# Towards strategic Kriegspiel play with opponent modeling

# (Extended Abstract)

Antonio Del Giudice, Piotr Gmytrasiewicz, Josh Bryan
Department of Computer Science, University of Illinois at Chicago
Chicago, IL, 60607-7053, USA
E-mail: a.delgiudice@gmail.com, piotr@cs.uic.edu, jbryan6@uic.edu

Kriegspiel is a chess variant belonging to the family of invisible chess that encompasses *partially observable* variants of the popular game. Playing Kriegspiel is difficult, first, because the player needs to maintain a belief over all possible board configurations. Second, the player needs to select his move, given his belief about the board configuration and given the likely responses of the opponent. Predicting the likely responses is crucial and has a long tradition in minimax approaches to fully observable games. Minimax assumes that the players are rational and have opposing preferences, which is common knowledge.[1] Further, minimax is applicable only to fully observable games. In partially observable games one needs to model not only the opponent's preferences, but also the opponent's belief about the board configuration, and possibly his belief about the original player. Further, the opponent's level of expertise may also be in question in realistic settings.

Our approach is based on interactive partially observable Markov decision process [2] (I-POMDPs). Like POMDPs, I-POMDPs provide a framework for sequential planning. However, they generalize POMDPs to multi-agent settings by including the models of the other agent in the state space. The models are used to form an informed prediction of the other agent's actions, which is then used during the move selection. Given the complications of maintaining the beliefs over the board configurations in Kriegspiel, the need to include the possible models of the other player further adds to the difficulty. We argue that without opponent modeling some important aspects of the game are necessarily neglected. In particular, without modeling the state of belief of the opponent the impact of moves which have the effect of supplying the opponent with information cannot be taken into account. As should be expected, I-POMDPs reduce to classical POMDPs if there are no other agents in the environment.

One of the obvious sources of complexity of solving I-POMDPs, apart from complexity inherent in POMDPs, is the need for considering all possible models of the opponent, and, if further levels of nesting are used, all possible models the opponent may have of the original agent, and so on. This paper proposes the use of the notion of quantal response [1, 4] to ameliorate the complexity. Quantal response allows one to use one model of the opponent to represent a whole ensemble of models with similar characteristics.

[1]We should remark that common knowledge is a very strong, indeed unrealistic, assumption. See, for example, [3].

**Cite as:** Towards strategic Kriegspiel play with opponent modeling, (Extended Abstract), Antonio Del Giudice, Piotr Gmytrasiewicz, Josh Bryan, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009),* Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 1265–1266


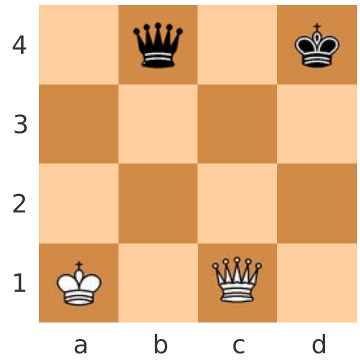


**Figure 1: Example of initial state chessboard**

Take an example of the configuration shown in Figure 1, if the White player were to attempt to move its Queen to the right (Qd1), the move would be considered legal, and executed with the referee's announcement "Check by File". Qb2 is also legal and referee would announce "Check by Long Diagonal". In particular, we compute the most desirable move for player $i$, given the assumption that both players know the locations of all pieces in this initial configuration. We show that $i$'s (White) best move is Queen to d1. This result could be computed both under the assumption that $j$ (Black) responds with a random move, and by modeling $j$ as maximizing its evaluation function. However, we also show that $i$'s move Qb2 is less preferable than Qd2. This is because the former move reveals a lot of useful information about the $i$'s Queen to the opponent. This insight is not possible if $j$'s state of information is not modeled.

I-POMPD framework [2] generalizes the concept of single-agent POMDPs to multi-agent domains. Player $i$'s belief update and decision making in I-POMDPs is formally derived in [2]. Applied to Kriegspiel, the belief update involves, first, updating probabilities of states given $i$'s moves. Second, updating $i$'s belief over $j$'s belief state given referee's announcements that $j$'s moves could generate. And, third, updating probabilities of states based on probabilities of $j$'s moves in its various possible belief states. $i$'s optimal move

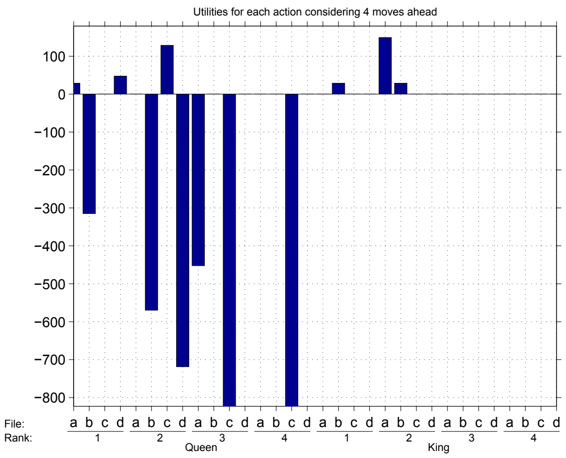


**Figure 2: The probabilities of each of $j$'s moves after hearing a "Silent" call from the referee under the assumption that $j$ considers the next 3 moves.**

is arrived at by exploring the utility of beliefs resulting from each of its moves.

Let us examine in further detail how $i$ (white) models $j$ (black) when only two moves are considered. In this case $i$ does not see far enough ahead to realize the utility of reducing $j$'s knowledge. If $i$ executes Ka2 or Qc2 the referee would respond with a "Silent" call. Since $j$ is modeled as only considering its next move, $i$ estimates that $j$'s most likely moves are Qc3 and Qb1. This is because after the "Silent" call, $j$ believes there is a chance that $i$'s queen is in b1 or c3, and therefore there is a chance to capture it. Since either of these moves would, at best, leave $i$ in check, $i$ instead opts for the safer move Qd1 despite the fact that it gives more information to $j$. Qd1 places $j$ in check without any risk of retaliation by $j$ on the next move.

Next, let us consider the case where $i$ models the next four moves. Figure 2 shows the computed probabilities of each of $j$'s moves after receiving a "Silent" response from the referee computed from the three layer model of $j$. Clearly the highest probability action for $j$ is Qa3. This is because the move Qa3 could potentially result in the capture of $i$'s queen had the "Silent" been generated by a Qa3 move by $i$. The second highest probability move for $j$ deserves some explanation. Qd4 has a high utility for $j$ even though it is aware that this would be an illegal move. In this case this is not because any information would be gained, but because $j$ knows that even if it does something illegal, it will still have a subsequent turn to perform an action that is useful. Other high ranking moves for $j$ include the three king moves Kc3, Kd3, and Kc4. These moves have a possibility of generating an "Illegal" call when moving into check, and thereby provide more information to $j$ about $i$'s positions. In the event that they don't produce an "Illegal" call, they would produce a "Silent" call and provide $i$ with little information about which move was performed.

With the above analysis of $j$'s most probable responses to a call of "Silent" by the referee, it becomes clear why $i$'s highest expected utility move is Ka2, if $i$ looks four moves ahead into the future. From this position, should $j$ follow through with the most probable move (Qa3), $i$ would be in a position to capture $j$'s queen. Furthermore, if $j$ chose a different move instead, each response generated from $j$'s move would allow $i$ to narrow the possible states of the board to at most two states. This increased information about $j$'s position would then allow $i$ to respond with reasonable confidence about the board state on the third move.

Our analysis of an example scenario shows that modeling the opponent using a less detailed approach, for example by assuming that the opponent will respond with a random move, is an approximation to the more sophisticated modeling approach. Also, as expected, assuming full observability, as in minimax, generates recommendations that neglect the value of moves that hide information.

In our future work we will explore I-POMDP approximation algorithms which are necessary to handle 8 by 8 chess boards.

# 1. REFERENCES

[1] C. F. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, 2003.

[2] P. Gmytrasiewicz and P. Doshi. A framework for sequential planning in multiagent settings. *Journal of Artificial Intelligence Research*, 24:49–79, 2005. http://jair.org/contents/v24.html.

[3] J. Y. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, July 1990.

[4] K. Ravikumar, A. Saroop, H. K. Narahari, and P. Dayama. Demand sensing in e-business, 2005.