

Towards Strategic Kriegspiel Play with Opponent Modeling

Antonio Del Giudice and Piotr Gmytrasiewicz

Department of Computer Science, University of Illinois at Chicago
Chicago, IL, 60607-7053, USA

E-mail: a.delgiudice@gmail.com, piotr@cs.uic.edu

Abstract

Kriegspiel, or partially observable chess, is appealing to the AI community due to its similarity to real-world applications in which a decision maker is not a lone agent changing the environment. This paper applies the framework of Interactive POMDPs to design a competent Kriegspiel player. The novel element, compared to the existing approaches, is to model the opponent as a competent player and to predict his likely moves. The moves of our own player can then be computed based on these predictions. The problem is challenging because, first, there are many possible world states the agent has to keep track of. Second, one may be unsure about the characteristics of the other player which could influence his behavior, such as his level of expertise or his evaluation function. To keep the size of the state space manageable we consider a scaled down version of Kriegspiel played on 4 by 4 chessboard with only a king and a queen on both sides. To deal with an opponent with uncertain characteristics we use the notion of quantal responses developed in behavioral game theory. This allows us to consider only one prototypical opponent while modeling a whole ensemble of possible opponents. We implemented our approach using influence diagrams, and discuss results in example situations.

Introduction

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 be smart about selecting its move, given its belief about the board configuration and given the likely responses of the opponent. Predicting the likely responses is, of course, crucial, and has a long tradition in Mini-Max approaches to fully observable games. Mini-Max assumes that the opponent has opposing preferences, but it is relatively easy to apply 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. 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 (0) (I-POMDPs). Like POMDPs,

I-POMDPs provide a framework for sequential planning. However, they generalize POMDPs to multiagent settings by including the models of the other agent in the state space.¹ 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, however, 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 crucial impact of moves which have the effect of supplying the opponent with information cannot be taken into account.

In previous work Parker et al. (0) use sampling to represent beliefs over the state of the board, and avoid modeling the opponent explicitly by assuming that it will move randomly. Russell and Wolfe (0) consider whether guaranteed wins exist in some end-game configurations and prove that, for that purpose, the opponent's state of belief does not matter. Parker et. al.'s work is particularly relevant to our approach because it can be viewed as an approximation. More precisely, the assumption that the opponent responds by executing a random move is an approximation to having a more informed prediction of the opponent's action obtained using a model of the opponent's preferences and beliefs about the state. Of course, one may model the opponent on a more detailed level by considering how it may model the original player, and so on. In I-POMDP framework (0) the nesting of models may be infinite, but finitely nested I-POMDPs are approximations which guarantee that the belief updates and solutions are computable. In our discussion below we illustrate how, for example, the assumption that the opponent will respond randomly approximates the solution obtained based on an explicit model of the opponent.

The improved quality of play based on explicit and more detailed models of the opponent comes at the cost of increased computational complexity. To manage this complexity this paper considers 4 by 4 chessboards with a king and a queen on both sides, resulting in less than 74 thousand possible board positions. Our player maintains its belief

¹We assume the presence of a single other player throughout the rest of the paper.

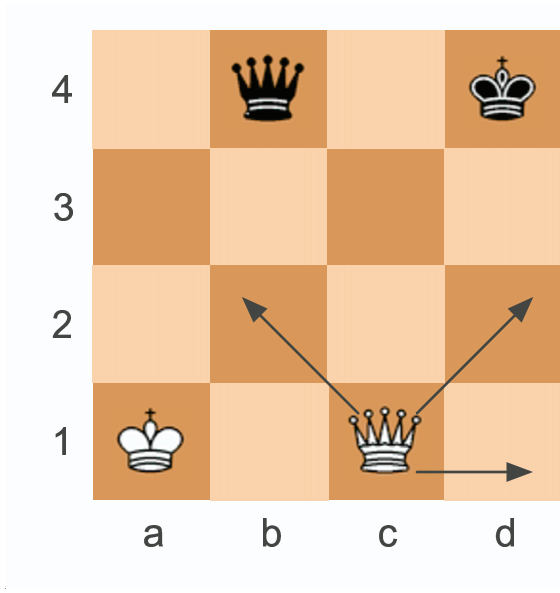


Figure 1: Example of initial state chessboard

over that space, and keeps track of the opponent’s possible beliefs. Further, instead of considering all possible models of the opponent’s evaluation functions and skill levels, we compute the desirability of the opponent’s moves based on one representative model. We then use the notion of quantal response (0; 0) to convert the opponent’s expected utilities to probabilities of its moves.

Kriegspiel

We briefly summarize the rules of Kriegspiel following (0; 0; 0; 0). The game of Kriegspiel involves two players, i (which will stand for White) and j (Black), and one *referee*. The two players can see only their own pieces; the opponent’s pieces are completely invisible. The referee can see pieces of both players. The rules governing moves of the pieces are identical as in chess. Every time a player, say i , proposes a move to the referee, the following happens:

- If the move is illegal (i.e., it does not comply with the rules of chess given all the pieces on the board), the referee announces “Illegal”, and the player i may propose another move.
- If the move is legal, it is executed, and the referee announces:
 - “Capture in X ”, if a piece is captured on square X .
 - “Check by Y ” if Black is in check. Y can have the values of Rank (row), File (column), Short Diagonal, Long Diagonal and Knight.
 - If Black has no legal moves: “Checkmate” if Black is in Check and “Stalemate” otherwise.
 - “Black to move” or “Silent” if none of the above happens.

For example, in 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”. We use the board configuration in Figure 1 as example in the rest of the paper. 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 best move is Qd1, as could be expected. 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 much less preferable than Qd2. This is because the former move reveals the exact position of i ’s Queen to the opponent. This insight is not possible if j is assumed to respond randomly.

Interactive POMDP

The I-POMDP framework (0) generalizes the concept of single-agent POMDPs to multiagent domains. An I-POMDP of agent i is

$$I\text{-POMDP}_i = \langle IS_i, A, T_i, \Omega_i, O_i, R_i \rangle \quad (1)$$

where:

- IS_i , the *interactive state*, is the cross product of the set of *physical* states S (in our case possible board configurations) and the set of possible models of the opponent, j . We will consider only *intentional* models (or types) here. A type of agent j , θ_j , consists of its belief state b_j and frame $\hat{\theta}_j$. As we explain further in the next section, the interactive states allow i to keep track of i ’s belief about the board configuration and about j ’s beliefs.

- A is the cross product of the actions agent i and opponent j can make.

- T_i , the *transition function*, defined as $T_i : S \times A \times S \rightarrow \{0, 1\}$. Thus, we assume that Kriegspiel is a deterministic domain and that agents’ actions are assumed to only influence the physical state part of the interactive state space (this is called a model non-manipulability assumption in (0).)

- Ω_i is the set of possible observations of i , here assumed to be the set containing all possible *referee*’s responses.

- O_i is an observation function $O_i : S \times A \times \Omega_i \rightarrow \{0, 1\}$. We assume that the referee’s responses are deterministic according to the rules of Kriegspiel explained above.

- R_i is the reward function $R_i : S_i \times A \rightarrow \mathbb{R}$. Both Kriegspiel and chess associate a reward only with terminal states of the game and win, lose or draw of each agent. Since we cannot search the game tree so far into the future, we use the the board evaluation function adopted from GNU chess to represent the utility of a state (i.e., the measure that the board configuration will lead to the agent’s win.)

In our implementation, we constructed an evaluation function in Equation 2 with four parameters weighted in way extrapolated from the GNU specifications (0).

$$U = \alpha \cdot (X' - X'') + \beta \cdot (Y' - Y'') + \gamma \cdot (Z' - Z'') \quad (2)$$

where

- X , Y and Z are functions in Table 1
- α , β , γ are weights in Table 1
- X' and X'' describe, respectively, the function X from player i 's or j 's perspective.

Weight	Symbol	Name
+ 36	X	King Centrality
+ 980	Y	Queen Presence
+ 50	Z	Check

Table 1: Evaluation parameters

In Table 1, *King Centrality* gives a reward for own king being in a central position in the board, *Queen Presence* acknowledges the importance of having own queen on the board, while *Check* shows when the opponent's king is under check.

Maintaining the Interactive Belief States and Decision Making

As in POMDPs, the agent i 's belief over interactive states is a sufficient statistic, that is, it fully summarizes its observable history. Each interactive state includes a physical state s and a model of j . An intentional model consists of j 's belief and its frame, θ_j , which contains other relevant properties of j (for example, j 's reward function. See (0) for details.)

i 's belief update and decision making in I-POMDPs is formally derived in (0). Applied to Kriegspiel, 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. For clarity we assume that i is certain which θ_j describes j (we relax this assumption using quantal response as explained later). i 's optimal move is arrived at by exploring the utility of beliefs resulting from each of its moves.

For specific case of Kriegspiel, we implemented the above using two kinds of decision networks (our implementation uses Matlab Bayes Net Toolbox, running on AMD 64-bit architecture). In Figure 2 we depicted the top-level decision network of agent i uses. It contains i 's own decision and utility nodes, and a random node representing the predicted actions of agent j . The dotted link between the referee's announcement due to i 's action and node A_j indicates that the referee announcement influences, although indirectly, the probability distribution associated with j actions. To compute this influence, the model of j in Figure 3 is used.

The network in Figure 3 models j 's decision making, but assumes that j is not further modeling i 's responses. The nesting of modeling could go into further levels of detail, with the usually associated tradeoff between decision quality and computational effort. Here, the model in Figure 3 is invoked many times, each time with the node representing

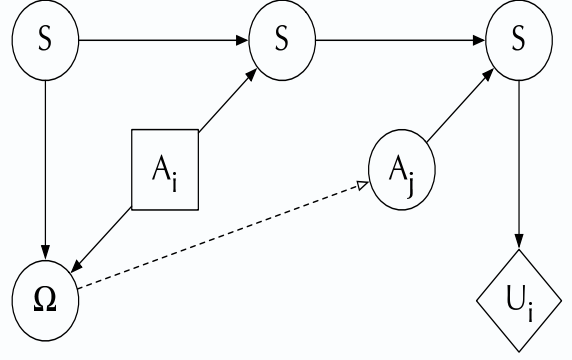


Figure 2: A top-level decision network of agent i .

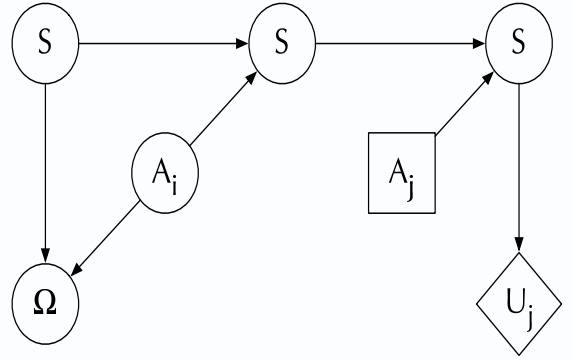


Figure 3: A decision network i uses to model j 's decision making.

the referee's announcement due to i 's action instantiated to a possible announcement. In each case, the state of j 's belief revised after the announcement is used to compute the expected utilities of all of j 's alternative moves.

The expected utilities of j 's actions, for each of j 's belief states resulting from various referee's announcements, are converted to probability distributions over j 's actions using quantal response explained below. The overall probability of j 's response is obtained as a probabilistic mixture of its responses for each belief state, weighted with the probability of a belief state (see (0) for formal details).

The notion of quantal response we use has been coined in the fields of behavioral game theory and experimental economics (0; 0). The idea is that decision makers can rarely be assumed to perfectly rational, and to compute their utilities and probabilities exactly. Hence, one may need to replace the postulate that a decision maker is surely going to maximize his expected utility with a postulate that likelihoods of various actions increase as the actions' expected utilities increase. The shape of this dependence quantifies the likelihoods of the decision maker's mistakes, as well as the errors inherent in the model of the decision maker.

Formally, the quantal response is defined by:

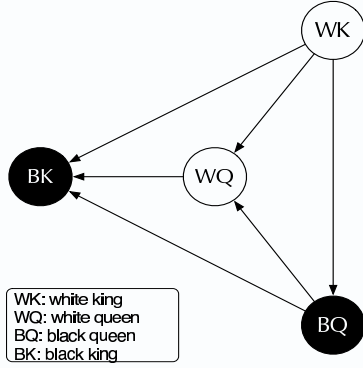


Figure 4: A Bayesian network depicting a detailed representation of the state of a chessboard with Queen and King pieces on both sides.

$$P(\alpha_j) = \frac{e^{\lambda U(\alpha_j)}}{\sum_{\alpha_j} e^{\lambda U(\alpha_j)}} \quad (3)$$

where $P(\alpha_j)$ is the probability for opponent j to execute action α_j and $U(\alpha_j)$ is the expected utility of an action computed by the model. The parameter λ quantifies the degree to which our model of j is correct. The high values of λ indicate that j is unlikely to perform an act that does not maximize its expected utility, as computed by the model.

In both Figure 2 and Figure 3, the state variable S is actually represented as a four-node bayesian network depicted in Figure 4.

Results

Now we discuss the results of i 's modeling j in the simple scenario depicted in Figure 1 (recall that, for simplicity, we assume that agents know positions of all pieces in this initial configuration.) Let us consider some of i 's plausible moves, and the possible responses that one can expect j to execute. In Figure 5 we depict how j may view the situation after the i 's moves Qd1 or Qd2 generate the referee's response "Check by File". After the announcement j knows that the white Queen is in d1 or d2, as computed by network in Figure 3. Now, three of j 's (Black) actions can be computed to have relatively high utilities. Black Qd2 has the expected utility of 490, and Black King moves Kc4 and Kc3 have the expected utilities of 0 and -7, respectively. All other moves have the expected utilities equal to -50. The probabilities of the moves are depicted in the right in Figure 5; we used the value of λ equal to 0.004.

In Figure 6 we depict the situation resulting from White executing Qb2. The referee would then announce "Check by Long Diagonal". The updated belief of j then leaves no doubt that White Queen is at b2. Now, four responses by Black stand out. By far the best is to capture White Queen by moving Black Queen to b2. With the value of λ as above, White computes the probability of this response as 0.7. The other three moves which remove the check from the Black

King are judged as equally good. Let us not that this analysis is limited in that Black taking White Queen in b2 exposes the Black Queen to capture by the White King. The model i uses of j , in Figure 3, would miss this danger since it does not model the White response.

Give the above analysis of Black responses to various moves by White, i (White) uses its top-level network to compute its own best move in scenario in Figure 1. The best White move is Q to d1, as could be expected, with the expected value of 20. The analysis of values of White moves Qd2 and Qb2 is also interesting. As we mentioned, modeling Black as intentional reveals that White Qb2 would result in very likely capture of White Queen, while the danger of capture as the result of White Qd2 much lower. This is due to the White Qb2 move providing valuable (and actionable) information to Black. This conclusion can be arrived only if Black is modeled as intentional. Assuming that Black's responses are always random would lead to the conclusion that both moves have equal value.

Conclusions and Future Work

We presented an approach to designing optimal control strategies in Kriegspiel based on the framework of Interactive POMDPs. I-POMDPs allow agents to model other agents as rational during an interaction, and derive optimal actions based on the predictions of other's behavior. We implemented the simplified 4 by 4 Kriegspiel domain with King and Queen on both sides. We further simplified the modeling of the opponent by using one representative model of its decision making and deriving the probabilities of opponent's responses using the notion of quantal response.

Our analysis of a simple 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. We expect that still more detailed models involving opponent's possibly modeling the original player result in still more informative estimates of the desired courses of action, but involve the usual tradeoff due to increased computational cost. In the scenario in Figure 1, for example, a further level of modeling could reveal that White Qb2 move is preferred to Qd1 since Black Queen would not respond with capturing the White Queen.

In our future work we will explore modeling opponent at the deeper levels of nesting, and to employ sampling approximations to Bayesian inference to handle 8 by 8 chess boards.

References

- C. F. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, 2003.
- P. Ciancarini, F. DellaLibera, and F. Maran. Decision making under uncertainty: a rational approach to kriegspiel. *Advances in Computer Chess* 8, 1997.
- I. Free Software Foundation. Heuristic descriptions for chess., 1987. [Online; accessed 8-February-2006].
- P. Gmytrasiewicz and P. Doshi. A framework for sequential planning in multiagent settings. *Journal*

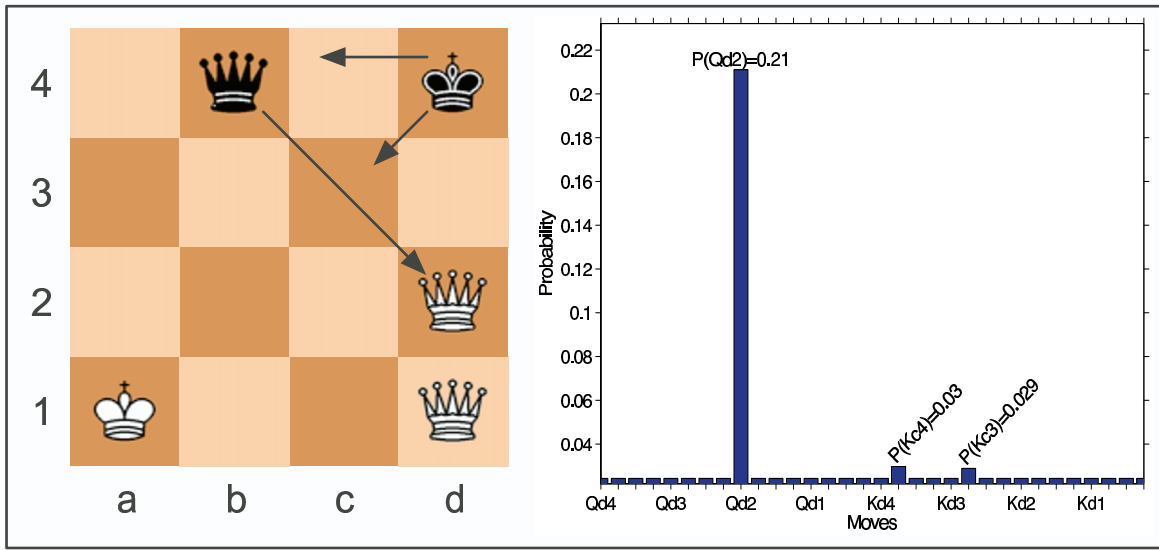


Figure 5: In this scenario, from the most likely to the less, we can see move Qd2 has probability 0.21 while move Kc4 and Kc3 have probability of 0.03 and 0.029. Other moves' probabilities are equal to 0.024.

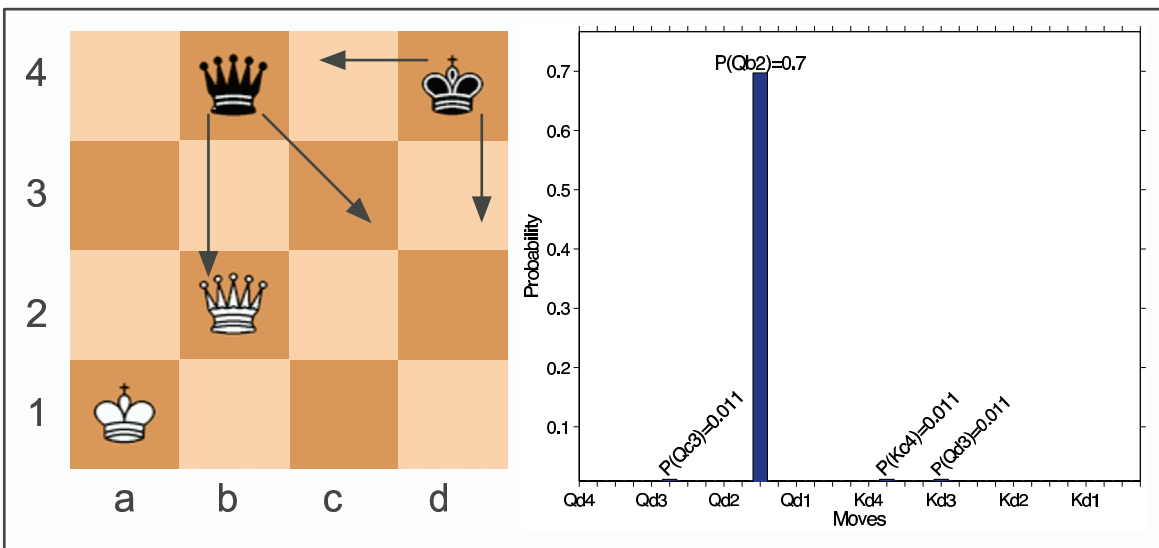


Figure 6: In this scenario, from the most likely to the less, we can see move Qb2 has probability 0.7 while moves Qc3, Kc4 and Qd3 have probability of 0.011. Other moves' probabilities are less than 0.01.

of Artificial Intelligence Research, 24:49–79, 2005.
<http://jair.org/contents/v24.html>.

A. Parker, D. Nau, and V. Subrahmanian. Game-tree search with combinatorially large belief states. *IJCAI*, 2005.

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

S. Russell and J. Wolfe. Efficient belief state and-or search, with application to kriegspiel. *IJCAI*, 2005.

Wikipedia. Kriegspiel (chess) — wikipedia, the free encyclopedia, 2005. [Online; accessed 15-January-2006].