Using Games to Learn Games: Game-Theory Representations as a Source for Guided Social Learning

Alan Wagner[1,2](✉)

[1] Georgia Institute of Technology, Atlanta, GA, USA
[2] Pennsylvania State University, State College, PA, USA
`azw78@psu.edu`

Abstract. This paper examines the use of game-theoretic representations as a means of representing and learning both interactive games and patterns of interaction in general between a human and a robot. The paper explores the means by which a robot could generate the structure of a game. In addition to offering the formal underpinnings necessary for reasoning about strategy, game theory affords a method for representing the interactive structure of a game computationally. We investigate the possibility of teaching a robot the structure of a game via instructions, question and answer sessions led by the robot, and a mix of instruction and question and answer. Our results demonstrate that the use of game-theoretic representations may offer new advantages in terms of guided social learning.

Keywords: Game theory · Social learning · Interactive games

1 Introduction

Social interaction often involves stylized patterns of interaction [1]. These patterns may dictate how and when a person interacts, what actions they choose, and how their goals and motivations change. Interactive games, such as Rock-Paper-Scissors and poker, often structure a person's interactive behavior in a predetermined manner conducive to the game. Recently artificial systems have become adept at both playing and learning how to play many different games [2]. Comparatively little attention, however, has been paid to the social aspects of game playing and game learning. For instance, how can a social robot or agent learn to play a game from a person offering only disorganized verbal instructions? How can a system teach a person to play a game using subtle social cues and questions to determine if and to what extent they understand? How can such a system be developed to cope with the differences in play, instructions, and learning that occur across ages and cultures?

This paper constitutes a preliminary examination of these questions. The overarching goal of this work is to develop a system that could learn a wide variety of games from the type of interactive instructions provided by a typical person. Hence, we strive for generality both with respect to the game and the instructor. Moreover, we believe that our approach can also work when the robot acts as the instructor, explaining how to play a game. An important initial step towards creating such a system is to determine how to computationally represent interactive games.

Game theory researchers have extensively studied the representations and strategies used in games [3]. The types of games examined as part of game theory, however, tend to differ from our common notion of interactive games. Games in game theory tend to encompass limited interactions over a small range of behaviors and are focused on a small number of well-defined interactions. The Ultimatum Game, for example, requires one individual to divide a valuable resource while the other individual in the game can accept the division and receive a share or reject the division and both players receive nothing. Moreover, game theory focuses on conceptualizations for strategic interaction. In contrast, interactive games like Monopoly and poker offer players several different actions as part of a sequential ongoing interaction in which a player's motives may change as the game proceeds or depend on who is playing.

We contend that learning a pattern of interactions, such as those used in most interactive games, is a critical component for human-robot interaction because many interpersonal interactions follow prescribed patterns [4]. Methods developed for learning the structure of an interactive game could potentially be applied to the human-robot interaction scenarios encountered in a wide variety of social environments. For example, in most western cultures when meeting a new person the expected pattern of interaction is to introduce oneself, to shake the other person's hand, and to then wait for the other person to state their name.

This paper investigates methods by which a robot could learn the structure of an interactive game from a person. We focus on direct instruction. In particular, this article demonstrates the use of written instructions and the use of questions by the robot that, when answered by a person, convey the structure of the game. Further, we show that the robot can use a game-theoretic representation to reason about and select specific probing questions with the intention of learning about unknown aspects of the game. Overall, our immediate goal is to highlight the potential advantages of this approach in terms of teaching a robot these stylized patterns of interaction. Our long-term goal is to develop the computational underpinnings that will allow a robot to learn new patterns of interaction from an inexperienced person's instructions. The remainder of the paper begins with a brief background discussion of game theory and interactive games, followed by experiments and results.

2 Background and Related Work

Game theory has been the dominant approach for formally representing strategic interaction for more than 80 years [3]. Game theory assumes that the players of a game will pursue a rational strategy. A game is a formal representation of a strategic interaction among a set of players. A solution to a game describes classes of strategies for how best to play a game. There are many different types of solution concepts in game theory, the Nash Equilibrium being the most famous example of a solution concept.

Several different categories of games exist [3]. Games in which players select actions simultaneously are typically represented as a normal-form game (Fig. 1 center). Formally, a normal-form game is defined as a tuple $\left(N, A^{1,\dots,N}, R^{1,\dots,N}\right)$ where N is the set of players, A^i is the action space of individual i, and $R^i\left(a^1, \dots, a^N\right) \rightarrow \mathfrak{R}$ is a payoff

function. Games in which players select actions sequentially are generally represented as extensive-form games (Fig. 1 left). In addition to the formal elements of a normal-form game, extensive-form games include a set of histories H for each player and function $P(h)$ for selecting the player whose turn is next. Perfect information games are a class of extensive-form games in which each player knows every player's history. In imperfect information games players do not know the actions chosen by other players. A stochastic game is a series of normal-form games in which the actions selected in one game probabilistically determine the subsequent game. Stochastic games include a transition function $T(q, a^1, \ldots, a^N, q') \to [0, 1]$. These games are generalizations of both normal-form and extended-form games. They also generalize Markov Decision Processes (MDPs) for multiple individuals. Stochastic games start at initial state s_0 and are played with each player selecting an action and possibly receiving a payoff. The game moves to stage s_{i+1} with probability determined by the distribution Q until reaching a termination state. A stochastic game may last either a finite or infinite number of stages.
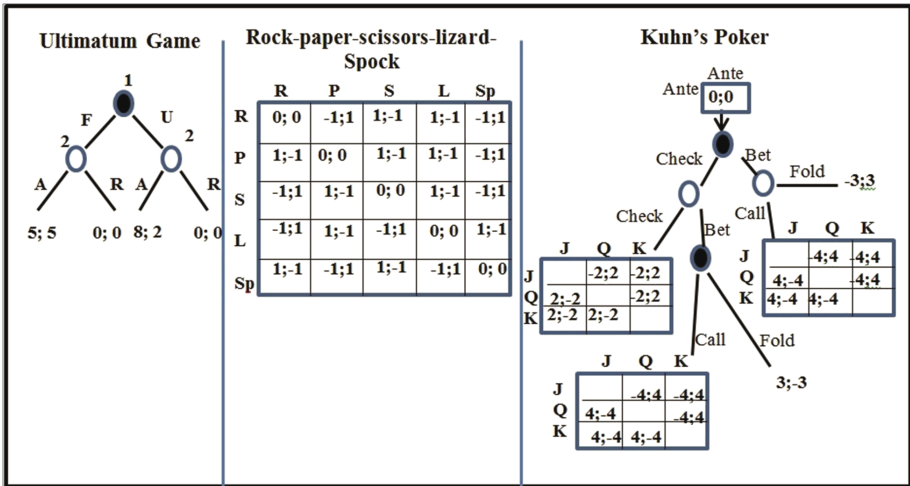

Fig. 1. Computational representations for the Ultimatum game, Rock-paper-scissors-lizard-Spock and Kuhn's poker are depicted above. The Ultimatum game is a sequential game between two players represented as an extensive-form game. Rock-paper-scissors-lizard-Spock is a simultaneous game represented as a normal-form game. Kuhn's poker is represented as a mixture of normal and extensive-form games. The selection of an action by a player results in a transition to the next stage of the game.

Most robotics related applications of game theory have focused on game theory's traditional strategy specific solution concepts [5]. Often, the structure of the game is preprogrammed and a game theory based controller is used to select the agent's actions. Recently this approach has resulted in tremendous advances in the quality of play in information imperfect games such as poker [6]. Nevertheless, this research is narrowly tailored to the development of agents that play optimally. In contrast, the work here is

not concerned with how well an agent or robot plays, but rather its ability to learn and represent different, unknown games.

Game theory has also been used as a means for controlling a robot [5, 7]. Game theory based robot control has similarly focused on optimization of strategic behavior by a robot in multi-robot scenarios. In particular, the use of Partially Observable Stochastic Games has been used as a means to control a robot team. In contrast to the prior work, we explore methods that will allow a human to teach a robot the structure of an interactive game such as Rock-Paper-Scissors and poker. Significant research has also explored the development of robots that learn games such as air hockey [8]. In contrast to strategic games, games such as air hockey tend to emphasize the physical and perceptual demands of play. Robot soccer, because of its dual physical and strategic demands, arguably represents the most challenging category of game. Research related to this game has explored both the physical demands [9] and the strategic demands [10].

Very little work has examined the use of game theory as a means for controlling a robot's interactive behavior with a human. Lee and Hwang attempt to develop a conceptual bridge from game theory to interactive control of a social robot [11]. Our own work has centered on the use of the normal-form game as a representation and means of control for human-robot interaction [12]. Yet, in this prior work we focused only on the use of the representation to control a robot's behavior and not the direct learning a game's interactive structure.

3   Representing Interactive Games

Game-theory representations have been used to formally represent and reason about a number of interactive games [13]. Games such as Snakes and Ladders, Tic-Tac-Toe, and versions of Chess have all been explored from a game theory perspective. The methods used to represent these games are well known.

The normal-form game and the extensive-form game serve as building blocks to represent a complete interactive game. Simultaneous stages of an interactive game are represented in normal-form as a matrix (Fig. 2). Each player's potential actions are listed along the dimensions of the matrix. Payoffs for selecting a particular set of actions are included as values within the matrix. Sequential stages of an interactive game are represented in extensive-form as a tree. A player's potential actions are denoted by the branches of the tree. Nodes of the tree indicate which player makes a decision at each particular stage of the game. Payoffs for selecting a particular set of actions are depicts at the stage in which the payoffs are received.

The cells in a normal-form game and the terminating branches of an extensive-form game direct the players to the subsequent stages of the game. Resembling a probabilistic-finite-state automaton (FSA), each state is a normal-form or extensive-form game representation and each transition occurs when arriving at a cell or a terminal tree node. Represented in this manner, the challenge of learning a new interactive game is reduced to learning the structure and underlying components of the game-theoretic representation. The section that follows investigates this challenge.
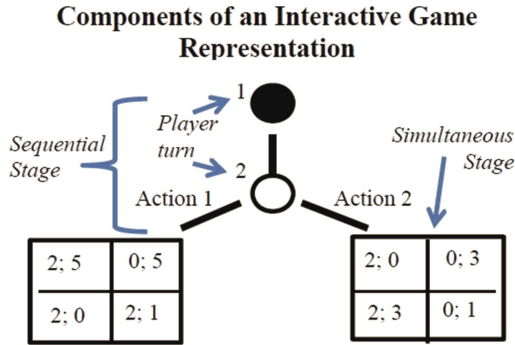
Components of an Interactive Game Representation

Fig. 2. Normal and extensive form games are used to represent the components of an interactive game. Sequential stages are represented in extensive form. Numbers are used to indicate a player's turn. Simultaneous stages are represented as a normal form game. Transitions connect components and denote the selection of an action.

4   Teaching a Robot an Interactive Game

The primary contribution of this work is to examine, present, and demonstrate techniques for learning the types of game-theoretic representations described above from the information provided by a human teacher. The most obvious and applicable methods for learning the structure of an interactive game are direct instruction and question and answer.

Direct instruction describes the explicit teaching of skills needed for some purpose. Some psychologists argue that direct instruction represents the most effective way to teach and to learn [14]. To directly instruct a robot to play a game, the human teacher simply communicates the underlying structure of the interactive game to be learned. This communication can be in the form of a list of spoken, written, or demonstrated instructions necessary for performing a task. Written instructions can be used in place of verbal instructions. In this case, an interactive game's set of instructions can be used to learn a new game.

The use of a game-theoretic representation requires that specific information is communicated to the robot. In general, for each stage of the game, the robot must know who is playing, what actions are available to each player, what reward or cost is associated with the selection of each action pair, whether actions are selected simultaneously or sequentially by the players, and which stage of the game results from the selection of an action pair. When direct instruction is used, these questions are addressed directly as a list of spoken, written, or demonstrated instructions.

A less obvious means for teaching a robot an interactive game is to allow the robot to ask questions about the game that the person answers. In this case, the robot acts as an inquisitor asking questions that allow it to build the game representation from the ground up. **The evolving game representation in this case determines what questions the robot must ask in order to flesh out the representation**. The first question asks how many players are participating. For each stage of the game, the robot then asks whether the players act sequentially or simultaneously. Next, the robot inquiries about

**Sample Question and Answer session for
learning an Interactive Game**

Q1: How many players for stage 0?
A1: *2*
Q2: Is stage 0 sequential or simultaneous?
A2: *sequential*
Q3: Which player acts at this stage?
A3: *1*
Q4: What actions are available to Player 1 at stage 0?
A4: *Action1*
Q5: If player 1 at stage 0 selections Action1 does a
sequential or simultaneous stage result?
A5: *sequential.*
Q6: Which player acts at this stage?
A6: *2*

Fig. 3. An example of a question and answer session for learning the first stage of the interactive game from Fig. 2.

the actions available to each individual at that stage. For each cell in a normal form game or terminal node in an extensive form game, the robot inquiries about the reward received and which, if any, subsequent stage results from the selection of the action pair. Figure 3 depicts an example of a question and answer session used to learn the first stage of the interactive game depicted in Fig. 2. We contend that a similar series of questions can be used in either a depth first or breadth first manner to learn the interactive structure of most games.

The proceeding techniques can be combined resulting in a system in which the robot builds the game representation from the instructions and then asks questions about any unknown or unclear parts of the representation. In this case the robot must determine which portions of the representation are unclear or unknown. In some cases the presentation of the instructions may afford measures of confidence with respect to the instructions. For instance, many natural-language-processing (NLP) algorithms provide confidence measures reflecting the system's estimation of accuracy. We conjecture that such a system could potential be used to assist the robot in determine if and what portions of the representation require follow up questions.

In some cases, the representation itself may suggest information that is missing. For example, the absence of reward or cost values in a matrix (e.g. the numbers 2; 5 from Fig. 2) is easily tested. In this case, the absence of expected reward values can prompt the robot to inquire about the value or cost of selecting particular actions during a stage of the game.

5   Experiments

Most applications of game theory evaluate the system's performance in terms of winning (e.g. [15]) or win related tasks such as scoring goals (e.g. [8]). In contrast, we argue that the best evaluation of game learning is to measure the system's ability to play a game

after being taught, regardless of whether it wins. Metrics such as illegal moves attempted, measure the accuracy of the robot's model of the game structure.

We evaluated the number of illegal moves attempted by the robot in three different games (Fig. 1): the Ultimatum game, Rock-paper-scissors-lizard-Spock, and Kuhn's poker. The Ultimatum game is a single stage sequential game in which one player chooses either a fair or unfair division of a resource and a second player either accepts or rejects the division. If the division is accepted then both players receive reward proportional to the division. Alternatively, if the division is rejected then both players receive nothing. Rock-paper-scissors-lizard-Spock is similar to the classic game rock-paper-scissors except with two additional actions. The lizard action defeats Spock and paper and is defeated by scissors and rock. The Spock action defeats scissors and rock and is defeated by paper and lizard. Figure 1 delineates which actions dominate other actions. In this game all players simultaneously make a hand sign representing one of the five namesakes. Finally, Kuhn's poker is a simplified version of Texas Hold'em poker. This game is played with only a jack, a queen, and a king. The game begins when each player bets 1 as an ante. Next each player receives a single card. Player 1 may check or bet. As depicted in Fig. 1 the actions available to player 2 depend on player 1's action. Each round of the game ends when a player either folds (resigns and forfeits their bets) or during a showdown stage each player's cards are revealed and the player with the higher card wins.

The robot learned each of the three games by direct instruction and mixed-direct instruction and question and answer. In the direct instruction condition, the robot was given a set of instructions (e.g. Fig. 3) describing how to play the game. The instructions were not in a natural-language format. Although, the challenge of translating from natural language to a game theory format is beyond the scope on this article, random errors were added to the instructions in an effort to roughly simulate the errors that would occur during translation. Each game instruction had a 15 % chance of being incorrect (translation error rate). This level of error was arbitrarily selected. Three different types of error occurred. Incorrect stage transitions occurred when the robot's representation erroneously indicated the stage that would result when a pair of actions was selected. Incorrect reward values inaccurately specified the amount of reward to be received at a stage of the when a pair of actions is selected. Finally, incorrect actions erroneously indicated which actions were available to the robot at a particular stage of the game.

The experimenter served as the robot's opponent. All of the experimenter's instruction and responses were predetermined to avoid bias. When the robot asked the human for missing information, the correct information was provided. The quantitative results that follow were obtained from an experiment conducted in simulation.

The data from the direct learning from instructions condition (Fig. 4) demonstrates that the robot selected illegal moves at a rate of 16, 11, and 17 % for the Ultimatum game, Rock-paper-scissors-lizard-Spock game, and Kuhn's poker respectively. These results indicate a rate of illegal moves which is approximately equal to the translation error rate. This rate is higher than expected. We hypothesized that translation errors related to the amount of reward would only impact strategy and not whether or not a move was illegal. We therefore believed that, in the first condition, the number of illegal moves would be significantly less than the translation error rate. We found, however,
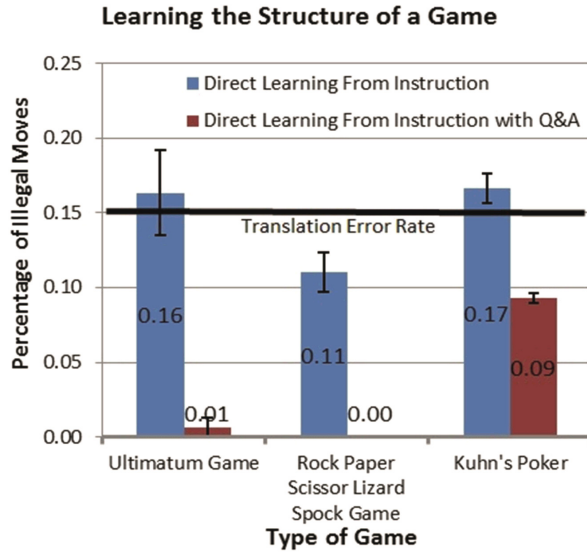
Learning the Structure of a Game

Fig. 4. Results from an experiment examining the possibility of a robot learning a game-theoretic representation of a game. The blue (left) columns depict a condition in which the robot learns the game from a set of imperfectly translated instructions. The red (right) columns depict a condition in which the robot is provided with instructions describing how to play the game which are missing information. The robot infers which information is missing and then asks the person to questions which allows it to complete the representation. (Color figure online)

that incorrect reward values can impact game structure by consistently guiding the robot towards illegal moves. In other words, translation errors can cause the robot to believe that it will obtain a large reward by performing an illegal move. This results in a strategy of using illegal moves which predominates. The data also shows that certain game structures appear to be more impacted by translation errors than other games. The Rock-paper-scissors-lizard-Spock game, for instance, was consistently found to result in fewer illegal moves when compared to the other games. Because this game consists of a single simultaneous stage, most errors do not result in illegal moves. Sequential games, on the other hand, afford multiple opportunities for selecting illegal moves.

In the second condition, the robot received instructions that had missing information in the form of reward values, potential moves, and game stage transitions. The robot then had the opportunity to ask the person questions about any information that it could identify as missing. Missing stage transitions could typically be inferred from the presence of stages not connected to the start state or some later stage of the game. Similarly, missing actions were often indicated by stage transitions without a requisite action pair. Missing reward values were easily inferred from the game-theoretic representation. The robot then asked the human to provide the missing information. The robot used question and answer to generate error-free representations of the game Ultimatum game and Rock-paper-scissors-lizard-Spock game. Because these games consisted of a single sequential or simultaneous stage, the robot could accurately infer which information was missing. Kuhn's poker, however, presented unique challenges in terms of inferring

missing information. Although missing reward values and transitions were identified, missing actions were seldom noticed. If a stage transition and an action during the stage were both missing, then inference that an action was missing was not possible. Actions that did not result in transitions were similarly not identified as missing. Overall, the results demonstrate that the game theoretic information does assist with inferring which information is missing from the game structure. Asking a person to provide missing information improves the robot's ability to play the game.

We tested the ability to learn and play these games on the NAO robot from Aldebaran. During this testing, the robot learned each of the games from written instructions and question and answer sessions with the experimenter. Question and answer sessions were conducted by typing answers to the robot's questions. The games were then played with the robot. The robot verbalized its actions instead of making physical actions. Each game was played 10 times with the robot. The NAO selected actions that were believed to be reward maximizing. We recorded each of the NAO's action selections. The robot was able to learn each of the games using both written instructions and question and answer session. However, because no error was introduced, game play was structurally perfect. Hence the robot's ability to play the learned games was confirmed although no quantitative results from these robot experiments are reported.

6 Conclusion

This article has examined the use of game-theoretic representations as a means of representing and learning interactive games involving a human and a robot. Our experiments demonstrate that written instructions and mixed instruction and question and answer can be used to learn different types of interactive games. We have shown that the use of game-theoretic representations of interaction offer several important features. First, and perhaps most importantly, the game representation affords a means for organizing the information needed by the robot to learn an interactive game. The computational representation of a game can be used to structure the information being received by a person and guide the robot when asking questions.

The research presented here could be an important step towards the development of a system for human-robot guided learning. Such a system might one day allow people to teach a robot the games that the person would like to play with the robot Before a fielded application could be realized, some assumptions would need to be addressed. For instance, we assumed that the robot already possessed the knowledge of how to perform all game related actions. We believe that learning these actions is related to game learning but best achieved by using learning from demonstration.

This work represents an initial investigation into the possibility of using game-theoretic representations to structure an interactive game. An important next step is to develop a system that learns a game from a naïve human subject. Such a system would require some competence in natural language understanding. Spoken or read instructions [16] would be used to broadly develop the interactive structure of the game, socially guided questions would then be used to rectify unclear or unknown portions of the game, learning by demonstration would be used to learn how to perform the actions, and

practice would result in the refinement of strategy. Although this paper has focused on learning how to represent these games, we believe that these representations could be used in many different interactive situations.

References

1. Kelly, H.H.: The theoretical description of interdependence by means of transition lists. J. Pers. Soc. Psychol. **47**, 956–982 (1984)
2. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M.: Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602 (2013)
3. Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press, Cambridge (1994)
4. Rusbult, C.E., Van Lange, P.A.M.: Interdependence, interaction, and relationships. Ann. Rev. Psychol. **54**, 351–375 (2003)
5. Emery-Montemerlo, R.: Game-theoretic control for robot teams. Ph.D. thesis, Carnegie Mellon University (2005)
6. Johanson, M., Bard, N., Burch, N., Bowling, M.: Finding optimal abstract strategies in extensive form games. In: Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI) (2012)
7. Bernstein, D.S., Hansen, E.A., Zilberstein, S., Amato, C.: Dynamic programming for partially observable stochastic games. In: AAAI Spring Symposium, Palo Alto, CA (2004)
8. Bentivegna, D., Ude, A., Atkeson, C.G., Cheng, G.: Humanoid robot learning and game playing using PC-based vision. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV (2002)
9. Grollman, D.H., Jenkins, O.C.: Learning robot soccer skills from demonstration. In: IEEE International Conference on Development and Learning (ICDL), London, UK (2007)
10. Ahmadi, M., Lamjiri, A., Nevisi, M., Habibi, J., Badie, K.: Using a two-layered case-based reasoning for prediction in soccer coach. In: Arabnia, H.R., Kozerenko, E.B., (eds.) International Conference on Machine Learning; Models, Technologies and Applications, CSREA Press, USA, pp. 181–185 (2003)
11. Lee, K., Hwang, J.-H.: Human-robot interaction as a cooperative game. In: Castillo, O., Xu, L., Ao, S.-L. (eds.) Trends in Intelligent Systems and Computer Engineering (IMECS 2007). Lecture Notes in Electrical Engineering, pp. 91–103. Springer, New York (2008)
12. Wagner, R.: Creating and using matrix representations of social interaction. In: Proceedings of the 4th International Conference on Human-Robot Interaction (HRI 2009), San Diego, CA (2009)
13. Berlekamp, E., Conway, J.H., Guy, R.: Winning Ways for your Mathematical Plays: Games in General. Academic Press, London (1982)
14. Kirschner, P.A., Sweller, J., Clark, R.E.: Why minimal guidance during instruction does not work an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educ. Psychol. **41**(2), 75–86 (2006)
15. Banerjee, B., Stone, P.: General game learning using knowledge transfer. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-07), Hyderabad, India (2007)
16. Branavan, S.R.K., Silver, D., Barzilay, R.: Learning to win by reading manuals in a Monte-Carlo framework. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 268–277. Association for Computational Linguistics (2011)