

Capturing The Information Conveyed By Opponents' Betting Behavior in Poker

Eric Saund
469 Clifton Avenue
San Carlos, CA 94070
saund@alum.mit.edu

Abstract—This paper develops an approach to the capture and measurement of the information contained in opponents' bet actions in seven card stud poker. We develop a causal model linking downcards with hand strength, thence to bet actions. The model can be inverted to infer probability distributions over possible downcards from bet actions, given knowledge of opponents' bet policies. For experimental purposes, we propose a simple yet plausible “default” bet policy including deceptive plays. In simulated games, this apparatus is used to compare the Kullback-Leibler information measure between inference of players' hand strength based on dealt cards *and* players' bet actions, versus inference of hand strength based on dealt cards only. We experimentally associate the K-L divergences with the win-lose rates for simulated players who either do or do not exploit knowledge of opponents' bet actions. Opponent inference carries up to a 36% information advantage over a cards-only player playing the same betting policy, and is worth on the order of .15 bets/hand.

Keywords: poker, information, stud, hand type, opponent model

I. INTRODUCTION

Simply by virtue of compounding complexity, natural and simulated mechanistic worlds present many unconquered challenges for modeling and reasoning by artificially intelligent systems. The challenges become vastly more difficult with the introduction of other intentional agents. If you think it's a challenge to keep weeds and bugs out of your garden, try fending off gophers, squirrels, and raccoons. A major goal for Artificial Intelligence in games is to develop ways to exploit the information conveyed by the behavior of intentional opponents. Opponents' actions are typically based on knowledge, beliefs, goals, and plans the subject player is not privy to. But with sufficient wisdom, these actions can be “read” to gain information about the opponents' hidden states.

The game of poker deals a prototypical example. The objective state of the game consists of possession of cards, some of which are held privately, and some of which are known to other players. Play decisions (bet/fold actions) are made on the basis of perceived relative hand strength; knowledge about opponents' hands beyond that objectively visible through displayed cards is of immense value. The structure of betting in poker is designed such that player actions convey information about their undisclosed cards. Stronger hands are incented to bet more heavily, but to do so broadcasts this information, so that opponents may exploit the telegraphed knowledge to better decide on their own plays. Hence the

most famous aspect of poker is the use of deception in the form of bluffing and slowplaying to mislead opponents about one's actual hand strength. Bluff and slowplay bet actions run counter to actual hand value, however. This leads to perplexing tradeoffs, efforts to outguess opponents, and all manner of psychology.

Poker has therefore been recognized as a model for broader classes of competitive situations involving uncertain belief about objective states, intentional opponents whose plans, goals, and belief states can only be inferred from partial and uncertain evidence, and promotion of *information* to the status of an asset to be managed along with objective ones. Examples include warfare [5], [6], and business [10].

This paper attempts to take one step toward the development of a theoretically sound and computationally practical framework for analyzing and exploiting information conveyed by intentional opponents in seven card stud poker. The form of poker enjoying by far the greatest public visibility and AI game interest is Texas Hold'em. We believe our formulation and results to be broadly applicable, but we focus on seven card stud because this game presents a particularly rich texture of possible outcomes and knowledge disclosure as players' individual hands evolve through successive rounds of dealing (known as “streets”), each accompanied by rounds of betting.

Our initial objective is simply to measure the information conveyed by bet actions, in comparison to the information offered by the visible cards alone. To do so requires the development of a great deal of apparatus modeling the relationship between dealt cards and sensible betting actions, and this necessarily involves modeling of rational players' decision-making processes to some rudimentary degree. The framework will accept more sophisticated opponent models as they are developed.

The paper proceeds as follows. Through the imaginary game of “face-up poker”, Section II reviews the logic of correct betting in poker, and it develops a forward causal model relating held cards to bet actions. The model extends directly to true poker in which some cards are hidden. Section III describes how the model can be inverted to infer probability distributions over opponents' possible downcards, given opponent models of those players' betting policies. Section IV introduces a simple form of such betting policies, and calls out two useful instances, the “honest player” who bets only by value, and a simple default deceptive player who

executes some degree of slowplaying and bluffing. Section V introduces a measure of information gained by reading opponents' bet actions in comparison with only observing dealt cards. Section VI presents experimental results of empirical measurements of this information gain for a corpus of simulated games. This section also ties this information gain with net win/lose rates for players who do or do not exploit knowledge of opponents' bet actions. Section VII concludes by discussing the results and their possible implications for live games.

II. THE LOGIC OF BETTING IN POKER

The logic of betting in poker is well described by Sklansky [12]. It is best understood by imagining a game of poker in which all cards are dealt face up, so that every player sees all of their opponents' cards as well as their own. Then, in principle every player can calculate their chances of having the best hand at showdown. Five-card hands are ranked by hand type, e.g. "Two pair, Tens and Fours with a Queen kicker." Given a partial hand and knowledge of cards remaining in the deck to be dealt, one may compute a probability distribution over the final hand achieved at showdown. Call this a *hand type probability distribution*, or *htpd* for short. This calculation can be performed or approximated by various means, including sampling simulated deals of the remaining cards, by enumeration[11], or by combinatoric analysis extending the reasoning of [1].

Given a set of *htpds* possessed by active players (players who have not folded their hands) the probability that player i 's final showdown hand will beat all others is the conjunction of events that his final hand type ht beats each other player j , summed over all hand types k , weighted by the probability $p_i(ht_k)$ that player i ends up with hand type ht_k :

$$p(\text{win}_i) = \sum_k p_i(ht_k) \prod_{j \neq i} \sum_{k'=0}^k p_j(ht_{k'}) \quad (1)$$

The final sum term in (1) assumes that hand types are rank ordered from worst ($ht_{k'_0} = 2-3-4-5-7$) to best ($ht_{k'_{max}} = \text{ROYAL-STRAIGHTFLUSH}$).

Figure 1 shows the *htpds* for two stages of the sample game whose game history is given in Figure 8.

Correct betting logic seems straightforward. Any player whose probability of showing the winning hand is greater than $1/N$ should bet or raise, where N is the number of active players. Any player whose probability of winning is greater than their effective odds should not bet or raise, but they should check or call. Effective odds e is the ratio of the amount a player will have to contribute to the pot, to the final pot. Money already in the pot justifies calls by players who have lower probabilities of winning. The more money already in the pot due to ante or previous betting rounds, the worse probability of winning a player may have and it still be worthwhile to call. Calculation of effective odds can be tricky, however, because it depends on predicting whether other players will bet, call, or fold as the game progresses. In this paper we employ a very simple model of effective

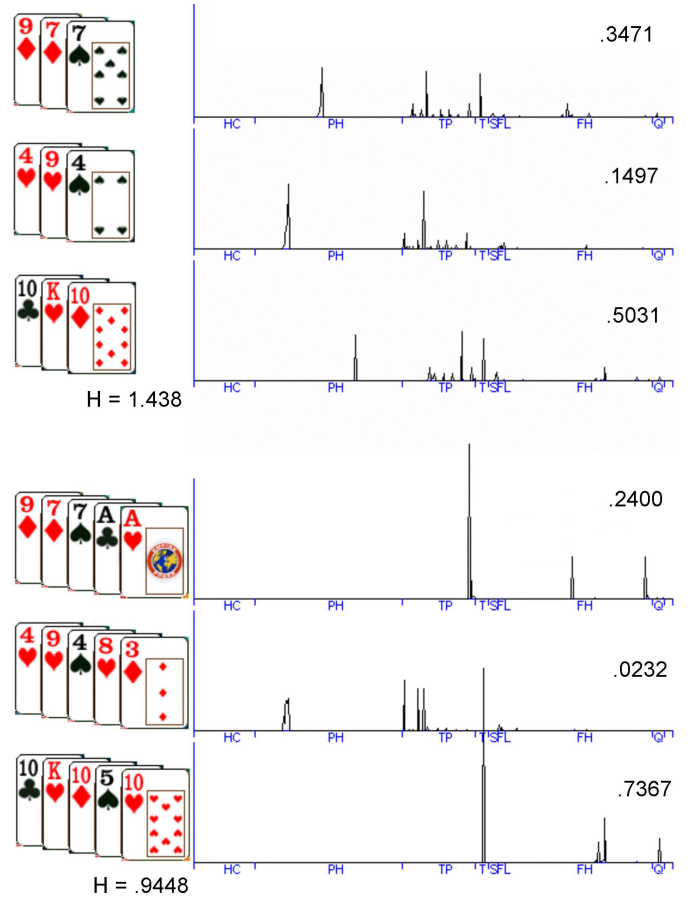


Fig. 1. Hand type probability distributions (*htpds*) showing the probability of achieving a final showdown hand, at stages 3B (following betting on 3rd street) and 5D (following the deal at 5th street), for the sample game of Figure 8. Only three *htpds* are shown at each stage because seats 1, 2, 4, and 6 folded at stage 3B. Possible hand types are ordered left to right from worst to best. Major hand categories listed are HC (High Card); PH (Pair-Highcard); TP (Two-Pair); T (Trips); S (Straight); FL (Flush); FH (FullHouse); Q (Quads). The numbers shown are the probabilities at these stages that each hand will win, and the entropies H at each stage.

odds which assumes that in addition to the current pot size, all currently active players contribute to the pot one bet per street, through successive streets to showdown.

Thus a model for the causal structure of betting in face-up poker is shown in Figure 2. A player's bet action depends on the effective odds, number of active players, and on their probability of winning at showdown. Probability of winning depends on their and their opponents' *htpds*. *Htpds* depend on cards held and cards available to be dealt.

This causal chain may be extended to true poker in which some cards are held privately. In seven card stud, the first two and the seventh street cards are dealt face-down. Figure 3 shows the extended model from the point of view of player i who knows his own downcards but not those of his opponents. Uncertainty about opponents' downcards can be represented in terms of a probability distribution over all possible combinations of downcards that the opponent may possess. For seven card stud this may be represented

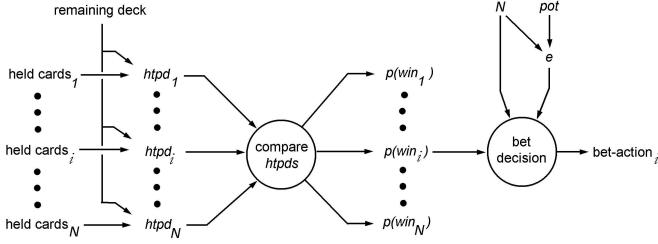


Fig. 2. Rational betting model for player i in face-up poker.

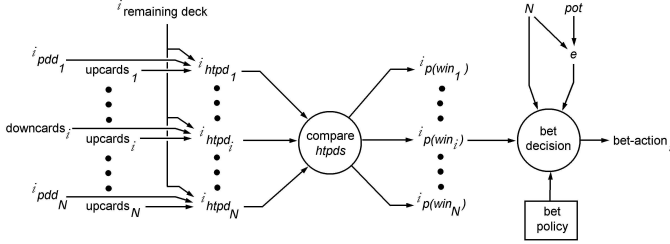


Fig. 3. Causal betting model for player i who knows his own downcards but represents opponents' downcards as the probability distributions pdd .

in a vector of length 52×51 , indexed by the variable, l . Call this a *possible-downcard-distribution*, or *pdd* for short. The notation, ${}^i pdd_j$ refers to the distribution of player j 's possible downcards from the point of view of what is known or believed by agent i , who may be a player or some other observer.

Some entries in the pdd vector may be zeroed out immediately, namely those downcard pairs that include any card that has been dealt face up to any player. Additionally, every player knows their own two downcards (or three at 7th street) which rule out their inclusion in any opponent's pdd . The goal of reading opponents' cards through their bet actions amounts to differentially weighing the remaining ${}^i pdd$ entries so as to reflect each opponent's apparent hand strength.

Given player j 's possible downcard distribution pdd_j , $htpd_j$ is computed by integrating the $htpds$ over possible downcard pairs l , weighted by each pair's probability $pdd_{j,l}$:

$$htpd_j = \sum_l p(pdd_{j,l}) htpd(pdd_{j,l}, upcards_j) \quad (2)$$

Obviously this operation can be computationally expensive so in practice it is important to have efficient implementation of the downcard-to- $htpd$ calculation, $htpd(pdd_{j,l}, upcards_j)$.

A second factor enters into the extension of Figure 2 to true poker. This is the addition of players' bet/call/fold policies. A basic strategy is to bet/call/fold based on estimates of probability of winning at showdown and effective odds, as described above. This is known as betting for value. But bet actions may be influenced by another reason, namely to induce other players to miscalculate one's own hand strength. Therefore, a player's bet strategy may incorporate deceptive plays which contradict the player's strictly value-based rationale for checking/betting or folding/calling/raising. Sklan-

sky's Fundamental Theorem of Poker states that one is advantaged to have one's opponents bet differently from the way they would bet if they knew one's downcards.

Optimal betting behavior including deceptive betting requires knowledge of how one's opponents will respond to the various bet actions one may take. These responses might be dependent on the opponents' beliefs about oneself. Even if opponents' beliefs and strategies were known precisely, optimal betting would then require forward chaining through many combinations of possible plays and responses. The conduct of this reasoning lies beyond the scope of this paper but is the topic of much of the poker AI literature [8], [2], [7]. Here we focus on trying to puzzle out opponents' $pdds$ based on relatively simple models of their betting policies.

Summary of Notation: as subscripts, the variables i and j index players in a game; as superscript prefixes they index agents who possess knowledge or belief, including players and other observers. The variable k indexes hand types. The variable l indexes possible downcard pairs (or triples at seventh street).

III. INVERTING THE CHAIN TO INFER DOWNCARDS

A key problem faced by a poker player is to make effective use of the information conveyed by opponents' betting behavior (check/bet and fold/call/raise actions). This amounts to inverting the forward model of opponents' betting in order to adjust beliefs over the opponent's possible downcards, represented in the pdd . In doing so, we must account for the possibility that opponent bet policies may include deceptive bluffs and slowplays.

Suppose that we know the opponent intimately, such that for any pair of downcards, plus observed upcards (both showing and folded) and remaining active players (we refer to this state information as the *table*, t), we know the probability that in this situation they will execute a particular bet action $b_j : b_j \in \{check, bet\}$ if $bet-to_j = 0$; $b_j \in \{fold, call, raise\}$ if $bet-to_j > 0$. In other words, if they hold downcards dc_l and the bet to them is zero, we know the probability that they will check versus bet, or, if an earlier player has already opened betting, we know the probability that they will fold vs. call vs. raise. Let us express this knowledge as

$$p_t(b_j | dc_l), \quad (3)$$

the probability that opponent j will perform bet action b_j given downcards dc_l , under the table circumstances t . We treat both opponent bet actions and belief about unobserved opponent downcards as random variables, while we treat knowledge of their conditional probability relation as being a known function which is contingent on the state of the table. This representation reflects the fact that opponent players may act nondeterministically, as is in fact recommended by game theory [4] as well as poker textbooks [12], [13].

When the opponent executes a bet action b_j , we may invoke Bayes' rule to perform inference about their downcards:

$$p_t(dc_l | b_j) = \frac{p_t(b_j | dc_l) p(dc_l)}{\sum_l p_t(b_j | dc_l) p(dc_l)} \quad (4)$$

The prior $p(dc_l)$ is the belief held that the opponent has downcards dc_l before we observed the bet action. This prior serves the role of carrying information forward from one street to the next. This calculation effectively performs a re-weighting of the possible-downcard-distribution by the likelihood of the bet action, followed by normalization.

Through implicit means, this mechanism achieves fairly subtle and complex reasoning. Opponents' actions of placing a bet (as opposed to checking or calling) tend to reweigh more heavily the possible downcard pairs that would offer that opponent a greater chance of winning given their upcards. Moreover, raises and re-raises weigh stronger downcards more heavily still, through an additional mechanism. Because the model has every player re-estimating the strength of every other players' hand after every action, when Player A bets, every other player will necessarily increase their belief that Player A has strong cards, which in turn decreases their beliefs in their own chances of winning. This narrows the pool of possible downcards that any player must hold to meet Player A's strength. So if Player B then goes on to raise or re-raise anyway, then for all of the players trying to estimate what Player B must be holding, (modulo bluffing) only the much stronger possible downcards for B will gain significant probability mass through the application of equation 4.

In a simpler game model and different network architecture, a Bayesian view of uncertainty and opponent modeling in poker was taken by Korb et. al. [9]. Following the tradition of Bayesian networks where conditional probabilities are straightforwardly represented by transition matrices, their work was designed for the probabilities to be acquired and modified by learning; a consequence however was a struggle with the curse of dimensionality due to the combinatoric complexity of the game.

For heads-up Hold'em games, Southey et. al. used Bayesian inference to select opponent models from a plausible prior distribution of models after relatively few observations [14]. Opponent hand strength was not modeled directly, but, for a simplified version of Hold'em it could be inferred from opponents' bet behavior after sufficient training. Because of the size of full heads-up Hold'em poker, extension to the full game required simplification of the model. Nonetheless, intelligent responses to differential opponent play of their partially hidden hands could be demonstrated.

IV. SIMPLE MODEL FOR RATIONAL BETTING BEHAVIOR

The opponent knowledge function (3) may be quite complex and difficult to discern. We propose to model it by appealing to the forward causal model for betting expressed in Figure 3. While the table situational factor t can be quite complex, significant elements will always be found in the two key parameters, probability of winning and effective odds. Generally, any halfway decent player will fold most of their losing hands (i.e. hands whose chances of winning are below the effective odds) (while perhaps bluffing with a few), raise their winning hands (i.e. hands whose chances

of winning are greater than $1/N$) (while perhaps slowplaying some of these) and call their intermediate hands. Under this reasoning, the opponent model (3) may be factored into two simpler components:

$$p_t(b_j|dc_l) \approx p_{e,N}(b_j|win_j)p(win_j|dc_l), \quad (5)$$

This factored opponent model employs the probability of opponent j winning at showdown, given the downcards they hold, as a random variable win_j that isolates their betting policy from their estimate of the overall strength of their hand. The complex situation embodied in the term, table, t in (3) decomposes now into two simpler terms, one containing effective odds and number of active players, and the other relating to the player's chances of winning at showdown according to the cards remaining to be dealt from the deck, and estimates of other players' hand strengths.

The term $p(win|dc_l)$ was discussed in Section II; this is the probability of winning under the *htpd* computed from the downcards dc_l , the upcards, and the remaining deck. All that remains to express the factored opponent model is to define the opponents' betting policy as a function of their probability of completing the winning showdown hand, the effective odds e , and number of other active players N . This form of representation for player betting policy is shown by example in Figure 4. The different regions of Figure 4a represent probability of check vs. bet, while the different regions of Figure 4b represent probability of fold vs. call vs. raise. Different styles of play may be interpreted as different shapes of these bet policy graphs. An interpretation of tight play would be a shift of the fold/call boundary to the right, corresponding to a requirement for a greater chance of winning to stay in the hand; aggressive play would shift the check/bet and call/raise boundaries to the left. "Honest" players who bet only for value would shrink to zero the bluff and slowplay probability regions, while very deceptive styles of play would increase these.

Clearly, this is a vast simplification of the betting strategy used by advanced players, and it is dumb in the AI sense that it relies heavily on calculation while it lacks strategy. Notably, this model fails to maintain a stance throughout a hand (e.g. a sustained bluff), or to decide how to bet based on anticipated responses of other players, such as planning and execution of check-raise maneuvers.

Nonetheless, we assert that the proposed factored betting policy model approximates a baseline default player model that is suitable for the purposes of this study, which is to gain insight into the quantity and value of information gained by exploiting knowledge of opponents betting behavior. More sophisticated modeling of betting behavior as represented by (3) may be substituted cleanly into the framework devised here, and is left for future work.

As a technical matter, it is useful to apply a simple transformation in the definition of the policy graphs. Default betting policy is expressed as a function of three variables, probability of winning, effective odds e , and number of active players, N . Instead of defining a separate pair of graphs

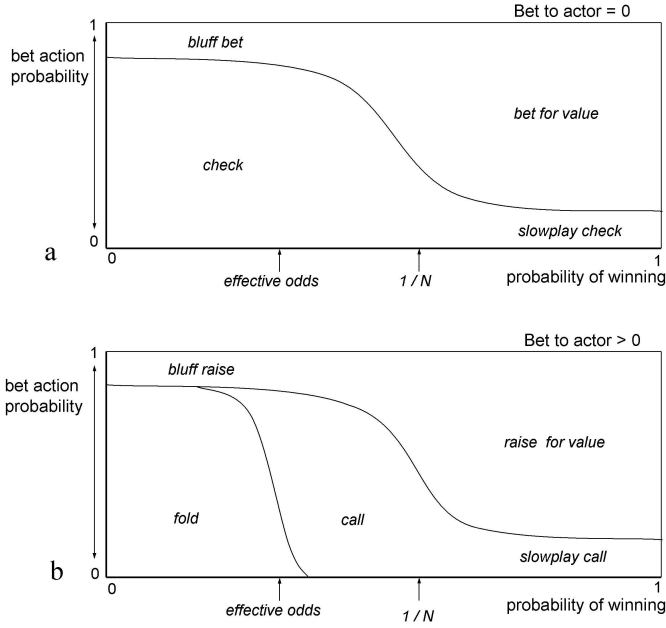


Fig. 4. Plausible betting policies for a deceptive poker player.

a	prob slowplay check	.2
b	prob bluff bet	.05
c	-log prob win offset check/bet	.1
d	prob slowplay call	.2
e	prob bluff raise	.05
f	-log prob win offset call/raise	.3
g	-log prob win offset call/raise	.2
h	-log prob win offset fold/call	.1

TABLE I
PARAMETERS OF THE DEFAULT BETTING MODEL
USED TO SIMULATE DECEPTIVE PLAYERS.

for every N , we apply the transform, $p' = -\log_N(p)$, that eliminates N as a degree of freedom in the graphs. For the experiments described in the following sections, we established a default player model with piecewise constant regions for each bet action, blended at their boundaries by linear interpolation in the \log_N transform space. Parameters for this betting policy are shown in Table I, and the corresponding bet policy graphs in Figure 5. These were chosen on an ad hoc basis over approximately 100 simulated games by adjusting parameters until the simulated players appeared to be making sensible checking, betting, calling, folding, and raising decisions.¹ A sampling of games under these parameters can be viewed at <http://www.saund.org/poker/sample-games.html>.

V. INFORMATION GAINED BY INFERENCE FROM BET ACTIONS

We are now in a position to experimentally measure the information gain and value of exploiting opponents' betting

¹All seven card stud poker games discussed in this work used the following fixed limit betting structure: Ante: .25; Bringin: .25; 3rd & 4th streets: 1.0; 5th, 6th & 7th street: 2.0; maximum four raises per street. There is no house rake.

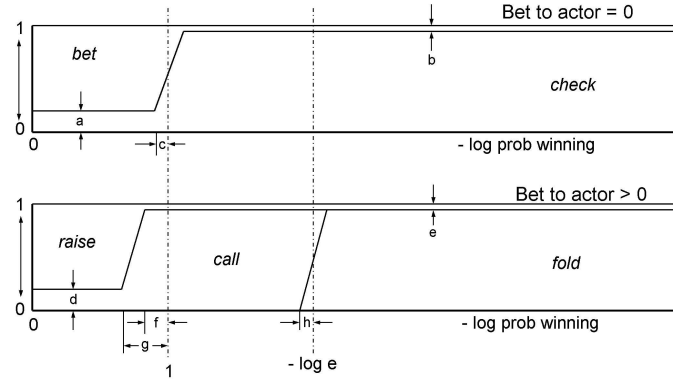


Fig. 5. Bet policies defining the default player model used by simulated players in determining their bet actions, and used to infer hand strength from opponents' bet actions. Note that probability of winning is expressed in the $-\log_N$ coordinate transform, where N is number of active players.

behavior in addition to knowledge of dealt cards in seven card stud. Let us consider $N + 2$ viewpoints on the dealt cards. Each of the N players knows all of the cards that have been dealt face up, plus their own two downcards (three at seventh street). The *public knowledge* player is like an observer on the sidelines; they know only what cards have been dealt face-up so are no longer in the deck. At the other extreme, the *omniscient* observer knows all of the cards that have been dealt to every player, whether face-up or face-down. The omniscient player cannot predict cards yet to be dealt at random from the remaining deck, but they are in the best position to predict the outcome of the game, in terms of eventual showdown hands.

We engineer simulated games in which each simulated player i maintains the following information resources:

- $^i htpd$ for his own hand, based on his current hand and cards still possibly remaining the deck, according to that player's knowledge.
- $^i pdds$ for each of his opponents. Opponents' possible downcards are successively pared as cards are dealt face up throughout the game. Additionally, $pdds$ are reweighed for opponents' betting actions according to equation (4).
- $^i htpds$ for each of the opponents, generated from the weighted $^i pdds$ according to equation (2).
- estimated probability of each player winning at showdown, calculated from the $^i htpds$ according to equation (1).

In the simulation, each player bets randomly according to the default betting model probabilities described in Section IV, and each player has a perfect opponent model, used in re-weighting the $pdds$, that every other player bets according to this betting policy.

The experiment is instrumented with the omniscient view of every player's downcards, hence their true $htpds$. The experimental subject is the public knowledge observer. The public knowledge observer maintains estimated $pdds$, $htpds$, and chances of winning for every player, but it lacks knowl-

edge of any downcards. Each player possesses slightly more information than the public knowledge observer (namely that player's two downcards), but the public knowledge observer constitutes a universal standpoint that does not depend on privileged information and is best suited to extending this analysis to real poker games observed from the sidelines.

We measure the information gained by exploiting observations of bet actions by comparing the public knowledge observer's probability estimate of each player winning, q , with that of the omniscient viewpoint, p . From omniscient knowledge, at any stage of the game the entropy H of the outcome probability distribution is

$$H = - \sum_i p_i \log_2 p_i. \quad (6)$$

One way of interpreting the entropy is this. For any game outcome, if the known probability of player i winning is p_i , then the Shannon theoretical optimum amount of information required to communicate that game's eventual outcome is $-\log_2(p_i)$. The entropy is the average of this, i.e. the average information required to communicate outcomes sampled from the distribution p .

If instead one possesses an imperfect estimated probability of winning distribution, q , then the average information cost of transmitting the outcome of games is $-\sum_i p_i \log_2 q_i$. The difference between this quantity and the actual entropy gauges the amount of information lost by the distribution q as compared to the true distribution p ; this is the Kullback-Leibler divergence,

$$KL = p \log(p/q). \quad (7)$$

If q represents any agents' imperfect estimates about the uncertain outcome of the game, the K-L divergence tells how far this estimate is from the optimal estimate reflected in the true entropy H .

VI. EXPERIMENTAL RESULTS

In simulated games, we may compute the K-L divergence between the omniscient probability for each player winning, p , and the estimated distribution q under two conditions. The *cards-only* condition updates public knowledge $pdds$ only by pruning possible downcards as they are dealt face up and hence removed from the deck. This condition gives rise to public knowledge probability of winning distributions q^c that ignore bet actions. The *bet-inference* condition prunes $pdds$ in this way, but additionally uses players' bet actions to reweigh the public knowledge $pdds$ as described in Section III giving rise to prob-win distributions q^b that are informed by bet actions and perfect opponent models.

Results for 1827 simulated games are plotted in Figure 6. The horizontal axis represents ten distinct information stages of a seven card stud game. Stages 3D, 4D, 5D, 6D, 7D measure information immediately following dealing of cards, while stages 3B, 4B, 5B, 6B, 7B occur following a round of betting. The thick solid green line (lower solid line) is the entropy of the probability of winning distribution p_i .

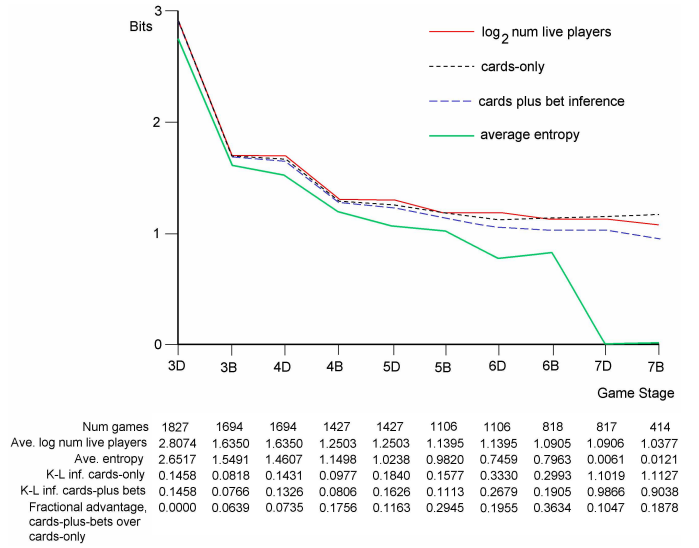


Fig. 6. Information gain results.

The thin solid red line is the \log_2 of the number of players, which corresponds to the average information cost when all active players are believed equally likely to win. The dashed lines are information measures for the cards-only and bet-inference conditions. These are simply the entropy added to the K-L distance for these conditions. The information advantage of exploiting players' bet actions is reflected in the lower positioning of the cards-plus-bet-inference curve with respect to the cards-only curve.

Figure 6 averages these measures over the 1827 simulated games. Game stages are included in the average only when they include at least two active players. To give a sense of the diversity of games over which the average is taken, Figure 7 plots the entropies of a subsample of individual games. In any individual game the entropy, or uncertainty about which player will win if they stay through showdown, tends to decrease. But by the luck of the cards, this can increase if a player suddenly catches a very good card. On average, however, the entropy decreases except at stage 6B. By 6th street, in most games, most players have folded. The simulated players are smart enough to fold if it appears clear that they have little chance of winning, that is, if the entropy for the game is probably low and they are on the losing end. Therefore most low entropy games are concluded by Stage 6B and the average entropy over remaining live games increases.

The numbers below the graph of Figure 6 tabulate the following quantities: the number of games still going at that stage so included in the average; average log number players; average entropy; average K-L distances under the two conditions; and fractional information gain obtained by exploiting opponents' bet actions, as opposed to calculating prob winning based only on dealt cards. The greatest percentage gain is at Stage 6B, immediately following the betting at sixth street, when the bet-inference public knowledge

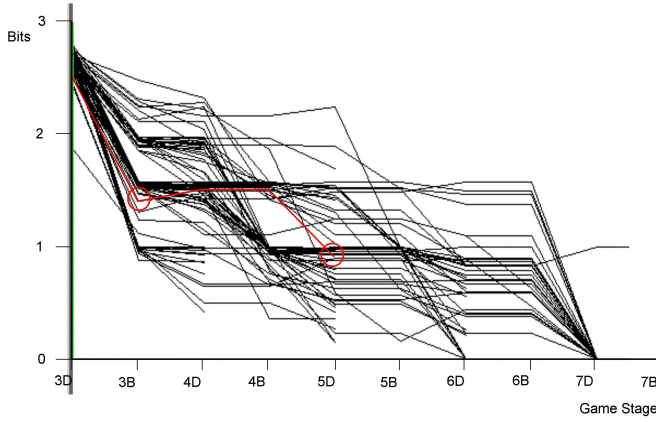


Fig. 7. Entropies for a sampling of 80 games played by the default simulated players. Circles identify two stages in the game whose history is shown in Figure 8. The *htpds* at these stages are shown in Figure 1.

	Bet activity @ street			Net won/lost	Cards
	3	4	5		
Seat 1	f			-0.25	4D JH 6S
Seat 2	f			-0.25	KC 2D 5C
Seat 3	c	.b	.bf	-2.50	9D 7D 7S AC AH
Seat 4	f			-0.25	2H QD 6C
Seat 5	.Bc	c	f	-1.75	4H 9H 4S 8H 3D
Seat 6	f			-0.25	8D 4C 5D
Seat 7	r	c	r	5.25	TC KH TD 5S TH
Final Pot: 8.75					

Fig. 8. Game history for a sample game whose entropy is plotted in red in Figure 7. Notation: “.” denotes lead actor at each street; “B”: Bring-in bet; “k”: check (no one checked in this particular game); “b”: bet; “f”: fold; “c”: call; “r”: raise.

observer gains a 36% information advantage over the cards-only observer. The percent advantage drops at Stages 7D and 7B simply because at this point all the cards have been dealt and the omniscient observer knows the outcome of the game. The optimal baseline entropy is zero here so the percentage gain of the bet-inference condition is smaller even though the magnitude of its information gain over the cards-only condition increases. (A nonzero entropy at Stages 7D or 7D indicate that a tie between two or more players occurred in a few games.)

An interesting feature of Figure 6 is that the cards-only condition for predicting game outcome actually performs worse than chance at seventh street. This is an indication that if a player remains in the game while their four upcards show a weaker hand than opponents’, then this player must have a strong hand hidden. The cards-only estimation of hand strength has no way of accounting for this, while the bet-inference condition successfully makes this inference in the course of the *pdd* reestimation procedure described in Section III.

It would be a mistake to read Figure 6 as suggesting that estimation of opponents’ possible downcards is of little value

simply because the cards-only and cards-plus-bet-inference curves look similar to the naive $\log_2 N$ curve in comparison to the true entropy. These curves were generated from simulated games whose players followed tight-aggressive deceptive bet policies dictated by Figure 5, and therefore the active players at each street had undergone a severe, informed self-selection procedure of folding perceived disadvantaged hands. Note also that success in poker often hinges on exploitation of relatively few big-pot hands; flat averages of information gain such as Figure 6 may not reflect this differential value of information.

How does this information advantage translate to win/loss rates? We performed a second experiment in which three players were constrained to be cards-only players by permitting them to use only their visible card knowledge in estimating their probability of winning, and hence in deciding their bet actions. In other words, the *pdd* re-estimation procedure exploiting opponents’ bet actions was omitted for these players. The remaining four players were provided this information; their opponent models used to infer hand strength from bet actions accurately reflected that four players were making use of the bet-inference public knowledge *htpds* in calculating their own chances of winning prior to every bet decision. These four players used the cards-only public knowledge *htpds* as a best available approximation to the beliefs held by the constrained, cards-only players, who know but obviously do not share their own downcard information.

It is well known that poker win/lose outcomes occur with high variance. Over 8977 simulated games, the resulting win/lose rates are shown in Figure 9. The four bet-inference players won on average .14 bets/game, while the three cards-only players lost on average .19 bets/game. This is clearly attributable to the cards-only players not folding when they should have. The bet-inference players bet an average of 1.57/hand and won pots at a average rate of 1.71/hand (netting .14/hand). The cards only players won significantly more pots, 2.50/hand, but at the cost of betting an average of 2.69/hand.

VII. DISCUSSION AND CONCLUSION

It is by no means surprising that it is advantageous to exploit information transmitted by opponents’ bet actions in poker. This paper has introduced a framework for doing so in a way that delineates the roles of exposed cards, calculation and comparison of possible hand outcomes, rational bet strategy, styles of play, opponent models, and knowledge and belief carried by players and observers. Using this apparatus, we have obtained experimental results quantifying information gain and its implications for win/lose rate by simulated deceptive players who possess perfect models of their opponents’ betting policies.

To extend these results to live poker games would raise several major challenges.

First, unless studies could be conducted from behind the House or game host’s omniscient viewpoint, in real games we would lack information about players’ downcards except

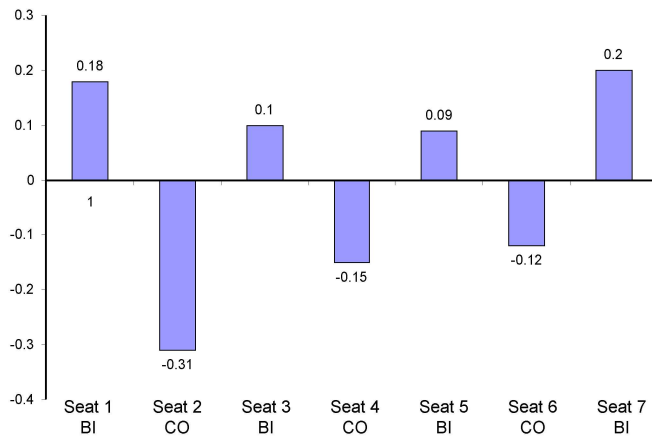


Fig. 9. Win/lose rates per hand for seven card stud players using the default betting model for players who do infer information about opponents' hand strengths (BI seats 1, 2, 5, 7) versus those who use only visible card information (CO seats 2, 4, 6). Averages are over 8977 simulated games.

when they stayed in to showdown. This limitation would prohibit accurate calculation of the omniscient probabilities of winning through the game. In good seven card stud games such disclosure happens relatively rarely. Moreover, the omniscient prob-win distribution requires knowledge of *all* dealt cards, not just those of players who reveal their downcards at showdown. This information is virtually never available. Conceivably, win/lose probabilities under different situations can be estimated from actual outcomes and extrapolated from whatever downcards do get exposed. It seems however that the sample size needed to approximate omniscient knowledge would be prohibitive. Therefore, it appears likely that Kullback-Leibler information measures based on omniscient knowledge can be pursued only under laboratory conditions.

Second, real games do not afford ready access to players' bet policies. Human players especially are likely to decide their bets on complex, variable, and contextually contingent criteria. Experienced poker players enjoy the process of observing other players and getting a fix on their styles of play. This translates to a very nice challenge for machine learning investigations, first to attempt to model and map the varieties of styles, and second to bring this knowledge to bear to infer particular opponents' habitation in the large space of playing styles, from a small number of observations. For example, Southey et. al. have experimented with sampling over prior distributions of possible opponents to enhance belief in those whose behaviors fit that of observed opponents [14].

This paper's experimentally observed benefits of opponent modeling are in a sense an upper bound because our simulated players possess perfect models of their opponents' betting policies. In more realistic scenarios, opponent models will be imperfect and players' policies may shift over time. The degradation in information advantage due to these factors is subject to further experimental investigation.

Finally, the use of artificial intelligence to offer real-time advice or automated play would require not only retrospective analysis of opponents' likely hand strength, but also forward reasoning about the expected value of potential bet actions. This is the subject of much of the work in AI for poker. One benefit of forward reasoning will be strengthening of the estimate of effective odds by better estimating the number of opponents to remain active through future rounds of betting. The effective odds calculation in the present study is quite rudimentary, although in the formulation presented systematic overestimates or underestimates in effective odds can be mitigated by adjustment of the parameters of the fold/call betting policy.

Poker is an important member of the class of games for which effective play lies not simply in out-calculating one's opponent with regard to the objective state of the game. Instead, poker is in a fundamental sense a game of minds against minds. This paper offers a glimpse of how we may cast in formal mathematical and algorithmic terms the processes of trying to figure out what intentional opponents know, what they believe, what opponents believe about what oneself believes, ad infinitum. Because of the myriad complexity and subtleties involved, poker would appear to offer a model system for investigations of the most perplexing epistemological questions of computational intelligence engaging intentional agents.

REFERENCES

- [1] B. Alspach; "7-Card Poker Hands", "http://www.math.sfsu.ca/~alspach/comp20", 2000.
- [2] D. Billings, L. Pena, J. Schaeffer, D. Szafron; "Using Probabilistic Knowledge and Simulation to Play Poker," *Proc. AAAI-99*, 1999.
- [3] D. Billings, A. Davidson, J. Schaeffer, and D. Szafron; "The Challenge of Poker," *Artificial Intelligence Journal*, Vol 134(1-2), pp 201-240, 2002.
- [4] D. Billings, "The First International RoShamBo Programming Competition," <http://www.cs.ualberta.ca/~darse/rsb-results1.html>, 1999.
- [5] K. Burns; "Heads-Up Face-Off: On Style and Still in the Game of Poker," *AAAI Fall Symposium on Style and Meaning in Language, Art, Music, and Design*, AAAI Technical Report FS-04-07, 2004.
- [6] K. Burns, "Pared-down Poker: Cutting to the Core of Command and Control," *Proceedings of the IEEE Symposium on Computational Intelligence and Games (CIG05)*, Essex University, Colchester, Essex, 2005.
- [7] T.S. Ferguson, C. Ferguson, and C. Gawargy, "Uniform(0,1) Two-Person Poker Models", "http://www.math.ucla.edu/~tom/papers/poker2.pdf", 2004.
- [8] D. Koller and A. Pfeffer, "Representations and Solutions for Game-Theoretic Problems," *Artificial Intelligence*, 94(1), pp. 167-215, 1997.
- [9] K.B. Korb, A.E. Nicholson, and N. Jitnah, "Bayesian Poker," *Proc. of Uncertainty in Artificial Intelligence*, pp. 343-350, Stockholm, Sweden, August, 1999.
- [10] J. McDonald, *Strategy in Poker, Business and War*, Norton, New York, 1950.
- [11] A. Prock; *Pokerstove.com* "http://www.pokerstove.com", 2004.
- [12] D. Sklanasky; *The Theory of Poker*, Two Plus Two Publishing, Henderson, NV, 1987.
- [13] D. Sklanasky, M. Malmuth, R. Zee; *Seven Card Stud For Advanced Players*, Two Plus Two Publishing, Henderson, NV, 1989.
- [14] F. Southey, M. Bowling, B. Larson, C. Piccione, N. Burch, and D. Billings; "Bayes' Bluff: Opponent Modeling in Poker," *Proc. 21st Conf. on Uncertainty in Artificial Intelligence (UAI '05)*, 2005.