Mediation among Advisors

Susan L. Epstein
Department of Computer Science
Hunter College of the City University of New York
695 Park Avenue
New York, NY 10021
SEHHC@CUNYVM.CUNY.EDU

Abstract
HOYLE is a system under development to explore game-playing expertise. Because many games induce notoriously large and complex search spaces, programs to play them have developed a broad spectrum of heuristics to curtail search. One aspect of HOYLE's preliminary success at a broad class of games is its alternative to search in these spaces: mediation among (usually disagreeing) Advisors. This paper describes HOYLE's metatheoretical approach to game playing and the three-level architecture that supports that metatheory. Subsequent sections explain how HOYLE's limited rationality has thus far provided a remarkably robust alternative to traditional, computationally expensive solution methods.

Games and Complexity

Informally, a game consists of some materiel (e.g., a board, playing pieces), a roster of one or more participants, and a list of rules. The rules describe how the participants are to take turns affecting the position and/or status of the materiel (e.g., removing pieces from the board or changing their location) until some rule terminates the process. When the process terminates, the rules declare either which participant has won or that the process was a draw. If all information about the game is disclosed and equally available to all the players (e.g., no closed hands, no uncertain outcomes as with dice), it is said to be a perfect information game. For the games considered here, there are exactly two participants: Player (the one who moves first) and Opponent. Player and Opponent are the roles in a two-person game.

Any game may 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 materiel. The participant who is not the mover in a state is referred to here as the non-mover. An edge in the game graph that goes from one (originating) state to another (resultant) state represents a move, the changes that occur when the mover in the originating state behaves (takes a turn) in any one manner permitted by the rules. Together, the materiel, participants, and game graph are called the elements of a game.

There are three kinds of nodes in the game graph: starts, finishes, and intermediate states. A start is an initial state of the materiel before any participant has taken a turn, with Player as mover. A finish is a node with no out edges, and is labeled either win (Player wins and Opponent loses), loss (Player loses and Opponent wins), or draw (no winner or loser). In the game graph, a path is a sequence $n_1e_1n_2e_2....n_ke_kn_{k+1}$ of nodes $n_1,n_2,...,n_k,n_{k+1}$ and edges $e_1,e_2,...,e_k$, such that e_i is an edge in the game graph from n_i to n_{i+1}, for i = 1, 2, ..., k. If there is a path n1e1n2e2....nkeknk+1 in a game graph, node nk+1 is said to be at depth k or k-ply from node n1. A path in a game graph from a start to a finish represents one complete experience of the game, and is called a contest.

Relatively brief paths beginning with a start are called openings ; relatively brief paths ending with a finish are called endgames. A path from the end of an opening to the beginning of an endgame is called a middlegame. The stage of a node is the path type (opening, middlegame, endgame) it is most likely to appear on. (In more complex games, such as chess, an endgame is distinguished by reduced materiel and "no reasonable chances to mount a direct mating attack on either King." (Mednis 1978) HOYLE has not yet attained this level of sophistication.)

During a contest, Player tries to arrive at a win and Opponent tries to arrive at a loss. Since from any state there are usually many legal moves, the challenge in play is for the mover to select the best move, i.e., one that will maximize the mover's opportunity to arrive at a desired state, while minimizing the non-mover's opportunity to do so. A theory for a game is one or more statements about the properties of and relations among its elements. For example, a possible theory statement for tic-tac-toe is "the center square is the key position." A heuristic for a game is an explicit directive to the mover for selecting a move. For example, "if the center square is vacant, move there." A heuristic is an operationalized version of a theory statement; there may be more than one way to construct such an operationalization. Finally, a strategy is a set of heuristics for a specific game with a control method for choosing among them. A strategy is thus a procedure that, given any state, returns a move for the mover. To learn to play a game well is to produce and refine a strategy for it, constructing, testing, and revising theories.

It is possible to represent a two-person, perfect-information game as a directed graph and "solve" the problem of perfect play by backing up the data from an exhaustive search of all possible sequences of moves (Nilsson 1980). Even for simple games, however, the game graph is quite large (tic-tac-toe, for example, has more than 5000 reachable nodes), and for difficult games the game graph is computationally intractable. (It has been estimated, for example, that there are 1020 nodes in the game graph for checkers, and 1043 in that for chess.) Although a rational player, one with extensive memory and great speed, could play perfect tic-tac-toe through exhaustive search of the game graph from the node representing the current state of the contest, more difficult games will not succumb to this approach.

A second rational approach to game playing is to consider the theoretical mathematical properties of a two-person, perfect-information game and exploit them. Recent mathematical research (Berlekamp, Conway, & Guy 1982), however, indicates that success in many of these games is determined by difficult, and often intractable, mathematical calculations involving abstract concepts called nimbers. Not enough is yet known about computing with nimbers to program for them. Thus the "rational" approaches to game playing, omniscience and theoretical computation, are really not open to computers.

Games and AI

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. These programs play checkers (Samuel 1963, Samuel 1967), Othello (Rosenbloom 1982, Lee & Mahajan 1988), backgammon (Berliner 1980) or chess (Ebeling 1986, Anantharaman, Campbell & Hsu 1988, Schaeffer 1988). The Chess 4.5 paradigm (Slate & Atkin 1977), on which many of them are based, advocates extensive forward search from the node in the game graph representing the current state of the game to a set of intermediate nodes, called tip nodes. These tip nodes are then heuristically evaluated and their values are backed up (Nilsson 1980, Berliner 1979, Palay 1982, Rivest 1987, McAllester 1988) to estimate the best move from the current state. Iterative deepening, selective extension (Anantharaman, Campbell, & Hsu 1988), and the null move heuristic (Goetsch & Campbell 1988) are refinements of this game graph search that have proved effective in chess. Such search is often supported by very large historical knowledge bases on openings, endgames, and other play experience (Schaeffer 1988, Schultz & De Jong 1988). Although people assuredly lack the speed, accuracy, and retention of their machine opponents, the outstanding human players at these games are still able to defeat the best of these one-game programs. There are, for example, hundreds of chess players who can still defeat the best of the chess-playing machines.

During chess competitions, outstanding players often have a team of experts with whom they discuss a contest. The hope is that members of the team will notice different aspects of the current game state hitherto overlooked, ones that will strengthen the decision-making process. In their consideration of alternative moves from the current state, people call upon their knowledge about game playing and problem solving. This knowledge can be contest-specific (i.e., related to a single experience of a game from a start to a finish), game-specific (a strategy for a particular game), or game-independent (metatheoretical). Different people may well have different viewpoints on alternative moves. HOYLE's design encourages the construction of different viewpoints as an alternative problem solving approach for a broad class of two-person, perfect-information games.

HOYLE's Architecture

Rather than explore deeply or recall perfectly, HOYLE approaches the many aspects of effective decision-making during game playing with a three-level architecture. On the first level, HOYLE captures and exploits the commonalities in its domain with a panel of Advisors, each of which takes a narrow, but quite rational view, of the move selection problem. Each Advisor is permitted only limited resources before it offers its recommendations. On the second level, HOYLE mediates among its Advisors, weighing their judgement against what it knows about the current contest and the game itself. On the third level, HOYLE learns about its Advisors and their general application in play.

The Advisors

At any given state in a contest, there is usually a variety of possible moves, each of which has something to recommend it (a positive comment) and/or advise against it (a negative comment). A comment is attributable to some particular view of the game state, possibly a very narrow one. In HOYLE, each comment cites supporting evidence and is advanced with some measure of strength.

Consider, for example, a tic-tac-toe game, where Player marks squares with X's and Opponent uses O's. In Figure 1, with Player as mover, one comment might be "move in square 6," to prevent a win by Opponent. This comment could cite as support the threat of the O's in 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. A second comment, of a similar nature, might be "move in square 8," since Opponent threatens to win in the second column after two turns. Such a comment must be weaker than the first, because it projects a longer contest. A third comment might be "move in square 2," because Player will win immediately.

Each Advisor in HOYLE epitomizes a different perspective on game playing. Panic, for example, looks to see whether the non-mover has a sure win on his next move. In Figure 1, it is Panic who insists that a move in square 6 is the only way to save the day. Victory is another Advisor; it looks for a sure win for itself on the current move. In Figure 1, it is Victory who insists that a move in square 2 is the best choice. At any state, an Advisor may make any number of comments, each with supporting evidence and a strength from 0 (adamant opposition) to 10 (insistent support).

A State in a Tic-tac-toe Trial

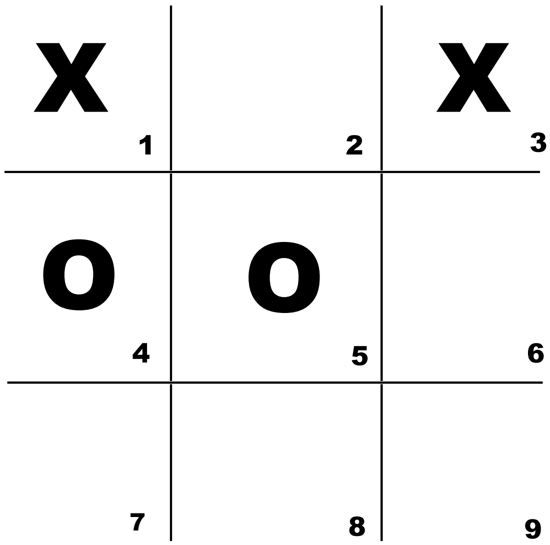

Figure 1

Some of the comments reference limited libraries constructed during the post-mortem after a contest is played. There is a library of openings, a library of significant states, and a library of previous contests. After HOYLE completes a contest, it extracts and caches the opening (the first 10% of the moves) and any significant states. A state is deemed significant if it is certain to lead to a win or loss. Consider the description in Figure 2 of a contest that Player wins on the 25th move, from state S-1 to state S. The second-to-last state, S-1, is cached as "win in 1" with the final move, M-25. HOYLE's Advisors, if they ever encounter S-1 again, will recall both the state and the winning M-25 move without search. The third-to-last state, S-2, is cached as "lose in 2" if and only if, under a limited resource allotment, exhaustive forward search from it fails to find a better alternative for the mover at that point. If the third-to-last state is cached, then the fourth-to-last state, S-3, is cached as "win in 3" with the winning M-23 move. If Opponent is the winner in state S, HOYLE caches S-1 as "lose in 1" with move M-25, and may, based on a search like the one described, cache S-2 as "win in 2" and S-3 as "lose in 3." Theoretically, this cycle of caching and searching significant states, analogous to a human post-mortem, can be carried back as far as the opening. Such search is costly, however, and HOYLE devotes only a finite amount of resources to it.

With access to these limited libraries, Advisors may warn against previously unsuccessful lines of play, and recommend those that were successful in the past. Advisors may also recommend simple plans and defend against a primitive model of the opposing player. There are even comments recommending psychological gambits. At the moment there are 12 Advisors, in various stages of implementation, descriptively named Victory, Panic, Sadder, Wiser, Open, Candide, Worried, Knot, Sneak, Feinter, Pitchfork, and Wild. Thus, HOYLE's first level deliberately chooses limited rationality with a variety of resource-limited, and often discordant, reasoners.

Mover	State before Move	Move
Player	S-3	M-23
Opponent	S-2	M-24
Player	S-1	M-25
Opponent	S	none

Calculating Significant States
Figure 2

Mediation

Based upon all this self-generated advice, HOYLE has to determine its next move. On the second level, HOYLE learns to mediate effectively among its (often disagreeing) Advisors for a specific game. HOYLE acquires and refines a strategy for a particular game from its experience in playing that game. This strategy is an instantiation of HOYLE's uniform strategy for all two-person, perfect information games.

Figure 3 shows HOYLE's uniform strategy frame and, in italics, a hypothetical instantiation for tic-tac-toe. (As HOYLE encounters more difficult games in its knowledge base, this frame will be modified.) The items in Figure 3 become a focus of attention simply because they appear in the frame, i.e., HOYLE's metatheory asserts that they are significant factors in playing any game.

It is an advantage to go first: *true*
List of best opening moves: *((center square))*
Under optimal play taking the Player/Opponent advantage it is possible to win: *false*
Optimal play generalization: *(Player takes the center square. Opponent takes ...)*
Most important territory to control: *(center, corners)*
Treatment of advice (equal or strength-based): *strength-based*
Weighted advice: *no*
Advisor weights: *none*

HOYLE's Uniform Strategy for Game Playing
Figure 3

When it is HOYLE's turn to move, it consults each of four tiers of Advisors in turn. HOYLE collects and weighs evidence for and against possible moves, and selects the move with the highest strength based on its current strategy for the game. The tiers capture the following commonsense rules :

- If you can move to a win now, do it.
- If you cannot win now and are about to lose at the next turn, block the move which threatens you.
- The more contests of the same game you play, the more you should rely on your previous experience.

- In long contests (longer, say, than the number of positions on the board) look for a safe cycle.
- When things look bad, play defensively.

HOYLE continues on to the next tier only when unable to make a clearly-substantiated decision from the information at hand.

Learning

Rather than calculate an ideal strategy from search of all possible contests, HOYLE revises its strategy over a set of contests. This enables HOYLE to learn from a kind of spiral curriculum: as HOYLE becomes more proficient at a game against a skilled opponent, the individual contests usually force the opponent to marshall more skill, and the contests provide HOYLE with more advanced and detailed information. It is important to note that HOYLE does not require thousands, or even hundreds, of practice games from which to learn. HOYLE's theory about a game evolves after each contest, during a human-like post-mortem. No game is ever assumed to be learned perfectly; HOYLE is always willing to play another contest.

Even from only a few contests at a particular game, HOYLE is able to extrapolate valuable knowledge. For now, HOYLE learns in an unstructured environment, i.e., it learns in random encounters against one or more opponents, without developing and testing new strategies on its own. Eventually, HOYLE will explore original game-specific strategies under the guidance of a Theoretician, an Experimenter, and a Critic. The Theoretician will postulate likely theories about a game, the Experimenter will operationalize each theory and explore the success of the resultant strategy in an experiment, and the Critic will revise the theories in light of the experimental results. This aspect is similar in spirit to LEX (Mitchell, Utgoff, & Banerji 1983) and (Falkenhainer & Rajamoney 1988).

A HOYLE experiment will be an examination of past contests or a tournament whose contests contribute to HOYLE's knowledge base for the game. For example, if the Theoretician suspects that the role of Player offers an advantage, the Experimenter could examine contests between two (human) experts and see which role wins most often. Currently, the list of openings is compiled from contests where an expert took the role of Player. (Eventually, the Theoretician might postulate good openings and have the Experimenter test them.) Whether or not it is possible to win and any optimal play generalizations will also be extracted from the contest library. In games involving territorial control, the library contests suggest the most important territory from expert openings (in games with uninvertible moves) and from unforced expert moves. (HOYLE could also calculate those positions that Player occupies most often in win nodes and/or Opponent occupies most often in loss nodes.) The Experimenter can evaluate the relative significance of the Advisors either by staging a tournament or by observing which Advisors gave/would have given the correct (winning or expert) advice in library contests.

Experimental design is part of HOYLE's metatheory. A sophisticated Experimenter should consider the number of different wins in the game graph, the nature of paths from starts (are they bushy? deep? cyclic?), the nature of the goals of the game (to possess territory, for example), and whether or not moves are invertible.

On the third level, HOYLE is intended to explore commonalities in game-playing. It is HOYLE's thesis that people have a metatheory for game-playing, an encapsulation of their experience that enables them to learn to play games many well. HOYLE's metatheory is a general, game-independent strategy, a set of techniques for learning and playing any two-person, perfect information game. Currently, HOYLE's metatheory consists of:

- the Advisors
- the uniform strategy frame
- the 4-tier consultation system

Eventually HOYLE is expected to learn its metatheory, just as it currently learns game-specific strategies.

Work in Progress

HOYLE is a system under development. It possesses a uniform declarative representation (an instantiable frame) for two-person, perfect-information games, and a uniform procedural representation whose application enables the correct play of any such game. The efficacy of this representation is well-supported by HOYLE's current ability to play a broad spectrum of culturally diverse games correctly under the direction of its metatheory. The games in the domain are 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 are popular in a variety of cultures, and therefore probably capture some aspects of game playing that people find particularly intriguing. Their game graphs certainly do not have the complexity of chess or Go, but they do offer the challenge of cycles and stage transitions, and at least one of the games has a game graph of billions of nodes.

Tic-tac-toe is well-understood, and HOYLE plays it optimally in either role, i.e., wins whenever possible and ties otherwise. Lose tic-tac-toe (Cohen 1972) has been solved mathematically, i.e., a strategy for both players has been clearly delineated and proved. Against this strategy, HOYLE plays flawlessly after only two trials, as either the first or the second player. In the rest of the games, HOYLE rarely loses after two trials, and regularly defeats most human opponents. The Advisors, with very little memory or forward search, have proved surprisingly powerful.

Current efforts are focused on the implementation and relative significance of the Advisors. In particular, Pitchfork, the Advisor on forking, is so successful that it now dominates its tier and guarantees optimal play without forward search into the game tree (Epstein 1989). From play, HOYLE should acquire and refine a reliable method of mediating among and applying its Advisors' recommendations to the particular game in question. Although chess, checkers, and Go are theoretically within HOYLE's domain (i.e., given the rules, HOYLE could play them correctly), none of them has been attempted. HOYLE's metatheory will continue to be developed as the games in its knowledge base become more difficult. The power of its Advisors will be evaluated from the prowess HOYLE displays and the speed with which it learns. In the meantime, HOYLE's mediation among its Advisors proves an effective, if limitedly rational, approach to game playing.

References

Anantharaman, T., Campbell, M., and Hsu, F. 1988. Singular Extensions: Adding Selectivity to Brute Force Searching. In *Proceedings of the AAAI 1988 Spring Symposium Series: Computer Game Playing*, 8-13. Stanford, California.

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

Berlekamp, E.R., Conway, J.H., and Guy, R.K. 1982. *Winning Ways for Your Mathematical Plays*, Volumes 1 and 2. London: Academic Press.

Berliner, H.J. 1979. The B*Tree Search Algorithm: A Best-First Proof Procedure. *Artificial Intelligence* 12: 23-40.

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

Cohen, D.I.A. 1972. The Solution of a Simple Game. *Mathematics Magazine* 45: 213-216.

Ebeling, C. 1986. All the Right Moves: A VLSI Architecture for Chess. Ph.D. diss., Department of Computer Science, Carnegie Mellon University.

Epstein, S.L. 1989. Deep Forks in Strategic Maps - Playing to Win. Submitted to IJCAI 1989.

Falkenhainer, B., and Rajamoney, S. 1988. The Interdependencies of Theory Formation, Revision, and Experimentation. In *Proceedings of the Fifth International Conference on Machine Learning*, 353-366. Ann Arbor, Michigan.

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.

Lee, K.-F., and Mahajan, S. 1988. A Pattern Classification Approach to Evaluation Function Learning. *Artificial Intelligence* 36: 1-26.

McAllester, D.A. 1988. Conspiracy Numbers for Min-Max Search. *Artificial Intelligence* 35: 287-310.

Mednis, E. 1978. *Practical Endgame Lessons*. New York: David McKay Company.

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, 163-190. Palo Alto, California: Tioga Publishing.

Nilsson, N.J. 1980. *Principles of Artificial Intelligence.* Palo Alto, CA: Tioga Publishing.
Palay, A.J. 1982. The B* Tree Search Algorithm - New Results. *Artificial Intelligence* 19: 145-164.
Rivest, R.L. 1987. Game Tree Searching by Min/Max Approximation. *Artificial Intelligence* 34: 77-96.
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.
Schaeffer, J. 1988. Learning from (Other's) Experience. *Proceedings of the AAAI 1988 Spring Symposium Series*: Computer Game Playing, 51-53. Stanford, California.
Schultz, A.C., and De Jong, K.A. 1988. Using Experience-Based Learning in Game Playing. *Proceedings of the Fifth International Conference on Machine Learning*, 284-290. Ann Arbor, Michigan.
Slate, D.J., and Atkin, L.R. 1977. Chess 4.5 - The Northwestern University Chess Program. In *Chess Skill in Man and Machine*, ed. P.W. Frey, 82-118. New York: Springer-Verlag.
Zaslavsky, C. 1982. *Tic Tac Toe and Other Three-in-a-Row Games, from Ancient Egypt to the Modern Computer*. New York: Crowell.