A Speculative Play against Semi-Random Self-Play

Yoichiro Kajihara, Makoto Sakuta and Hiroyuki Iida
Department of Computer Science, Shizuoka University

Jos W.H.M. Uiterwijk and H. Jaap van den Herik
Department of Computer Science
Universiteit Maastricht

Abstract

In previous study we proposed a semi-random self-play, which is a search strategy using random decision and look-ahead search. This paper explores a speculative play against the semi-random self-play. It is a kind of opponent-model search against a player who follows look-ahead search with random decision in some cases. We implemented such a speculative play in our program of TicTacToe, then experiments have been performed. The experimental results confirm the effectiveness of the proposed speculative strategy .

1 Introduction

The opponent-model search (OM-search) for game-playing in two-person games has been investigated, which was followed by a recent study on (D, d)-OM search, which is a speculative strategy while using a model of the opponent, in which difference in search depths is explicitly taken into account[2]. We recognize that OM-search or (D, d)-OM search is a speculative strategy when one has perfect (or reliable) knowledge of the opponent's model (evaluation function, search depth, search strategy and so on). In actual game-playing, it is difficult to know the opponent's model exactly, i.e., a reliable estimate of the search depth and evaluation function of the opponent. One player will only have a tentative model of his opponent, and as a consequence this will lead to a risk if the model is not in accordance with the real opponent's thinking process.

The different models constructed by the semi-random self-play may reflect such actual game playing[1]. For example, experimental results of semi-random self-play using TicTacToe suggest that a player using look-ahead search by six ply even with random decision may correspond to a top-level (or perfect) player, while one-ply look-ahead player with random decision may correspond to a beginner or amateur level, and so on. Our current interest is to consider whether there exists any speculative strategy against a player using look-ahead search with random decision (i.e., semi-random self-play), since in this case it is impossible to gain reliable knowledge of the opponent's model as it is in actual game playing.

In this paper, we first give a short sketch of semi-random self-play. We then propose a speculative play which anticipates the error of the opponent whose search depth is smaller. Some experiments of the speculative play vs. semi-random self-play using TicTacToe are performed and its results are shown.

2 Semi-random Self-play

In this section we give a definition of semi-random self-play. For clarity, the two players in a two-person game are distinguished as a max player and a min player. Let us assume that the max player plays a move at a given position.

Definition 1 *A semi-random self-play is a search strategy to choose a move after look-ahead searching by a given search depth, defined by* **R1** *and* **R2**.

R1: *Generate all possible moves to build a game-tree search of a given height (i.e., search depth). If there is a winning move (by which the max player is able to reach a winning position, irrelevant to the min player's response), then choose it. If not, go to* **R2**.

R2: *Remove all losing moves (after which the min player is able to reach his winning position, irrelevant to the max player's response) from the list of candidates at a position considered. If the list is not empty, select a move among the list at random. Otherwise, select a move at random among all possible moves.*

Following the idea of semi-random self-play, several player's models with different skill can be obtained as a function of look-ahead depth. We then call a player **Pi** who looks ahead by i-th ply while following **R1** and **R2**.

Table 1: Results of semi-random plays in TicTacToe between **Pi** and **Pj**.

W\B	P0	P1	P2	P3	P4	P5	P6
P0	59.06	39.88	8.77	9.00	4.64	4.33	0
	12.74	8.15	22.20	19.98	21.15	20.32	22.63
	28.20	51.97	69.03	71.02	74.21	75.35	77.37
P1	81.59	67.81	13.84	14.36	7.58	7.13	0
	6.18	4.45	20.37	17.86	19.98	19.32	22.63
	12.23	27.74	65.79	67.78	72.44	73.55	77.37
P2	89.28	88.68	31.45	31.13	18.40	16.73	0
	9.07	8.78	51.26	44.82	58.60	55.97	83.60
	1.65	2.54	17.29	24.05	23.00	27.30	16.40
P3	93.93	93.55	52.61	52.52	30.80	29.78	0
	4.95	4.79	34.59	30.45	49.19	46.60	83.60
	1.12	1.66	12.80	17.03	20.01	23.62	16.40
P4	93.97	93.56	52.67	52.24	30.43	29.68	0
	5.09	5.09	37.32	34.37	54.49	50.80	88.67
	0.94	1.35	10.01	13.39	15.08	19.52	11.33
P5	96.51	96.47	76.62	76.53	67.74	67.67	0
	3.09	2.98	20.62	20.30	28.56	26.61	88.67
	0.40	0.55	2.76	3.17	3.70	5.72	11.33
P6	96.60	96.60	77.62	77.62	67.80	67.80	0
	3.40	3.40	22.38	22.38	32.20	32.20	100
	0	0	0	0	0	0	0

The number in each column represents the ratio of the outcome, the upper for the winning ratio, the middle for the draw ratio and the lower for the losing ratio. 'W' means White who is to play first, similarly 'B' means Black who is to play second.

Below we list two important observations we made through experiments of semi-random self-play using TICTACTOE.

- A player who is to play first and looks ahead by a deeper ply outperforms another player looking ahead by a smaller depth [3].
- When one is to play second, there are several exceptions. For example, see P5-P2 (27.30: P5's win) and P6-P2 (16.40: P6's win).

3 A Speculative Strategy

We give the definition of a speculative play against semi-random self-play below.

Definition 2 *Let d_x be a search depth for player X who performs a speculative play proposed. Similarly, let d_y be a search depth for player Y who is to follow semi-random self-pay. Under the condition $d_x \geq d_y + 2$, a speculative play against semi-random self-play is performed by selecting a move which maximizes the ratio of X's winning nodes to the all leaf nodes in each subtree with the root that is a successor at the first ply in a game tree.*
The leaf node is a terminal node or a node that can be expanded.

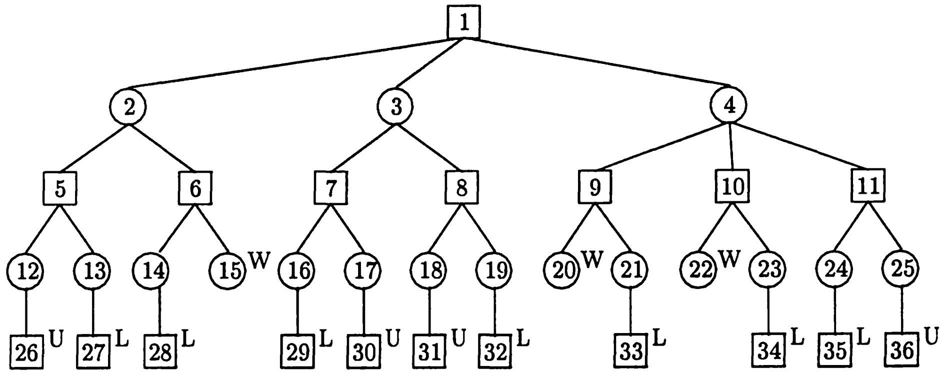

Figure 1: An Example Tree of Semi-random Speculative Play.
'W' represents a win of the max player while 'U' and 'L' represents an unclear position and a loss of the max player, respectively.

We should note that the ratio of X's winning nodes to the all leaf nodes in each subtree is possibly different from X's point of view and Y's, respectively.

Let us show, in Figure 1, an example game tree of speculative play. In this example, the max-player is able to look ahead by four-ply and performs the speculative strategy described in Definition 2, while the min-player is to follow the semi-random self-play with one-ply described in Definition 1. Where, the max player is to move at the root-node position in the game tree. Note that the min-player looks ahead with one-ply at the one-ply nodes (#2, 3, 4) in the game tree of Figure 1, while the max-player can see all nodes in the game tree.

In the game tree of Figure 1, the left-hand subtree contains 1 winning nodes (#15) among 4 leaf nodes (#26, 27, 28 and 15), the center subtree 0 among 4, and the right-hand subtree 2 among 6. Since the ratio of winning nodes to all leaf nodes in the right-hand subtree is larger than other two subtrees, the max player thus selects a move $1 \rightarrow 4$ based on the proposed speculative strategy.

4 Experiments and Results

We have performed experiments of speculative play, in which a computer who follows the speculative strategy described in the previous section plays with a computer using the semi-random self-play over 10,000 games as White and Black respectively. Below we show the results in Table 2. From these data, some observations are made.

- P6's winning ratio in P6 vs. Pi ($i < 6$) is less than P5's one in P5 vs. Pj ($j < 5$). A set of unclear positions which P6 (and P5) might select during the look-ahead is small than that Pj ($j < 5$) might select during his look-ahead. Therefore, there would be less potential to anticipate the opponent's error to gain better results, while he might take less risk.
- Following the speculative play, P5 becomes grade up to the level of P6's semi-random self-play. This is because the possibility of P5's losing disappears when P5 has (not so much but reliable enough) knowledge of the opponent model.

5 Discussion

Let us show, in Table 3, the comparison (increasing ratio) of the winning ratio of the semi-random self-play and the proposed speculative strategy. From these data, we made the following observations.

- In some cases (P4-P2, P2-P4, P6-P2 and P6-P3) the speculative strategy does not improve the winning ratio of semi-random self-play. Here we may give a reason with focus on P4-P2. Even though P2 is unable to notice his loss beyond three ply, two-ply later he has still a chance to prevent from his loss. This means that P4 is anticipating such P2's error but cannot gain it.
- There are some cases in which P6's winning ratio decreases when P6 is to play second.

Table 2: Results of speculative play Pi vs semi-random self-play Pj in TicTacToe.

W\B	P0	P1	P2	P3	P4	P5	P6
P0	59.06	39.88	10.47	9.08	2.70	2.24	0
	12.74	8.15	18.27	9.67	13.61	9.72	17.56
	28.20	51.97	71.26	81.25	83.69	88.04	82.44
P1	81.59	67.81	13.84	12.95	4.92	3.76	0
	6.18	4.45	20.37	8.92	13.12	9.60	17.08
	12.23	27.74	65.79	78.13	81.96	86.64	82.92
P2	89.56	88.68	31.45	31.13	21.95	13.33	0
	9.08	8.78	51.26	44.82	55.04	58.62	86.38
	1.36	2.54	17.29	24.05	23.01	28.05	13.62
P3	97.38	97.35	52.61	52.52	30.80	20.53	0
	1.90	1.33	34.59	30.45	49.19	52.42	85.80
	0.72	1.32	12.80	17.03	20.01	27.05	14.20
P4	97.73	97.33	52.39	52.24	30.43	29.68	0
	1.76	1.71	37.30	34.37	54.49	50.80	85.39
	0.51	0.96	10.31	13.39	15.08	19.52	14.61
P5	99.45	99.36	87.39	87.54	67.74	67.67	0
	0.55	0.64	12.61	12.46	28.56	26.61	88.67
	0	0	0	0	3.70	5.72	11.33
P6	99.54	99.46	87.49	87.27	67.00	67.80	0
	0.46	0.54	12.51	12.73	33.00	32.20	100
	0	0	0	0	0	0	0

The number in each column represents the ratio of the outcome, the upper for the winning ratio, the middle for the draw ratio and the lower for the losing ratio. 'W' means White who is to play first, similarly 'B' means Black who is to play second.

Table 3: The increasing the winning percentage by the speculative play.

	P0	P1	P2	P3	P4
P2	0.28				
	2.23				
P3	3.45	3.80			
	10.23	10.35			
P4	3.76	3.77	-0.28		
	9.48	9.52	0.01		
P5	2.94	2.89	10.77	11.01	
	12.69	13.09	0.75	3.43	
P6	2.94	2.86	9.87	9.65	-0.80
	5.07	5.55	-2.78	-2.20	3.28

The number in each column represents the ratio of the outcome, the upper for the case where the player using a speculative strategy is to play first, the lower for another case (second).

6 Concluding Remarks

There exists a speculative strategy against the semi-random self-play, by which one player gains a better result than non-speculative play. There are a few exceptions where one player has to take a risk while obtaining a worse result than non-speculative play. From the standpoint of non-losing percentage (win + draw), the proposed speculative strategy is better than non-speculative play.

As future works, we will consider the variations of speculative strategy, e.g., one is to improve the winning percentage and another one is to improve the non-losing percentage.

References

[1] Kajihara Y. Sakuta M., and Iida H. Semi-Random Self-Play in Game Playing, to appear in *Proceedings of Joint International Conference on Advanced Science and Technology (JICAST'99).* (1999).

[2] Xinbo G., Iida H., Uiterwijk J.W.H.M. and Herik H.J.v.d. A Speculative Strategy, in Jaap van den Herik and Hiroyuki Iida (Eds.) Proc. Internat. Conf. on Computers and Games, CG'98, *Lecture Notes in Computer Science*, vol.1558, Springer, Heidelberg, pp.74-93 (1999).

[3] Newborn M. A Hypothesis Concerning The Strength Of Chess Programs, *ICCA Journal*, 8(4), pp.209–215. (1985).