An Empirical Exploration of Learning to Use Memory

Nicholas Gorski & John Laird
University of Michigan

May 20th, 2010 - 30th Soar Workshop

A Brief Historical Review

- Soar 8.6.3
  - Declarative knowledge in WM, procedural knowledge as rules
- Soar 9
  - Short-term declarative knowledge in WM
  - Long-term declarative knowledge in EpMem & SMem
  - Procedural knowledge as rules, tuned by RL

- Conventional use case: RL learns strategy in environment
  - Knowledge from memory tested on LHS of RL rules
  - Hand-coded rules specify how & when to retrieve from memory

- Moving towards: RL learns strategy over *internal* actions, too
  - RL selects memory retrievals, storage to memory, …
  - When is it computationally feasible for RL learn to use memory?

Learning to Use Memory

- Two senses of "using memory":
  - Testing knowledge from memory
  - Performing encoding, storage, maintenance & retrieval actions
- Learning to use memory involves both senses
  - Learn according to state of world, but also *internal* state
  - Learn when to perform actions over memory

Last Year: Learn Target Behavior

- Develop tasks to elicit specific behaviors
  - Use RL to learn *specific* cognitive capabilities using EpMem
- Well World tasks
  - Virtual sensing
  - Remembering past actions
- Problems:
  - Difficult to design tasks that *require* learning to use a specific memory mechanism
  - Difficult to require that learning a specific cognitive capibility is best
  - Even when you do… conclusions don't necessarily generalize

Now: Large Empirical Study

- No longer trying to learn specific behaviors
- Instead, combine RL + memory + task
  - Study what behaviors emerge
  - Answer whether the task is learnable for RL + memory

- No longer in Soar, instead using a lightweight framework
  - Simpler memory mechanisms
  - Easier to make modifications and explore different architectural commitments
  - Less overhead means we can study many tasks (environments)
  - BUT: want to relate our results back to Soar

Tasks X Memories

- Big empirical study is cross product of task & memory dimensions (two passes)

1. Look at all combinations of tasks and memories
   - Task dimensions X memory dimensions
     - Identify characteristics of task, study them independently
     - Identify representative memory models
   - Analysis
     - Quantitative: time to convergence, % optimal, value of policies
     - Qualitative: classifying behaviors of agents during learning
   - Answer *what* differences exist

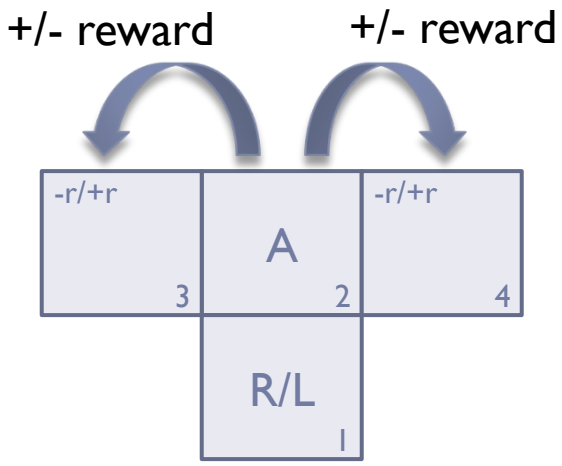
2. In response to differences, modify memories to confirm understanding
   - Answer *why* the differences exist
   - Modify memory mechanisms, provide partial background knowledge, …

Dimensions of Task

- Explore bottom-up: each dimension, independently
- Parameterize a simple task along each dimension, then measure how learning performance scales
- Task characteristics:
  1. Temporal delay between acquiring knowledge & using it
  2. Number of actions that depend on salient knowledge
  3. Amount of salient knowledge that must be maintained
     - E.g. items on shopping list
  4. Number of types of salient knowledge
     - E.g. vocabulary of symbols
  5. Second-order knowledge
  6. Distracting environment features
  7. Distracting action space
  8. Relative cost of internal actions to external actions

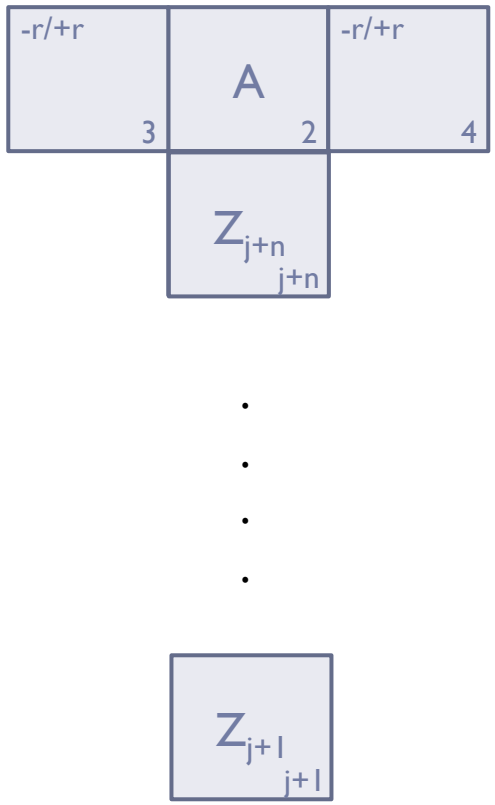
Extensible T-mazes

- Base configuration very simple
- Partially observable
  - Requires memory to solve
- Small number of
  - States
  - Features
  - Actions
- Easily extendible along the task dimensions that we care about



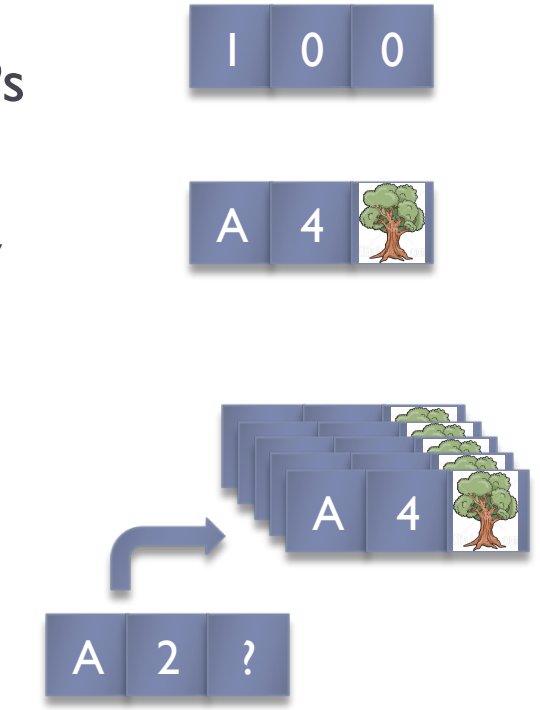
Temporal Delay of Salient Knowledge

- Salient knowledge must persist in memory over time
- How does the delay between acquiring knowledge and when it is used affect learning to use memory?
- Vary the delay between
  - where salient knowledge is obtained
  - where salient knowledge must be used



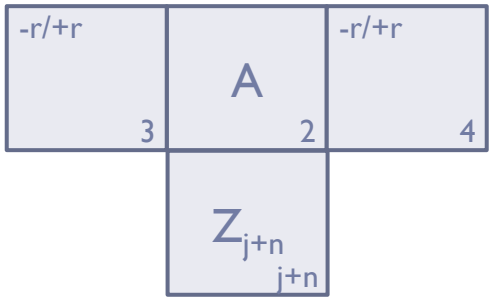
Dimensions of Memory

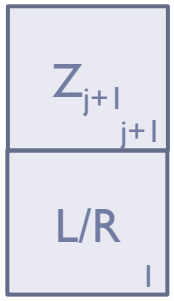
- Space of possible memory models is large: top-down
- Bit Memory
  - Inspired by early work in RL with POMDPs
- Gated Working Memory
  - Perceptual symbols maintained in memory
- Associative Memory
  - "Episodic-like" but without temporal indexing
  - Cue-based retrieval
  - Dimensions: which biases determine best match



Current Progress

- Framework up and running
  - Bit memory & gated working memory learn base TMaze
  - Completed initial parameter sweeps
- Gathering results from temporal delay task
  - Gated working memory scales better than expected
  - No qualitative analysis yet





+

(gated WM)

Desirable Contributions

1. Better understanding of how tasks imply which memory models are most applicable

2. Better understanding of how traits of memory models affect how an agent may learn to use them

3. Rough, empirical bounding of space of tasks that are computationally tractable to learn to use memory in

4. Identifying and classifying behaviors that emerge in the course of learning to use memory

5. Ideally, analytical results
   - describing what combinations of task + memory are learnable
   - describing how task + memory scale along dimensions

Nuggets & Coal

- Empirical work is straightforward…
- Potentially applicable to
  - Soar,
  - AI,
  - CogSci communities
- Analytical results would be great
- …analysis is the hard part
- Soar and its memory mechanisms aren't being used; need to make extra effort to be relevant
- Empirical results might not be exciting