Transfer Learning in Spatial Reasoning Puzzles

Baylor Wetzel

Department of Computer Science and Engineering, University of Minnesota

wetzel@cs.umn.edu

1 Introduction

Transfer learning is the process of using knowledge gained while solving one problem to solve a new, previously unencountered problem. Current research has concentrated on analogical transfer – a mechanic is able to fix a type of car he has never seen before by comparing it to cars he has fixed before. This approach is typical of case-based reasoning systems and has been successful on a wide variety of problems [Watson, 1997]. When a new problem is encountered, a database of previously solved problems is searched for a problem with similar features. The solution to the most similar problem is selected, adapted and then applied to the new problem. Similar methods exist for adapting reinforcement learning policies [Taylor and Stone, 2009].

We refer to the above approaches as solution adaptation algorithms – a pair of problems are matched on similarity and the solution to the first problem, after some adaptation, is applied to the second problem. The solution adaptation approach requires three things. First, the two problems must be substantially similar in surface or structural features. Second, there must exist a clear method of adapting one solution to another. This is typically done through manually-authored feature mappings or adaptation rules. Third, problem similarity must imply solution similarity.

The "similar feature, similar solution" assumption does not hold for all domains. One such domain is tower defense. Tower defense (TD) is a broad category of spatial reasoning puzzles that share a common theme – a set of agents follow a path through a maze (the map) and the player must prevent them from making it to the exit by placing "defense towers" at strategic locations. There are many different tower defense games, each with their own set of towers. The majority of these games have pieces that fit into one of five archetypes: slow but strong, fast but weak, area of effect (AoE), damage over time (DoT) and slowing. The specific traits of each tower and their relative strengths vary from game to game.

There are hundreds of TD games, most with multiple maps, making it a good domain for investigating multiple levels of transfer learning. It is not, however, a good candidate for solution-adaptation algorithms. Problem (maze) similarity does not imply solution (piece placement) similarity. Small differences in mazes can lead to qualitatively different solutions and there is no clear way to adapt one maze's solution to another.
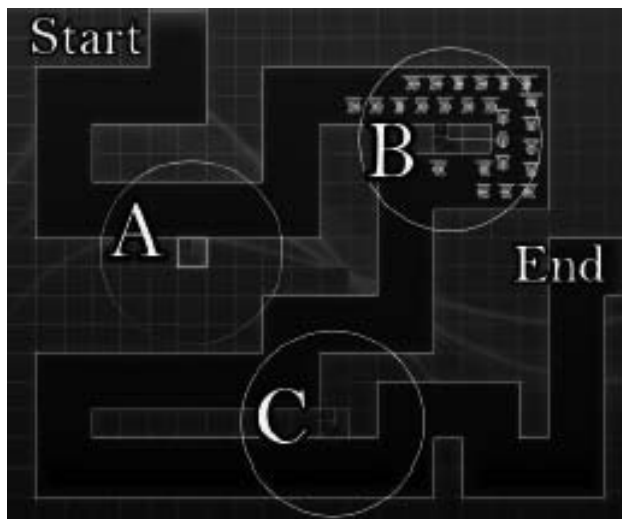

Figure 1: Training map in GopherTD. Yellow objects are agents trying to move from Start to End. Circles show range of three towers in three qualitatively different positions.

2 A New Approach to Transfer Learning

Through human studies we have shown that significant transfer learning is possible in tower defense. Human subjects were given two training problems, neither of which subjects could initially solve. After subjects announced they had learned to solve these problems, they were given six new problems, three equally difficult and three harder. 75% of subjects solved all equally difficult problems on their first attempt and half solved at least one hard problem.

During training, subjects learned which properties of the problem to pay attention to, the opportunities they present and strategies for exploiting them. In this domain, properties were high-level, moderately abstract properties and features such as U-turns, path length, region capacity, areas of significant overwatch and degree of temporal separation of nearby path segments. When presented with a new problem, subjects generated novel solutions by re-representing the problem in terms of the features they had learned, identifying threats and opportunities, selecting a set of general purpose and opportunity-specific strategies and then using those

strategies to make the set of decisions, either locally or holistically, that constituted the solution. The process of applying strategies included merging non-competing strategies, arbitrating conflicts between competing ones, mapping abstract to concrete actions and making tradeoff decisions in the face of limited resources. As new problems were solved, new representations and strategies were added to the existing set.

The ability to learn strategies during problem solving and apply them to new, qualitatively different problems is, in computer science, an unexplored method for transfer learning. We are using a combination of human studies and computer models to investigate how this might be implemented in a computer.

3 Work to Date

In the first phase of our work, we collected data from human subjects on solutions, strategies, features, representations and performance as well as between and within subject variability. Subjects solved problems under a think aloud protocol and were then interviewed in an after-action review. Problems were presented in GopherTD, an instrumented tower defense game created for this research.

We were not surprised to see that a subject's strategies, features and performance were relatively stable while their solutions were not – when asked to repeat a given solution, subjects were unable to recreate their previous solutions exactly, although the new solutions were qualitatively similar in that solutions had roughly the same shape, distribution and density of pieces and that placement decisions were explained with the same justifications.

What we did not expect was the variance between subjects. Although we had believed the task to have a single dominant strategy, eight subjects came up with eight substantially different sets of strategies, some of which solved problems we had previously believed to be unsolvable. There was moderate overlap in features, goals and strategies but overall solutions differed significantly.

In the second phase, we investigated how well a human could recognize a problem solver from their solution and how well they could infer the strategies used to generate a solution. The answer to both questions is quite well. This information was required to evaluate the results of the third phase.

Subject descriptions were often not directly implementable and certainly not complete. In the third phase, we created a set of computational representations that we believed necessary to support the reported strategies. We attempted to find the minimal set of features necessary to cover all strategies reported by all subjects, balancing the abstraction necessary for generalization and transfer against the concreteness necessary to make the data actionable. To test these representations, we manually created agents that simulated specific subjects and used a Turing-like test (double-blind classification task) to determine whether a human expert could distinguish a subject's solutions from ones generated by a computer imitating the subject. For the subjects modeled so far, the answer is no.

4 Negative Transfer

An unexpected finding was that 25% of subjects performed at or below chance. Further research determined that this poor performance was due to transfer learning.

Subject had no previous experience with GopherTD but had mastered one or more other tower defense games. All of the subjects who did well in GopherTD used five or more slowing towers. Subjects who did poorly did not. These subjects had previous experience with TD games where slowing towers were less useful and as a result transferred the high-level knowledge "slowing towers are weak" to GopherTD. Although the slowing towers in each game are different, their high-level abstractions ("a tower that slows enemies") are not, leading to negative transfer.

Although there are many benefits to transfer learning, we believe that, outside of a researcher's carefully controlled experiments, it is just as likely to harm as help. Before transfer learning can find widespread success, there are many questions that need to be answered such as how to determine when transfer learning is appropriate, whether performance is benefitting or being harmed by transfer learning and how to selectively apply (and stop applying) knowledge.

5 Future Work

We continue to bring in new subjects, collecting data and modeling subjects.

Our current work focuses on finding representations that allow for knowledge transfer. In addition to being comprehensive and actionable, these representations must allow for learning, meaning that they must serve as a learning target.

Our next step is to model strategy inference - given a set of solutions, the computer must infer what strategies the solver used in enough detail that the computer can predict what solutions that solver would generate on related but qualitatively different problems (e.g., on different maps).

6 Conclusions

We believe this approach to be quite novel and consequently unexplored. Despite the amount of work we have done, there are numerous open questions and attendant risk. The rewards, however, are significant. For example, we believe that the strategy learning and application approach described here is capable of achieving true level eight transfer performance on DARPA's 10-level transfer learning scale [DARPA, 2005].

References

[DARPA, 2005] DARPA. Transfer learning proposer information pamphlet for broad agency announcement 05-29, 2005.

[Taylor and Stone, 2009] Matthew E. Taylor and Peter Stone. Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10(1):1633–1685, 2009.

[Watson, 1997] Ian Watson. *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Morgan Kaufman, 1997.