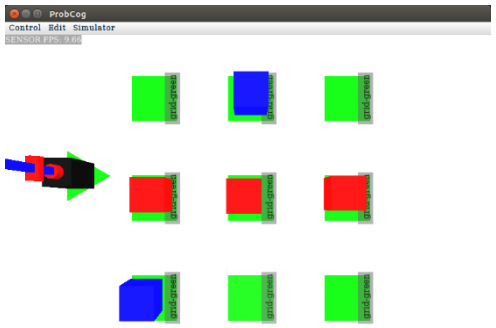
red(O_3)

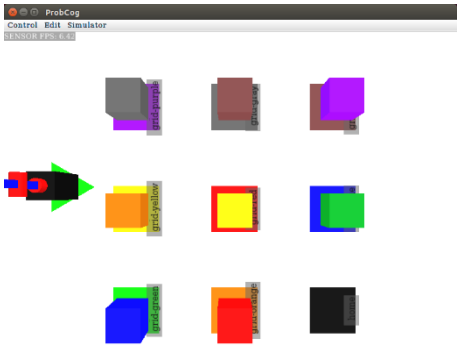
linear(L_1,L_5,L_9)

Goal State Refinement

- Rosie can be instructed to *ignore* or *attend* to specific predicates (by name) and objects (by description)
- Objects and predicates are added and removed from the goal state
- Examples
  - *Attend the blocks* (Tower of Hanoi)
  - *Ignore the blue blocks* (Tic-Tac-Toe)
  - *Ignore below* (Tic-Tac-Toe, Tower of Hanoi)
  - *Ignore near* (Eight puzzle)







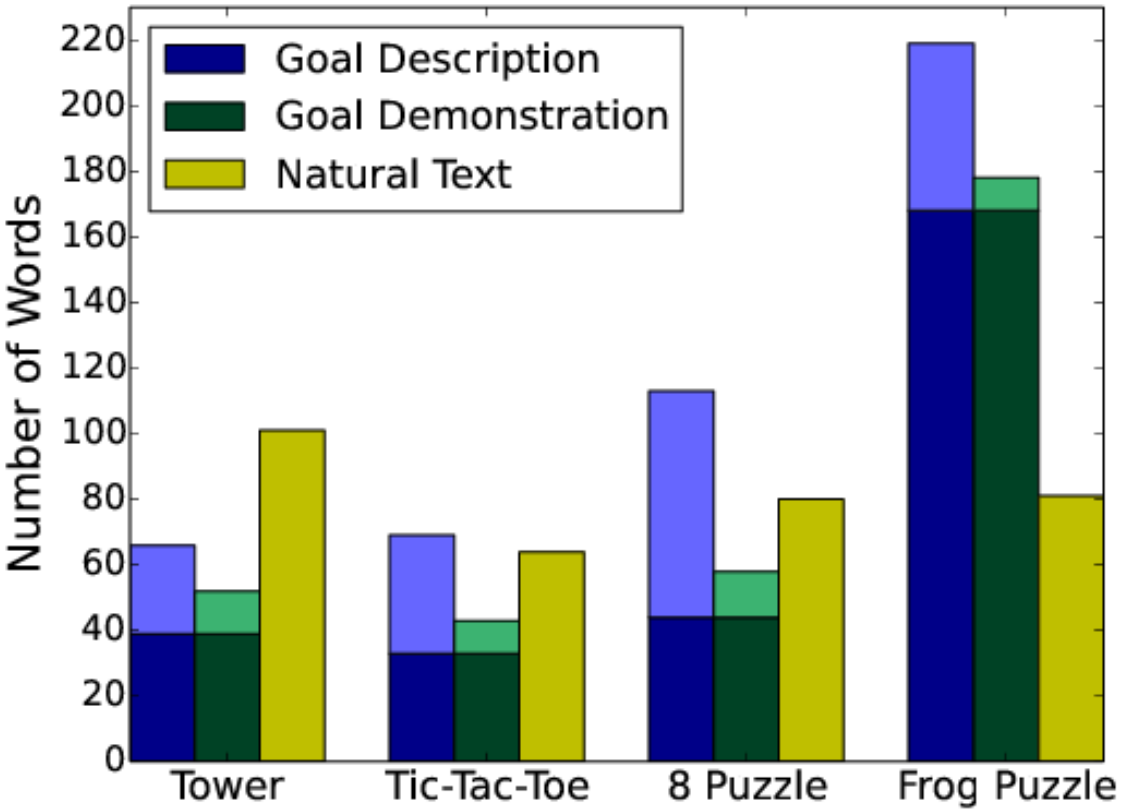
Action Goal Demonstrations

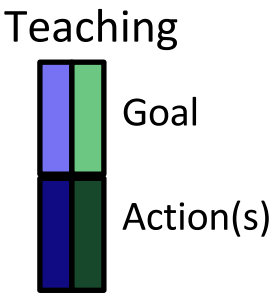
- Rather than starting with the goal, the instructor only provides the sequence of actions to the goal
  - Same algorithm used to estimate goal state between implied initial and final states
- To refine the goal state a different approach is used
  - Rosie keeps track of objects, predicates that were part of action execution sequence
  - Irrelevant new predicates created by the actions are ignored
  - Rosie will only attend to those objects and predicates
  - Also will attend to predicates that changed and then changed back
    - Ex: in the action store, the pantry is closed in the initial and final states

Evaluate both Approaches

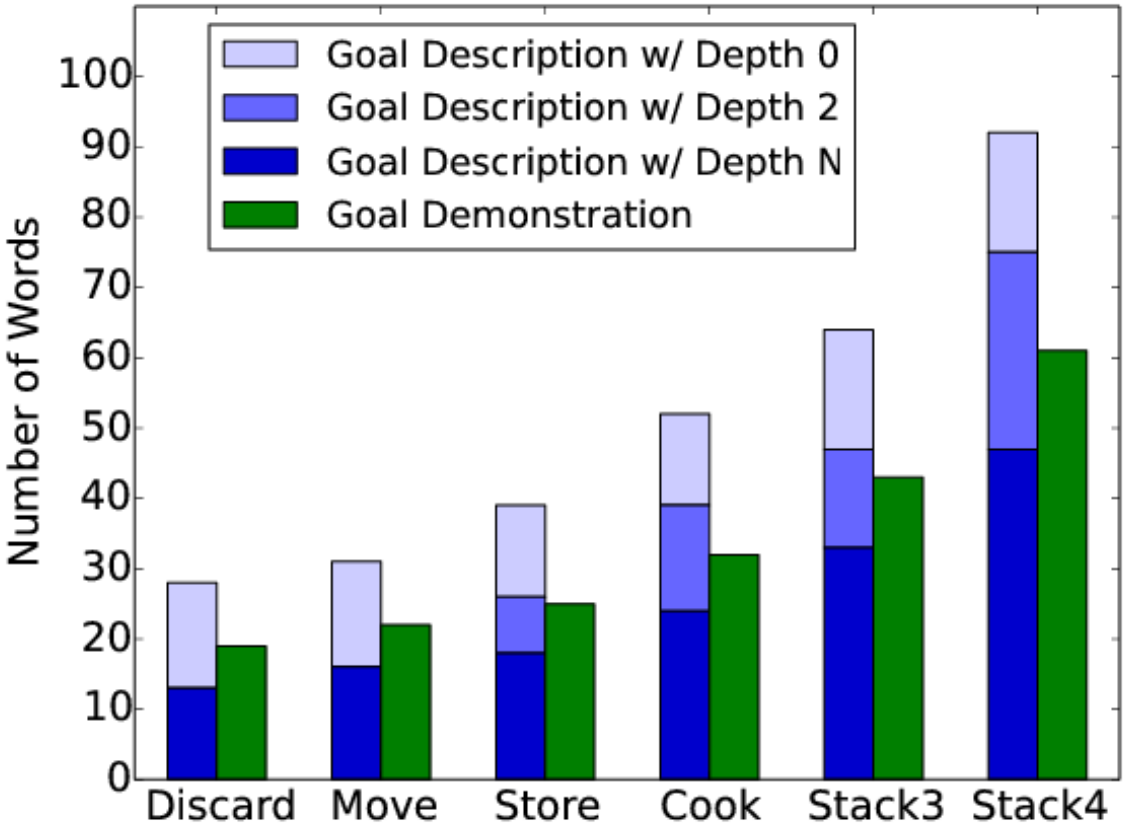
- Evaluate efficiency based on number of words used (by teacher) in teaching interactions
- Game Learning
  - Evaluated on 4 games
  - Compare demonstrations of goal vs. description
  - Natural Language descriptions results as comparison
- Action Learning
  - Evaluated on 6 actions
  - Only can use search when goal is described (not demonstrated)
  - Compare
    - Goal Description with unlimited search, limited, and no search
    - Goal Demonstration

Game Learning Efficiency





Action Learning Efficiency



Nuggets and Coals

Nuggets

- Demonstrates effectiveness/generality over many games and actions
- Improves efficiency of interactions (with some tradeoff)
- Expands accessibility of agent (more ways of communicating)

Coals

- No multi state demonstrations to clarify
- No support for disjunctive conditions in goal
- Teacher cannot access agent's estimations or state of knowledge easily

Questions?

Interactive Task Learning

- New Grand Challenge problem for AI (See 2014 NSF Workshop)
- Interactive
  - Real time
  - Natural (language, gestures, demonstrations)
  - Situated in a shared environment
- Task
  - Policies for solving efficiently
  - The problem formulation (action preconditions, goals, failure conditions, etc.)
- Learning
  - Acquires all knowledge necessary to understand, solve, and perform the task
- **Not**
  - Programmed to handle new tasks, conditions, situations
  - Limited to a specific set or type of tasks
  - Reliant on offline batch processing
  - Using pseudocode-like language specifications

Interactive Task Learning Agents

- Interactive Task Learning agents
  - Dynamically extend tasks that can be performed
  - Interact with a human teacher in a shared environment
  - Accumulate knowledge over many different tasks
  - Applications: service robots, computer assistants, virtual agents
- What are the desired criteria?/How do we evaluate them?
- Desiderata
  - Task Competent
  - Continuous, Accumulative Learning
  - Efficient Execution
  - Task General
  - Efficient Communication
  - Accessible Communication