福岡工業大学　機関リポジトリ

FITREPO

Title	Proposal and analysis of stochastic variants of Quantum Tic-tac-toe
Author(s)	Makoto Sakuta ; Yusei Tomioka
Citation	福岡工業大学研究論集　第50巻 第1号（通巻77号）P1-P7
Issue Date	2017-10
URI	http://hdl.handle.net/11478/705
Right	
Type	Research Paper
Textversion	publisher

Fukuoka Institute of Technology

Proposal and analysis of stochastic variants of Quantum Tic-tac-toe

Makoto SAKUTA (Department of Information and Systems Engineering)
Yusei TOMIOKA (Graduate School of Information and Systems Engineering)

Abstract

The stochastic variants of Quantum Tic-tac-toe are proposed as more quantum-mechanical metaphor, and the theoretical outcomes as expected values are computed with a stochastic modification of the alpha-beta search. The outcomes vary from complete win to draw with increasing the advantageous probability for the second player. In addition, new zero-sum scoring with a positive score for the second player in a drawn game is introduced and analyzed. The results show the outcome becomes zero when the advantageous probability is around 0.75, which is considered as the evenly-matched interesting settings. Finally, the availability of playing system of the variants is shown.

Key words: *Quantum Tic-tac-toe, stochastic game, expected value, window search*

1　Introduction

Though Tic-tac-toe is a well-known two-person game with perfect information, it is so easy that the upper bound of the state-space size is only $3^9 = 19,683$. However, recently, the game of Quantum Tic-tac-toe (denoted as QT below) based on Tic-tac-toe was introduced by Goff *et al.*[1,2] as a metaphor for quantum mechanics, and analyzed by Ishizeki and Matsuura.[3,4] It is notably interesting that QT is such a relatively complex game that the upper bound of the search-space size is 10^{30}, while the classic Tic-tac-toe is a very simple game the solution of which is easily analyzed.

It should be noted that the rules of QT are elaborately contrived. However, we claim that there exists a reasonable modification of rules. This paper proposes the stochastic variants of QT as more quantum-mechanical metaphor, and analyzes the game characteristics of the variants. In addition, the availability of playing system of the variants is discussed.

2　Original rules and move formalization

2.1　Original rules

Below the rules of original QT are shown with the formulation by Ishizeki *et al.*[4]

Rule 1 (Board)
　　The board of 3 by 3 squares are used. The squares are

Figure 1: Numbering of squares

numbered from one to nine as in Fig. 1.

Rule 2 (Usual move)
　　The first and second players are X and O. Each player plays by turns. At each turn, the player selects two indefinite squares (e.g. the squares x and y) and places "spooky" marks. The move is denoted as x–y, which means the player's actual move is either x or y. The player cannot mark the definitely occupied squares. The game state becomes a composition of actual positions, that is, a *superposition*. The squares that link via spooky marks are called entangled. Entanglement is represented as the form of the graph of marked squares.

Rule 3 (Collapse)
　　Cyclic entanglement is a cycle in the entanglement graph. If there occurs the cyclic entanglement after a move by the player A (X or O), the state is recognized as being in the collapse. Then all of the entangled marks that are connected to the two squares of the last move change into the definite moves. Since there exist two ways for settling the collapse, the superposition splits into two superpositions.

Principal variation:

X1.1-9 O2.1-3 X3.1-3
selA O4.5-7 X5.5-7 selB
O6.2-8 X7.2-4 O8.2-4
selA

Final (super)position: narrow win of the first player

X3	O8	O2
X7	O4	
X5	O6	X1

Figure 2: Principal variation and final (super)position

Rule 3a (Selection)

The opponent of the player A selects one superposition as the game state.

Rule 4 (Special final move)

If the player is to move at the position in which eight squares are definite and only one empty square x is remained, the player makes a move x-x, which is equivalent to a definite move x.

Rule 5 (Game endings)

A winning row is the row of three definite marks of one player either horizontally, vertically, or diagonally. If the player A (X or O) has made a winning row and the other player has not, the game ends in the *complete win* of the player A. If the player has made a winning row and the other player has also made another winning row, through checking the maximum ply of the moves for each winning row, the player that has the smaller maximum ply wins. Then the game ends in the *narrow win*. If all squares are definitely occupied but there exist no winning rows for both sides, the game ends in the *draw*.

Rule 6 (Double win)

The original paper[2] proposed that the case that one player has made two winning rows could be treated specially. We call this case the *double win* and treat it as an optional rule.

2.2 Solution and move formalization

Ishizeki *et al*. showed QT ends in the narrow win of the first player with the initial move 1-9 but did not show the winning line. We have analyzed the games of original rules and obtained one of the longest principal variations shown in Fig. 2.

In this paper, all the moves are normalized in ascending order of squares. Each move in a move line is preceded by the player mark (X or O), the ply of the move, and a dot. Moreover, 'selA' and 'selB' in a move line indicate that the

X1	O2	X3		O2	X3
					X1

Figure 3: Superposition in collapse after X1.1-9 O2.1-3 X3.1-3

X3		O2		O2		X3
		X1				X1

selA   selB

Figure 4: Two possible (super)positions

previous move is settled into the definite move of the first and second square of the spooky marks, respectively. Note that the selection moves, selA and selB, do not increase the ply of the game. For instance, in the above principal variation, the first collapse occurs after X1.1-9 O2.1-3 X3.1-3 shown in Fig. 3. Then, there exist two possible superpositions as the settlement of the collapse, which are shown in Fig. 4. Using the last move X3.1-3 and selA, X3 is settled into the first square of 1-3, that is, the square 1.

3 Claims and new rules with scoring

There exist the reasonable modifications for the rules of QT. Two arguing points are shown below. The first point lies in Rule 3a. Considering the world of quantum mechanics where God does play dice, the concepts of quantum mechanics should be based on the probabilistic distribution and the information should be obtained by observation. Therefore, as a metaphor for quantum mechanics, it is better that the selection is given by the *System*, rather than the player selects one of two superpositions.

The second arguing point is that QT is too advantageous to the first player. Theoretically, QT ends in the narrow win of the first player. Moreover, as for playing by humans, the games are considered very tough for the second player.

Table 1: Four deterministic games

	selection	theoretical outcome
QT (original)	The player that has not caused the collapse selects one superposition.	narrow win of the first player
QT with Rule 3a″	The player that has caused the collapse selects one superposition.	complete win of the first player
SQT $(p, q) = (1, 0)$	The System always selects the superposition advantageous to the first player.	complete win of the first player *
SQT $(p, q) = (0, 1)$	The System always selects the superposition advantageous to the second player.	draw *

∗ Shown in Section 4.3

Table 2: Original scoring

	first	second
draw	0	0
	winner	loser
narrow win	1	1/2
complete win	1	0
double win (proposal)	$1\frac{1}{2}$ or 2	0

Table 3: Zero-sum scoring

	first	second
draw	0	0
	winner	loser
narrow win	10	-10
complete win	20	-20
double win (optional)	40	-40

Table 4: New scoring of drawn games

	first	second
draw	-5	5

The rest of values are the same as those in Table 3.

In respond to the above arguing points, this paper proposes the modification of Rule 3a below.

Rule 3a′ (Probabilistic selection)

If there occurs the cyclic entanglement after the player's move, the System selects one of two superpositions according to the probabilities below.

under the probability p
　　The superposition advantageous to the first player is selected.

under the probability q $(=1-p)$
　　The superposition advantageous to the second player is selected.

The probabilities p and q are adjustable in the range from zero to one. In particular, the random selection is performed at $p=q=0.5$. By changing Rule 3a to 3a′, the game becomes a game of chance, that is, a stochastic game. Below, the stochastic variants of QT are denoted by the Stochastic Quantum Tic-tac-toe (abbreviated as SQT). Since the System is actually the computer program that is able to calculate the values of child nodes, a computer is required to play SQT.

Incidentally, Ishizeki et al. introduced and analyzed QT with the selection in which the player that has caused the collapse selects one superposition (Rule 3a″) in place of Rule 3a, where the game ends in the complete win of the first player.[3,4] Note that the game of SQT at $(p, q)=(1, 0)$ or $(p, q)=(0, 1)$ is not stochastic but deterministic. Here, as the variants of QT, there are four deterministic games as shown in Table 1. All of them have the distinct rules.

Since QT and SQT are the two-person competitive games, it is natural to adopt the zero-sum settings of scoring. The original paper showed the scoring as listed in Table 2 with the proposal that the double win is valued at 1.5 or 2 points. In this paper, the zero-sum scoring is adopted as shown in Table 3. The values are calculated by multiplying the original scores by 40 and balancing the gains and losses, with simply setting the original score of double win at 2 points. Moreover, in order to reduce the advantage of the first player, we propose new scoring of the drawn game shown in Table 4.

4　Analysis of theoretical outcome

4.1　Search algorithm

Since SQT is a two-person zero-sum complete-information stochastic game, the theoretical outcome of a game is given as the expected value. The outcome can be obtained by the perfect lookup, that is, the search of state space. However, differently from the non-stochastic two-person game, the search is not so efficient. Though the special algorithms such as the ∗-minimax search[5,6] are known for searching in stochastic games, the efficiencies are inferior to those of the alpha-beta search in non-stochastic games because the requirement of expected values reduces tree-pruning.

In addition, there exists a difficult point specific to this

game. In a chance node, let p and q be the probability advantageous to the first and second player respectively, and v_1 and v_2 be the values of child nodes after selA and selB respectively, the expected value v is given using either one of the following two formulae.

$$v = p\,v_1 + q\,v_2$$
$$v = p\,v_2 + q\,v_1$$

In order to determine the formula to be applied, the true magnitude relation between v_1 and v_2 is required. This difficulty obstructs the application of the well-studied algorithms such as the $*$-minimax search.

4.1.1　Naive search

The naive method to obtain the true values of v_1 and v_2 is searching with no window as follows, where $\mathrm{search}(s, x, y)$ is the alpha-beta search of the superposition s with the lower bound x and upper bound y, s_1 and s_2 are the superpositions after selA and selB respectively, and LB and UB are the lower bound and upper bound for all nodes respectively.

$v_1 \leftarrow \mathrm{search}(s_1, \mathrm{LB}, \mathrm{UB})$

$v_2 \leftarrow \mathrm{search}(s_2, \mathrm{LB}, \mathrm{UB})$

However, the amount of pruning is considered very small in this method because the searches with no window are performed. Therefore, the high performance of this naive search cannot be expected.

4.1.2　Modified search

Now, in SQT, there exist the branches of two chance moves only when the cyclic entanglement occurs, and the rest can be computed within the framework of alpha-beta search. Making the most of this characteristics, the stochastically modified alpha-beta search is introduced and adopted in this study.

In the alpha-beta search, a node is visited with the lower bound a and the upper bound b. Let the lower and upper bound for searching the next child node be a_c and b_c respectively. Firstly, (a_c, b_c) is assigned to $(-b, -a)$ if using the negamax method, though b_c may be changed in the searches before the target child node. In a chance node, firstly, the values of child nodes after selA and selB, v_1 and v_2, are obtained using the alpha-beta searches with the normal window (a_c, b_c).

$v_1 \leftarrow \mathrm{search}(s_1, a_c, b_c)$

$v_2 \leftarrow \mathrm{search}(s_2, a_c, b_c)$

Let v_f be one of v_1 and v_2 advantageous to the first player, and v_s be the other one advantageous to the second player, the provisional expected value v_0 is given by the formula below.

$$v_0 = p\,v_f + q\,v_s$$

If v_1 or v_2 is out of the window (a_c, b_c), the re-searches with different window are performed under some conditions, the details of which are shown in Table 5. If no re-searches are performed, v_0 is adopted as the return value. Otherwise, v_0 is recalculated using the new v_f and v_s and returned. It is ensured that the expected value returned is one of the following.

$$\begin{cases} \text{a certain value} \leq a_c \\ \text{a true value} \quad > a_c \text{ and } < b_c \\ \text{a certain value} \geq b_c \end{cases}$$

Though the proposed method performs at most two re-searches, it is expected that the total efficiency is enhanced because most searches have the small window.

4.2　Experimental settings

The basic part of the search in this study is a simple alpha-beta search, which does not use an evaluation function of non-terminal nodes. The same as the search method in the previous study[4], considering the board symmetry, the initial moves are restricted to the following eight moves: 1-2, 1-3, 1-5, 1-6, 1-9, 2-5, 2-6, and 2-8.

First, without and with double-win option, increasing the probability q from zero to one with the step 0.05, the outcomes were computed for all restricted initial moves. In addition, using the new scoring of drawn games, the outcomes were computed with varying the probability q similarly.

Secondly, in order to check the search efficiencies, the number of searched nodes and the elapsed time were measured for solving SQT at $q = 0.75$ (a typical probability advantageous to the second player) and $q = 0.05$ (the most time-consuming probability in the above experiments). For comparison, the similar data were measured for solving
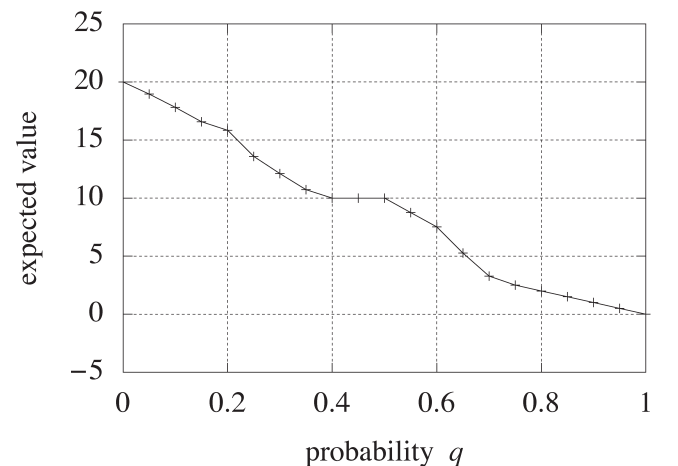

Figure 5: Expected values (double win: no, scoring: Table 3)

Table 5: Conditional re-search

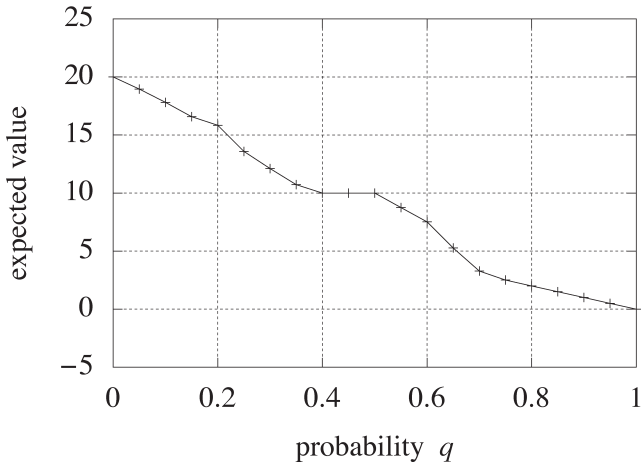|	$v_1 \leq a_c$	$a_c < v_1 < b_c$	$v_1 \geq b_c$
$v_2 \leq a_c$	—	if $v_0 > a_c$ $v_2 \leftarrow \text{search}(s_2, \text{LB}, v_2)$	$v_1 \leftarrow \text{search}(s_1, v_1, \text{UB})$ $v_2 \leftarrow \text{search}(s_2, \text{LB}, v_2)$
$a_c < v_2 < b_c$	if $v_0 > a_c$ $v_1 \leftarrow \text{search}(s_1, \text{LB}, v_1)$	—	if $v_0 < b_c$ $v_1 \leftarrow \text{search}(s_1, v_1, \text{UB})$
$v_2 \geq b_c$	$v_1 \leftarrow \text{search}(s_1, \text{LB}, v_1)$ $v_2 \leftarrow \text{search}(s_2, v_2, \text{UB})$	if $v_0 < b_c$ $v_2 \leftarrow \text{search}(s_2, v_2, \text{UB})$	—

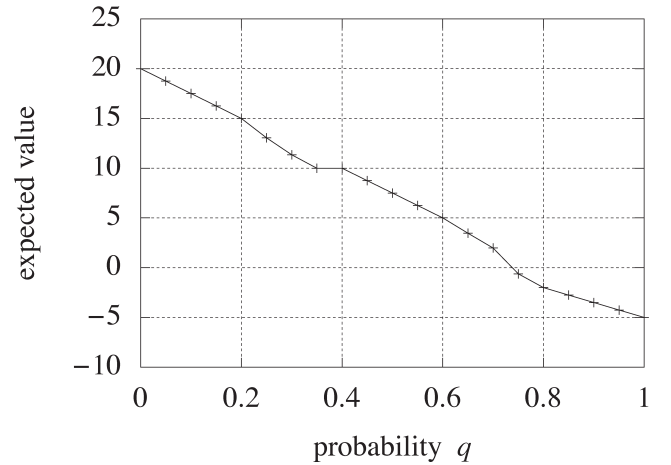Figure 6: Expected values (double win: yes, scoring: Table 3)

Figure 8: Expected values (double win: no, scoring: Table 4)

QT.

The experiments were performed on a computer with Intel Xeon 3.5 GHz under CentOS 7. Programs written in Java were executed using OpenJDK.

4.3　Results and discussions

The results using the scores in Table 3 without and with double-win option are shown in Fig. 5 and Fig. 6. The outcomes are: starting with the expected value of 20 at $q=0$, decreasing as q gets larger, ending with the expected value of 0 at $q=1$. For the perfect-lookup players, the first player has the advantage over the second player unless $q=1$. The game at $q=1$ is considered evenly-matched, but the game is not stochastic.

The outcomes at $q=0$ without and with double win are both exactly 20, which means the complete win of the first player. Thus the contribution of double win cannot be found even in the completely advantageous settings for the first player. Since the graphs in Fig. 5 and Fig. 6 are nearly the same, the influences of the double-win option are very subtle for the perfect players. However, the situation may be different for human players, since the double win is considered achievable only when the second player makes

blunders.

For each restricted initial move, the outcomes from $q=0$ to $q=1$ are computed. The results without double-win option are shown in Fig. 7. It is interesting that the order of initial moves changes depending on the q value. In addition, note that the expected values of the initial moves except 1-3, 1-5, and 1-9 go negative as q gets larger. It is expected that these characteristics enhance the interest of playing for humans.

The results using the new scores of drawn games in Table 4 without double-win option are shown in Fig. 8. The advantage of the first player is weakened and the expected value reaches zero around $q=0.75$, which is considered as the evenly-matched game for the perfect players. In addition, the similar results are obtained in the case with double-win option.

The number of searched nodes and the elapsed time for solving SQT at $q=0.75$ and $q=0.05$, and solving QT are shown in Table 6. Though comparing the data of QT and SQT is not so appropriate because there is essential difference between the search in the non-stochastic game and that in the stochastic game, the node counts in SQT seem to be several times larger than those in QT because of less prun-
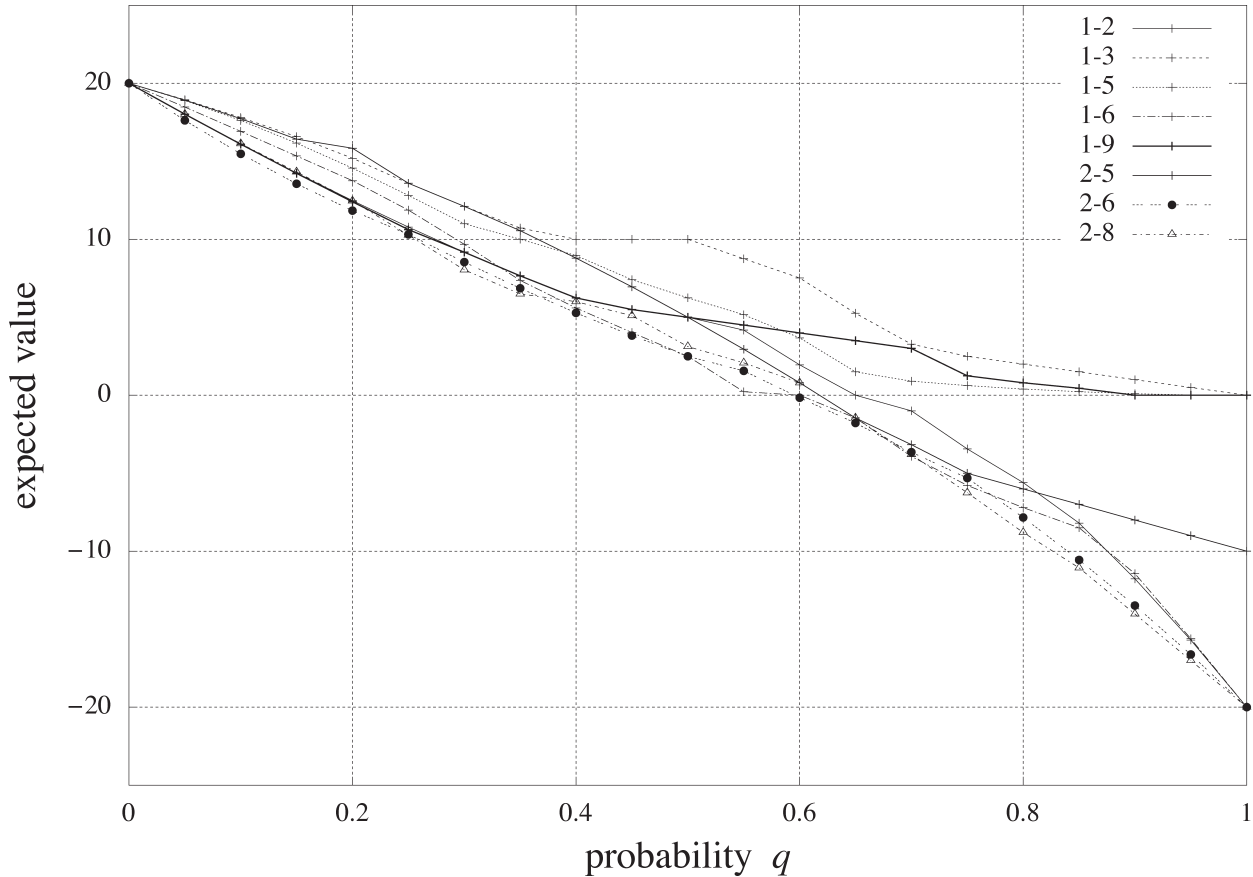
Figure 7: Expected values grouped by initial moves (double win: no, scoring: Table 3)

ing.

As for the difference between the naive search and the modified alpha-beta search in SQT, the node count in the modified method is about 3% reduced compared with that in the naive method at $q=0.75$, while 0.5% increased at $q=0.05$. These data indicate the trade-off between the losses of re-search costs and the gains by small-window searches. Nevertheless, the elapsed times are over 20% reduced using the modified method. This shows the effectuality of the modified method introduced.

5 Availability of the playing system of SQT

In SQT, when the collapse occurs, the System requires the information about which selection is better for the player. However, fairly hard computing is required to obtain the theoretical values of both child nodes, especially in the short-ply superpositions. This could obstruct the quick responses of the playing system for human players.

As the preparatory experiment, the child-node values, v_1 and v_2, were obtained for all the 2, 3, 4-ply collapses, with varying the probability q from zero to one with the interval 0.05 (double win:no, scoring: Table 3). When the time

threshold is set at 0.5 seconds, there are around 8,000 collapses (from 1,846 to 10,346 collapses depending on the value q) at which the time for calculating the child values exceeds the threshold. The maximum times for obtaining v_1 and v_2 at 2, 3, 4-ply collapses are 83.4, 15.5, and 2.4 seconds respectively. Considering the decreasing rate, we do not require the data of collapses at the plies over four because calculating v_1 and v_2 should be sufficiently quick. From the above, if the system utilizes the collapse database of child values at 2, 3, 4-ply collapses and the interpolation if necessary, the response of the system can be sufficiently quick for human players.

Incidentally, since SQT requires the computer system, playing SQT is not possible only by humans. However, using one or two dice to decide which player selects one of two superpositions, we can play the games roughly corresponding to SQT.

6 Concluding remarks

The stochastic variants of Quantum Tic-tac-toe are proposed as more quantum-mechanical metaphor, and the theoretical outcomes as expected values are computed by the

Table 6: Node count and elapsed time

	QT	SQT ($q = 0.75$)		SQT ($q = 0.05$)	
		naive	modified	naive	modified
node count	87,833,388	140,389,750	137,103,153	355,203,626	357,285,830
elapsed time (sec)	533	505	377	1,307	1,007

stochastically modified alpha-beta search. Using the new zero-sum scoring with a positive score for the second player in a drawn game, the outcome becomes zero when the advantageous probability is around 0.75, which is considered as the evenly-matched interesting settings. Moreover, by setting the probability appropriately, we can play games with variety of advantages or handicaps to both players. It is noted that human players may feel somewhat different from the theoretical data for the perfect players.

Though the stochastically modified alpha-beta search has reduced the solving time to a certain degree, it is considered that there should be some room for improvements. For instance, if we use the good evaluation function for non-terminal nodes, the search efficiency could be enhanced.

The actual playing system of SQT should be available with utilizing the collapse database, which is the remained subject for the near future.

References

1 ） Allan Goff, Dale Lehmann, and Joel Siegel. Quantum Tic-Tac-Toe, spooky-coins & magic-envelopes, as metaphors for relativistic quantum physics. AIAA Paper 2002-3763, American Institute of Aeronautics and Astronautics, 2002.

2 ） Allan Goff. Quantum tic-tac-toe: A teaching metaphor for superposition in quantum mechanics. *American Journal of Physics*, 74(11): 962–973, 2006.

3 ） Takumi Ishizeki and Akihiro Matsuura. Analysis of Quantum Tic-Tac-Toe. In *The 15th Game Programming Workshop 2010*, pages 101–107, 2010. (in Japanese, with English abstract).

4 ） Takumi Ishizeki and Akihiro Matsuura. Solving Quantum Tic-Tac-Toe. In *Proc. of the International Conference on Advanced Computing and Communication Technologies (ACCT 2011)*, pages 330–334, 2011.

5 ） Bruce W. Ballard. The ＊-minimax search procedure for trees containing chance nodes. *Artificial Intelligence*, 21(3): 327–350, 1983.

6 ） Thomas Hauk, Michael Buro, and Jonathan Schaeffer. Rediscovering ＊-minimax search. In *Computers and Games CG 2004 Revised Papers*, pages 35–50, Berlin, 2006. Springer-Verlag.