Graph-Based Analysis of Human Transfer Learning Using a Game Testbed

Diane J. Cook and Lawrence B. Holder
Washington State University
Pullman, WA 99164
{cook,holder}@eecs.wsu.edu

G. Michael Youngblood
The University of North Carolina at Charlotte
Charlotte, NC 28223
youngbld@uncc.edu

Index terms: data mining, graph algorithms, transfer learning, games

Abstract

The ability to transfer knowledge learned in one environment in order to improve performance in a different environment is one of the hallmarks of human intelligence. Insights into human transfer learning help us to design computer-based agents that can better adapt to new environments without the need for substantial reprogramming. In this paper we study the transfer of knowledge by humans playing various scenarios in a graphically realistic urban setting which are specifically designed to test various levels of transfer. We determine the amount and type of transfer that is being performed based on the performance of human trained and untrained players. In addition, we use a graph-based relational learning algorithm to extract patterns from player graphs. These analyses reveal that indeed humans are transferring knowledge from one set of games to another and the amount and type of transfer varies according to player experience and scenario complexity. The results of this analysis help us understand the nature of human transfer in such environments and shed light on how we might endow computer-based agents with similar capabilities. The game simulator and human data collection also represent a significant testbed in which other AI capabilities can be tested and compared to human performance.

1. Introduction

Learning capabilities in biological systems still greatly exceed the capabilities of machine learning algorithms. This is partly due to their ability to transfer knowledge acquired over the course of their lives. Simon [28] defined learning as the ability to leverage experience in previous tasks into improved performance on new tasks *from the same population*. This represents a very basic type of learning. What sets humans apart is the ability to leverage experience in previous tasks into improved performance on tasks from *different* populations: tasks that superficially have little or nothing in common with the previous ones, but share some deep structure or subtasks that make it possible to apply previously-acquired knowledge and skills.

The ability to identify these deeper, subtler connections, what we term *transfer learning*, is the hallmark of human intelligence. Byrnes [6] defines transfer learning as the ability to extend what has been learned in one context to new contexts. Thorndike and Woodworth [34] first coined this term as they explored how individuals transfer learning in one context to another context that shared common features. The importance of transfer learning for educational methods has been studied extensively [9], and evidence of transfer learning for tasks such as learning how to use text editors is provided in the literature [5]. Barnett and Ceci provide a taxonomy of features that influence transfer learning in humans [2].

Inspired by biological transfer learning, machine learning researchers have studied methods for implementing transfer learning in computational settings, including reinforcement learning [1,30,32], cognitive architectures [13], Bayesian networks [22], genetic algorithms [31], neural networks [35], inductive logic programming [36], and logistic regression [21]. Achieving transfer learning on a variety of tasks remains an elusive goal for these researchers. Others have

demonstrated the use of transfer learning in humans for cognitive rehabilitation [10] and to improve the health of businesses [17].

In this paper, we analyze the ability of humans to perform transfer learning. In particular, we hypothesize that humans are able to perform transfer learning at several levels of complexity. We analyze this ability using graph-based relational learning applied to data collected from humans playing a military-realistic game. We first observe the amount of transfer that occurs by comparing learning curves between trained and untrained human players. Next, we take a closer look at the knowledge itself that is transferred from source to target scenarios. Employing unsupervised discovery and supervised learning techniques, we identify patterns in logged games that indicate the amount and nature of the knowledge that is transferred in a variety of game transfer scenarios. The results of this analysis help us understand the nature of human transfer and shed light on how we might endow computer-based agents with similar capabilities.

2. Urban Combat

To support our study, we have developed the Urban Combat Testbed (UCT) [39]. UCT is a high-fidelity Military Operations in Urban Terrain (MOUT) simulation built on top of the commercial, open sourced Quake 3 Arena [26] first-person shooter (FPS) game engine. UCT provides a high-bandwidth interface for computer-based agents to perceive the environment and execute actions in real time. UCT also provides a standard FPS graphics interface for human agents to play the various scenarios defined for UCT. In either case, the perceptual and command execution information is logged in enough detail to replay the game. This logged information provides valuable data for performing behavioral analysis of human, and ultimately computer, players.

We perform a series of experiments to determine what types of knowledge humans transfer in these game scenarios, which types of players transfer more or less information, and how per-

formance is affected by the transfer of learned knowledge. Each of the experiments is conducted in the context of UCT. Researchers argue that transfer of learning occurs primarily when previous knowledge is applied to solving a problem in a new situation [24]. This makes a gaming environment particularly appealing for analyzing human transfer learning.

The player's goal in a UCT scenario is to move through the environment, moving past barriers when possible, in order to find and diffuse an Improvised Explosive Device (IED). We divide the environment into a set of polyhedrons. Figure 9 shows the regions for the environment depicted in Figure 2. These regions may contain a variety of other elements such as buildings, vegetation, electrical hazards, and barriers, some of which can be overcome (see Figure 1).

Figure 1. The Urban Combat Testbed includes scenario elements such as IEDs (left), jumpable barriers (middle), and climbable barriers (right).

For our data collection, we asked volunteers to play a set of UCT scenarios. The volunteers represented a wide range of previous gaming experience and job backgrounds. The scenarios are grouped into eight levels, designed to demonstrate different amounts and types of knowledge transfer. Each of the eight levels contains five source scenarios (with the exception of Level 6, which has ten source scenarios) and five target scenarios. The general methodology is to compare the performance of two players on the target scenarios, one player having been trained by first playing the source scenarios, and one player having no training. The difference between source and target scenarios depends on the particular transfer learning level. Thorndike [33] hypothesized that the degree of transfer in humans depends upon the match between elements across the two contexts. While some effort has been spent on formalizing the notion of related-

ness between contexts or tasks [3], characterizing the amount of similarity and predicting expected transfer largely remains an open question. These levels were painstakingly defined by the DARPA Transfer Learning program to include a spectrum of near and far transfer elements [8]. Alternative approaches to characterizing levels of transfer have also been considered [20].

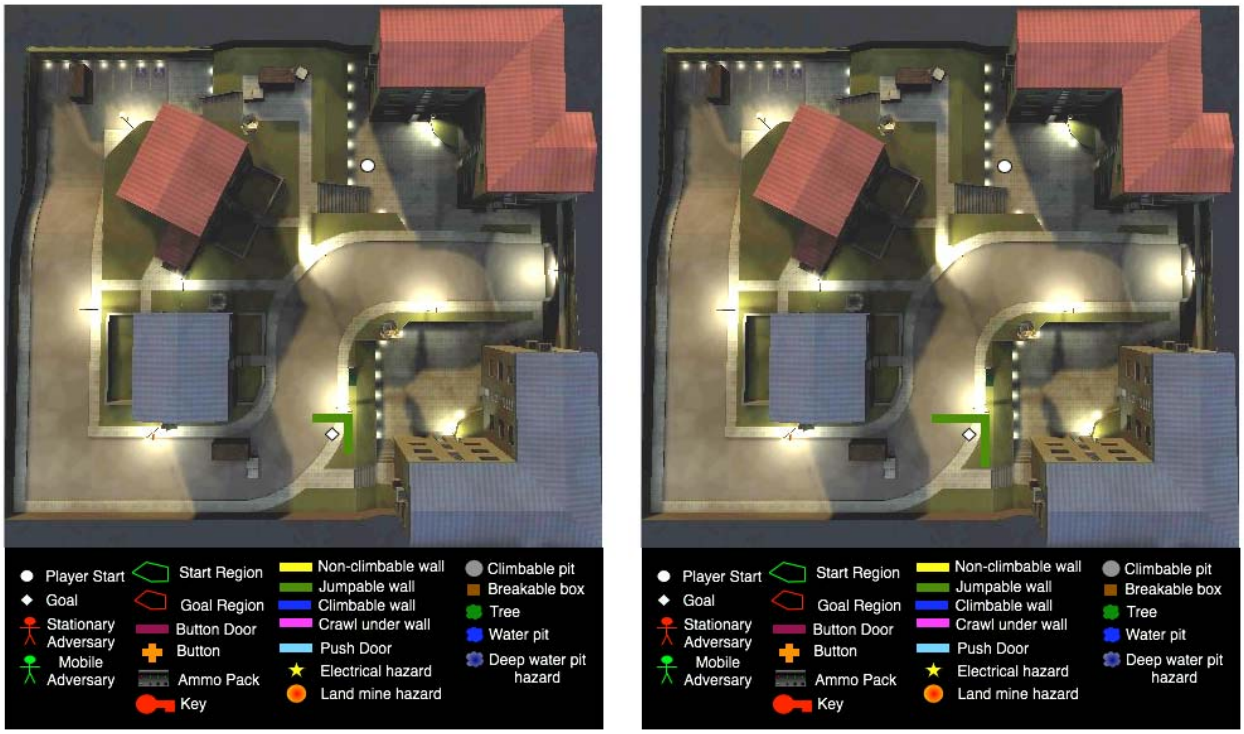

Figure 2. In these Level 1 scenarios, the player starts in region 29 (marked by a white oval). The goal of the source (left) and target (right) games is to find the IED in region 42 (marked by a white diamond). The scenario components are the same except that the jumpable barrier surrounding the goal location in the target scenario is longer than in the source scenario.

Transfer Learning Level 1: Parameterization. This is the simplest transfer level. Target scenarios in level 1 use the same components in the same configurations as in the source problems. The goal location (location of the IED) is in the same region as the goal location in the source problem, but may be in a different location within the region. Parameter values of scenario components may also vary between source and target. For example, Figure 2 shows a source / target scenario pair from level 1. The goal location is the same in both scenarios, but the jumpable barrier has an extended length in the target scenario. This type of change should re-

flect what is known as "near transfer" in human players [25]. For our experiments, 16 volunteers played the source scenario followed by the target scenario for each of 5 scenarios. An additional 17 volunteers played the 5 target scenarios without having played the corresponding source scenarios. We thus have data for (16*2*5) + (17*5) = 245 logged games.

Transfer Learning Level 2: Extrapolating. Like Level 1, the components are the same in source and target scenarios and the parameter values for the components may change from source to target. Unlike Level 1, the parameter value changes may force qualitative changes in a solution. For example, an insurmountable barrier may be extended from source to target in such a way that a different path to the goal must be considered. A total of 15 volunteers played all 5 source and target scenarios, while an additional 16 volunteers played just target scenarios, for a total of 230 games. The volunteers who played the games at Level 2 were not the same individuals who played the games at Level 1 – no volunteer played games at more than one level.

Transfer Learning Level 3: Restructuring. Level 3 forces players to piece together solution subsequences in different orders. Solutions to target problems reuse subsequences of actions found in solutions to source problems but in a different order. For example, one of the target scenarios specifies the same start and goal locations as the source. The target contains the same types of barriers as well, but the locations of the jumpable barrier and the insurmountable barrier have switched. 18 volunteers played all 5 source and target scenarios, and 19 volunteers played just the target scenarios, yielding a total of 275 logged games.

Transfer Learning Level 4: Extending. In this level, solutions to target problems require repeated application of elements found in solutions to source problems. As an example, Figure 3 shows a scenario in which the start location, goal location, and barrier elements are the same for source and target, but the coverage of the obstacles doubles from source to target. 16 volunteers

played all 5 source and target scenarios, and an additional 16 volunteers played just target scenarios, for a total of 240 logged games.

Figure 3. Example Level 4 source (left) and target (right) scenarios. The start and goal locations are the same, but the source scenario has two layers of obstacles (walls around the start and goal locations) while the target contains four layers of obstacles.

Transfer Learning Level 5: Restyling. In this level source and target scenarios will contain the same set and number of components, but they may be formulated differently. These changes require recognition of the commonalities by the player in order to transfer learned knowledge. For example, one of the target solutions is identical to the corresponding source solution except instead of obtaining two keys to open a door, the player must push two buttons to open the door. 21 volunteers played all 5 source and target scenarios, and 24 volunteers played just the target scenarios, yielding a total of 330 logged games.

Transfer Learning Level 6: Composing. Unlike the other levels, players at this level will receive training from two distinct source scenarios before playing the target scenario. In this level, target scenarios will contain all of the components found in the combination of the source scenarios. For example, the route from start to goal may be reused from one source to the target,

while the other source will train the individual on bypassing types of barriers that are also found in the target. 23 volunteers played two source games and one target for 5 scenarios, while another 23 volunteers played just the target scenarios, yielding a total of 460 logged games.

Transfer Learning Level 7: Abstracting. In this level, solutions to the target scenarios reuse the structure of solutions found in the corresponding source scenarios, but the details differ. In one source scenario, the player must use a key to open the front (steel) door of a building because bars over the window prevent entry there, and then must traverse through the first floor to reach the back courtyard where the goal is located. In the corresponding target scenario, players must first acquire ammunition for a pistol in order to shoot open the front (wood) door of the same building (there are bars over the window), and then must traverse the same first floor to reach the goal in the back courtyard. This is an example of "far transfer" [5] in which structural elements transfer between contexts, so the human player must abstract learned information to apply it to the new context. 23 volunteers played the source and target games for 5 scenarios, while another 23 volunteers played just the target games, yielding a total of 345 logged games.

Transfer Learning Level 8: Generalizing. At this level, solutions to target problems reuse actions found in solutions to the source problems, but players must apply them to different objects. In one scenario, the source game contains the goal IED surrounded by breakable boxes, while in the target game the IED is surrounded by breakable trees. 22 volunteers played all 5 source and target games, while an additional 24 volunteers played just the target games, yielding a total of 340 games. *We are thus analyzing captured data for a total of 2,465 games.*

We evaluate player performance based on a time-score metric [40]. Time is measured in seconds to completion of the goal, but 2 seconds are added for hitting minor hazards (e.g., glass), 30 seconds are added for hitting major hazards (e.g., landmines), and 2 seconds are added for each

ammunition round that is used. The goal for the player is thus to minimize the total score in number of seconds. This player performance provides a method to measure the amount of knowledge that is learned and transferred from the training (source) scenarios to the test (target) scenarios. While each individual is playing a game, their status is continuously sampled and stored to a file. The player location and status is recorded as <*x, y, z, yaw, pitch, roll, speed, time (msec), time (sec), health, shots, score, defuses*>. The game ends when the player defuses the IED (defuses = 1) or quits the game (defuses = 0). We are able to identify the region for any <x,y,z> location, which provides more abstract position information. The sequence of locations, time, and score represent the data we use to perform our analyses.

3. Analyzing Transfer Learning Curves

We are interested in analyzing the ability of humans to transfer game-playing knowledge in the context of UCT games. Singley and Anderson [29] measure the ability to transfer knowledge by calculating performance improvement on a new task, and state that the number of shared procedural elements between contexts is a predictor of the amount of transfer. We note that while the transfer levels are designed to test a particular type of transfer, the human players we are analyzing may in actuality employ a different type of transfer that also results in improved perform on target problems. Because of this inherent assessment difficulty, Greeno, et al. [14] advocate scrutinizing the actual strategy used on a new task for indications of influence from previous activities. We will apply both types of assessment in our analysis.

Quantitatively, human transfer learning can be observed through improved performance on a given task after the human gains experience with a similar task. This improved performance may be manifested in three ways, as shown in Figure 4 [8]: through improved initial performance on a new task (Type 1), a rate of performance improvement (Type 2), and an asymp-

totic/maximal achievable performance increase (Type 3). These three measurements are consistent with findings in education and cognitive psychology, where researchers argue for the existence of each of these types of transfer in humans [5]. In addition, some transfer influences may be negative. For example, the initial effect of transfer may be negative because previously-learned routines must be changed to deal with new settings. This is referred to as the "J curve effect" [5]. In our analysis, Type 1 performance improvement (a jump start from having addressed a similar problem in the past) is measured as the difference in the y-intercepts between a learning curve for a human that has received training on a similar problem and a human who has received no training. We measure Type 2 performance by calculating the area between the learning curves for trained and untrained humans. To measure Type 3 performance, we calculate the difference in y values between trained and untrained humans on the last instance of the new task.

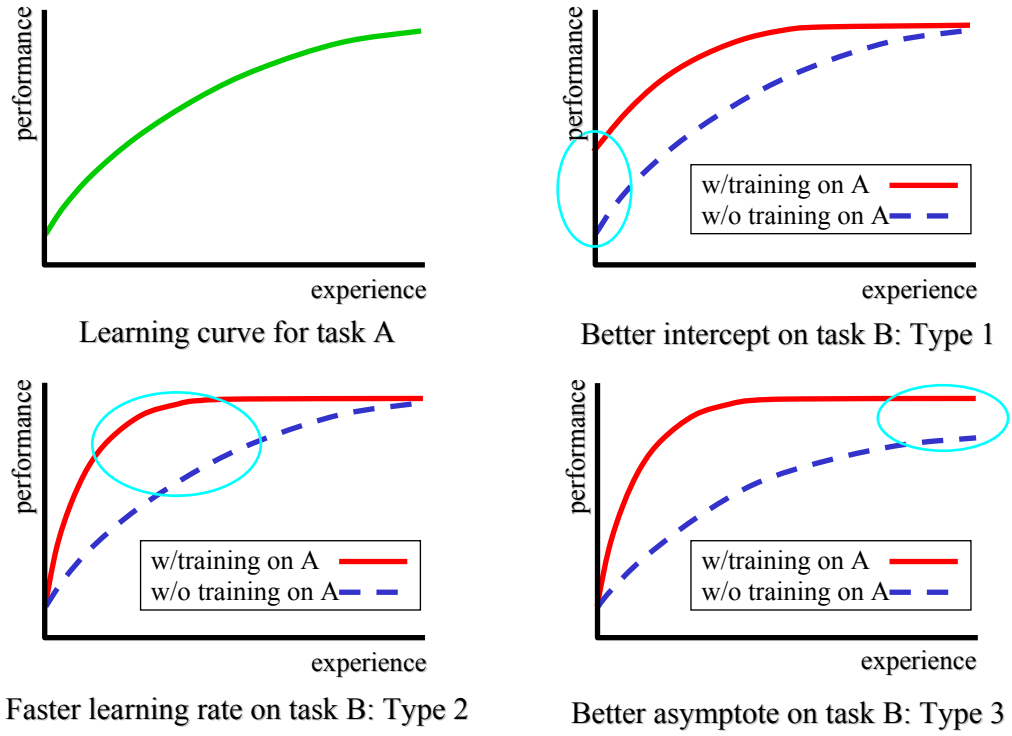

Figure 4. Three performance measures are used to determine the amount and type of observed transfer learning.

To measure the amount of Type 1, Type 2, and Type 3 transfer that occurs in human UCT players, we randomly pair human players for each transfer learning level. For each pairing, one of the humans receives training on a set of source scenarios and the other does not. Although the target scenarios are presented in random order between player pairs, the trained and untrained human within a pair see the target scenarios in the same order. We then calculate the average of the differences in performance among all the player pairs.

Transfer Learning Level		**TL Metrics**		
		Jump start (Type 1)	**Transfer difference (Type 2)**	**Asymptotic advantage (Type 3)**
Level 1	Score	-0.65	163.09	8.94
	p	0.54	0.00	0.11
Level 2	Score	-42.33	104.67	34.40
	p	0.81	0.12	0.09
Level 3	Score	-14.95	193.97	40.42
	p	0.69	0.12	0.09
Level 4	Score	104.75	264.97	108.88
	p	0.00	0.07	0.00
Level 5	Score	121.63	843.42	79.37
	p	0.08	0.00	0.01
Level 6	Score	86.39	709.07	175.74
	p	0.08	0.02	0.00
Level 7	Score	-21.48	207.41	9.22
	p	0.71	0.00	0.37
Level 8	Score	41.24	209.95	87.48
	p	0.24	0.03	0.02

Figure 5. Amount of human Type 1, Type 2, and Type 3 transfer for each UCT transfer learning level. Differences in performance between trained and untrained players are averaged over all player pairs, and significance values are calculated for the set of differences.

The results of this analysis are shown in Figure 5, with p-values highlighting the statistical significance between trained and untrained player performance. There are a few observations to make. First, the amount of jump start transfer (Type 1), particularly in the lower levels, is sometimes negative. This is due to the fact that human players assigned to scenarios randomly. In many cases a player with no gaming experience received training on source scenarios, while a

player with extensive experience received none. Even though the player being trained may witness a performance improvement from source to target games, the difference in performance on target games between the two humans may be negative. The expert has in a sense already received training from many similar gaming scenarios in the past. This behavior is only observed with Type 1 transfer. The average difference in Type 2 performance (area between the two curves) and Type 3 performance is always positive, showing that trained human players do learn strategies for each transfer learning level faster, and by the time the last scenario is played human players benefit from training on similar problems.

In addition, note that the amount of transfer tends to increase, particularly through level 6. This is because the scenarios in these levels are increasing in complexity. As a result, the players benefit even more from training on similar problems. The Type 2 transfer difference p-value for five of the eight transfer learning levels is less than 0.05, lending strength to the hypothesis that the human players are indeed transferring expected components from source to target games.

Figure 6 plots the performance of trained and untrained players for each level. Because the scenarios are presented in random order, the trends reveal the amount of increasing ability with each scenario that is played, rather than indicating the difficulty of a particular scenario. Performance is plotted as the negative of the time-score, so a line higher on the graph indicates favorable performance. The plots show that while performance may vary with the first scenario, trained players gain more in performance with each scenario than untrained players, and the total performance for trained players is better across all levels than for those who have not received training, and therefore cannot transfer learned knowledge from source to target scenarios.

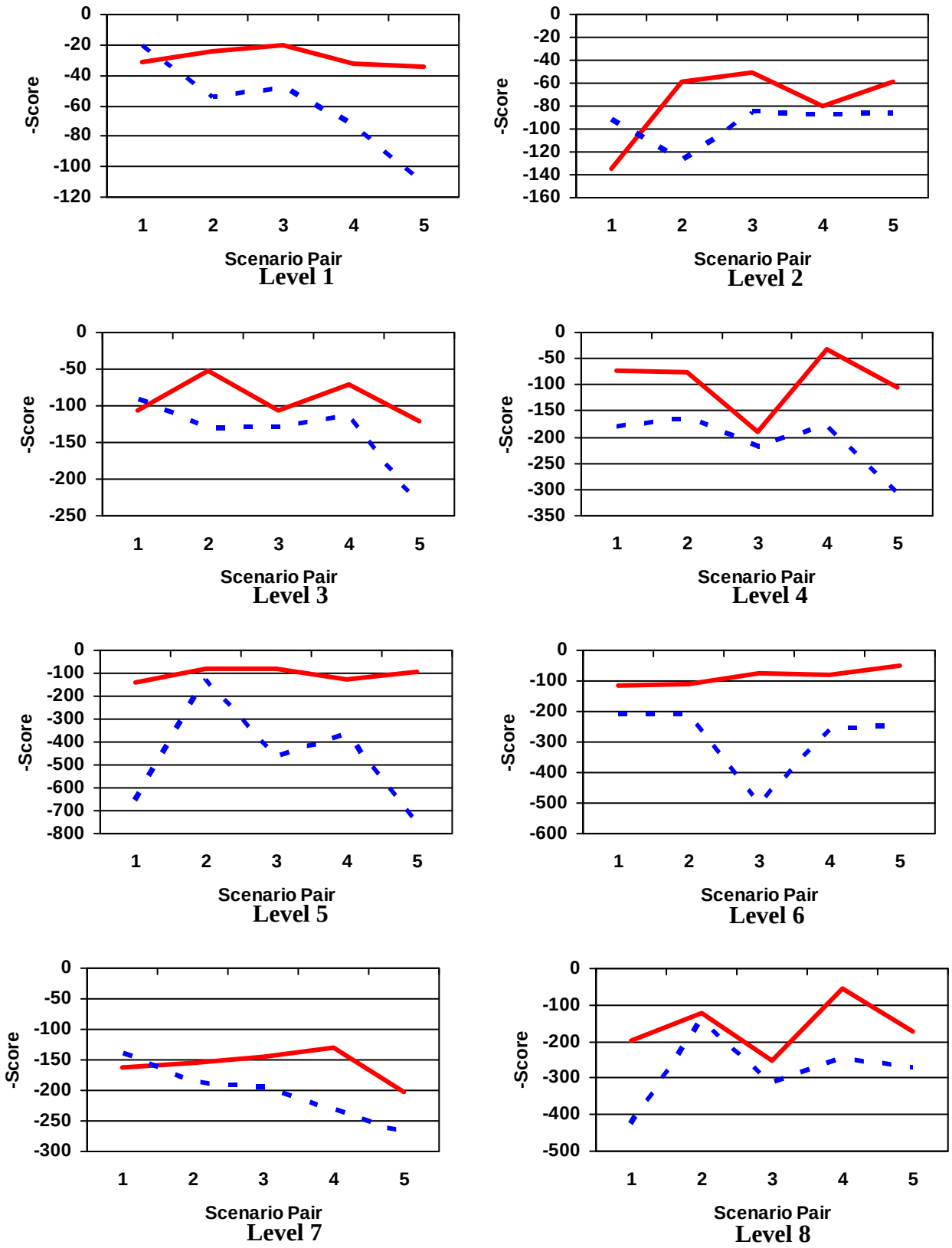

Figure 6. Plots of performance for trained (solid red lines) and untrained (dashed blue lines) players for each transfer learning level. The negative time scores are averaged over the player pairs and are plotted for the sequence of five scenario pairs that are presented.

Now that we have ascertained transfer is in fact occurring in human game playing, we will take a closer look at the type of knowledge that is transferred from source to target scenarios, according to transfer learning level as well as player experience. This type of analysis is consistent with the approach recommended by Singley and Anderson [29] for ascertaining if humans have transferred learned information to a new context. For our analysis, we make use of a graph-based relational learner. We describe the learning tool and analysis of player data in the next sections.

4. Substructure Discovery

The goal of our analysis is to determine what types of knowledge humans transfer in the context of the Urban Combat Testbed. This analysis can provide insight on human learning and reasoning. In addition, the findings may help guide the creation and assessment of automated approaches to learning and transferring knowledge in similar gaming and military environments.

There are a number of data mining algorithms that could be useful in analyzing this data. Because we want to identify common patterns in the data, supervised learning algorithms will not be sufficient for this task. In addition, the data is inherently structural. While individual player locations may be interesting, it is the sequence of locations, and patterns formed by these sequences, that we anticipate will provide the greatest insights on player strategies and transferred knowledge. Thus we need a structural discovery, or graph-based relational learning, approach.

Graph-based data mining is the task of finding novel, useful, and understandable graph-theoretic patterns in a graph representation of data. Several approaches to graph-based data mining identify frequently-occurring subgraphs in graph transactions, i.e., those subgraphs meeting a minimum level of support [18, 19, 23, 37].

We distinguish graph-based relational learning (GBRL) from graph-based data mining in that GBRL focuses on identifying novel, but not necessarily most frequent, patterns in a graph representation of data [16]. Only a few GBRL approaches have been developed to date. Two specific approaches, Subdue [7] and GBI [38], take a greedy approach to finding subgraphs maximizing an information-theoretic measure. Subdue searches the space of subgraphs by extending candidate subgraphs by one edge. Each candidate is evaluated using a minimum description length (MDL) metric [27], which measures how well the subgraph compresses the input graph if each instance of the subgraph were replaced by a single vertex. GBI continually compresses the input graph by identifying frequent triples of vertices, some of which may represent previously-compressed portions of the input graph. Candidate triples are evaluated using a measure similar to information gain. Kernel-based methods have also been used for supervised GBRL [11]. Because we want to be able to perform unsupervised discovery as well as supervised learning on our structural data, and because we want to be able to easily interpret the results, we use the Subdue algorithm for this analysis.

The Subdue algorithm [7] encompasses several approaches to graph-based learning, including discovery, clustering and supervised learning. Subdue uses a labeled graph $G = (V,E,L)$ as both input and output, where $V = \{v_1, v_2, \ldots, v_n\}$ is a set of vertices, $E = \{(v_i, v_j) \mid v_i, v_j \in V\}$ is a set of edges, and L is a set of labels that can appear on vertices and edges. The graph G can contain directed edges, undirected edges, self-edges (i.e., $(v_i; v_i) \in E$), and multi-edges (i.e., more than one edge between vertices v_i and v_j). The input graph G need not be connected, but the learned patterns must be connected subgraphs (called substructures) of G.

Unsupervised Discovery. Inputs to Subdue's discovery algorithm include the input graph (or a set of graphs), the beam length, and a limit on the total number of substructures considered by

the algorithm. Subdue searches for a substructure that best compresses the input graph. A substructure in Subdue consists of a subgraph definition and all its occurrences throughout the graph. The initial state of the search is the set of substructures consisting of all uniquely labeled vertices. The only operator of the search is the *ExtendSubstructure* operator, which extends a substructure in all possible ways by a single edge or an edge and a neighboring vertex.

Subdue uses a beam search to identify candidate substructure concepts by applying the *Extend-Substructure* operator to each substructure in the current state. The substructures are kept on a beam-limited queue and are ordered based on their description length (sometimes referred to as value) as calculated using the MDL principle.

The search terminates upon reaching a user-specified limit on the number of substructures extended, or upon exhaustion of the search space. Once the search terminates and Subdue returns the list of best substructures found, the graph can be compressed using the best substructure. The compression procedure replaces all instances of the substructure in the input graph by single vertices, which represent the substructure definition. Incoming and outgoing edges to and from the replaced instances will point to, or originate from the new vertex that represents the instance. The Subdue algorithm can be invoked again on this compressed graph. This procedure can be repeated a user-specified number of times, and is referred to as an *iteration*.

Subdue's search is guided by the minimum description length (MDL) [27] principle. The evaluation heuristic based on the MDL principle assumes that the best substructure is the one that minimizes the description length of the input graph when compressed by the substructure. The description length of the substructure S given the input graph G is calculated as $DL(S) + DL(G|S)$, where $DL(S)$ is the description length of the substructure, and $DL(G|S)$ is the description length of the input graph compressed by the substructure. Description length is calculated as

the number of bits in a minimal encoding of the graph. Subdue seeks a substructure S that maximizes compression, calculated as $Compression = \frac{DL(G)}{DL(S) + DL(G \mid S)}$.

As an example, Figure 7 shows the four instances that Subdue discovers of a pattern S_1 in the example input graph and the resulting compressed graph, as well as the pattern S_2 found in this new graph and the resulting compressed graph. To allow slight variations between instances of a discovered pattern (as is the case in Figure 7), Subdue applies an error-tolerant graph match between the substructure definition and potential instances.

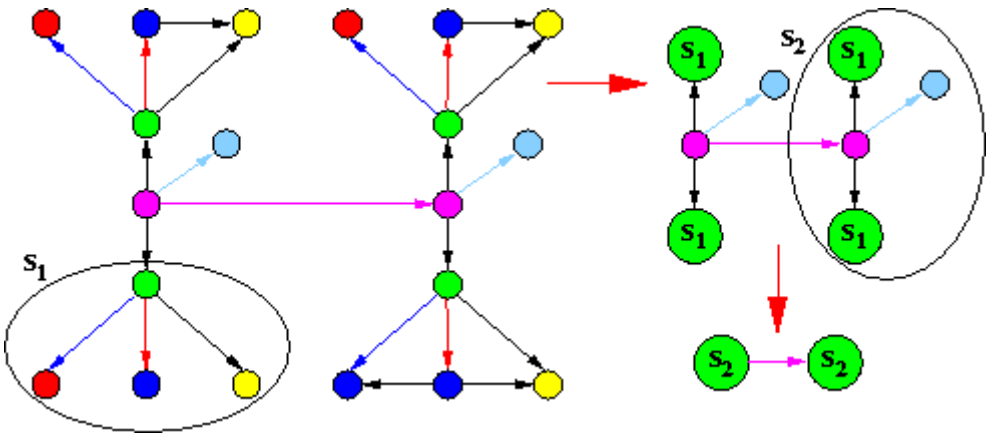

Figure 7. Example of Subdue's unsupervised discovery algorithm. A repetitive subgraph (S_1) is identified and used to compress the graph. New discoveries (in this case, pattern S_2) are made in subsequent iterations of the algorithm.

Supervised Learning from Training Graphs. Extending a graph-based discovery algorithm to perform supervised learning introduces the need to handle negative examples (focusing on the two-class scenario). The negative information can come in two forms. First, the data may be in the form of numerous small graphs, or graph transactions, each labeled either positive or negative. Second, data may be composed of two large graphs: one positive and one negative.

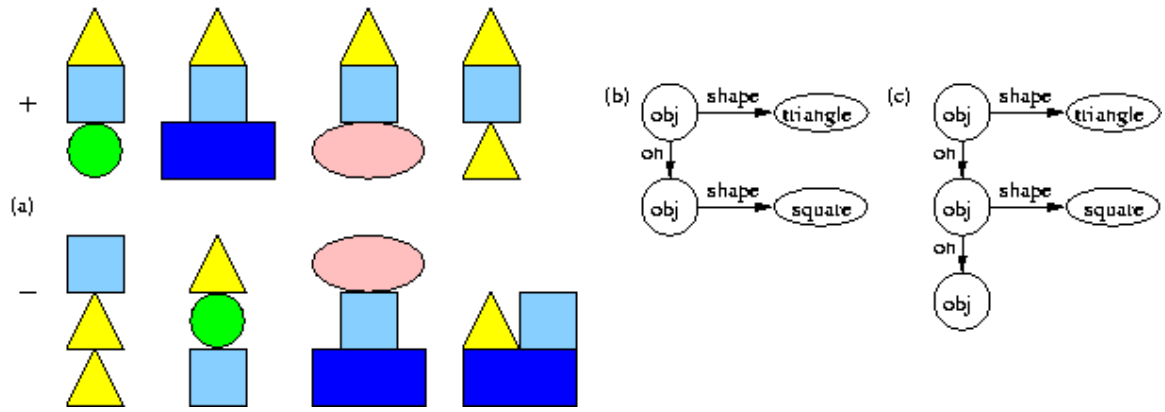

Figure 8. Visualization of graph-based data with (a) four positive and four negative examples, and (b,c) two possible graph concepts learned from the examples. Graphs that contain the learned concept are labeled as examples of the positive class, other graphs are members of the negative class.

The first scenario is closest to the standard supervised learning problem in that we have a set of clearly-defined examples. Figure 8 depicts a set of positive (G^+) and negative (G^-) examples. One approach to supervised learning is to find a subgraph that appears in many positive graphs, but in few negative graphs. This amounts to replacing the compression-based measure with an error-based measure. For example, we would find a subgraph S that minimizes the value $\frac{|\{g \in G^+ \mid S \not\subseteq g\}| + |g \in G^- \mid S \subseteq g\}|}{|G^+| + |G^-|} = \frac{FN + FP}{P + N}$, where $S \subseteq g$ means S is isomorphic to a subgraph of g (although we do not need to perform a subgraph isomorphism test during learning). The first term of the numerator is the number of false negatives, and the second term is the number of false positives.

This approach will lead the search toward a small subgraph that discriminates well, e.g., the subgraph shown in Figure 8b. However, such a subgraph does not necessarily compress well, nor represent a characteristic description of the target concept. We can bias the search toward a characteristic description by using the compression-based measure to look for a subgraph that compresses the positive examples, but not the negative examples. If *DL(G)* represents the description length (in bits) of the graph *G*, and *DL(G|S)* represents the description length of *G* compressed

by subgraph *S*, then we look for an S that minimizes $DL(G^{+}|S)+DL(S)+DL(G^{-})-DL(G^{-}|S)$, where the last two terms represent the portion of the negative graph incorrectly compressed by the subgraph. This approach will lead the search toward a larger subgraph that characterizes the positive examples, but not the negative examples, e.g., the subgraph shown in Figure 8c.

Finally, this process can be iterated in a set-covering approach to learn a disjunctive hypothesis. Using the error measure, any positive example containing the learned subgraph would be removed from subsequent iterations. Using the compression-based measure, instances of the learned subgraph in both the positive and negative examples (even multiple instances per example) are compressed to a single vertex.

5. Analysis of Patterns

For our analyses of human transfer learning data, we represent logged game information as a graph and use Subdue to identify common patterns. The raw player data is recorded using actual <x,y,z> locations. To facilitate the identification of more abstract patterns, we map the 3D location into a region of the scenario map. A volumetric map of the scenario shown in Figure 2 is given in Figure 9 (left). The figure indicates how the scenario is divided into labeled regions.

The input to Subdue is a labeled graph. Every region that is visited during the game is indicated by a vertex in the graph, labeled with the region ID number. Two vertices are connected with a directed edge labeled "next" when the player moves from the source region to the destination region. Only one vertex is created for each visited region. This way, if a player visits a region more than once, the loop will be evidenced as a cycle in the graph representation. Applied to interactive player-based environments, the resulting graph of captured spatial interaction is called a player-graph [40].

For our first analysis, we fed data from the entire collection of games to Subdue. Our goal was to identify player movement patterns that were common among all scenarios and players. The most prevalent pattern (referred to as SUB_1) is the sequence of regions highlighted with dashed lines in Figure 9, 102→96→86→62→75→47→35. As the corresponding picture in Figure 2 shows, this pattern represents a player moving down the sidewalk between two buildings. All of the scenarios contain the same buildings, sidewalks, and roads, so this movement pattern will be common for all players who are exploring the space or using the simplest path to move across the space. Subdue found other patterns of this type as well. All of the most common patterns reflected movement patterns across the main sidewalks/roadways in the game environments.

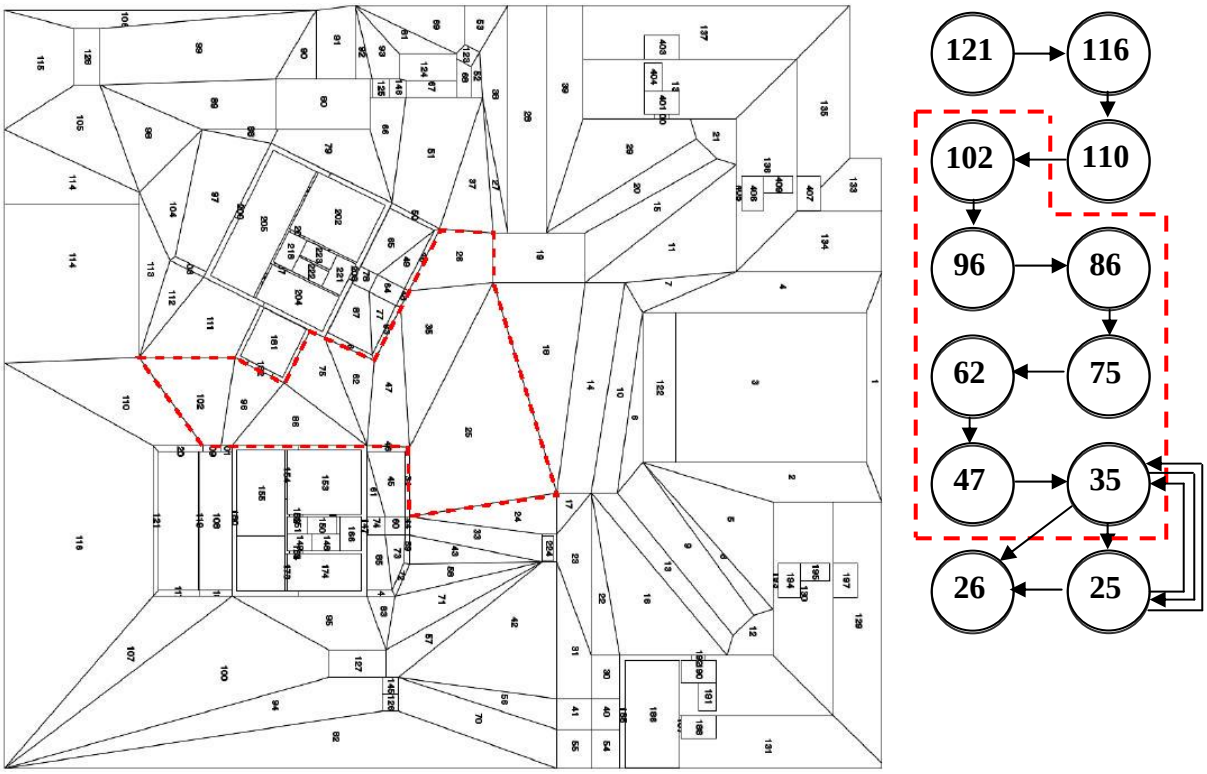

Figure 9. Volumetric map of Scenario 1 for Level 1 (left). The graph representation for a portion of a logged game is also shown (right). A region in the scenario that a player visits is represented by a node in the graph, and the order of the visited regions is indicated by directed edges between the nodes, all labeled "next".

In later iterations of Subdue, a pattern that emerges is SUB_1 connected to itself multiple times. SUB_1 refers to the first substructure, described above. After the first iteration, Subdue

compresses the graph by replacing every instance of the substructure by a single vertex labeled SUB_1. This discovered pattern thus represents a player exploring the main road and returning several times to already-visited regions on the main road. For players who are unfamiliar with the game, the scenario, or even the specific goal of a given scenario, this type of exploring will be necessary to find the goal, to identify the obstacles, and to find paths that move the player beyond the obstacles.

This initial analysis provided several insights. First, we were able to verify that the region graph representation adequately represented the game logs and that the patterns could be interpreted in a straightforward manner. Second, we were able to identify some activities common to all players, trained and untrained. The main goal of this work, however, is to identify knowledge that players transfer from one game to another. We focus on this task in the next section.

6. Global Analysis of Knowledge Transfer

As a first step in identifying transferred knowledge, we use Subdue to learn a concept that distinguishes game logs of players who were trained on source games (trained players) from game logs of players who were not trained on source games. Only target games are used in this analysis, and logs from all scenarios at all levels are included. For this analysis, we use the MDL-based supervised learning evaluation measure, because this provided the most insightful results.

We initially allowed the "trained" games to represent the positive examples of the concept and "untrained" games to represent the negative examples. However, very few high-value patterns emerged. This is because trained players will generally not execute many extra movements, but will take the most efficient path to the goal.

To test this theory, we ran Subdue again as a supervised learner, this time labeling the untrained games as positive examples and trained games as negative examples. As anticipated, this

analysis resulted in higher-valued patterns. The initial top concept learned by Subdue (SUB_1) is highlighted in Figure 10, which is a simple path on the main road. The next most-interesting substructure (SUB_2) contains a simple path down another section of the main road. Interestingly, however, once the graph is compressed using these patterns, the next concepts that emerge are highly cyclic. Of the top twenty patterns that were discovered, almost half included multiple occurrences of SUB_1 or SUB_2, the top two patterns. The third highest-value pattern contained 16 occurrences of SUB_1 connected to itself and to a small sequence of other regions, and another top-twenty pattern contained 24 occurrences of SUB_1 connected to itself. These are very large action sequence patterns. Similarly, other top patterns contained 8 or more occurrences of SUB_2 connected to itself. Because a majority of the actions in the sequence are repetitive (e.g., moving along the main road 16 times), we infer that these are exploratory sequences in which the player is becoming familiar with the surroundings and looking for the goal IED.

What is interesting about these findings is that while both the trained players and the untrained players contain the SUB_1 and SUB_2 sequences, the untrained players have more repetitive occurrences of these sequences connected to themselves. This indicates that untrained players cycle through regions of space looking for a pathway or looking for the goal itself. Because trained players transfer key pieces of knowledge, these cycles do not exist in their graphs. This analysis indicates that in general, untrained players perform more repetitive, exploratory movements in their games. This finding helps to validate the hypothesis set forward by Greeno, et al. [14] that during initial learning humans may acquire an action schema that is responsive to an action opportunity, and will apply this schema in a new situation. Applying the schema of moving directly down a main path toward the goal is a common theme in our captured data, and in turn accounts for the overall improved performance of trained players over untrained players. These re-

sults support the hypothesis that *humans are able to transfer learned game strategy that reduces repetitive movement patterns in new, previously-unplayed scenarios.*

Figure 10. Visualization of learned patterns. The top sequence is indicated with the solid arrow moving left and down the path, while the next highest-valued sequence is indicated with the dashed arrow moving down from the building then right to the edge of the environment. The top substructure occurred in 107 untrained and 51 trained games, while the second substructure occurred in 140 untrained and 77 trained games.

7. Detailed Analysis of Knowledge Transfer

Next, we perform a level-by-level analysis of knowledge transfer. For each transfer knowledge level, we use Subdue to learn concepts that distinguish trained from untrained players, letting the trained games initially represent positive examples, then repeat the process with untrained players representing positive examples. Because there are many results from each piece of the analysis, we only include the highlights of the results in this discussion.

Results from analysis of Level 1 are shown in Figure 11. The patterns in the left figure are common among trained players and not among untrained, while the patterns in the right are common for untrained players. Recall that Level 1 tests the ability of humans to parameterize

game information. These results indicate that trained players take a more direct route to the goal, while the untrained players cycle through a series of regions seeking a way to reach the goal.

These results are similar to those for the global analysis. The differences between source and target scenarios for the first several levels are fairly minor, so the discovered patterns generally reflect the same principles. For example, the learned concept for trained players at Level 2 contains a straightforward sequence of regions, while the learned concept for untrained players is a substructure containing 32 vertices and 53 edges (indicative of a large number of cycles). The next three highest-ranked concepts for untrained players contains (10 vertices, 31 edges), (79 vertices, 119 edges), and (10 vertices, 43 edges), respectively, giving additional evidence to support this hypothesis. Similar results can be found for Level 3. The results for these three levels

Figure 11. Visualization of learned patterns for Transfer Learning Level 1. The picture on the left highlights (with arrows) the learned concept from trained players, and the picture on the right highlights the learned concept from untrained players.

Unlike the earlier three levels, the learned concept for Level 4 trained games does involve some cycles. Upon closer inspection of the scenarios, this makes sense. Some of the scenarios at this level require the player to push buttons that open a door along the path to the goal. Play-

ers need to maneuver a bit to find and push the buttons, which is consistent with the patterns that are learned by Subdue. In contrast, the cycles that are found in the untrained games cover much larger regions, which is consistent with the exploration strategy that must be employed when the players do not have knowledge to transfer. *From these experiments we conclude that humans are able to parameterize, extrapolate, restructure, and extend knowledge from one game to another new game scenario.* We see from the data that the exploration required to find key elements (e.g., the buttons to push or the obstacles to climb) is eliminated from source to target games. These players are thus demonstrating an "extending" type of knowledge transfer. However, some of the details in handling these challenges (e.g., manipulating the button itself) is not perfectly learned after one source game and requires additional training.

In many of the Level 5 scenarios, players must maneuver around a building to find the IED. Target scenarios often place the IED in a similar setting but in a different building, which tests the ability of humans to reuse learned plans in isomorphic settings. Once the mapping between source and target buildings is identified, the source plan can be reused in the new building.

Figure 12. Visualization of a learned pattern for games played by trained players at Level 5 (left, highlighted by the arrow) and the area inside a scenario building that precipitates exploration loops for both trained and untrained players (right).

The concepts that are learned by Subdue reveal some interesting insights. The trained players utilize fewer cycles around the environment and building to find the IED. This lends evidence to the hypothesis that *human players are adept at restyling solutions to a new layout*. However, the learned patterns also indicate that even trained players perform initial exploration in the new building before moving to the goal. *We hypothesize that this initial exploration is needed to build the mapping between the building layout in the source scenario and the isomorphic layout in the target scenario.* The source and target buildings look similar, and the initial exploration is useful to determine their location in the new building. This activity likely reinforces the learned mapping and reduces subsequent activity in the target scenarios. After the initial exploration, the number of cycles identified by Subdue in the trained games is fewer than in the untrained games. Figure 12 visualizes one of these learned patterns and shows the region inside the building that is traversed multiple times by trained and untrained players.

The higher transfer learning levels tend to place increasingly-great transfer demands on the player. However, Subdue easily found patterns to distinguish between trained and untrained human players at Level 6. Untrained players continued to explore in several loops (the leading pattern contains 55 vertices and 90 edges), while trained players moved directly to the goal (the average pattern size is 4 vertices and 3 edges). This may be due to the fact that trained players at this level received two source scenarios. The solutions to the target scenarios represent concatenations of the solutions to the two corresponding source scenarios. The dramatic difference in patterns and in performance between trained and untrained players at Level 6 is consistent with theories espoused by Gick and Holyoak [12] and by Bjork and Richardson-Klavhen [4] that transfer across contexts is more likely to occur when a subject is taught in multiple contexts rather than only a single context.

The example shown in Figure 13 can be used to distinguish trained from untrained players. In this scenario, the goal is behind a climbable fence. All of the untrained players roam about the space before climbing the fence. In one case, a player saw the IED behind the fence yet initially tried to break into the building to explore it before finally jumping the fence and reaching the IED. In contrast, the trained players already climbed a fence to solve a similar problem, and so quickly use this solution for the target problem. The player data and Subdue-discovered patterns support the hypothesis that *humans can transfer composing knowledge between games*. This knowledge helps them to combine subproblem solutions into a solution to a new target game.

Figure 13. Visualization of a target scenario for Level 6 (left) and a picture of the IED in the target scenario, highlighted in white, that is located behind a climbable fence (right).

Subdue had a more difficult time finding strong (high-valued) concepts in Levels 7 and 8 than in the lower levels. This may be due to the fact that the knowledge to be transferred in these levels is more abstract. Hayes and Simon [15] have observed that subjects usually do not find deep structure similarity unless the relationship is explicitly shown, and we observe the same phenomenon in this analysis. Another interesting observation is that while patterns found in the untrained games do involve many cycles as before, more of the patterns found in the trained games also include cycles. Despite the overall lack of prevailing patterns, there are a few pieces of evi-

dence that indicate the human players are continuing to transfer pieces of learned knowledge. As an example, one of the learned patterns for trained games, shown in Figure 14, is found in five of the games. This makes sense, because one of the five scenarios at Level 7 requires the human to use the key (highlighted with a circle in Figure 14) to open the door to the building on the right. No instances of this pattern are found in the untrained games. The untrained players either time out or eventually use the key as part of a lengthier and more complex task solution.

Figure 14. Visualization of learned concept for trained games at Level 7. The player grabs the key (circled in the figure) and uses it to open the door to the right.

8. Analysis of Knowledge Transfer by Expertise Level

Before each volunteer played their assigned games, we asked them to provide information about themselves. This information included the amount of experience and expertise playing Quake, playing FPS games, and playing any computer games. We hypothesize that *humans will transfer different amounts and different types of knowledge based on their expertise on the task.*

To validate this hypothesis, we partitioned players according to their expertise. If they had some experience with Quake and FPS, they were partitioned into the expert category. Individu-

als with little or no Quake experience but at least a fair rating on FPS games were partitioned into the intermediate category, and individuals with little or no experience with Quake or FPS games were partitioned into the beginner category.

We measured amount of transfer as improvement in the final score from the source to the target game. As expected, beginners and intermediate players saw greater improvement from source to target games (12.24 seconds on average, based on the time-score metric described earlier) than experienced players (8.24 seconds on average). Experienced players have in essence been trained on many games in the past and can transfer abstract strategies to these new scenarios, resulting in less improvement from additional training games.

In order to analyze the type of knowledge that is transferred, we need to perform supervised learning on the three classes. We accomplish this with a series of binary concept learning tasks, distinguishing beginner from intermediate or expert, followed by distinguishing beginner or intermediate from expert. The highest-valued concepts resulted from learning patterns that distinguish beginners from intermediate or expert players. Interestingly, the learned patterns typically involved many cycles of regions. This is surprising, because we are only analyzing games for players who received training on source games. With our earlier analyses, such cycles primarily appeared in untrained games.

Upon closer inspection, we see that these cycles do not reflect large movements throughout the environment. Instead, they are cycles back and forth between small regions that correspond to picking up a key and opening a door, finding and pushing a button, breaking boxes to reach the goal, and other fine-grained tasks. Many of these tasks differed between source and target scenarios. Experienced players would be familiar with these manipulations from earlier games, but beginners need to reposition the player several times in order to perform the more subtle func-

tions required by some of the scenarios. A similar case occurs for Scenario 2 in Level 8, where the player must learn to jump over the hazard (a pool of water) to reach the goal. The beginner players spend much time trying to swim across the body of water before they attempt to jump over the hazard. The visualized pattern and hazard are shown in Figure 15. Therefore, players with less gaming learn more knowledge in the source games that can be transferred to the target games; whereas the players with gaming experience already possess this knowledge.

Figure 15. Visualization of a learned concept (left) and the corresponding water hazard the player is learning to overcome (right) in Level 8 scenarios.

9. Conclusions

In this study we analyze the amount of learned information that is transferred to new gaming scenarios by human players. To more closely analyze the type of knowledge that is transferred, we use graph-based relational learning to identify patterns of transfer learning in humans. Using data collected from logged Urban Combat games, we recognized common movement strategies. We also identified patterns that distinguished players who were trained on similar games to those who were playing target games with no prior training. From the evidence that was produced by the graph-based relational learner, we conclude that humans are able to perform transfer of learned knowledge at multiple levels of complexity. We observe that beginners benefit the most

from training in the game scenarios, and as a result transfer the greatest amount of learned knowledge to new situations.

There are several influential factors in our analysis that warrant greater attention in future experiments. Due to the small sample size that was available for each of the many scenarios, we randomly paired human players. Future experiments could control for differences in prior experience by pairing players with similar gaming expertise to provide a clearer assessment of performance that is due to transfer. Future work is also needed to more precisely quantify the similarity between source and target scenarios, and to use this as a basis for determining the expected amount of transfer between scenarios. Barnett and Ceci's taxonomy of far transfer [2] includes a temporal context dimension, and they argue that depending on the task the ability to transfer learned knowledge may peak immediately after training or may peak much later. This is an aspect that would be interesting to analyze as part of our future work.

Results from this study indicate that data mining algorithms, such as the Subdue graph-based relational learning system, can be used to effectively analyze human behavior and capabilities. We would like to extend this analysis to human transfer learning on a greater range of tasks. We anticipate that this analysis will also form the basis for analyzing the ability of artificial agents to perform transfer learning, and to compare strategies for task completion as well as transfer of learned task strategies between artificial agents and biological agents.

The patterns we found in the human-based results can be used to instill the same patterns in artificial agents in order to improve their human consistency. This is important for providing both challenging opponents in games and as the basis of realistic simulation. For example, if we want to simulate an artificial agent finding IEDs, we will want the agent and the simulation to be con-

sistent with how we observe humans performing such a task in the real world. This is an important direction for continued research.

Acknowledgements

This material is based on research sponsored by DARPA under agreement number FA8750-05-2-0283. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA or the U.S. Government.

References

1. M. Asadi, V.N. Papudesi, & M. Huber. Learning skill and representation hierarchies for effective control knowledge transfer. *Proceedings of the ICML Workshop on Structural Knowledge Transfer for Machine Learning,* 2006.

2. S.M. Barnett & S.J.Ceci. When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4):612-637, 2002.

3. S. Ben-David & R. Schuller. Exploiting task relatedness for multiple task learning. *Proceedings of the Conference on Learning Theory*, pages 567-580, 2003.

4. R.A. Bjork & A. Richardson-Klavhen. On the puzzling relationship between environment context and human memory. In Current Issues in Cognitive Processes: The Tulane Flowerree Symposium on Cognition, Elrbaum, 1989.

5. J.D. Bransford, A.L. Brown, & Rodney R. Cocking, editors. How people learn: Brain, mind, experience, and school. The National Academy of Sciences, 1999.

6. J.P. Byrnes. Cognitive Development in Instructional Contexts. Allyn & Bacon, 1996.

7. D. Cook & L. Holder. Graph-based data mining. *IEEE Intelligent Systems*, 15(2):32-41, 2000.

8. DARPA. Transfer Learning (BAA 05-29).

9. D. Detterman & R. Sternberg, editors. Transfer on Trial. Ablex Publishing, 1993.

10. C. Fidoplastis, C. Stapleton, J. Whiteside, C. Hughes, s. flore, G. Martin, J. Rolland, & E. Smith. Human experience modeler: Context-driven cognitive retraining to facilitate transfer of leraning. *CyberPsychology and Behavior*, 9(2):183-187, 2006.

11. T. Gartner, Kernel methods for mining graph data. In *Mining Graph Data* (D. Cook & L. Holder, eds.), Wiley, 2006.

12. M.L. Gick & K.J. Holyoak. Scheme induction and analogical transfer. *Cognitive Psychology*, 15:1-38, 1983.

13. N. Gorski & J. Laird. Experiments in transfer across multiple learning mechanisms. *Proceedings of the ICML Workshop on Structural Knowledge Transfer for Machine Learning*, 2006.

14. J.G. Greeno, D.R. Smith, & J.L. Moore. Transfer of situated learning. In D. Detterman & R. Sternberg (eds.), Transfer on Trial, Ablex Publishing, 1993.

15. J.R. Hayes & H.A. Simon. Psychological differences among problem isomorphs. In N.J. Castellan, D.B. Pisoni, & G.R. Potts (eds.), Cognitive Theory, Erlbaum, 1977.

16. L. Holder & D. Cook. Graph-based relational learning: Current and future directions. ACM SIGKDD Explorations, 5(1):90-93, 2003.

17. E. Holton, III & T. Baldsin. *Improving Learning Transfer in Organizations.* Wiley, 2003.

18. A. Inokuchi, T. Washio, & H. Motoda. Complete mining of frequent patterns from graphs: Mining graph data. *Machine Learning*, 50(3):321-354, 2003.

19. M. Kuramochi & G. Karypis. Finding frequent patterns in a large sparse graph. *Data Mining and Knowledge Discovery*, 2005.

20. N. Love. Transfer learning evaluation with relational nets. *Proceedings of the ICML Workshop on Structural Knowledge Transfer for Machine Learning*, 2006.

21. Z. Marx, M.T. Rosenstein, & L.P. Kaelbling. Transfer learning with an ensemble of background tasks. *Proceedings of the NIPS Workshop on Inductive Transfer: 10 Years Later*, 2005.

22. A. Niculescu-Mizil & R. Caruana. Learning the structure of related tasks. *Proceedings of the NIPS Workshop on Inductive Transfer: 10 Years Later*, 2005.

23. S. Nijssen & J. Kok. A quickstart in frequent structure mining can made a difference. *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, 2004.

24. J.E. Ormrod. Human Learning. Pearson Education, 2004.

25. D.N. Perkins & G. Salomon. Transfer of learning. In *International Encyclopedia of Education*, Pergamon Press, 1992.

26. Quake III Arena. http://www.idsoftware.com/games/quake/quake3-arena.

27. J. Rissanen. Stochastic Complexity in Statistical Inquiry. World Scientific, 1989.

28. H. Simon. Why should machines learn? In Machine Learning: An Artificial Intelligence Approach (R. Michalski, J. Carbonell, & T. Mitchell, eds.), Tioga Publishing, pages 25-37, 1983.

29. Singley & Anderson. The Transfer of Cognitive Skill. Harvard University Press, 1989.

30. E. Talvitie & S. Singh. An experts algorithm for transfer learning. *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1065-1070, 2007.

31. M. Taylor, S. Whiteson, & P. Stone. Transfer learning for policy search methods. *Proceedings of the ICML Workshop on Structural Knowledge Transfer for Machine Learning*, 2006.

32. M. Taylor & P. Stone. Behavior transfer for value-function-based reinforcement learning. AAMAS, pages 53-59, 2005.

33. E.L. Thorndike. Educational Psychology. Columbia University Press, 1913.

34. E.L. Thorndike & R.S. Woodworth. The influence of improvement in one mental function upon the efficiency of other functions (I). *Psychological Review*, 8:247-261, 1901.

35. S. Thrun. Is learning the N-th thing any easier than learning the first? *Advances in Neural Processing Systems*, 8:640-646, 1996.

36. L. Torrey, J. Shavlik, T. Walker, & R. Maclin. Relational skill transfer via advice taking. *Proceedings of the ICML Workshop on Structural Knowledge Transfer for Machine Learning*, 2006.

37. X. Yan & J. Han. gSpan: Graph-based substructure pattern mining. *Proceedings of the International Conference on Data Mining*, 2002.

38. K. Yoshida, H. Motoda, & N. Indurkhya. Graph-based induction as a unified learning framework, *Journal of Applied Intelligence*, 4:297-328, 1994.

39. G. Michael Youngblood, Billy Nolen, Michael Ross, & Lawrence Holder. Building test beds for AI with the Q3 Mod Base. *Proceedings of the Second International Conference on Artificial Intelligence and Interactive Digital Entertainment*, pages 153-154, 2006.

40. G. Michael Youngblood & Lawrence B. Holder. Evaluating human-consistent behavior in a real-time first-person entertainment-based artificial environment. *Proceedings of the International Conference of the Florida Artificial Intelligence Research Society*, pages 32-36, 2003.