Goal Problems in Gambling Theory*

Theodore P. Hill
Center for Applied Probability and
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0160

Abstract

A short introduction to goal problems in abstract gambling theory is given, along with statements of some of the main theorems and a number of examples, open problems and references. Emphasis is on the finite-state, countably-additive setting with such classical objectives as reaching a goal, hitting a goal infinitely often, staying in the goal, and maximizing the average time spent at a goal.

1. Introduction

In the classical gamblers' ruin problem (cf. Feller [F]) a gambler beginning with x_0 dollars bets \$1 at each play, winning a dollar with probability p and losing a dollar with probability $q = 1 - p$, until he either goes broke (hits 0) or else reaches a given goal such as \$1000. The successive wins/losses are independent, and he has no control over the game or decisions to influence his likelihood of winning.

In many real-life gambling situations, however, the gambler is not forced to make the same wager at each step, but rather he is free to select from a variety of strategies. For example, in red-and-black (cf. [DSa]) the gambler may at each step bet any amount b less than or equal to his current fortune x and he then wins b dollars with probability p and loses b dollars with probability q (i.e., he moves to $x + b$ with probability p and to $x - b$ with probability q).

It is this freedom of selection of strategies which is the essence of abstract mathematical gambling theory, and the purpose of this article is to give a brief introduction

---

to this rich and deep sub-field of discrete-time stochastic control, with emphasis on finite-state (hence countably-additive) gambling problems with fixed goals. As such, the ideas reviewed here are a small part of the general abstract theory of gambling which includes general state spaces (including questions of measurable and general non-measurable selection strategies) and general payoff or utility functions. Dubins and Savage's classic treatise [DSa] is the basis for many of the ideas presented here; for a more extensive introduction to the subject, and its application to the field of stochastic games, the reader is referred to Maitra and Sudderth's ([MS1], [MS2]) recent articles. Abstract gambling theory, in turn, is part of the mathematical theory of stochastic processes, and is closely related to Blackwell's (e.g., [Bl]) theory of dynamic programming.

2. Gambling Problems and Strategies

The purpose of this section is to give a formal mathematical description of gambling problems and strategies; notation will generally follow that in [MS1].

A *gambling problem* is a triple (S, Γ, u) where: S is the *state space* (in this exposition always a nonempty finite set), which represents the gambler's possible fortunes; Γ is a function from S to nonempty sets of probability measures defined on (the sigma algebra of all subsets of) S, where $\Gamma(x)$ represents the bets or gambles available to the player when his fortune is $x \in S$; and u is a bounded function (called the *payoff function*) from $S^{\mathbb{N}}$ to the real numbers which the gambler is trying to maximize.

Example 2.1. (Red-and-black) $S = \{0, 1, \ldots, 2g - 2\}$, $\Gamma(x) = \{\delta(x)\}$ for $x \geq g$, and for $x < g$, $\Gamma(x) = \{p\delta(x + j) + q\delta(x - j) : j = 0, 1, \ldots, x\}$, where $\delta(a)$ is the dirac delta measure with mass 1 on the singleton $\{a\}$. The goal here is to reach at least g, so u is the function $u(x_1, x_2, \ldots) = 1$ if $x_j \geq g$ for some j, and $= 0$ otherwise. Note that $\Gamma(0) = \{\delta(0)\}$ (so 0 and all states $x \geq g$ are absorbing states) and that for this classical formulation of red-and-black, a sequence of fortunes which never reaches g is worthless, no matter how close to g it comes.

A *strategy* available to the gambler who has initial fortune x_0 is a selection rule σ

which associates to each finite sequence of fortunes $x_0, x_1, x_2, \ldots, x_n$ a single gamble $\sigma(x_0, x_1, \ldots, x_n) \in \Gamma(x_n)$. Thus an initial gamble $\sigma(x_0) \in \Gamma(x_0)$ is selected, and then for every subsequent stage of play, one single gamble is identified from the set of gambles available at the current fortune. In general, strategies may be heavily dependent on the past histories of fortunes, but two important natural and comparatively simple classes of strategies are those of *stationary strategies*, which depend only on the current fortune, and *Markov strategies*, which depend on the current fortune and time. Thus a stationary strategy satisfies $\sigma(x_0, x_1, \ldots, x_n) = \gamma(x_n)$ for all $x_0, \ldots, x_n \in S$, where for each $x \in S$, $\gamma(x) \in \Gamma(x)$ is a fixed gamble which will be used whenever the current fortune is x; each such strategy corresponds to a single Markov chain on S with stationary transition probabilities (selected to be in Γ). Similarly, a Markov strategy corresponds to a Markov chain on S whose transition probabilities at time n are independent of past successions of states, formally, $\sigma(x_0, x_1, \ldots, x_n) = \sigma(x_0, x_1', x_2', \ldots, x_n')$ whenever $x_n = x_n'$. (Note that by definition all stationary strategies are Markov.)

Example 2.2. (*Timid play* in red-and-black). Define the strategy σ_t by $\sigma_t(x_0, x_1, \ldots, x_n) = p\delta(x_n + 1) + q\delta(x_n - 1)$ for all $0 < x < g$. This is the stationary strategy which selects the smallest possible nonzero bet at each (non-absorbing) state.

Example 2.3. (*Bold play* in red-and-black). Define the stationary strategy σ_b by $\sigma_b(x_0, x_1, \ldots, x_n) = p\delta(x_n + k) + q\delta(x_n - k)$ where for $x_n < g$, $k = \min\{x_n, g - x_n\}$ is the largest available bet which will not overshoot the goal. So in particular if $g = 1000$, σ_b prescribes a bet of exactly of 3 dollars if the current fortune is either 3 or 997, etc.

Example 2.4. (A nonstationary Markov strategy in red-and-black). Define σ_M by $\sigma_M(x_0, x_1, \ldots, x_n) = \sigma_b(x_0, \ldots, x_n)$ if $n < 10$, and $= \sigma_t(x_0, \ldots, x_n)$ if $n \geq 10$. This simple Markov strategy switches from bold play to timid play at time 10.

Example 2.5. (A non-Markov strategy in red-and-black). Define σ_N by $\sigma_N(x_0, x_1, \ldots, x_n) = \sigma_b(x_0, x_1, \ldots, x_n)$ if $x_1 < 3$, and $= \delta(x_n)$ otherwise. This somewhat irrational strategy uses bold play at time n if the fortune at time 1 was < 3, and otherwise stagnates forever.

Each strategy σ induces a countably-additive probability measure P_σ on the Borel subsets of $S^{\mathbb{N}}$ in a natural way as follows. First, $P_\sigma(x_1 \times S^{\mathbb{N}}) = \sigma(x_0)(\{x_1\})$, i.e., the probability that the sequence starts with x_1 is the measure of x_1 under the initial gamble chosen, $\sigma(x_0) \in \Gamma(x_0)$. Next $P_\sigma((x_1, x_2) \times S^{\mathbb{N}}) = \sigma(x_0)(\{x_1\}) \cdot \sigma(x_0, x_1)(\{x_2\})$, etc. In other words, if $X_1, X_2, \ldots$, represent the fortunes at times $1, 2, \ldots$, then X_1 has distribution $\sigma(x_0)$, and the conditional distribution of X_2 given $X_1 = x_1$ is $\sigma(x_0, x_1)$, etc. Thus a given strategy σ determines the complete joint distribution P_σ of the stochastic process $X_1, X_2, \ldots$ of states via the standard extension from the measures of those cylinder sets.

For example, in bold play σ_b for red-and-black (with $g = 100$, say, and $x_0 = 10$), $P_{\sigma_b}(X_1 = 10) = 0$, $P_{\sigma_b}(X_1 = 20, X_2 = 40) = p^2$, $P_{\sigma_b}(X_1 = 20, X_2 = 0) = pq$, etc.

3. The Optimal Return Function and the Value of a Strategy

Recall that the payoff function u is a bounded real-valued function on $S^{\mathbb{N}}$, the set of all infinite sequences of states (often called *histories*), and that each strategy σ determines a unique probability measure on the Borel subsets of $S^{\mathbb{N}}$.

The *value of strategy* σ is simply the expected payoff $V(\sigma)$ with respect to the probability distribution P_σ defined by σ, that is

$$V(\sigma) = \int_{S^{\mathbb{N}}} u \, dP_\sigma$$

Example 3.1. In red-and-black, where the objective is simply to reach the goal g (i.e., $u(x_1, x_2, \ldots) = 1$ if $x_i \geq g$ for some i, and $= 0$ otherwise), the value $V(\sigma)$ of the strategy σ is simply

$$V(\sigma) = P_\sigma(X_i \geq g \text{ for some } i \geq 1),$$

and since the timid-play strategy $\sigma_{t,x}$ starting at x corresponds to the classical gambler's ruin problem, its value is simply the probability that a simple random walk

4

which starts from x and moves to the right with probability p will reach g before reaching 0, and the classical solution (cf. [DSa, p. 170]) is

$$V(\sigma_{t,x}) = \left(\frac{p}{q}\right)^{g-x} \frac{1 - \left(\frac{p}{q}\right)^x}{1 - \left(\frac{p}{q}\right)^g}$$

Example 3.2. The corresponding value for bold play in red-and-black starting at x, $\sigma_{b,x}$, is more complicated, but for small state spaces may be calculated easily as follows in the case $g = 3$. $V(\sigma_{b,0}) = 0$ and $V(\sigma_{b,g}) = 1$ by definition, and $V(\sigma_{b,1})$ and $V(\sigma_{b,2})$ satisfy $V(\sigma_{b,1}) = qV(\sigma_{b,0}) + pV(\sigma_{b,2})$, $V(\sigma_{b,2}) = qV(\sigma_{b,1}) + pV(\sigma_{b,g})$, which implies $V(\sigma_{b,1}) = p^2/(1 - qp)$ and $V(\sigma_{b,2}) = p/(1 - qp)$. (Note that for $g = 3$, $\sigma_t = \sigma_b$.)

The *optimal payoff function* V is the real-valued function on S which is the optimal expected payoff over all possible strategies starting at x

$$V(x) = \sup\{V(\sigma) : \sigma \text{ is a strategy starting at } x\}.$$

That is, $V(x)$ represents the best a gambler can do starting at x.

An *optimal strategy* starting at x is a strategy at x satisfying

$$V(\sigma) = V(x).$$

In general, the values of general strategies and optimal payoff functions are difficult to obtain explicitly, and a major contribution of [DSa] is a general theory which characterizes optimal payoff functions.

For red-and-black, however, it is known that for favorable games (i.e., $p > 1/2$) timid play is optimal [Br], and for unfavorable games ($p \leq 1/2$) bold play is optimal [DSa, Theorem 5.3.1]. This corresponds to the intuitively plausible idea that for favorable games, the random walk will tend to drift to the right, so taking steps as small as possible will make this drift look almost deterministic, making the probability of reaching the goal high. For unfavorable games however, the drift is to the left, and small bets will be a poor strategy for exactly the same reason; playing just the opposite by maximizing bets (bold play) is much better.

4. Goal Problems

In a goal problem there is a fixed set of *goal states* $G \subset S$, and the payoff function u reflects various objectives associated with G such as reaching G at least once, hitting G infinitely often, staying in G forever, or maximizing the average time spent in G. This section will survey some of the basic theorems about goal problems and give examples to illustrate them.

Reaching a Goal. If S is finite and $\Gamma(x)$ is finite for each $x \in S$ (that is, each state has only a finite number of gambles available at it), then there is always a stationary optimal strategy [DSa, Theorem 3.9.1] for the payoff

$$u(x_1, x_2, \ldots) = 1 \text{ if } x_i \in G \text{ for some } i \geq 1, \text{ and } = 0 \text{ otherwise.}$$

In red-and-black, bold play σ_b is one such optimal stationary strategy if $p \leq \frac{1}{2}$, and timid play σ_t if $p > \frac{1}{2}$. If $\Gamma(x)$ is infinite for some x, there may exist no optimal strategies.

Example 4.1. $S = \{0, 1, g\}$, $\Gamma(0) = \{\delta(0)\}$, $\Gamma(1) = \Gamma(g) = \{p\delta(g) + (1-p)\delta(0) : p < 1\}$. Then $V(0) = 0$, $V(1) = 1$, but $V(\sigma_s) < 1$ for any stationary strategy σ_s starting at state 1.

As is seen in the above example, however, there do exist arbitrarily good stationary strategies (i.e., for each $\epsilon > 0$ there is a stationary strategy σ_{s_ϵ} with $V(\sigma_{s_\epsilon}) \geq 1 - \epsilon$). This is always the case for finite-state goal problems with objective reaching the goal, which is a special case of [DSa, Theorem 3.9.2], and which was extended in [O, Theorem B] to countable S, and in [Su, Theorem 2.3] to a much larger class of problems including many with uncountable S and finitely additive transition probabilities. Generalizing in other directions, for countable S there is always a stationary strategy which both (nearly) maximizes the probability of reaching the goal and (nearly) minimizes the expected time to the goal [DH1, Theorem 4.2]. For finite S there is always a Markov strategy which is monotonically improving and optimal is the limit along every history [HvW, Theorem 5.15], but such a strategy may not

6

always be constructed by simply switching to a "better" gamble at each successive return to a state; in some cases it is necessary to use certain gambles for arbitrary long time periods, then switch to a different gamble for an even larger time, and so on, as can be seen Example 4.2 below.

Hitting a Goal Infinitely Often. In some goal problems, the objective is not just to reach the goal once, but to hit the goal infinitely often with as high probability as possible, so

$$u(x_1, x_2, \ldots) = \limsup_{n \to \infty} I_G(x_n),$$

where $I_G(\cdot)$ is the indicator function of the goal set G. Even if S has as few as three states, there may be strategies which hit the goal infinitely often with high probability, but where each *stationary* strategy is worthless for this objective (Example 4.1 above). On the other hand, in all finite-state goal problems, there always exist Markov strategies which nearly maximize the probability of hitting the goal infinitely often [Hi, Theorem 8.1], but the structure of such good Markov strategies in general is much more complicated than just using a better gamble at each time period.

Example 4.2. $S = \{0, 1, g\}$, $\Gamma(0) = \{\delta(0)\}$, $\Gamma(g) = \{\delta(1)\}$, $\Gamma(1) = \{3^{-n}\delta(0) + 2^{-n}\delta(g) + (1 - 3^{-n} - 2^{-n})\delta(1) : n \in \mathbb{N}\} \cup \{\delta(1)\}$. Note that state 0 is absorbing, state g is reflecting, and each of the sets of gambles $\Gamma(x)$ are closed. It can be checked using the conditional version of Borel-Cantelli (Lévy 0-1 Law) that there exist Markov strategies which are arbitrarily good (and which necessarily use successive gambles for long time periods, roughly the order of 2^n times), that no Markov strategy using each gamble at most a fixed number of times N is good, and that limits of good strategies are not good, since every good strategy necessarily uses gambles closer and closer to $\delta(1)$, itself a worthless gamble.

Staying in the Goal. Analogous to the lim sup objective of hitting a goal infinitely often, the payoff function

$$u(x_1, x_2, \ldots) = \liminf_{n \to \infty} I_G(x_n)$$

7

reflects the objective of maximizing the probability that the process eventually stays in the goal set G forever. For countable S with this $\liminf$ payoff (and in fact for a much larger class of payoffs), there always exist arbitrarily good Markov strategies [HP, Theorem 4.1]. That neither good stationary strategies nor optimal Markov strategies exist in general can be seen in Example 4.1 above.

Maximizing the Average Time at a Goal. An objective function which is between that of reaching a goal set once and eventually staying in the goal set forever is the objective of maximizing the average time spent in the goal, that is

$$u(x_1, x_2, \ldots) = \limsup_{n \to \infty} n^{-1} \sum_{j=1}^{n} I_G(x_j)$$

(for finite S, $\limsup$ and $\liminf$ are equivalent [DH2, Corollary to Theorem 2]).

If S is finite and $\Gamma(x)$ is a *closed* set (total-variation norm) of transition probabilities for each x, then there always exist nearly-optimal stationary strategies for the objective of maximizing the average time in the goal [DH2, Theorem 1], in contrast to the objective of hitting a goal infinitely often, as seen in Example 4.2 above. If Γ is arbitrary, there always exist nearly-optimal Markov strategies for this objective [DH2, Theorem 2]. As with previous objectives, the limit of nearly-optimal strategies may be worthless, as can be seen modifying Example 4.2 by letting $\Gamma(g) = \delta(g)$; all good strategies starting at 1 use gambles close to $\delta(1)$, which is itself worthless for the average-time objective as well.

Other Goal Payoffs. Other common payoffs in goal problems include those of discounted rewards ($u(x_1, x_2, \ldots) = \sum_{n=1}^{\infty} \beta^n I_G(x_n)$ for $0 < \beta < 1$ fixed), finite-horizon rewards (e.g., $u(x_1, x_2, \ldots) = \sum_{n=1}^{T} I_G(x_n)$ for fixed T), supremum rewards, product rewards, etc.; the interested reader is referred to [HP].

The proofs and applications of the above-mentioned results span a wide spectrum of standard probability theory including Markov chains, martingales, 0-1 laws, strong laws, randomizations, stochastic matrices, and dynamic programming, as well as intricate *ad hoc* arguments and a broad and deep general theory of inequalities

8

for stochastic processes in [DSa]; the references below are starting points to this interesting and powerful subject.

5. Open Problems

The fundamental gambling-theory text by Dubins and Savage [DSa] contains a well-indexed list of open problems in this field; the following are several additional ones.

Q1. (cf. [Hi]) Do good Markov strategies exist in all countable-state goal problems with objective of hitting the goal infinitely often?

Q2. (cf. [HvW]) Do monotonically-improving limit-optimal strategies exist in all finite-state goal problems with average reward payoff?

Q3. (cf. [St]) For what general class of goal problems do good measurable strategies always exist?

Q4. (cf. [Ho]) What are efficient algorithms for computation of optimal or good strategies in various classes of goal problems?

References

[Bl]   Blackwell, D. (1965) Discounted dynamics programming, *Ann. Math. Statist.* **36**, 226–235.

[Br]   Brieman, L. (1961) Optimal gambling systems for favorable games, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* **1**, 65–78, Univ. Calif. Press.

[DH1]   Demko, S. and Hill, T. (1981) Decision processes with total-cost criteria, *Ann. Probab.* **9**, 293–301.

[DH2]   ———, (1984) On maximizing the average time at a goal, *Stoch. Processes and Their Applic.* **17**, 349–357.

[DSa]   Dubins, L. and Savage, L. (1976) *Inequalities for Stochastic Processes: How to Gamble if You Must*, Dover, New York.

[DSu]   Dubins, L. and Sudderth, W. (1977) Persistently ϵ-optimal strategies, *Math. Operations Research* **2**, 125–135.

[F]   Feller, W. (1968) *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd edition, Wiley, New York.

[Hi]   Hill, T. (1979) On the existence of good Markov strategies, *Trans. Amer. Math. Soc.* **247**, 157–176.

[HP]   Hill, T. and Pestien, V. (1987) The existence of good Markov strategies for decision processes with general payoffs, *Stoch. Processes and Their Applic.* **24**, 61–76.

[HvW]   Hill, T. and van der Wal, J. (1987) Monotonically improving limit-optimal strategies in finite-state decision processes, *Math. Operations Research* **12**, 463–473.

[Ho]   Howard, R. (1960) *Dynamic Programming and Markov Process,* Technology Press, Cambridge, MA.

[MS1]   Maitra, A. and Sudderth, W. (1995) An introduction to gambling theory and its applications to stochastic games, preprint.

[MS2]   ———, (to appear) *Discrete Gambling and Stochastic Games.*

[O]   Ornstein, D. (1969) On the existence of stationary optimal strategies, *Proc. Amer. Math. Soc.* **20**, 563–569.

[St]   Strauch, R. (1967) Measurable gambling houses, *Trans. Amer. Math. Soc.* **126**, 64–72.

[Su]   Sudderth, W. (1969) On the existence of good stationary strategies, *Trans. Amer. Math. Soc.* **135**, 399–415.