Learning to Control a Blackboard System for Game Playing*

Susan L. Epstein
Department of Computer Science
Hunter College of the City University of New York

Abstract

HOYLE is a program designed to learn to play any two-person, perfect information game. HOYLE derives much of its power from a set of knowledge sources, called Advisors, that recommend and advise against moves and plans during play. This paper describes HOYLE's task, the nature of its Advisors, its blackboard architecture, and its control mechanisms. The discussion highlights important issues in learning to control such a blackboard system. At its current stage of development, HOYLE learns to control the blackboard to play seven games expertly. Empirical results are provided on five of these games for four different control paradigms.

The AI approach to game playing has thus far been to build a program that plays only a single game, and plays it very well. (For example, Samuel, 1963, 1967; Rosenbloom, 1982; Berliner, 1980; Anantharaman, Campbell, & Hsu, 1990; Chen, 1989.) HOYLE, the program under development described here, differs from traditional game playing programs in three ways:

- it is able to play any of a broad class of games correctly, i.e., according to the rules
- it improves its performance through a variety of learning paradigms
- it avoids extensive forward search into the game tree.

Instead of search, HOYLE uses a broad range of instantiable, game-independent approaches to the move selection problem. HOYLE's initial success is directly attributable to expert knowledge sources, called Advisors, and their interface through a blackboard system (Nii, 1986). The blackboard uses a variety of control mechanisms to integrate these diverse problem solvers into a common framework. The first section of this paper describes HOYLE's domain and task. Section 2 sketches HOYLE's knowledge sources; Section 3 explains the blackboard architecture and its control mechanisms. Section 4 highlights important issues in learning to control such a blackboard system, and provides

---

* This work was supported in part by NSF DCR-8514395, NSF 9001936, and PSC-CUNY 668287.

empirical results for four control paradigms. The final section summarizes the current results.

1. The Task

Informally, a *game* consists of some material (e.g., a board, playing pieces), a roster of participants, and a list of rules. One complete experience at a game, from the initial board state to the termination of the process, is called a *contest*. The rules describe how the participants are to take turns affecting the position and/or status of the material (e.g., removing pieces from the board or changing their location) until some rule terminates the contest. When the contest terminates, the rules declare which participant has won or that the contest was a draw.

For the games in HOYLE's domain, there are three fundamental assumptions. First, there are exactly two participants: *Player* (the one who moves first) and *Opponent*. Player and Opponent are the *roles* in a two-person game. Second, all information about the game is disclosed and equally available to all the participants (e.g., no closed hands, no uncertain outcomes as with dice), i.e., it is a *perfect information* game. Finally, the game can be represented as a finite, directed graph (the *game graph*) in which a node (*state*) both identifies the participant whose turn it is (the *mover*) and describes a possible arrangement of the material. The participant who is not the mover in a state is referred to here as the *non-mover*.

An expert at a game should be able to take the role of either Player or Opponent and consistently maximize that role within the game graph. Consistency is important; the expert must perform well over a long sequence of contests (a *tournament*). This does not mean that the expert should always win, or even never lose; many popular games (e.g., tic-tac-toe) result in a draw when played by experts. Rather, an expert should consistently win when possible, failing that draw, and lose as rarely as the nature of the game permits. In addition, in games where there is a final score for each participant, an expert should win by a large margin, and lose by a small one. This definition of expertise deliberately ignores the skill level of any other participant, and instead defines the strength of a strategy with respect to the game itself. *True expertise* is perfectly backed up knowledge from the entire game graph.

HOYLE is intended to solve several different kinds of problems. On the highest level, HOYLE is a program to learn to play any two-person, perfect information game. The program learns about an individual game by playing contests of it, against either HOYLE itself, a programmed expert, or a human participant. HOYLE has two goals: to learn as much as possible about the nature of the game, and to play expertly. During a contest, HOYLE is repeatedly faced with the task of selecting its next move. The smallest unit for solution elements is a move; moves may be organized into plans, and future research will organize the plans into higher-level structures. Implicitly, a solution interval is a contiguous portion of a contest.

The input to HOYLE for a game is a set of procedures that define the game by describing and displaying its material and rules. The program has no heuristics for symmetry. HOYLE's task is to learn to play each game very well, ideally with true expertise.

2. The Advisors

HOYLE was originally inspired by the team of consultants that chess grandmasters sometimes confer with during important matches. These people discuss and analyze a contest with the grandmaster; their role is to provide a variety of perspectives. Although each of them has studied chess for many years and they share much common knowledge, often one consultant will see a contest differently. The consultants are collaborating experts; they share and argue their views with the expectation that they will improve performance both in the specific contest and at the game in general.

Like the grandmaster, HOYLE has a panel of knowledge sources, called *Advisors,* that participate in play. No Advisor is a full-fledged consultant, however. Instead, each one epitomizes a different, specialized perspective on game playing in general, a weak theory for the domain. An Advisor takes a fairly narrow, but rational, view of the move selection problem, one found generally applicable by experienced game players. Each Advisor is implemented as a procedure with a time limit. An Advisor can post only plans (sets of moves) or opinions about moves on the blackboard. Advisors have access to each other's plans and opinions only through the blackboard. At the moment there are 20 game-independent Advisors in various stages of implementation. Their names are descriptive: Victory, Panic, Sadder, Wiser, Don't-Lose, Enough-Rope, Pitchfork, Open, Candide, Worried, Shortsight, Sneak, Feinter, Anthropomorph, Not-Again, Cyber, Spotter, Challenge, Greedy, and Wild.

Between contests of a particular game, HOYLE seeks and selectively caches information from play traces to aid move selection. HOYLE makes this information available to its Advisors in libraries that are gradually instantiated based on its experience. Thus an Advisor may make different comments about a state in a later contest from the comments it made about the same state in an earlier contest. (See (Epstein, 1989) for a more detailed discussion.)

3. The Blackboard, Priority Classes, and Control

HOYLE's blackboard contains an encoding of the current game board, the identity of the mover, and the HOYLE libraries for the game. When it is HOYLE's turn to move in a contest, one of the game-defining procedures computes all the legal moves from the current state and posts them on the blackboard. Once the legal moves are posted, Advisors may be called upon to post plans (sets of moves) and comments about moves on the blackboard.

For any state where HOYLE is mover, an Advisor may make any number of comments, each with a strength from 0 (adamant opposition) to 10 (insistent support). Comment strengths are only intended to reflect relative value within an individual Advisor's expertise; they are not intended to reflect any absolute standard with respect to the opinions of other Advisors. Consider, for example, the state in Figure 1 from a tic-tac-toe contest where HOYLE as Player (X's) is the mover. The legal moves from this state are the squares 2, 6, 7, 8, and 9. Each Advisor evaluates these alternatives independently, from its own narrow perspective. The Advisor Panic looks to see whether the non-mover has a sure win on his next move. In Figure 1, Panic would insist that a move in square 6 is the only way to prevent a win by Opponent. This comment could cite as support the threat of the O's in

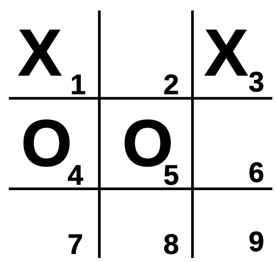

Figure 1 . A state in a tic-tac-toe trial. Player (X's) is mover. The Advisors are called upon for move recommendations on this state.

squares 4 and 5, and would have maximum strength, since a win by a reasonably observant Opponent across the second row on the next turn is virtually certain. Another Advisor, Worried, looks for longer-range wins for the non-mover. In Figure 1, Worried would suggest moves in both square 2 and square 8, since Opponent could win in the second column after two turns. Worried also suggests moves in squares 7, 8, and 9, to prevent Opponent's eventual win in the third row. The comments about the third row would have a lesser strength than those about the second column because they project a longer contest. Victory is another Advisor; it looks for a certain win for itself on the current move. In Figure 1, Victory would insist that a move in square 2 is the best choice, because Player will win immediately. Some of the more sophisticated Advisors employ iterative deepening, selective extension (Anantharaman, Campbell, & Hsu, 1990), and the null move heuristic (Goetsch & Campbell, 1988), recommend plans, suggest psychological gambits, and defend against primitive models of the opposing participant.

As in the example of Figure 1, the Advisors rarely agree, but HOYLE must select a move based upon the advice they post on its blackboard. (Like a person, HOYLE need not have a particular strategy, or even a plan, when it makes a move.) HOYLE's control structure has two segments: a weak theory for the domain of two-person, perfect information games and a learning module.

One aspect of HOYLE's weak theory recognizes that all Advisors are not created equal, i.e., that in all games some Advisors are always more important than others. The Advisors are therefore grouped into priority classes called *tiers*. When it is HOYLE's turn to move, it consults the tiers in turn. If any tier produces an absolute recommendation or reduces the posted legal moves to a single one, none of the subsequent tiers will be consulted. This *hierarchical consultation system* encapsulates general knowledge about game playing, knowledge that does not have to be relearned for each game. The tiers capture such commonsense rules as "If you can move to a win immediately, do it" and "If you cannot win now and are about to lose on the other participant's next turn, block the move that threatens you."

The classification of an Advisor is determined by the nature of its problem solving expertise. Tier 1 Advisors must always be quickly and absolutely correct in their opinions. This restricts them to the ability to win (Victory) or avoid sure loss (Don't-Lose) on a single move, or to recognize the current state and recall its certain outcome from previous experience (Sadder and Wiser). Tier 2 Advisors are permitted somewhat more time in which to produce absolutely correct opinions, based on a two-ply lookahead in the game tree. They search for nearby wins and losses (Panic, Shortsight), or leave opportunity for the other participant to err (Enough-Rope). The difference between Tier 1 and Tier 2 is primarily in the resource requirements of the individual Advisors. Within both of these tiers the Advisors are consulted in a fixed, game-independent order that is part of HOYLE's weak theory.

The control learning module refers only to the broad spectrum of heuristic Advisors in Tier 3, whose relationships may not be game-independent. Some of the Advisors (e.g., Pitchfork) may not be relevant to every game. To speed the construction of their opinions, some Advisors (e.g., Candide) deliberately overlook data that should be relevant. Some are partially (e.g., Pitchfork) or completely (e.g., Opening) dependent on the game libraries. All of the Tier 3 Advisors have the opportunity to offer opinions before any decision is made. At this time, control learning is directed only to move selection, as discussed below.

Recall that an Advisor can make one or more comments on a state, that a comment can be for a move or against it, and that every comment has a weight associated with it. For any given state, denote the opinion of Advisor A about move m as A(m), where

$$0 \leq A(m) \leq 10.$$

A *smoothed* comment, denoted as $\sigma(A(m))$, has intent rather than strength. This transformation is defined by

$$\sigma(A(m)) = \begin{cases} -1 \text{ if } A(m) < 5 \\ 0 \text{ if } A(m) = 5 \\ +1 \text{ if } A(m) > 5 \end{cases}$$

Smoothing is a form of scaling that permits meaningful summation of comments on a move, e.g., a group of comments against a move is expressed as the sum of negative numbers.

If a move has not been chosen prior to Tier 3, all the advice from the tier is combined, according to the frequency and weight of the comments for and against each move. HOYLE currently has four paradigms to integrate the expertise available on the blackboard in Tier 3.

• *Paradigm 1:* Every Advisor expresses every comment with its associated strength. The selected move m is the one whose A(m) sum across the Advisors is a maximum, i.e.,

$$\max_m \left[\sum_A A(m) \right]$$

• *Paradigm 2*: All comments are smoothed, i.e., are -1, 0, or +1. The selected move m is the one whose $\sigma(A(m))$ sum across the Advisors is a maximum, i.e.,

$$\max_m \left[\sum_A \sigma(A(m)) \right]$$

• *Paradigm 3:* Every Advisor A gets a single weighted vote v(A). For each Advisor A, comments A(m) for all moves m are ranked by strength and the *strongest* (furthest from 5 in absolute value) is selected. If there is more than one strongest comment for an Advisor, one is chosen at random as v(A). The selected move is the one whose v(A) sum across A is a maximum, i.e.,

$$\max_m \left[\sum_A A(m)\chi_m(v(A)) \right] \text{ where } \chi_m(v(A)) = \begin{cases} 1 \ \text{ if } v(A) = m \\ 0 \ \text{ otherwise} \end{cases}$$

• *Paradigm 4:* Every Advisor A gets a single smoothed vote $\sigma(v(A))$. For each Advisor A, comments A(m) for all moves m are ranked by strength; a strongest is then selected and smoothed. If there are multiple top-ranking comments for an Advisor, one is chosen at random as $\sigma(v(A))$. The selected move is the one whose $\sigma(v(A))$ sum across A is a maximum, i.e.,

$$\max_m \left[\sum_A \sigma(A(m))\, \chi_m(v(A)) \right] \text{ where } \chi_m(v(A)) = \begin{cases} 1 & \text{if } v(A) = m \\ 0 & \text{otherwise} \end{cases}$$

Each of these approaches has some justification. Paradigm 1 thoroughly considers every carefully constructed expert opinion. The Advisors, however, do not purport to separate game-playing knowledge into mutually exclusive classes. To the extent that the Advisors overlap, they may comment the same move for the same underlying reason; Paradigm 1 will overemphasize any such intersection in expertise. Paradigm 2, although it retains the multiplicity of comments from Paradigm 1, is deliberately formulated to ignore the comments' strengths. Advisors can only support (+1), oppose (-1), or be indifferent to (0) a move. Although the strengths constructed within a single Advisor carry important intra-Advisor knowledge (e.g., threats further away are of less concern to Worried than closer ones), there is no reason to expect that they will scale properly for inter-Advisor debate in every game. Thus Paradigm 2 captures only the overall sense the Advisors have about each move. Paradigm 3 retains the comment strengths, but forces each Advisor to express a single preference. This approach is intended to minimize the potential distortion in Paradigm 1 due to overlap of expertise, but it creates another difficulty: when an Advisor has both positive and negative advice to give, one kind of comment is certain to be discarded. Finally, Paradigm 4 restricts each Advisor to a single comment with direction rather than strength. This approach smooths what may be inaccurately scaled and repetitive data, and constructs a coarse survey of the Advisors.

4. Learning Control

The games in HOYLE's current testbed are culturally diverse and quite varied: tic-tac-toe, lose tic-tac-toe, two versions of three-dimensional tic-tac-toe, tsoro yematatu (Zaslavsky 1982), pong hau k'i, and achi (Bell 1969). They were chosen because they span a variety of cultures, and therefore probably capture some aspects of game playing that people find particularly intriguing. Their game graphs lack the complexity of chess or Go, but some of them offer the challenge of cycles and stage transitions, and at least one of the game graphs has over a billion nodes.

HOYLE learns to play well through competition in a tournament against an expert for each game. Many of HOYLE's Advisors can be instantiated by this play, i.e., the program selectively extracts knowledge from the contests and stores it for their subsequent reference. HOYLE's initial research (Epstein, 1989) used Paradigm 1 for control of Tier 3 in all its tournaments, and learned to play all seven games extremely well.

4.1 Issues

Recent research has identified three significant issues in learning to control HOYLE's blackboard:

• *What is the best control strategy for learning?* One of the paradigms might learn faster than another, or not at all. There may, for example, be hidden biases in the assigned weights that favor one Advisor over another once the comments are forwarded to the blackboard.

• *Is there a difference between a good control strategy for learning a task and a good control strategy for executing that task?* Once the program has extracted knowledge from a tournament and has learned to play very well under one paradigm, another paradigm might then play as well or even better. This is particularly likely if the learning was incomplete initially, i.e., HOYLE did not learn to play perfectly.

• *In a broad domain, does the relative significance of expert perspectives vary from one instance to another?* For specific games, some Advisors' comments may be more significant than others'. Perhaps a comment should be weighted not only by the strength its Advisor assigns, but also by the value that Advisor has had in other contests of the same game. (Note that this is different from "other contests at any game," which formulated the tiers.) Such credit assignment has ample precedent in the machine learning literature (See, for example, Langley, 1985; Mitchell, Utgoff, & Banerji, 1983; Pazzani, 1987.)

The empirical study described below has addressed these questions for five games: tic-tac-toe, lose tic-tac-toe, achi, pong hau k'i, and tsoro yematatu.

4.2 Experiment 1

To address the first issue, that of the best control strategy for learning, each of the paradigms was used to learn against an *expert program,* one that offers a randomized variety of high-quality play without exhaustive search. Because HOYLE learns by selective memorization and generalization of its experience, there is no way to determine at what point all the knowledge to play a particular game with true expertise has been acquired. A behavioral standard is set instead: for the purpose of this experiment, a game is considered *learned* when the outcome in 10 consecutive contests against an expert program matches true expertise. Because the expert program is non-deterministic, learning speed may vary. Therefore, five runs for each paradigm are averaged to provide a measure of learning speed. (Examination of HOYLE's libraries indicates that under different paradigms, and even on different runs, different knowledge is acquired for the same game.) The results of Experiment 1 appear in Table 1.

	Tic-tac-toe	Lose tic-tac-toe	Achi	Pong hau k'i	Tsoro yematatu
Paradigm 1	21.6	73.2	10.0	10.0	10.0
Paradigm 2	19.8	143.2	25.2	10.0	10.0
Paradigm 3	28.6	230.2	10.0	10.0	10.0
Paradigm 4	24.2	211.2	16.0	10.0	10.0
Average	23.6	164.45	15.3	10.0	10.0

Table 1 . Experiment 1. Number of contests HOYLE played until it learned each game under four different control paradigms. A game is considered learned when HOYLE plays as an expert would in 10 consecutive contests.

Clearly some paradigms are far faster at learning these games than others. In particular, Paradigm 1 seems to be a generally effective mechanism for learning these five games. Table 1 also demonstrates that the learning of even a relatively easy game can be notably delayed by the choice of a control strategy. (Note the performance for achi under Paradigm 2, for example.) An unanticipated result was the ability to distinguish *difficult*, i.e., requiring on average many contests to learn, games from easy ones. Although the researchers intuitively knew which games were more difficult, particularly without knowledge of symmetry, Experiment 1 confirms that lose tic-tac-toe is by far the most difficult to learn of the five.

4.3 Experiment 2

To study the possibility that the best learning strategy may not be the best playing strategy, two versions of HOYLE using different control paradigms are pitted against each other, in tournaments of 100 contests each, for every game under consideration. Because this experiment is intended to measure post-learning performance, both versions share a common knowledge base, one acquired during the first experiment, but left unchanged by experience in the second experiment. The results appear in Tables 2 through 6. The entry "a, b, c" for the row labeled i and the column labeled j means that Paradigm i won a matches, drew b matches, and lost c matches against Paradigm j. For example, for two perfect players in a game where true expertise draws, the entry should be 0, 100, 0.

Experiment 2 demonstrates that a strategy apparently successful in the first experiment may still have much to learn in the second, i.e., may not yet be a true expert. Inspection of the traces reveals that such uneven play is caused by paradigms whose decisions force the exploration of different, possibly less promising, regions of the game tree. For example,

Tic-tac-toe	**Paradigm 2**	**Paradigm 3**	**Paradigm 4**
Paradigm 1	0, 100, 0	2, 92, 6	4, 95, 1
Paradigm 2	-	3, 94, 3	1, 97, 2
Paradigm 3	-	-	4, 93, 3

Table 2 . Experiment 2. Testing playing control strategies against each other in tic-tac-toe.

in its learning phase against the expert program in tic-tac-toe, Paradigm 1 only encounters openings in the center or in a corner. When confronted by another paradigm in this experiment with a different opening, a weak opening the training program and human experts would avoid, Paradigm 1 loses. Under normal circumstances, HOYLE would learn from each loss, but in this experiment, where relative strengths are tested, the numbers gauge post-learning robustness in the face of imperfect opposition. In tic-tac-toe, the various versions of HOYLE present these new, potential, but unrealized, learning experiences to each other fairly infrequently.

Lose tic-tac-toe	**Paradigm 2**	**Paradigm 3**	**Paradigm 4**
Paradigm 1	45, 17, 38	46, 11, 43	41, 14, 45
Paradigm 2	-	47, 16, 37	33, 22, 45
Paradigm 3	-	-	33, 24, 43

Table 3. Experiment 2. Testing playing control strategies against each other in lose tic-tac-toe.

The extreme results for lose tic-tac-toe in Table 3, however, where only 17% of all contests end in a draw as perfect play would, suggest that the behavioral standard for learning was insufficient, i.e., that Experiment 2 was run prematurely. Lose tic-tac-toe demands a complex strategy that, without knowledge of symmetry, requires a tournament experience of at least 90 contests. The non-deterministic nature of the learning process clearly did not provide HOYLE with enough experience; it had played almost 200 contests before Experiment 2 began, but they had not provided all the data HOYLE needed to play expertly.

Pong hau k'i, on the other hand, is quite a simple game for HOYLE. Its small cyclic game graph is readily dealt with, typically by the Advisors. Table 4 indicates that HOYLE found nothing more to learn about pong hau k'i in this experiment.

Pong hau k'i	**Paradigm 2**	**Paradigm 3**	**Paradigm 4**
Paradigm 1	0, 100, 0	0, 100, 0	0, 100, 0
Paradigm 2	-	0, 100, 0	0, 100, 0
Paradigm 3	-	-	0, 100, 0

Table 4. Experiment 2. Testing playing control strategies against each other in pong hau k'i.

Table 5 shows that achi, whose first stage is quite similar to tic-tac-toe, also had an incomplete learning experience in Experiment 1. In particular, the programmed achi expertalways opens in the center; only Paradigms 3 and 4 risk the other openings, to the surpriseof Paradigms 1 and 2. The results are not as dramatic for those of lose tic-tac-toe, but not as robust as those for tic-tac-toe.

Achi	**Paradigm 2**	**Paradigm 3**	**Paradigm 4**
Paradigm 1	0, 100, 0	12, 72, 16	19, 61, 20
Paradigm 2	-	9, 74, 17	16, 73, 11
Paradigm 3	-	-	20, 64, 16

Table 5 . Experiment 2. Testing playing control strategies against each other in achi.

Table 6, for tsoro yematatu, is identical to that for pong hau k'i. This was an unanticipated result, because most people find this a far more challenging game than pong hau k'i. The table indicates that, once learned, any of the paradigms is as good as any other for tsoro yematatu.

Tsoro yematatu	**Paradigm 2**	**Paradigm 3**	**Paradigm 4**
Paradigm 1	0, 100, 0	0, 100, 0	0, 100, 0
Paradigm 2	-	0, 100, 0	0, 100, 0
Paradigm 3	-	-	0, 100, 0

Table 6 . Experiment 2. Testing playing control strategies against each other in tsoro yematatu.

In general, results in Experiment 2 may be somewhat exaggerated by the fact that learning is turned off during these tests. Thus the various versions of HOYLE may be repeatedly confounded by weaker playing that they would, under normal circumstances, quickly learn to combat.

4.4 Experiment 3

The third issue, whether expert perspectives have different relative importance in different games, is more complex. Only a simple study of it is offered here. Using the traces of the contests from the first experiment, HOYLE reviews its cached knowledge and determines which of the 13 Advisors in Tier 3 would have made any meaningful contribution to successful play. In tic-tac-toe, for example, Candide, Worried, and Greedy ranked as the most consistently *relevant* (having a comment to make) and *significant* (consistently agreeing with the move choices of the non-losing participants). After the first experiment, each of the paradigms challenges the expert program once again, but this time with the irrelevant Advisors eliminated and the comment weights of the three most significant ones doubled. Table 7 displays the results of such rematches in tournaments of 20 contests each against the same expert programs. HOYLE was permitted to learn during these tournaments. Elimination of irrelevant Advisors provided the expected substantial computational speed-up, since each Advisor has its own time allotment. In most cases, doubling of the comment weights appears, to have no significant post-learning impact on play performance against an expert. In lose tic-tac-toe, however, under Paradigm 4 HOYLE appears to lose too much information, either because the game is difficult or because it is partially learned. Traces indicate that under Paradigm 4 in this experiment for lose tic-tac-toe HOYLE attempted a variety of unusual openings that necessarily led to its defeat.

	Tic-tac-toe	**Lose tic-tac-toe**	**Achi**	**Pong hau k'i**	**Tsoro yematatu**
Paradigm 1	20	20	20	20	20
Paradigm 2	20	20	20	20	20
Paradigm 3	18	20	18	20	20
Paradigm 4	20	13	20	20	20

Table 7 . Results of Experiment 3. After learning, the number of contests HOYLE played as an expert would have in 20 consecutive contests against an expert program.

The overall strength of the results in Table 7, and the variation in the choice of most significant Advisors from one game to the next, supports the following inductive statements about the fundamental nature of these games. For tic-tac-toe the most significant are Candide, Worried, and Greedy. This suggests that tic-tac-toe can be played naively: attempt to forward (Candide) multiple (Greedy) plans and block (Worried) simple ones, without elaborate strategic contingencies. Even the opening is properly dictated this way. For lose tic-tac-toe, pong hau k'i, and achi, the three most significant Advisors are

Pitchfork, Open, and Shortsight. This suggests that simple lookahead (Shortsight) and the correct opening (Open) are key in these games, and that the ways plans overlap (Pitchfork) are important. Finally, the three most significant Advisors for tsoro yematatu are Candide, Worried, and Anthropomorph. This suggests that simple imitation of observed expert behavior (Anthropomorph) is most important, along with naive planning and anticipation, to expert play at this game.

5. Conclusion

Against a behavioral standard, HOYLE has learned to play seven different games expertly with a blackboard system. Four paradigms are postulated here for control of this blackboard. Empirical evaluation on five of these games indicates:

• Each game is learnable under at least one blackboard control paradigm against an expert. All the games are readily learnable under Paradigm 1, but some are learned as quickly, and apparently as well, under simpler controls.

• Once learned, each game can be controlled expertly either by the paradigm under which it was learned or under others. Expertise can be enhanced by playing different control structures against each other, that is, the program can continue to explore different portions of the game graph and continue to learn by varying its control strategy in play against itself. This suggests that "learned" might be better defined by knowledge stability, rather than by a behavioral standard, and that expertise should be developed first against an expert and then against players of lesser skill.

• All the control paradigms may be safely improved through the elimination of Advisors that did not contribute to the decision-making process during the learning phase of a particular game.

• Different games rely more heavily on different knowledge sources. The relevant knowledge sources do not necessarily reflect human intuition about the games.

• The performance of control paradigms with primitive weighting of significant Advisors provides no visible performance change for a well-learned game. For at least one game, however, restricting Advisors to a single, smoothed comment substantially reduces post-learning playing skill.

Future research will include the consideration of alternative control paradigms and the automation of these experiments as part of HOYLE's discovery process.

References

Anantharaman, T., Campbell, M. S. and Hsu, F.-H. 1990. Singular Extensions: Adding Selectivity to Brute-Force Searching. *Artificial Intelligence* 43 (1): 99-110.

Bell, R.C. 1969. *Board and Table Games from Many Civilizations*. London: Oxford University Press.

Berliner, H.J. 1980. Backgammon Computer Program Beats World Champion. *Artificial Intelligence* 14: 205-220.

Chen, K. 1989. Group Identification in Computer Go. In *Heuristic Programming in Artificial Intelligence - The First Computer Olympiad*, ed. D. Levy and D. Beal. New York: John Wiley.

Epstein, S. L. 1989. The Intelligent Novice - Learning to Play Better. In *Heuristic Programming in Artificial Intelligence - The First Computer Olympiad*, ed. D. Levy and D. Beal. New York: John Wiley.

Goetsch, G., and Campbell, M. 1988. Experiments with the Null Move Heuristic in Chess. In *Proceedings of the AAAI 1988 Spring Symposium Series: Computer Game Playing*, 14-18. Stanford, California.

Langley, P. 1985. Learning to Search: From Weak Methods to Domain-Specific Heuristics. *Cognitive Science* 9: 217-260.

Mitchell, T.M., Utgoff, P.E., and Banerji, R. 1983. Learning by Experimentation: Acquiring and Refining Problem-Solving Heuristics. In *Machine Learning: An Artificial Intelligence Approach*, ed. R.S. Michalski, J.G. Carbonell, and T.M. Mitchell. Palo Alto, California: Tioga Publishing.

Nii, H. P. 1986. Blackboard Systems: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures. *AI Magazine* 7 (2): 38-53.

Pazzani, M. J. 1987. Inducing Causal and Social Theories: A Prerequisite for Explanation-based Learning. In *Proceedings of the Fourth International Workshop on Machine Learning*, 230-241. Ed. P. Langley. Los Altos, CA: Morgan Kaufmann.

Rosenbloom, P.S. 1982. A World-Championship Level Othello Program. *Artificial Intelligence* 19: 279-320.

Samuel, A.L. 1963. Some Studies in Machine Learning Using the Game of Checkers. In *Computers and Thought*, ed. E.A. Feigenbaum and J. Feldman, 71-105. New York: McGraw-Hill.

Samuel, A.L. 1967. Some Studies in Machine Learning Using the Game of Checkers. II - Recent Progress. *IBM Journal of Research and Development* 11: 601-617.

Zaslavsky, C. 1982. *Tic Tac Toe and Other Three-in-a-Row Games, from Ancient Egypt to the Modern Computer*. New York: Crowell.