

A Markov Decision Process-based handicap system for tennis

T.C.Y. Chan^{1,2} and R. Singal^{*,1,3}

¹*Department of Mechanical and Industrial Engineering, University of Toronto*

²*Centre for Healthcare Engineering, University of Toronto*

³*Department of Industrial Engineering and Operations Research, Columbia University*

January 3, 2017

E-mail addresses: chan@mie.utoronto.ca (T.C.Y. Chan), rs3566@columbia.edu (R. Singal)

*Corresponding author.

Abstract: Handicap systems are used in many sports to improve competitive balance and equalize the match-win probability between opponents of differing ability. Recognizing the absence of such a system in tennis, we develop a novel optimization-based handicap system for tennis using a Markov Decision Process (MDP) model. In our handicap system, the weaker player is given β “free point” or “credits” at the start of the match, which he can use before the start of any point during the match to win the point outright. The MDP model determines two key features of the handicap system: (1) Fairness: the minimum value of β required to equalize the match-win probability, and (2) Achievability: the optimal policy governing usage of the β credits to achieve the desired match-win probability. We test the sensitivity of the handicap values to the model’s input parameters. Finally, we apply the model to real match data to estimate professional handicaps.

Keywords: Markov Decision Process, Optimization, Fairness, Dynamic handicap

Contents

1	Introduction	3
2	Tennis scoring background	6
3	Markov Decision Process model	7
3.1	Server-independent model	7
3.2	Value function interpretation	10
3.3	Server-specific model	11
3.4	Computing handicaps	13
3.4.1	Server-independent model computations	13
3.4.2	Server-specific model computations	15
3.5	Model sensitivity to point-win probabilities	18
3.5.1	Sensitivity of the server-independent model	18
3.5.2	Sensitivity of the server-specific model	19
4	Handicapping the “Big Four”	20
5	Conclusion	23

1 Introduction

Handicap systems are used in many sports and games to improve competitive balance and equalize the probability of winning a match between two competitors. Golf has perhaps the most well-known handicap system (USGA, 2015), but many other sports and games such as horseracing (BHA, 2015), chess (Hooper and Whyld, 1992), and go (BGA, 2015) have well-established handicapping rules. Handicap systems are often used at the grassroots or amateur level to make competition more enjoyable. They also have natural and widespread applications in sports betting, and are closely connected to ratings systems.

Tennis does not have an official handicap system, although several ideas have been proposed. For example, Sir Richard Branson developed “Richard’s Rules”, a set of handicapping rules that give the weaker player “credits” that can be used to purchase advantages throughout the match (Branson, 2015). The credits can be used to purchase points directly or to provide an advantage such as eliminating the opponent’s second serve. These rules are actually used in the Necker Cup, a pro-am tournament hosted by Branson. Another example comes from the Tencap rating system, which uses the (integral) ratings difference between two tennis players to determine how many points the weaker player can claim at any time in a set (Tencap, 2015). The simplest idea for handicapping involves giving the weaker player an initial lead in the match. The first two examples represent *dynamic* handicap systems, where an advantage by the weaker player can be claimed at any time in the match. In contrast, the initial-lead handicap system is *static*.

Two questions arise when designing handicap systems in general: 1) *Fairness*: How can skill differences be rigorously mapped to a handicap in order to properly equalize the match-win probability? 2) *Achievability*: How should the handicap be implemented during the match to actually achieve the desired competitive balance? Fairness is an issue in both dynamic and static systems, whereas achievability is an issue primarily for dynamic systems. The Branson and Tencap systems described above specify how a handicap should be calculated, but do not quantify the fairness or achievability of that calculation. Before

describing the approach we take to address these issues, it is important to understand the “hierarchical” nature of scoring in tennis, which complicates the design of a tennis handicap system.

The scoring systems of different sports can generally be classified into one of two types: non-hierarchical and hierarchical. Sports with non-hierarchical scoring (*non-hierarchical sports*) include soccer, basketball, baseball, hockey, and football. Non-hierarchical scoring means that the final score in a match is simply the sum of the points scored throughout the match. Sports with hierarchical scoring (*hierarchical sports*) include tennis, volleyball, badminton, and match play golf. The hierarchy refers to the fact that a match is decomposed into several subunits, each with its own scoring system. For example, a tennis match is composed of sets, which are composed of games, which are composed of points.

In non-hierarchical sports, because the final score is the sum of the points scored throughout, each point contributes equally to the final score. However, this is not the case in hierarchical sports. If a tennis player wins a point within a game, but ultimately loses the game, the point won by the player does not contribute to the match score. In fact, it is possible for a player to win more points than her opponent but still lose the match.

It is well-known that in hierarchical sports, some points are “worth” more than others, in terms of their contribution to the match outcome (Barnett et al., 2004; Morris, 1977; O’Donoghue, 2001). For example, assuming point outcomes are independent and identically distributed (iid) (Klaassen and Magnus, 2001), winning a set-clinching point in tennis has a much bigger impact on the match-win probability than winning the very first point in a set. A related feature of hierarchical scoring is that small differences in point-win probabilities may translate into large differences in the match-win probability (Fischer, 1980).

We now return to the two questions posed earlier and illustrate their relevance through the following hypothetical tennis match. Suppose Player A is weaker than Player B. To make the match more competitive, Player B decides to give β “free points” or “credits” to Player A at the start of the match. A credit can be used at any time during the match before the

start of a point; whenever Player A uses one, he wins the point outright. Thus, β represents a handicap that helps the weaker player make the match more competitive. However, the true value of the handicap depends on how close β is to equalizing the match-win probability (i.e., fairness) and when Player A decides to employ the β credits (i.e., achievability). Naturally, these two issues are linked: Player A’s advantage grows with a larger value of β , but shrinks with a worse strategy for claiming the free points.

In this paper, we develop a credit-based, dynamic handicap system for tennis using a Markov Decision Process (MDP) model that addresses both the fairness and achievability criteria described above: it determines both the smallest value of β needed by the weaker player to reach a match-win probability of at least 0.5 and the corresponding optimal policy (i.e., strategy that prescribes an action in every score state) governing credit usage to achieve this equalizing probability. In our system, credits can only be used to purchase points. The MDP model has additional benefits such as generalizing the initial-lead handicap system, and providing novel insight into effort/energy allocation in hierarchical sports.

Our three specific contributions are as follows. We develop the first optimization-based handicap system for tennis. Our MDP approach models a point-by-point decision-making process that explicitly integrates hierarchical scoring and is applicable to matches of general length and structure. The model is tractable using standard methods like policy iteration. It is also flexible, with the ability to accommodate non-iid effects and server-specific point-win probabilities.

Second, we establish that the value function of the MDP represents the state-specific probability of winning the match by the weaker player. As a result, a range of desired match-win probabilities can be generated through an appropriate choice of β , with the handicap defined as the smallest value, β^* , required for the weaker player to meet or exceed a match-win probability of 0.5 (i.e., fairness). We show that even a few credits can have a large impact on the match-win probability. Moreover, the optimal policy represents a strategy for credit use throughout the complete horizon of the match to maximize the match-win probability

(i.e., achievability). In addition, we show that the optimal policy in general employs a “conservative” strategy of credit usage. Finally, we produce a state-dependent definition of “important points”, extending existing literature and forming the basis for a comprehensive decision-making model to guide effort/energy allocation in hierarchical sports.

Third, we demonstrate the model’s application using real match data to estimate handicaps between the “Big Four” mens tennis players (Djokovic, Federer, Murray, Nadal) on different court surfaces, over different years, and using server-specific point-win probabilities.

2 Tennis scoring background

A tennis match is composed of sets, which are composed of games, which are composed of points. Matches are usually a best-of-five or best-of-three sets competition. To win a set, a player must win six games by a margin of at least two games. If the score is 5-games all, then a player needs to win two more games to win the set (final set score of 7-5). If the score in the set reaches 6-6, then a tiebreaker is played to determine the winner of the set, at which point the final set score will be 7-6 for the winner. To win a tiebreaker, a player must win seven points by a margin of at least two points. If the score is 6-points all in the tiebreaker, then the tiebreaker continues indefinitely until one player has a two-point lead. To win a game, a player must win four points by a margin of at least two points. If it is tied at 3-points all, play continues indefinitely until one player has a two-point lead. Instead of counting points won in a game as 1, 2, 3, the points are called “15”, “30”, “40”. Thus, a 3-1 point score within a game is actually referred to as 40-15, and 40-40 is “deuce”.

Different tournaments employ slightly different scoring rules. For example, the four Grand Slam tournaments (biggest, most lucrative, and most competitive) employ a best-of-five sets format for the men, while lower tier tournaments typically employ a best-of-three sets format. All women’s tournaments are best-of-three sets. Three of the four Grand Slams (Australian Open, French Open, Wimbledon) do not employ a tiebreaker in the final set,

and instead the match ends when a player wins the final set with at least a 2-game lead. Only the U.S. Open has a tiebreaker in the final set.

For a full description of the official rules of tennis, including scoring, we refer the reader to [ITF \(2015\)](#).

3 Markov Decision Process model

First, assuming that the point-win probability in tennis does not depend on the server, we formulate an MDP model ([Puterman, 1994](#)) to model the sequential decision-making problem of when to use the credits. Second, we present several properties of the value function. Third, we show how to extend our basic model to include server-specific point-win probabilities. Fourth, we present illustrative handicap computations using the two MDP models. Finally, we explore model sensitivity to the point-win probabilities.

3.1 Server-independent model

The MDP model is based on the classical Markov Chain model for tennis ([Kemeny and Snell, 1960](#)). For ease of exposition in model development, we assume the point-win probability, p , is server-independent in this section. We also make the common assumption that point outcomes are iid, noting that some research has shown that “deviations from iid are small” ([Klaassen and Magnus, 2001](#)). These two assumptions can easily be relaxed in a model with server-specific transition probabilities (see [Section 3.3](#)) and an augmented state space, respectively.

States: The state of the MDP is generally written as $s = (\sigma, \beta)$, where σ refers to the current score in the match and β , a nonnegative integer, refers to the number of free points remaining. Let S be the set of all possible states in the match. Both σ and β are recorded from the perspective of the weaker player, denoted Player A. The score σ is written as a triple (x, y, z) , where x is the match score in sets, y is the set score in games, and z is the

game score in points. For example, the state $((1-0, 2-3, 30-30), 5)$ represents the state of the match where Player A has won the first set, Player A is trailing in the second set by a score of two games to three, the current score in the sixth game of the second set is tied at 30-all, and Player A has five credits remaining in the match. For any score σ , it will be useful to define two other states, σ^+ and σ^- , which denote the scores that result when Player A wins or loses, respectively, the point at hand. We denote by W the “win” score state and L the “lose” score state, corresponding to Player A winning or losing the match, respectively. There is one absorbing state, Δ , corresponding to the end of the match (β is not included in this state definition).

Actions: At each state (σ, β) such that $\beta > 0$, Player A has two possible actions: use a credit ($a = 1$) or not ($a = 0$). For states (W, β) , (L, β) , Δ , or where $\beta = 0$, the only possible action is $a = 0$. The function $\pi : S \rightarrow \{0, 1\}$ that maps all states to corresponding actions is called a *policy*. In other words, a policy dictates whether a free point should be used in any given state.

Transition probabilities: Let $p_{s,s'}^a := P(s'|s, a)$ denote the transition probability from state s to s' when action a is used. We slightly abuse notation and write $p_{\sigma,\sigma'}^a := P(\sigma'|\sigma, a)$ when a focus on score states is more convenient. When a credit is used ($a = 1$), we assume the only possible transition is from score state σ to σ^+ , that is, $p_{\sigma,\sigma^+}^1 = 1$. For action $a = 0$, we assume that $p_{\sigma,\sigma^+}^0 = p$ and $p_{\sigma,\sigma^-}^0 = 1 - p$. (Since the state is recorded from the perspective of the weaker player, $0 < p < 0.5$.) Thus, there are only three possible states that can be reached from any non-absorbing state $s = (\sigma, \beta)$ where $\sigma \neq W, L$: $(\sigma^+, \beta - 1)$, (σ^+, β) , or (σ^-, β) . At states (W, β) and (L, β) , the only possible transition is to Δ . Given the structure of the state space and the transition probabilities, it should be clear that the score states W and L are reachable from any previous score state with sufficiently many point wins or losses, respectively.

Rewards (or utility): We assign a reward of 0 to all states irrespective of the action chosen, except for states of the form (W, β) , which receive a reward of 1. That is, the reward function r satisfies $r(s, a) = 1$ if $s = (W, \beta)$, $a = 0$ and 0 otherwise. Note that “reward” is standard terminology in MDP models, and in this context is equivalent to “utility,” with which certain readers may be more familiar.

Bellman equation: The value of employing a particular policy π can be computed using the standard Bellman equation

$$V^\pi(s) = r(s, \pi(s)) + \sum_{s' \in S} p_{s, s'}^{\pi(s)} V^\pi(s'), \quad (1)$$

where $V^\pi(s)$ is the expected “utility” or value of the rewards associated with policy π at state s , capturing both the immediate value of a particular action in a particular state through the reward function as well as the expected value over all future sample paths. As we will show in Theorem 1, $V^\pi(s)$ is the probability the weaker player wins the match when in state s and using policy π . The objective of the weaker player (the decision-maker in our model) is to maximize her match-win probability. Accordingly, an optimal policy can be determined by the optimality equation

$$V^*(s) = \max_a \left\{ r(s, a) + \sum_{s' \in S} p_{s, s'}^a V^*(s') \right\}. \quad (2)$$

Specializing equation (2) to the tennis context described above, we get the following optimality equation at non-absorbing states where $\sigma \neq W, L$ and $\beta \geq 1$

$$V^*(\sigma, \beta) = \max\{V^*(\sigma^+, \beta - 1), pV^*(\sigma^+, \beta) + (1 - p)V^*(\sigma^-, \beta)\}, \quad (3)$$

where the first term in the max corresponds to $a = 1$ and the second term corresponds to $a = 0$. When $\beta = 0$, only second term remains in equation (3). When $\sigma = W$, $V^*(\sigma, \beta) = 1 + V^*(\Delta)$ and when $\sigma = L$, $V^*(\sigma, \beta) = V^*(\Delta)$. Lastly, $V(\Delta) = V(\Delta)$.

This MDP is an expected total reward model with non-negative rewards, also known as a positive bounded model (Puterman, 1994). For such models, it is known that the minimal solution to the optimality equations is the desired solution, which in our case sets $V^*(\Delta) = 0$ and correspondingly implies $V^*(W, \beta) = 1$ and $V^*(L, \beta) = 0$.

3.2 Value function interpretation

The value function $V^\pi(\sigma, \beta)$ has a natural interpretation in the context of tennis and possesses several intuitive properties, which when taken together establish the existence of the handicap β^* .

Theorem 1. *$V^\pi(\sigma, \beta)$ equals the probability the weaker player wins the match when using policy π given score σ and remaining credits β .*

Proof. Given a policy π , the Bellman equations for $V^\pi(s)$ are exactly the system of equations describing the probability of absorption in a Markov Chain (Bertsekas and Tsitsiklis, 2008) via the win state. Thus, $V^\pi(s)$ is the corresponding probability under policy π . \square

Corollary 1. *$V^*(\sigma, \beta)$ equals the probability the weaker player wins the match under the optimal policy given score σ and remaining credits β .*

Given that $V^*(\sigma, \beta)$ represents the probability Player A wins the match under the optimal policy, the following properties of the value function are immediate.

Lemma 1. *Let σ_0 denote the score at the start of the match. Then,*

1. $V^*(\sigma_0, 0) < 0.5$,
2. $V^*(\sigma, \beta) = 1$ for β sufficiently large and $\sigma \neq L$,
3. $V^*(\sigma, \beta)$ is non-decreasing in β for all σ .

Proof. We only include a formal proof of the third claim. The case of $\sigma = W$ or L is trivial since $V^*(\sigma, \beta)$ is constant in β . Next, consider an arbitrary state (σ, β) where $\sigma \neq W, L$,

and let the corresponding optimal action be $\pi^*(\sigma, \beta)$. It suffices to show that $V^*(\sigma, \beta) \leq V^*(\sigma, \beta + 1)$. Consider a policy π that at state $(\sigma, \beta + 1)$ copies the optimal action at state (σ, β) , i.e., $\pi(\sigma, \beta + 1) = \pi^*(\sigma, \beta)$, which is feasible since there is at least one credit remaining at $(\sigma, \beta + 1)$. Then, $V^*(\sigma, \beta) = V^\pi(\sigma, \beta + 1) \leq V^*(\sigma, \beta + 1)$, where the equality follows by construction of π and the inequality follows by optimality. \square

The intuition behind these properties is straightforward. Property 1: By definition the weaker player has $p < 0.5$ and is thus more likely to lose in the absence of a handicap; Property 2: With enough free points, Player A can win from any state in the match (e.g., $\beta = 72$ is enough to win a best-of-five set match from σ_0); Property 3: More free points cannot lower the probability of winning the match.

It immediately follows from Lemma 1 that there exists a unique handicap β^* that is the smallest handicap the weaker player needs to achieve a match-win probability at least 0.5, assuming credits are used according to the optimal policy.

Theorem 2. *There exists a unique positive integer β^* such that $V^*(\sigma_0, \beta^*) \geq 0.5$ and $V^*(\sigma_0, \beta^* - 1) < 0.5$.*

3.3 Server-specific model

In Section 3.1, we assumed that the point-win probabilities do not depend on the server. This assumption was made to facilitate intuition building while developing the model, but is unrealistic in practice. We now relax this assumption and present a server-specific extension to the basic MDP model. We use p_1 and p_0 to denote the point-win probability of Player A when she is serving and receiving, respectively. Actions and rewards are not affected, so we do not discuss them in this section. For the remaining model components, we only discuss the differences from before and do not repeat any development that is unchanged by the addition of server-specific point-win probabilities.

For completeness, we need to redefine the weaker player (Player A) in the extended model. Earlier, Player A was defined to have $p < 0.5$. Instead, Player A in the extended

model is the player with $p_1 + p_0 < 1$ (or equivalently, the player with the lower point-win probability on her serve). Note that the model in Section 3.1 is a special case of the extended model presented below when $p_1 = p_0 = p$. Moreover, it is straightforward to show that the theoretical results from the server-independent model, appropriately adjusted, extend to the server-specific case.

States: The state of the server-specific MDP is now $s = (\sigma, \beta, \psi)$, where σ and β are the same as before and $\psi \in \{0, 1\}$ denotes whether Player A is receiving ($\psi = 0$) or serving ($\psi = 1$). We define ψ^+ and ψ^- in a similar manner as we defined σ^+ and σ^- . For example, when Player A has a game point on her serve ($\psi = 1$), then $\psi^+ = 0$ since winning the point will result in Player B serving on the next point. On the other hand, $\psi^- = 1$ since losing the point will mean that Player A keeps serving.

Transition probabilities: The definition of $p_{s,s'}^a := P(s'|s, a)$ remains the same, given the expanded state definition $s = (\sigma, \beta, \psi)$. As before, we use $p_{\sigma,\sigma'}^{a,\psi} := P(\sigma'|\sigma, a, \psi)$ when a focus on score states is more convenient. The possible transition probabilities are similar to before. Specifically, when $a = 1$, $p_{\sigma,\sigma^+}^{1,\psi} = 1$ for $\psi \in \{0, 1\}$. When $a = 0$, then $p_{\sigma,\sigma^+}^{0,\psi} = p_\psi$ and $p_{\sigma,\sigma^-}^{0,\psi} = 1 - p_\psi$.

Bellman equation: The Bellman equation (1) and the optimality equation (2) still hold. Specializing equation (2) to the server-specific model, we get the following optimality equation at non-absorbing states where $\sigma \neq W, L$ and $\beta \geq 1$

$$V^*(\sigma, \beta, \psi) = \max\{V^*(\sigma^+, \beta - 1, \psi^+), p_\psi V^*(\sigma^+, \beta, \psi^+) + (1 - p_\psi)V^*(\sigma^-, \beta, \psi^-)\}, \quad (4)$$

where the first term in the max corresponds to $a = 1$ and the second term corresponds to $a = 0$.

3.4 Computing handicaps

Given the point-win probabilities, which encode the skill difference between two players, we can determine the handicap β^* by solving the Bellman equations and determining the smallest β such that $V^*(s_0) \geq 0.5$, where s_0 denotes the state at the beginning of the match, including who is serving in the server-specific model. Recall that we do not need to examine values of β beyond 72 for a five-set match. It turns out that policy iteration solves this problem efficiently. For example, in the server-specific model, there are roughly 800,000 states for a best-of-five sets match with $\beta = 72$ and policy iteration solves this problem in under 10 minutes on a standard laptop computer (1.8 GHz Intel Core i5 processor with 4 GB of memory) using Matlab. In Sections 3.4.1 and 3.4.2, we present illustrative computations using the server-independent and server-specific models, respectively.

3.4.1 Server-independent model computations

Figure 1 illustrates the impact of β on the match-win probability assuming a server-independent point-win probability and a best-of-five sets match. It is evident that even a few credits can substantially increase the match-win probability, especially when compared to the initial-lead handicap system giving a β -point head start at the start of the match to Player A. For example, with a point-win probability of 0.49, three credits suffice to bring the match-win probability from 0.36 to 0.50. Under the initial lead system, 12 free points, equivalent to a 3-0 lead in the set or half a set, are needed to raise the match-win probability to 0.50. When $p = 0.45$, a handicap of 15 is needed under the MDP model to equalize the match-win probability, whereas the initial-lead system requires over 50 points, i.e., a two-set lead. In general, heuristic policies will be suboptimal and unlikely to achieve fairness given the optimal handicap β^* (associated with the optimal policy) derived from the MDP. The initial-lead policy is one example of a suboptimal policy for credit usage and Figure 1 illustrates how a poor policy regarding credit usage can be far from achieving fairness given the same number of credits used by the optimal policy. It would be

worthwhile to further study the impact of suboptimal credit usage, but this is outside the scope of the present paper, which focuses on the rigorous development of the handicap system. Finally, we note that our MDP model can take any policy, even suboptimal ones, and evaluate it using the Bellman equation (cf. Theorem 1). In fact, the act of evaluating a policy simply reduces the MDP to a Markov Chain with the corresponding transition probabilities induced by the given policy.

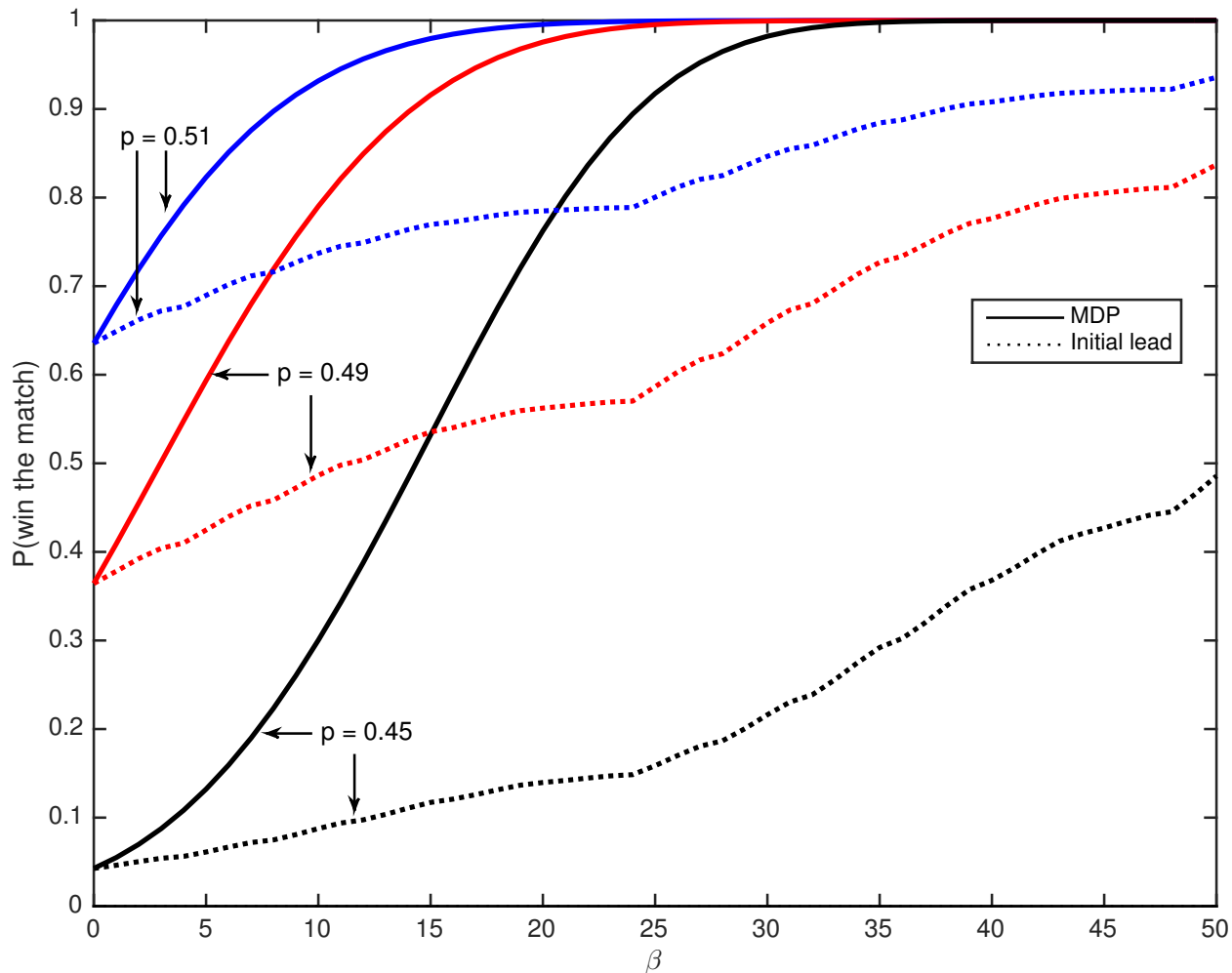


Figure 1: Match-win probability as a function of the number of credits for various server-independent point-win probabilities under our handicap system and the initial-lead system.

Figure 2 shows the optimal policy for a one-set, no tiebreaker match with $\beta = 2$ and $p = 0.48$. No tiebreaker means that a set score 6-5, 7-6, etc. is equivalent to 5-4. We see that the optimal policy is “conservative”, recommending credit use primarily when it would

secure the game or bring the game to deuce. A conservative policy avoids wasting credits. Of course, the policy recommends using a credit when Player A is two points from winning the match (i.e., when the set score is 5-0, 5-1, ..., 5-4 and the game score is 30-0, 30-15, or 30-30). A credit is also used when losing the point would end the match.

		Player B's score																			
		0-0		0-1		0-2		0-3		0-4		0-5									
		0	15	30	40	0	15	30	40	0	15	30	40	0	15	30	40	0	15	30	40
Player A's score	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b
	15	0	0	0	0	15	0	0	0	15	0	0	0	15	0	0	0	15	0	0	b
	30	0	0	0	0	30	0	0	0	30	0	0	0	30	0	0	b	30	0	0	b
	40	0	0	0	0	40	0	0	0	40	0	0	0	40	0	0	0	40	0	0	0
	1-0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b
	15	0	0	0	0	15	0	0	0	15	0	0	0	15	0	0	0	15	0	0	b
	30	0	0	0	0	30	0	0	0	30	0	0	0	30	0	0	b	30	0	0	b
	40	0	0	0	0	40	0	0	0	40	0	0	0	40	0	0	0	40	0	0	0
	2-0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b
	15	0	0	0	0	15	0	0	0	15	0	0	0	15	0	0	0	15	0	0	b
	30	0	0	0	0	30	0	0	0	30	0	0	0	30	0	0	b	30	0	0	b
	40	0	0	2	0	40	0	2	0	40	0	0	0	40	0	0	0	40	0	0	0
	3-0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b
	15	0	0	0	0	15	0	0	0	15	0	0	0	15	0	0	0	15	0	0	b
	30	0	0	0	0	30	0	0	0	30	0	0	0	30	0	0	b	30	0	0	b
	40	0	0	b	0	40	0	0	b	40	0	0	b	40	0	2	0	40	0	0	0
	4-0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b
	15	0	0	0	0	15	0	0	0	15	0	0	0	15	0	0	0	15	0	0	b
	30	0	0	0	0	30	0	0	0	30	0	0	0	30	0	0	2	30	0	0	b
	40	0	0	b	0	40	0	0	b	40	0	0	b	40	0	0	b	40	0	0	b
5-0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	0	0	0	0	15	0	0	0	15	0	0	0	15	0	0	0	15	0	0	0	
30	2	2	2	0	30	2	2	2	30	2	2	2	30	2	2	2	30	b	b	b	
40	b	b	b	2	40	b	b	b	40	b	b	b	40	b	b	b	40	b	b	b	

Figure 2: Optimal policy in a one-set, no tiebreaker match with $\beta \in \{1, 2\}$ and $p = 0.48$. The set score is the number in the upper left corner of each tableau and the game scores are in the row and column labels of each tableau. “1” indicates the states in which it is optimal to use a free point when $\beta = 1$ but not optimal when $\beta = 2$. “2” indicates the states in which it is optimal to use a free point when $\beta = 2$ but not optimal when $\beta = 1$. “b” indicates the states in which it is optimal to use a free point for both $\beta = 1$ and $\beta = 2$.

3.4.2 Server-specific model computations

Figure 3 is similar to Figure 1 but considers server-specific point-win probabilities $(p_1, p_0) = (p + c, p - c)$ for $p = 0.45$. As c increases from 0 to 0.15, β^* decreases from 15 to 9. Note that β^* depends on ψ_0 (whether or not Player A is the initial server). However, as shown in Table 1, the sensitivity of β^* to ψ_0 is small, with a slight first-server advantage, as expected.

Consequently, unless stated otherwise going forward, we fix $\psi_0 = 1$ for simplicity.

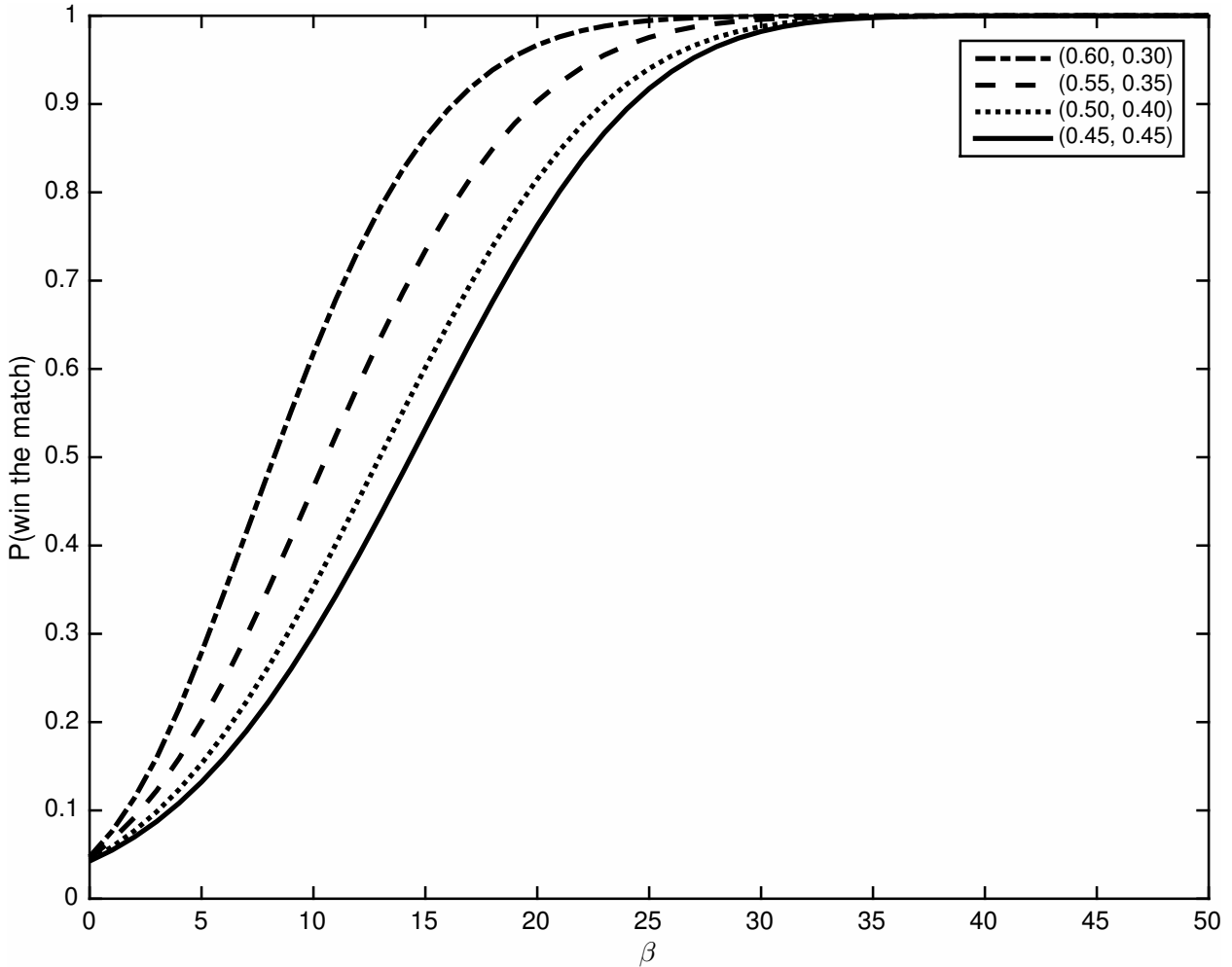


Figure 3: Match-win probability as a function of the number of credits for various server-specific point-win probabilities under our handicap system assuming Player A is the initial server.

Next, we examine the structure of the optimal policy given server-specific point-win probabilities $(p_1, p_0) = (p + c, p - c)$ with $p = 0.48$ and $c = 0.2$. Figures 4 and 5 display the optimal policy assuming Player A is receiving ($\psi = 0$) and serving ($\psi = 1$) in that state, respectively. Although not identical, the optimal policies depicted in Figures 4 and 5 retain the conservatism shown in Figure 2, the server-independent case. Under the realistic assumption that the point-win probability is higher for the server, credits are used more often on return (Figure 4) than on serve (Figure 5) and typically to generate (or save) a service break.

Table 1: Handicap values for various combinations of ψ_0 and (p_1, p_0) . The handicaps are fractional because they are linearly interpolated between β^* and $\beta^* - 1$ to achieve a match-win probability of exactly 0.5.

ψ_0	(p_1, p_0)			
	(0.45, 0.45)	(0.50, 0.40)	(0.55, 0.35)	(0.60, 0.30)
0	14.35	13.18	10.79	8.39
1	14.35	12.97	10.58	8.24

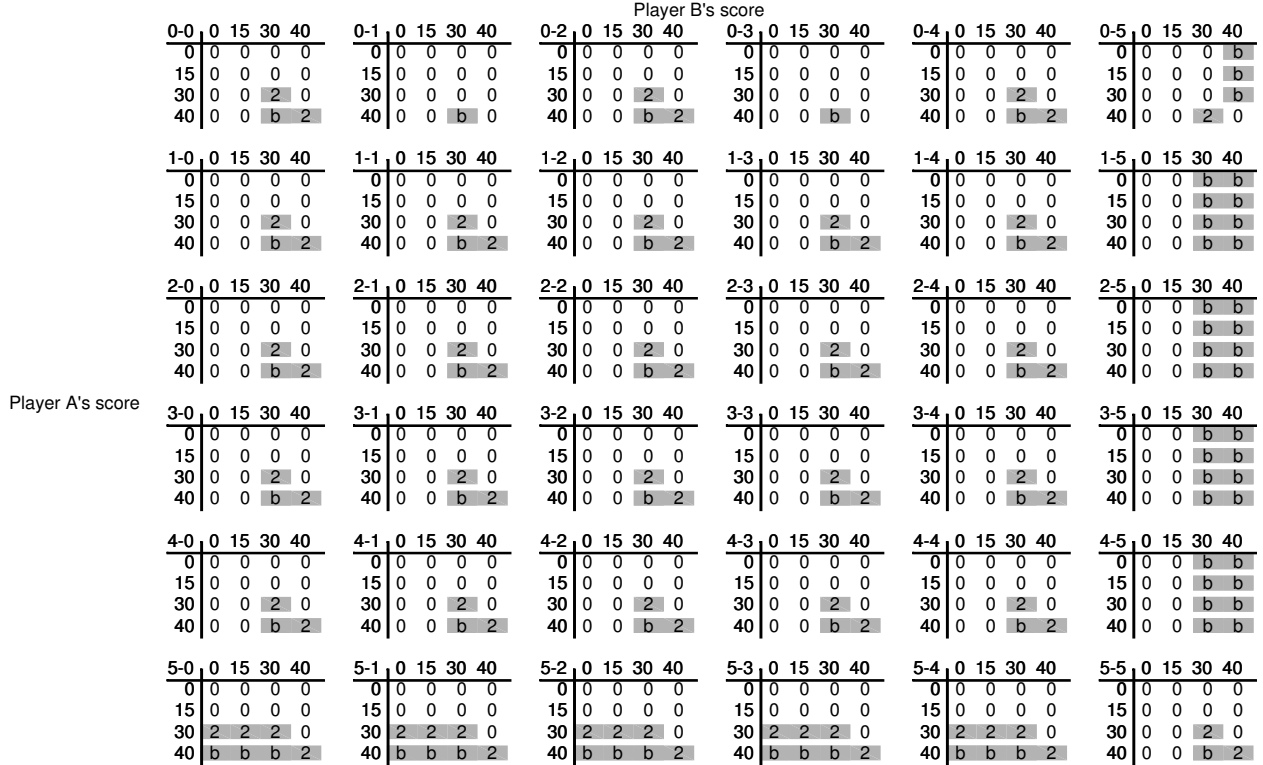


Figure 4: Optimal policy in a one-set, no tiebreaker match with $\beta \in \{1, 2\}$, $\psi = 0$, and $(p_1, p_0) = (0.68, 0.28)$. “1, 2, b” are defined as before.

The example optimal policies shown in Figures 2, 4, and 5 also illustrate the concept of “important points”, typically considered to be the points with the largest difference in the resulting match-win probability if they are won versus lost. The optimal policy identifies when it is worthwhile to immediately claim a point even if it reduces by one the budget of credits for future use. Thus, the optimal policy provides a novel state-dependent definition of important points. Furthermore, the action $a = 1$ can be seen as an extreme case of a player exerting extra effort to win a point. Thus, the optimal policy also provides insight into when it is most beneficial to expend that energy.

		Player B's score												
		0-0	0 15 30 40	0-1	0 15 30 40	0-2	0 15 30 40	0-3	0 15 30 40	0-4	0 15 30 40	0-5	0 15 30 40	
Player A's score	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0
	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0
	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0
	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0
	1-0	0 15 30 40	1-1	0 15 30 40	1-2	0 15 30 40	1-3	0 15 30 40	1-4	0 15 30 40	1-5	0 15 30 40		
	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0		
	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0		
	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0		
	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0		
	2-0	0 15 30 40	2-1	0 15 30 40	2-2	0 15 30 40	2-3	0 15 30 40	2-4	0 15 30 40	2-5	0 15 30 40		
	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0		
	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0		
	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0		
	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0		
	3-0	0 15 30 40	3-1	0 15 30 40	3-2	0 15 30 40	3-3	0 15 30 40	3-4	0 15 30 40	3-5	0 15 30 40		
	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0		
	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0		
	30	0 0 0 0	30	0 0 0 0	30	0 0 0 1	30	0 0 0 0	30	0 0 0 0	30	0 0 0 0		
	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0		
	4-0	0 15 30 40	4-1	0 15 30 40	4-2	0 15 30 40	4-3	0 15 30 40	4-4	0 15 30 40	4-5	0 15 30 40		
0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0			
15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0			
30	0 0 0 0	30	0 0 0 0	30	0 0 0 b	30	0 0 0 b	30	0 0 0 2	30	0 0 0 2			
40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0	40	0 0 0 0			
5-0	0 15 30 40	5-1	0 15 30 40	5-2	0 15 30 40	5-3	0 15 30 40	5-4	0 15 30 40	5-5	0 15 30 40			
0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0	0	0 0 0 0			
15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0	15	0 0 0 0			
30	2 2 2 0	30	2 2 2 0	30	2 2 2 0	30	b b b b	30	b b b b	30	0 0 0 2			
40	b b b 2	40	b b b 2	40	b b b 2	40	b b b b	40	b b b b	40	0 0 0 0			

Figure 5: Optimal policy in a one-set, no tiebreaker match with $\beta \in \{1, 2\}$, $\psi = 1$, and $(p_1, p_0) = (0.68, 0.28)$. “1, 2, b” are defined as before.

3.5 Model sensitivity to point-win probabilities

Since point-win probabilities are the key input to the MDP model, we explore the sensitivity of the outputted handicap values to these probabilities in this section. We wish to understand the extent to which uncertainty in the estimated probabilities may translate into variability in β .

3.5.1 Sensitivity of the server-independent model

Figure 6 illustrates how the handicap β and match-win probability vary as a function of p from 0 to 0.5. The handicaps are fractional because they are linearly interpolated between β^* and $\beta^* - 1$ to achieve a match-win probability of exactly 0.5. Over this entire range, clearly β and the match-win probability vary a lot. However, we argue essentially by contradiction that the uncertainty in p is likely to be captured in a small range. First, note that small changes in p , especially in the range $p \in [0.4, 0.6]$ translate to a huge difference in match-win

probability. Suppose match-win probability can only be estimated to within 40 percentage points, from 0.3 to 0.7, for example, which we argue is conservative. Even in this case, p resides in the small interval $[0.48, 0.52]$, corresponding to a handicap between 0 and 5 given to one of the two players. So, even though the model is sensitive to p over a large range, over realistic ranges of uncertainty in p , the impact on handicap variability should hopefully be small.

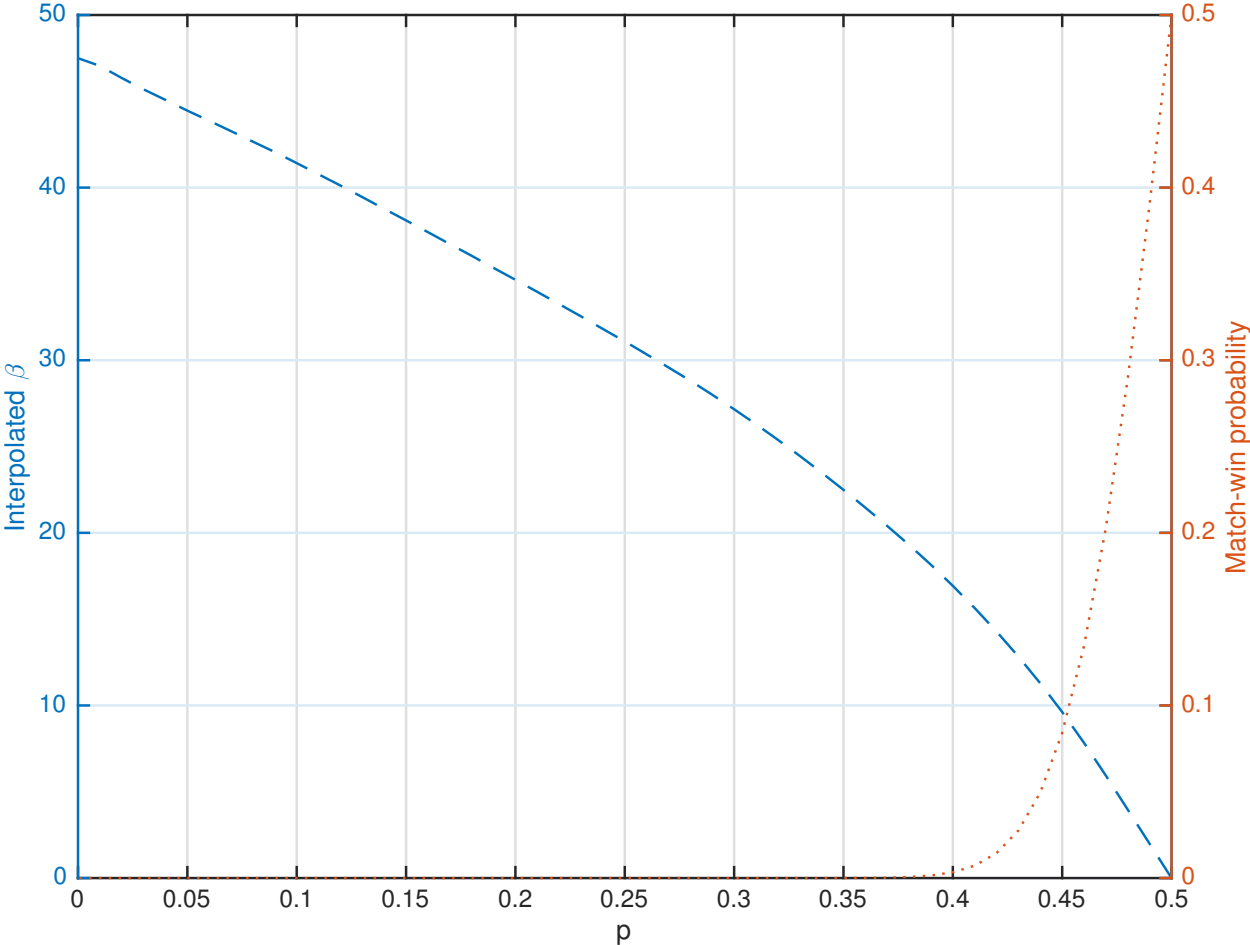


Figure 6: Handicap and match-win probability as a function of the point-win probability in a best-of-three sets match with tiebreakers.

3.5.2 Sensitivity of the server-specific model

In Table 2, we show match-win probability and handicap as a function of p_1 and p_0 . We evaluate a small range for both probabilities. Since the average point-win probability on

Table 2: Sensitivity of the match-win probability and handicap as a function of (p_1, p_0) for $p_1 \in \{0.63, 0.64, \dots, 0.67\}$ and $p_0 \in \{0.33, 0.34, \dots, 0.37\}$ in a best-of-three sets match with tiebreakers. Entries (x, y) in the table: x = match-win probability, y = interpolated handicap. Positive (negative) handicap indicates that the player receives (gives) the corresponding number of credits from (to) the other player.

p_1	p_0				
	0.33	0.34	0.35	0.36	0.37
0.63	(0.31, 1.87)	(0.35, 1.43)	(0.40, 0.97)	(0.45, 0.50)	(0.50, 0.00)
0.64	(0.36, 1.35)	(0.40, 0.91)	(0.45, 0.47)	(0.50, 0.00)	(0.55, -0.50)
0.65	(0.40, 0.86)	(0.45, 0.44)	(0.50, 0.00)	(0.55, -0.47)	(0.60, -0.97)
0.66	(0.45, 0.42)	(0.50, 0.00)	(0.55, -0.44)	(0.60, -0.91)	(0.65, -1.43)
0.67	(0.50, 0.00)	(0.55, -0.42)	(0.60, -0.86)	(0.64, -1.35)	(0.69, -1.87)

serve is around 0.65 for men (Klaassen and Magnus, 2001), we do sensitivity analysis around $(p_1, p_0) = (0.65, 0.35)$. We see that even for a relatively large range in match-win probability, over 35 percentage points, the variability in the handicap is less than four credits. We can also gain some insight from Table 2 into how sensitive the handicap calculation is to the value of $p_1 + p_0 - 1$, which is the difference in point-win probabilities on serve between the two players. A difference of 0.01 corresponds to about a 0.4 to 0.5 difference in handicap. Furthermore, even for the same value of $p_1 + p_0 - 1$, there can be slight differences in the induced handicap, since the specific p_1 and p_0 values may be different. This result is evident when examining the diagonals (bottom left to top right) in Table 2, which correspond to cases with the same value of $p_1 + p_0 - 1$.

4 Handicapping the “Big Four”

The “Big Four” refers to Novak Djokovic, Roger Federer, Andy Murray, and Rafael Nadal, players who have dominated men’s tennis over the last decade. From 2004-2015 inclusive, these four have won 42 out of the 48 Grand Slam tournaments. In this section, we illustrate the application of our method to calculate handicaps between these players on different surfaces and over different years. Data from <http://www.tennisabstract.com> is used to calculate server-specific point-win probabilities.

First, we determined the server-specific point-win probabilities overall and by court surface based on point-level data from all matches played between the two players to the end of 2015 (Figure 7). Specifically, we estimate the $\mathbb{P}(\text{player } i \text{ wins a point against player } j \mid \text{player } i \text{ is serving and court surface is } k)$ as the proportion of points won by player i against player j when player i was serving and the court surface was k . Other methods leveraging betting markets data or advanced statistical models can also be used to estimate point-win probabilities, but we take a simple approach here since the probability estimation is not the main focus of our paper.

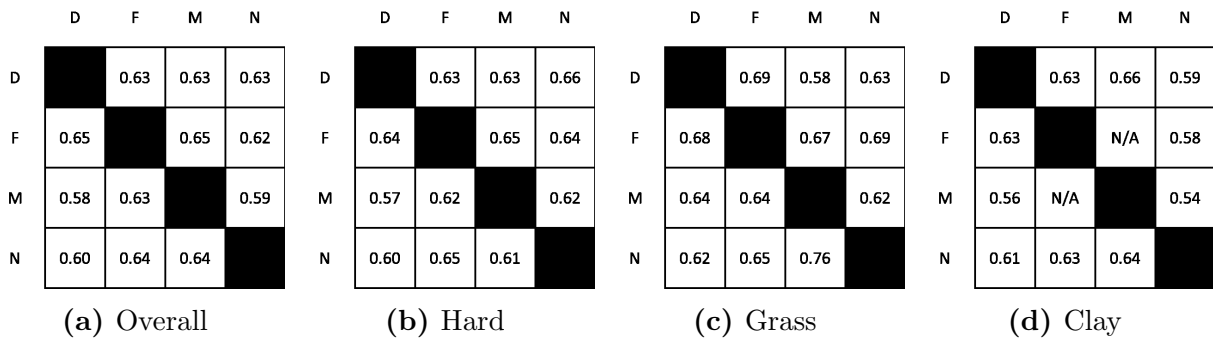


Figure 7: Server-specific point-win probabilities between the Big Four overall (over surfaces and years) and by surface (over years). Numbers in cell refer to probability the row player wins the point when serving to the column player. D = Djokovic, F = Federer, M = Murray, N = Nadal. Point-win probability was calculated as total points won divided by total points played.

Using the estimated point-win probabilities and assuming a best-of-five sets match, we determined the corresponding match-win probabilities and handicap using our MDP model (Figure 8). The handicaps are fractional because they are linearly interpolated between β^* and $\beta^* - 1$ to achieve a match-win probability of exactly 0.5. To read Figure 8, notice that below diagonal match-win probabilities greater than 0.5 indicate the row player is stronger, and thus, the above diagonal handicap is positive, meaning his opponent (the new row player) receives credits. As a helpful mnemonic, from the row player’s perspective, “red = receive” and “green = give” in terms of credits. Alternatively, via a “betting the spread” lens, one can view a positive value as a handicap “taken” by the weaker player and a negative value as a handicap “given” by the stronger player, from the row player’s perspective.

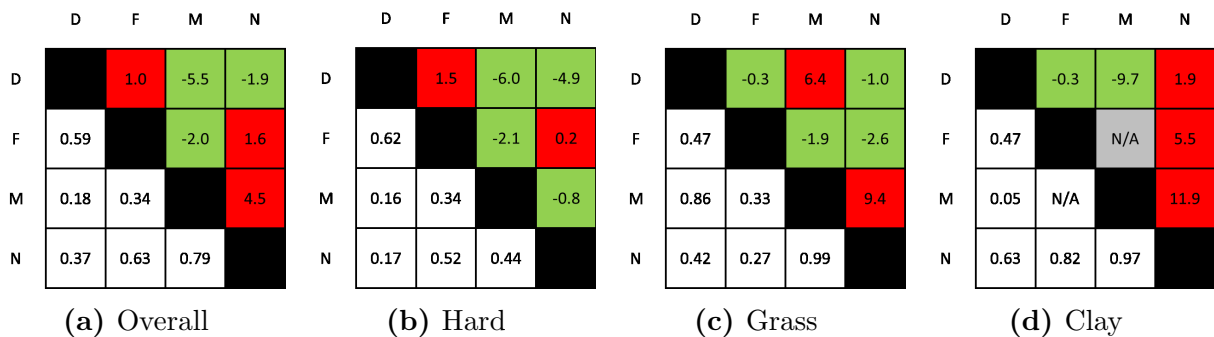


Figure 8: Match-win probabilities and handicaps between the Big Four overall (over surfaces and years) and by surface (over years). Numbers below diagonal refer to probability the row player wins the match against the column player. Numbers above diagonal refer to the interpolated handicap taken or given by the row player: positive/red indicates row player is weaker (taken) and negative/green indicates row player is stronger (given). D = Djokovic, F = Federer, M = Murray, N = Nadal.

We see that overall (Figure 8a), the closest rivalry involves Djokovic and Federer, with Djokovic receiving one credit from Federer to balance competition, historically. Also, there is a “triangle” relationship where Federer gives Djokovic credits, Djokovic gives Nadal credits, and Nadal gives Federer credits. Murray, on the other hand, has generally been dominated by the others, receiving credits from all three.

However, the handicap relations change when we examine each court surface separately. For example, while Nadal gives 1.6 credits to Federer overall, Federer’s historical dominance on grass shows up as 2.6 credits given to Nadal (Figure 8c). Nadal’s dominance on clay is evident as all his opponents receive credits from him on that surface (Figure 8d). The Djokovic-Murray rivalry is quite different based on surface, with Djokovic dominating on hard courts (Figure 8b) and clay (Figure 8d), and Murray dominating on grass (Figure 8c).

In Figure 9, we illustrate how the handicap between Djokovic and the other three players has changed from year to year on hard courts, the surface with the most match data between these players. We see that Djokovic has generally enjoyed an advantage over Murray and Nadal on hard courts, with the advantage being firmly in his favor since 2011. On the other hand, he has experienced much more balanced competition against Federer over the last decade. It may seem unintuitive to see Djokovic receiving a handicap against Federer in

2014 and 2015 when Djokovic has enjoyed considerable success at the Grand Slams (winning four of the eight, including both hard court Slams in 2015) often beating Federer in the finals. However, when examining all hard court matches played between the two, Federer has had an ever-so-slight advantage on (server-specific) point-win probabilities. Thus, the incongruence between expectation and results in Figure 9 is largely explained by the fact that Djokovic has been more dominant at the larger events and during critical junctures in the match. For example, during the 2015 U.S. Open Final, Djokovic converted six of 13 break chances, while Federer only converted four of 23.

Lastly, we comment on the intuitive and concrete interpretability of our handicap metric for a tennis follower. Our handicap metric quantifies the difference in performance between the two players as a number of points won/lost – precisely, the minimum number of points won/lost at the “right” times in the match – which is done by no other metric. In other words, a handicap provides a unique quantification of how close the players are: winning a few well-timed points is often the difference between winning and losing the match.

5 Conclusion

In this paper, we developed a novel credit-based, dynamic handicap system for tennis with a Markov Decision Process at its core. Given a skill difference encoded by a point-win probability that may be server-specific, the MDP model computes the smallest handicap required by the weaker player to achieve at least a 0.5 probability of winning the match, along with the optimal policy governing credit usage. We demonstrated that a few credits can have a large impact on the match-win probability, unlike the naive handicap system that simply gives the weaker player an initial lead in the match. We provided examples of optimal policies, highlighting their conservatism towards credit usage.

To illustrate our method, we applied it to match data between the men’s Big Four: Novak Djokovic, Roger Federer, Andy Murray, and Rafael Nadal. We quantified the lifetime handicaps between the players over different surfaces and overall. As an example of handicap

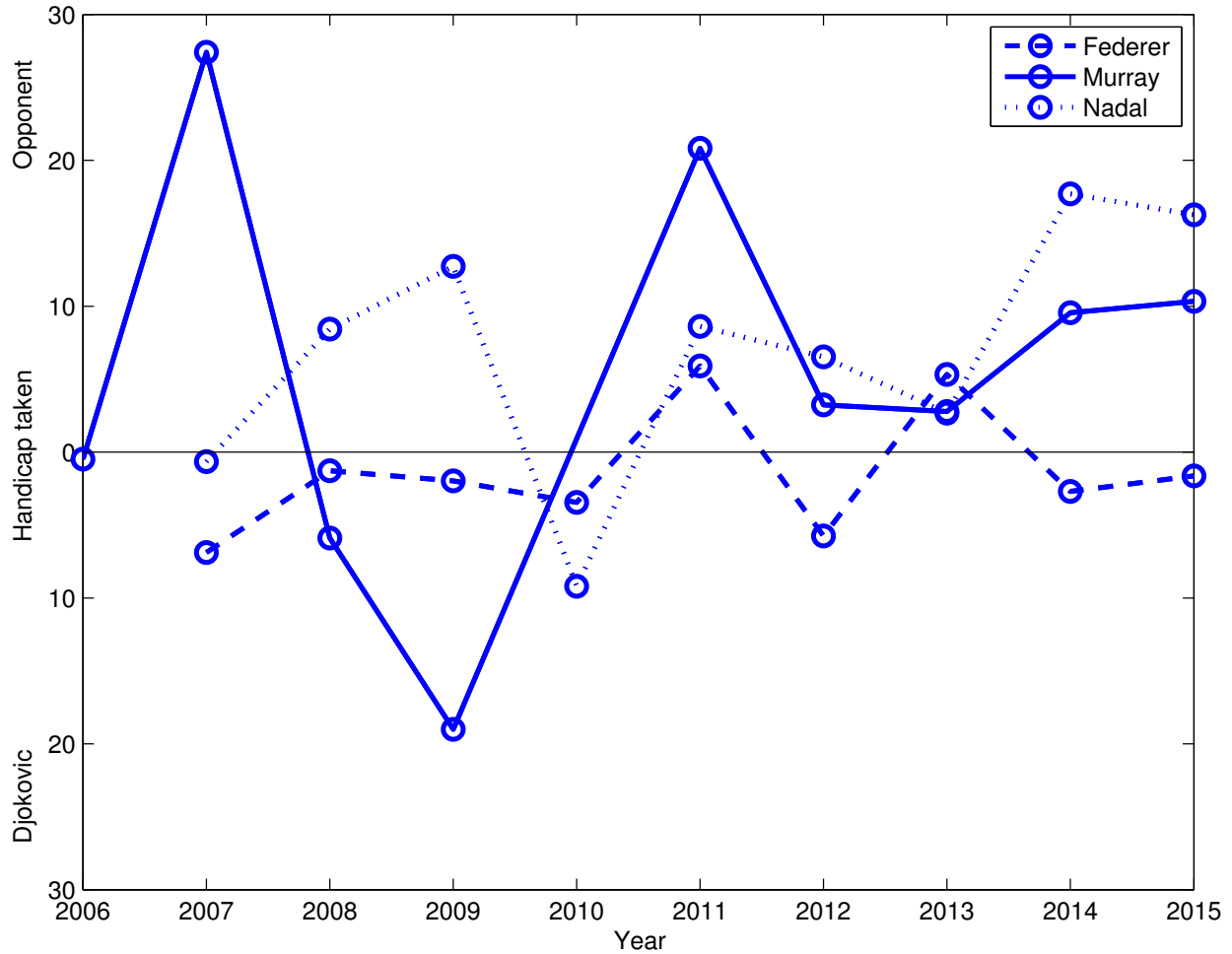


Figure 9: Evolution of handicap between Djokovic and rest of Big Four on hard court.

evolution, we also illustrated how Novak Djokovic’s handicap against the other three on hard courts changed over the course of his career, in congruence with his overall career trajectory.

Our system and analysis can be extended to other hierarchical sports, since it is based on a Markov Decision Process, which is a general framework for decision making under uncertainty. Furthermore, the optimal policy associated with a particular handicap provides an optimization-driven perspective on effort allocation and point importance. Future research should consider mapping our handicap method to established player rating systems used by social tennis players, to enable uptake without the need for specialized math knowledge. The fractional part of an interpolated handicap can also be explored further to equate its value with advantages like taking away an opponent’s second serve,

based on similar analysis of its effect on the match-win probability.

References

- Barnett, T., A. Brown, and S. R. Clarke (2004). Optimal use of tennis resources. In *Seventh Australasian Conference on Mathematics and Computers in Sport. Massey University: Palmerston Nth, New Zealand*, pp. 57–65.
- Bertsekas, D. and J. Tsitsiklis (2008). Introduction to Probability (Athena Scientific). *Nashua, NH*.
- BGA (2015). British Go Association, How to Play. <http://www.britgo.org/intro/intro2.html>. [Accessed: 22-November-2015].
- BHA (2015). British Horseracing Authority, Handicapping. <http://www.britishhorseracing.com/bha/what-we-do/handicapping/>. [Accessed: 22-November-2015].
- Branson, R. (2015). Shake up tennis with ‘Richard’s Rules’. <http://www.virgin.com/richard-branson/shake-up-tennis-with-richards-rules>. [Accessed: 22-November-2015].
- Eisenhauer, J. G. (2003). Regression through the origin. *Teaching Statistics* 25(3), 76–80.
- Fischer, G. (1980). Exercise in probability and statistics, or the probability of winning at tennis. *American Journal of Physics* 48(1), 14–19.
- Hooper, D. and K. Whyld (1992). *The Oxford Companion to Chess*. Oxford University Press.
- ITF (2015). International Tennis Federation, Rules of Tennis. <http://www.itftennis.com/officiating/rulebooks/rules-of-tennis.aspx>. [Accessed: 22-November-2015].

- Kemeny, J. G. and J. L. Snell (1960). *Finite Markov Chains*, Volume 356. van Nostrand Princeton, NJ.
- Klaassen, F. J. and J. R. Magnus (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association* 96(454), 500–509.
- Knottenbelt, W. J., D. Spanias, and A. M. Madurska (2012). A common-opponent stochastic model for predicting the outcome of professional tennis matches. *Computers & Mathematics with Applications* 64(12), 3820–3827.
- Morris, C. (1977). The most important points in tennis. *Optimal Strategies in Sports* 5, 131–140.
- O’Donoghue, P. G. (2001). The most important points in grand slam singles tennis. *Research Quarterly for Exercise and Sport* 72(2), 125–131.
- Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.
- Tencap (2015). Tencap Sports, Tencap’s Rating System. <http://www.tencapsports.com/rating.aspx>. [Accessed: 22-November-2015].
- USGA (2015). United States Golf Association, Handicap System Manual. <http://ctan.tug.org/tex-archive/info/lshort/english/lshort.pdf>. [Accessed: 22-November-2015].
- USTA (2015). United States Tennis Association, National Tennis Rating Program. <https://www.usta.com/Adult-Tennis/USTA-League/ntrp/>. [Accessed: 23-November-2015].
- UTR (2015). Universal Tennis, Universal Tennis Rating System. <http://universaltennis.com>. [Accessed: 23-November-2015].