Categorizing Transfer for Reinforcement Learning

MATTHEW E. TAYLOR AND PETER STONE

taylorm@usc.edu
Computer Science Department
University of Southern California

pstone@cs.utexas.edu
Department of Computer Sciences
University of Texas at Austin

Based on the forthcoming JMLR article: Transfer Learning for Reinforcement Learning Domains: A Survey

### Algorithm Differences

<table>
<thead>
<tr>
<th>Method</th>
<th>Source Task</th>
<th>Target Task</th>
<th>Cooperative</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MB model based learner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRL relational reinforcement learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Selection of Open Questions

**Theoretical results**
- Majority of results are empirical
- Guarantee improvement for pair of tasks
- Define relationship between amount/quality of knowledge and improvement
- Find an optimal inter-task mapping

**Negative transfer**
- Transfer can be harmful for a pair of tasks
- Identify incompatible pairs of tasks (per TL method)
- Identify when transfer is harming learner (on-line) in target task

**Concept drift**
- In transfer, new tasks are typically announced and changes are discrete
- What if tasks change gradually over time?
- What if agent is not told when it enters a new task?

### Task sequence construction

Given a target task, one may construct/select a sequence of source tasks
- Goal: reduce total training time
- What is the best way to select this sequence?
- Meta-planning problem

### New Directions

- Transfer in repeated normal form games or stochastic games?
- Transfer in POMDPs?
- Learn multiple RL tasks simultaneously (MTL)?
- Develop a domain-independent metric for TL performance?

### Acknowledgements

This work has been placed in the Learning Agents Research Group (LARG) at the Artificial Intelligence Laboratory, The University of Texas at Austin. LARG research is supported in part by grants from the National Science Foundation (CNS-0613014), DARPA (FA8750-05-2-0283 and FAMOS/CNCW2), the Federal Highway Administration (DTH01-07-H-00030), and General Motors.

### Distinctions from Other Settings

**Transfer Learning (TL)**
- Use source task knowledge to learn target task
- Goal 1: Learn target task(s) better with past knowledge
- Goal 2: Learn sequence of tasks better than directly learning final task

**Multi-task Learning (MTL)**
- Fixed (often known) distribution over tasks
- Goal: learn n + 1 th task better

Many goals and metrics are used: no standard.

### Related paradigms

- Lifelong learning: Tasks may be spatially (and temporally) separated; agents identify new tasks autonomously
- Imitation Learning: Observe an outside actor rather than reuse own knowledge
- Human Advice: Human integrated in the loop to give on-line feedback
- Shaping: Human directs training process (e.g., reward shaping)