Methods for Transfer Learning Using Soar

26th Soar Workshop

May 24, 2006

Nicholas Gorski, John Laird, Taylor Lafrinere

What is Transfer Learning?

- Transfer Learning: using previously learned knowledge to improve performance on later, related tasks
- Important because traditional machine learning techniques learn over narrow problem-spaces
- TL approaches should be general enough to apply to a wide variety of domains

Developing TL Approaches

- Our strategy is to develop multiple approaches and compare their performance on various tasks
- Combining approaches to leverage strengths of each will achieve good transfer

- Soar is a good platform for this comparison study: it provides multiple learning mechanisms in a single architecture

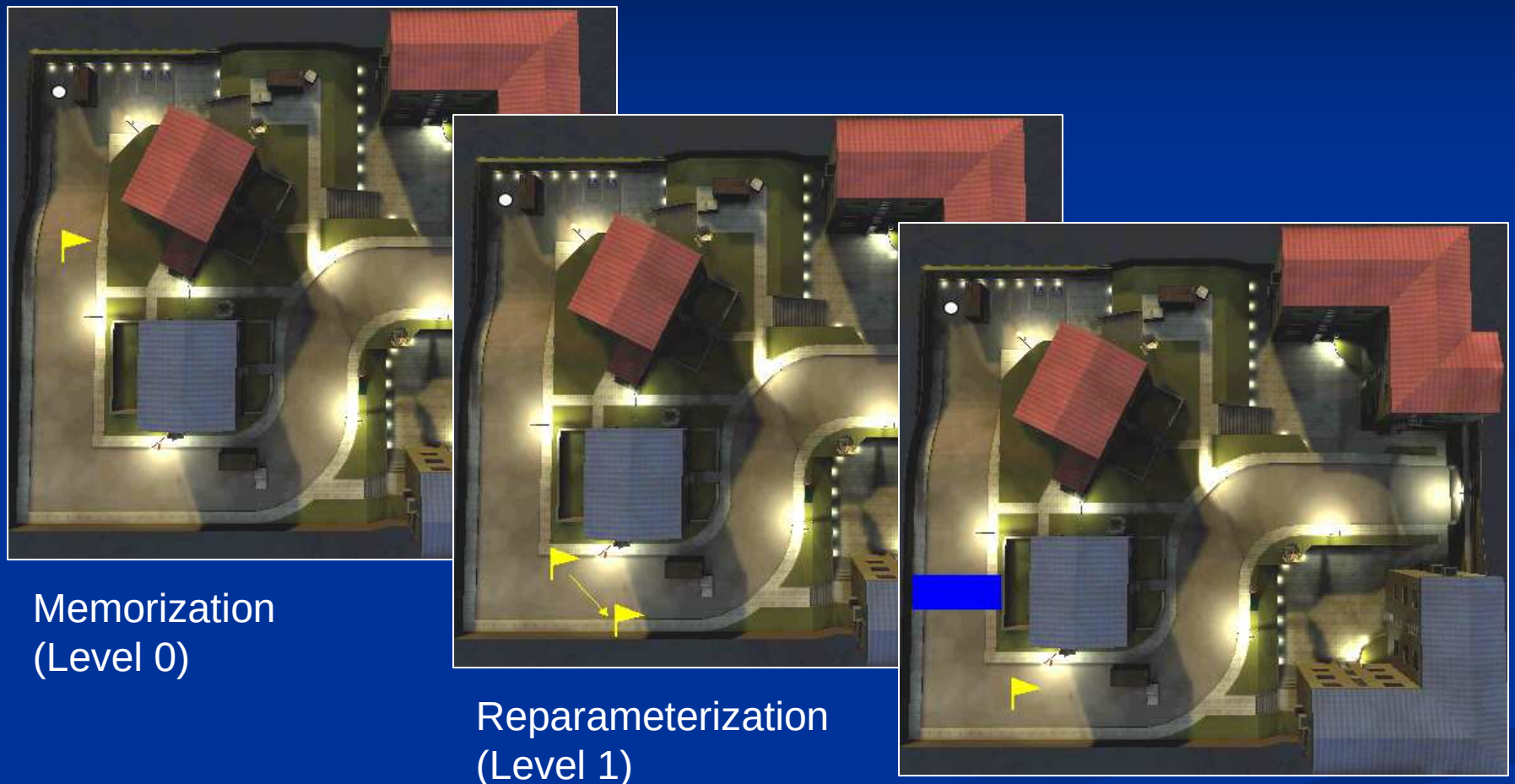
Urban Combat Testbed

- Multi-agent FPS real-time video game
- Built on Quake 3 engine
- UCT exposes shared memory interface, interfaced to Soar kernel via SML

Transfer Learning in UCT

- Many different types of transfer, divided into 10+1 levels
- First-to-flag scenarios created for UCT to test each level of transfer
  - Each scenario consists of a source and target problem
  - Knowledge transferred from the source improves performance on the target
- Declarative knowledge that can be transferred includes the location of a flag, a map, and routes

Transfer Learning in UCT



Memorization
(Level 0)

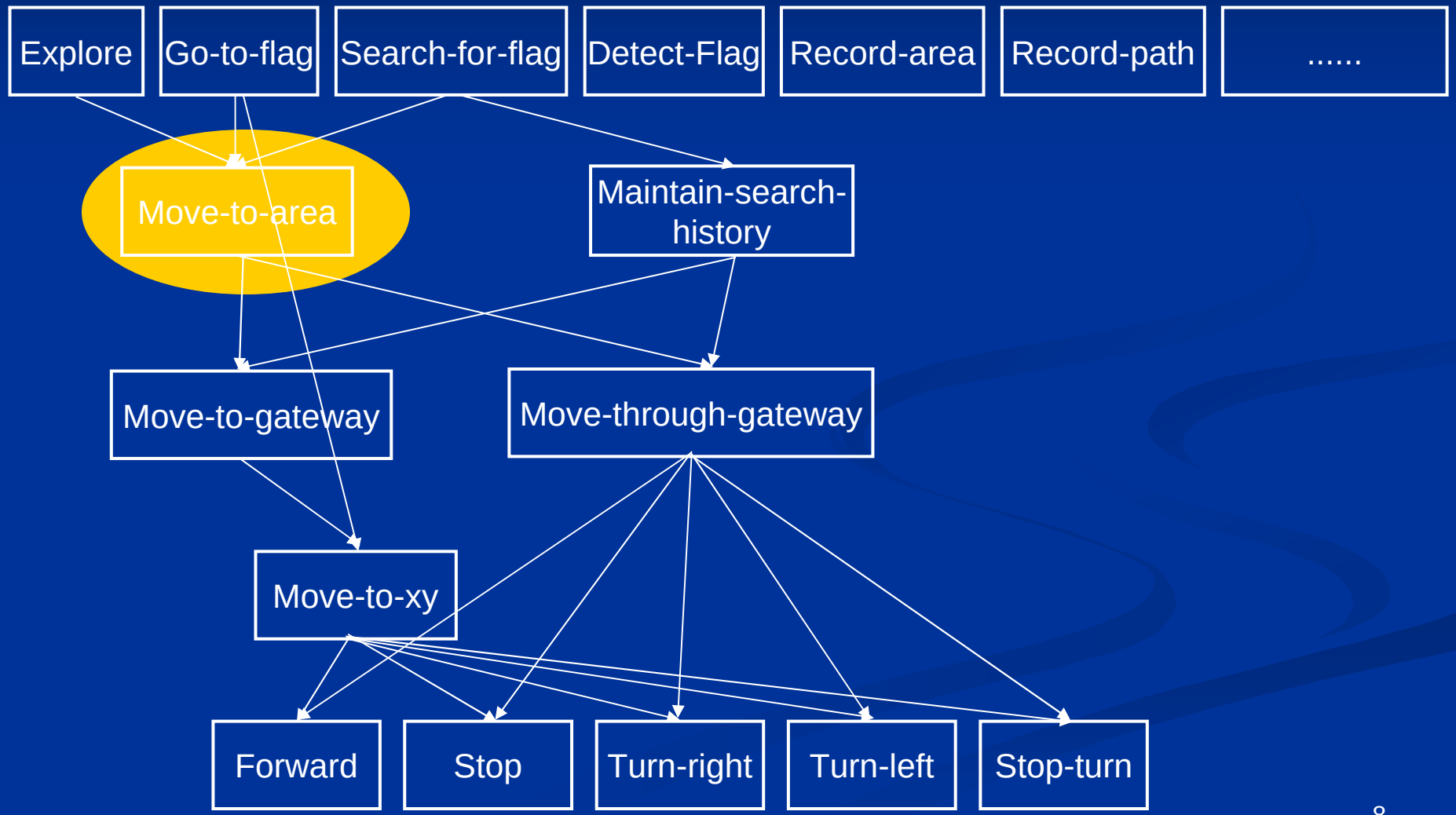
Reparameterization
(Level 1)

Extrapolation
(Level 2)

Approaches

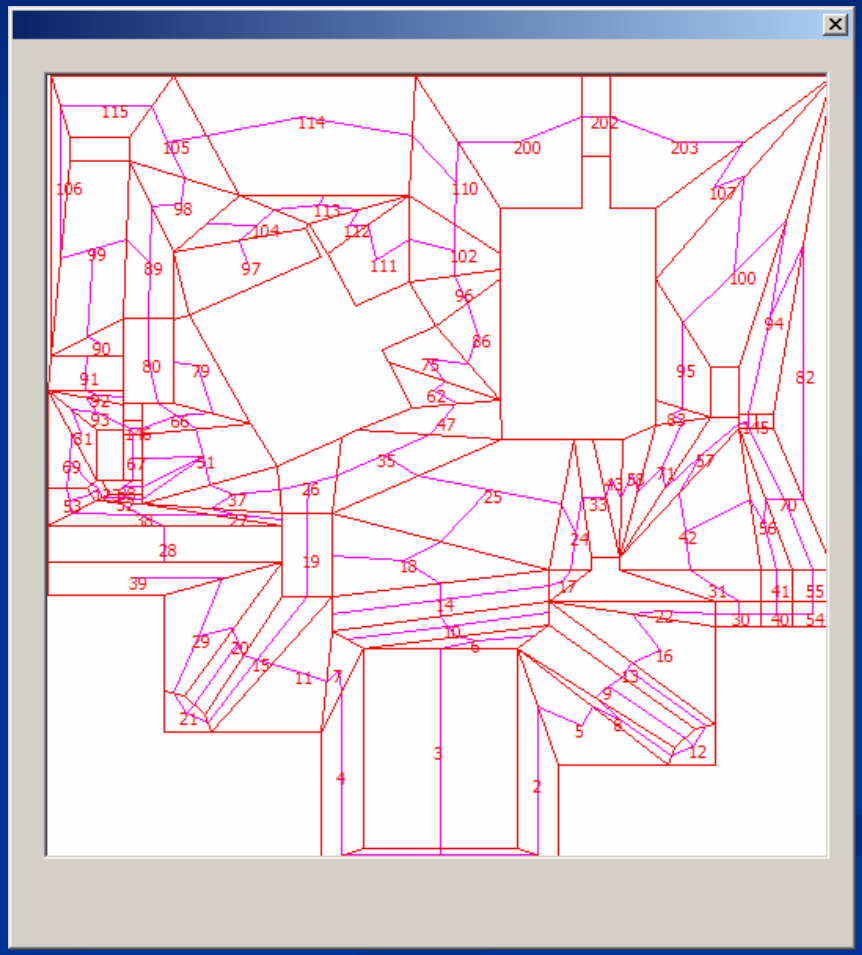
- Multiple approaches developed that take advantage of multiple learning mechanisms in Soar
  - Memory-based
  - Search-based
  - Reinforcement Learning

Urban Combat Agent - UCBot



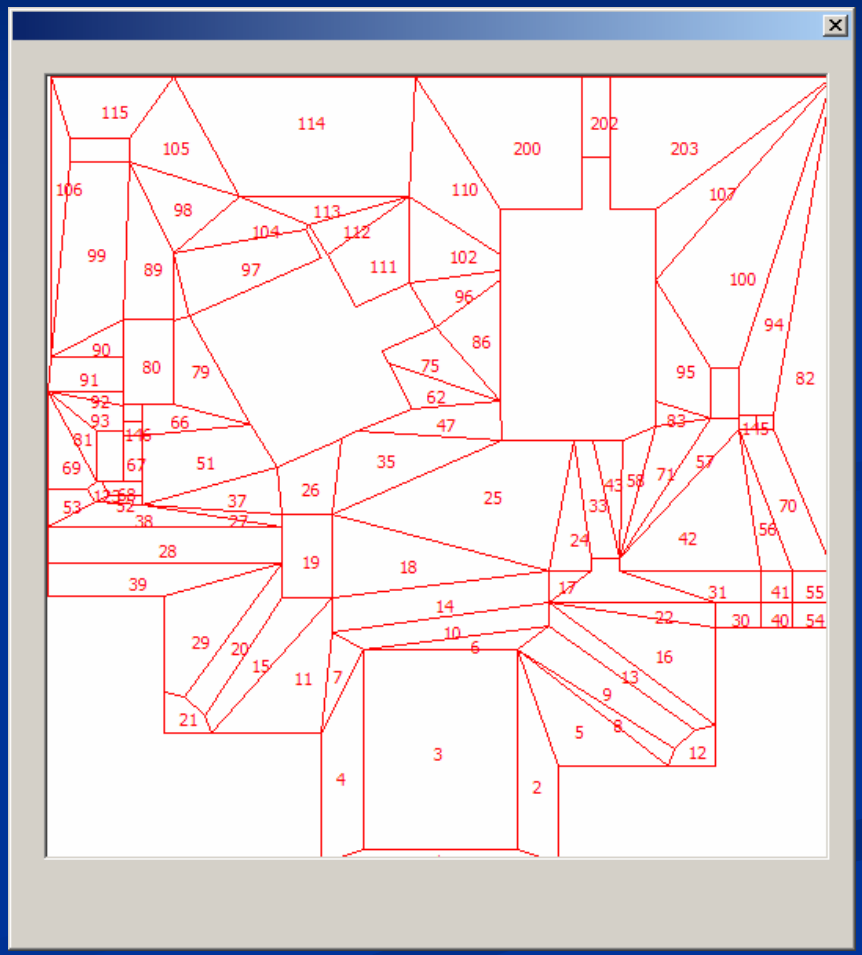
Memory-based

- Trades memory for fast execution
- Stores *O(N)* areas & *O(N²)* paths in working memory
- When the agent needs a path, no computation is required
- Abuses working memory and affects the Rete matcher



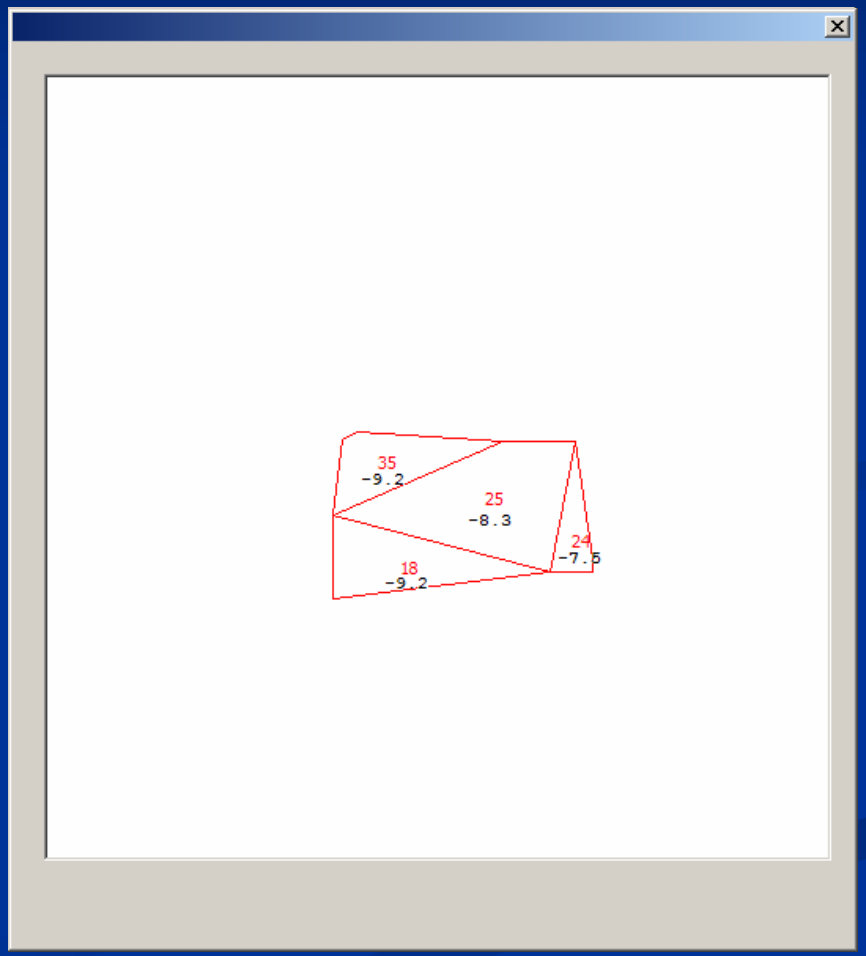
Search-based

- Still maintains N areas, but no path information in working memory
- When the agent needs a path, it performs a model-based search using a nearest-gateway heuristic
- Chunks results of search: agent stores path information as procedural knowledge



Reinforcement Learning

- Stores no spatial knowledge in working memory
- Learns value of moving to an area, stored as numeric preferences on operator proposals
- Significantly longer training time required
- However, route finding UCT is essentially deterministic



Reinforcement Learning

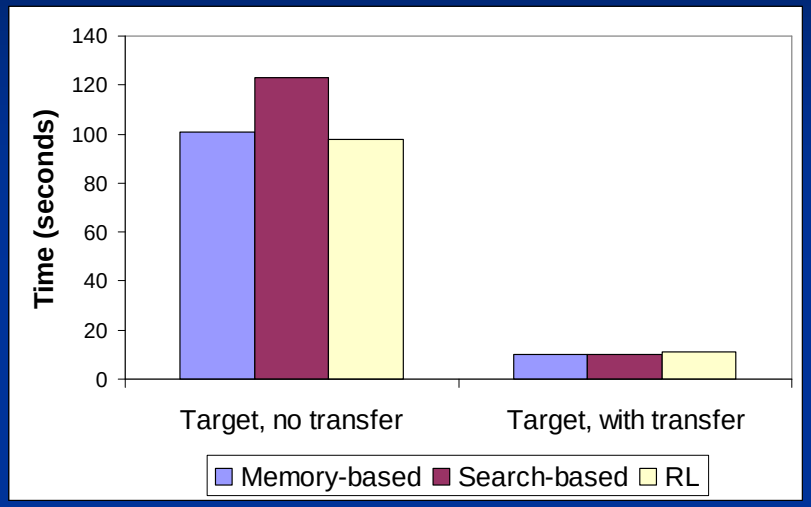
- Location of flag and paths are stored implicitly as expected values
- Captured as RL rules, where numeric preferences specify the expected value of moving to an area

```
sp {|RL-115|
   (state <s1> ^name go-to-flag ^top-state <s2> ^operator <o1> +)
  -{ (<s2> ^map <m*1>)
     (<m*1> ^areas <a*1>)}
   (<s2> ^parameters <p1>)
   (<p1> ^strategy |RL|)
   (<o1> ^name move-to-area ^type connected-area ^path-to-take <f1>)
   (<f1> ^to-area <a1>)
   (<a1> ^id 115)
-->
   (<s1> ^operator <o1> = 0.00385)
}
```

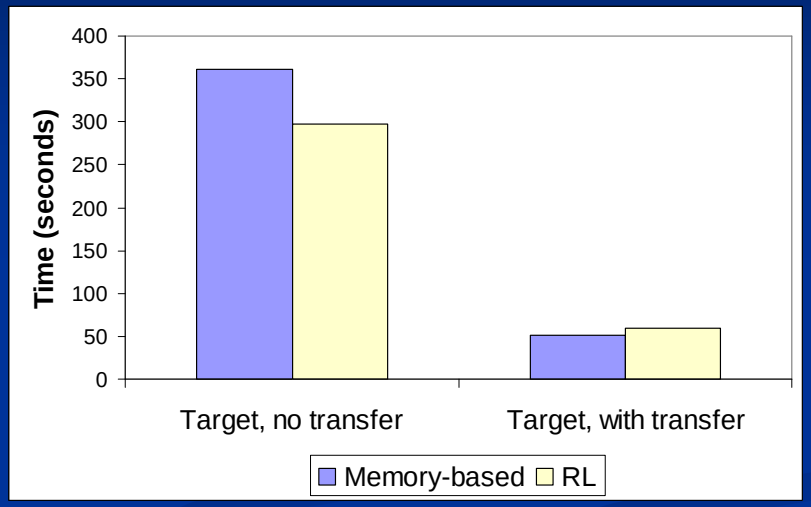
```
sp {|RL-107|
   (state <s1> ^name go-to-flag ^top-state <s2> ^operator <o1> +)
  -{ (<s2> ^map <m*1>)
     (<m*1> ^areas <a*1>)}
   (<s2> ^parameters <p1>)
   (<p1> ^strategy |RL|)
   (<o1> ^name move-to-area ^type connected-area ^path-to-take <f1>)
   (<f1> ^to-area <a1>)
   (<a1> ^id 117)
-->
   (<s1> ^operator <o1> = 1.0)
}
```

Evaluation

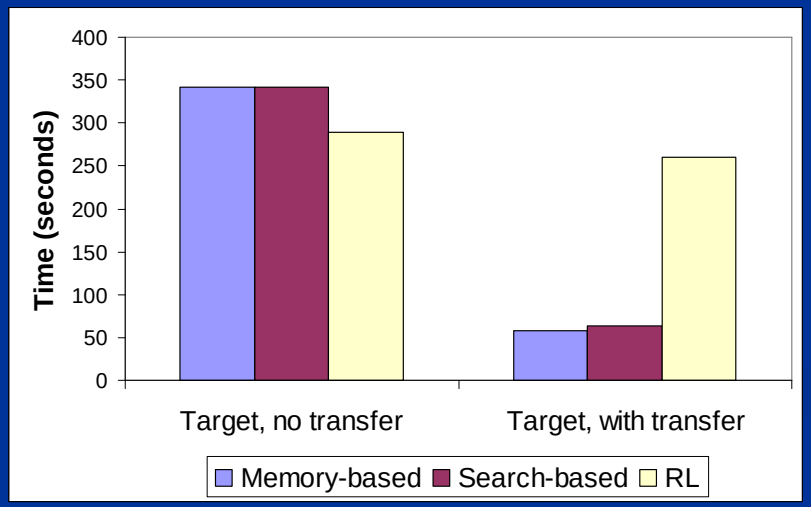
Memorization (Level 0)



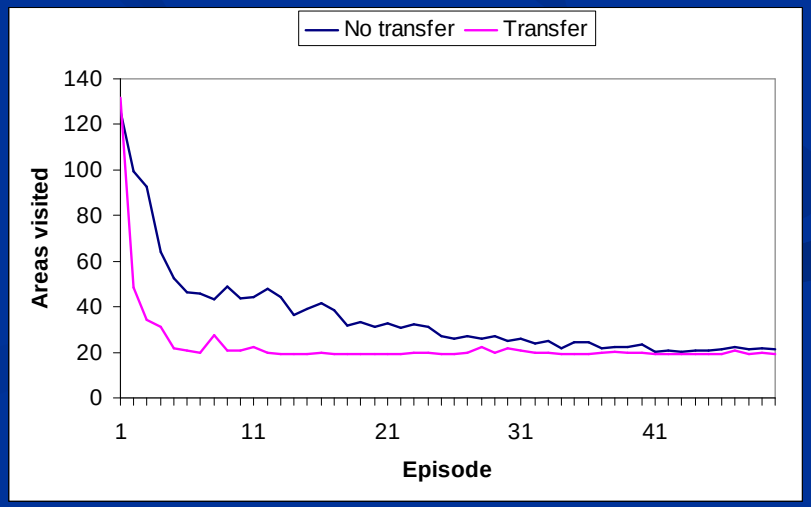
Reparameterization (Level 1)



Extrapolation (Level 2)



RL agent's level 2 learning curve



Generalizing the Approaches

- How do these approaches generalize across domains?
  - Memory-based: pre-compute all solutions and store in working memory
  - Search-based: store domain knowledge in working memory, but search for solutions rather than pre-computing all possible solutions
  - Reinforcement Learning: transfer of statistical knowledge
- Strongest transfer will be achieved when all approaches are combined, leveraging the strengths of each

Nuggets & Coal

- Existing UCBot should easily extend to higher levels of transfer
- Comparison of memory/search based methods to RL highlight strengths and weaknesses of each
- Results for only one domain, and UCT is not an ideal TL testbed (yet)
- Haven't investigated transferring more complicated semantic knowledge