

Q学習を用いたゴルフの攻略ルート支援に関する研究

菅原 翔悟^{†1,a)} 川村 秀憲^{†1,b)} 鈴木 恵二^{†1,c)}

概要：ゴルフに関する研究は様々な分野において行われてきたが、多くは飛距離や正確性などのスキルに注目しており、クラブやボールの落としどころを選択する方針である戦略についての研究はほとんど行われていない。本研究はゴルフのスコアアップにつながる戦略を獲得するためのフレームワークを開発することを目的とし、ゴルフをボールの位置が確率的に決定されるマルコフ決定過程の性質を満たす確率モデルとして定義し、それをもとに期待スコアを最小化する最適化問題を定義して、Q学習により最適化を行った。最後に、獲得された戦略を用いたシミュレーションを行うことにより、スキルとコースに応じた戦略が獲得されることを示し、フレームワークの応用例として、ゴルフ戦略における攻略ルートの支援ツールであるバーチャルゴルフコーチの機能と支援例について記述した。

SHOGO SUGAWARA^{†1,a)} HIDENORI KAWAMURA^{†1,b)} KEIJI SUZUKI^{†1,c)}

1. はじめに

ゴルフは、クラブという道具を用いてボールを打ち、カップと呼ばれる地面にあいた穴にボールを入れることを目的とするスポーツである。ゴルフではボールを打った回数（打数）がスコアとなり、できるだけ少ない打数でカップインを目指す。少ない打数でのカップインのため、ゴルフファーの多くはより遠く、より正確にボールを打つことを考え、ゴルフ練習場に通ったり、コーチのレッスンを受けたり、スイングについての教則本を買ったりする。

そのようなゴルフファーのスキルについての関心と比例して、ゴルフのスキルについての研究が数多く行われてきている。Friedらは、プロゴルフの大会の統計データ（スタッツと呼ばれる）を収集し、ゴルフファーの運動能力とプレッシャー下でのパフォーマンスを評価するための指標を示し [1]。鈴木らは、ゴルフクラブ・ボールの性能評価を目的としたゴルフスイングロボを開発し、スイングの三次元動的モデルによるスキルの分析を行った [2]。また、Broadieらはショット後のボール位置が確率的に決定されるシミュレーションモデルを開発し、ティーショットにおいて飛距離と方向性のどちらがスコアに影響を与えるか定

量化して、方向性の方が重要であると結論付けた [3]。

スキルと対照的に、ゴルフの戦略についての研究はほとんど行われていないのが現状である。スポーツとしてゴルフと似た特徴を持つビリヤードでは、Archbaldらがシミュレーションによりスキルと戦略の相互作用を実験的に調査し、狙った通りにボールを打てないような不完全なスキルの場合に、知能的な戦略が勝率に寄与していることを最も識別可能であると論じており、ゴルフでも同じことが言える可能性があると言及している [4]。ゴルフの戦略については、「球聖」 Bobby Jones が「ゴルフは耳と耳の間にあるもの（頭を使うスポーツという意味）」と残しているように、スコアに戦略が寄与していることは一般的に認識されている。ゴルフ場によってはコースの特徴、狙いどころ、危険な場所についてまとめたコースガイドがあったり、ゴルフファーにアドバイスをするキャディを雇えたり、戦略的なアドバイスを受けることができるが、全てのゴルフファーがお金のかかるキャディを雇えるわけでもなく、ゴルフファーのスキルによっては役に立たない戦略的なアドバイスもある。

本研究は、スコアアップにつながる戦略を獲得するためのフレームワークを開発することを目的とし、ゴルフのシミュレーションモデルを開発して、マルコフ決定過程において用いられる最適化の手法である Q 学習を適用することで、期待スコアが最小となるような戦略を獲得できるフレームワークを新たに提案した。実験として、スキルとコースに応じた戦略が獲得されたかシミュレーションによ

^{†1} 現在、北海道大学大学院情報科学研究科
Presently with Graduate School of Information Science and
Technology, Hokkaido University

a) sugawara@complex.ist.hokudai.ac.jp

b) kawamura@complex.ist.hokudai.ac.jp

c) suzuki@complex.ist.hokudai.ac.jp

り示し、フレームワークの応用例として、ゴルフの戦略における「攻略ルート」を決めるうえで有用な情報を提示するためのツールである「バーチャルゴルフコーチ」による支援例を提示する。

2章ではゴルフのシミュレーションモデルについて述べ、ゴルフを構成する要素を列挙し、本研究でモデル化した範囲について説明する。最後に、戦略・スキル・コースを条件とするスコアの条件付き期待値を定式化し、それを最小化する最適化問題を定義する。3章ではQ学習をどのようにゴルフに適用するかを説明し、4章で獲得した戦略について実験を行った。5章でバーチャルゴルフコーチの機能と攻略ルート支援例を説明し、6章で本稿のまとめを行う。

2. モデリング

ゴルフコース	ゴルフショット
<ul style="list-style-type: none"> ○コースレイアウト ・ティーインググラウンド・林 ・フェアウェイ ・ラフ ・グリーン 	<ul style="list-style-type: none"> ○パッティング ・パッティングスキル
<ul style="list-style-type: none"> ○起伏 ○風 ○気温・気圧・湿度 	<ul style="list-style-type: none"> ○ショット ・ストレートボール ・スライス、フェード ・フック、ドロー ・インテンショナルハイ・ローボール ・ミスショット ・トップ、ダフリ・チヨロ、テンブラ ・ラン ・ショットスキル

青:実装済み 黄:部分的な実装 赤:未実装

図1 ゴルフの要素と実装状況

この章では、ゴルフゲームの要素であるゴルフコースとゴルフショットについてどのようなモデル化を行ったかを述べる。図1はゴルフを構成する要素と本研究での実装状況についてまとめた図である。以降でそれぞれの項目について説明する。また、章の最後で戦略・スキル・コースを条件とするスコアの条件付き期待値を定式化し、それを最小化する最適化問題を定義する。

ゴルフのシミュレーションモデルは過去にも研究が行われており [3][5]、それを参考にモデリングを行っているが、本研究では期待スコアを定義するために参考文献より詳細に定式化を行った。

2.1 ゴルフコース

本稿では、ゴルフの競技を行うために定められた区域をコース(ホール)、全てのホールを含めたゴルフ場の総称をゴルフコース、グリーン上にある穴のことをカップと記述する。コースは以下の区域に分けられる。

- (1) ティーインググラウンド (Teeing ground) 1打目を打つスタート地点。
- (2) フェアウェイ (Fairway) 芝が短く刈り込まれた部分。
- (3) ラフ (Rough) 芝が長くなっている部分。
- (4) バンカー (Bunker) 砂が敷き詰められたくぼ地。
- (5) ウォーターハザード・ラテラルウォーターハザード

(Water Hazard, W.H.) 池など水に関わる危険区域。ボールを打ち込んだ場合、打数に1を加えることで(1打罰)ボールを移動する救済措置を受けることができる。

- (6) グリーン (Green) パターによるショット(パッティング)のための区域。グリーン上ではパターを使ってボールを転がし、カップにボールを入れる(カップイン)。
- (7) OB (Out of Bounds) プレーが出来る区域の外。OBとなった場合は1打罰を受けるとともに、ボールを打つ前の地点から打ち直しとなる。
- (8) 林 空間的な障害物。木より高い軌道でボールを打てば衝突することなく越えることができる。

上記の区域をコース区分(図2)と呼び、コース区分の集合 L を以下のように定義する。

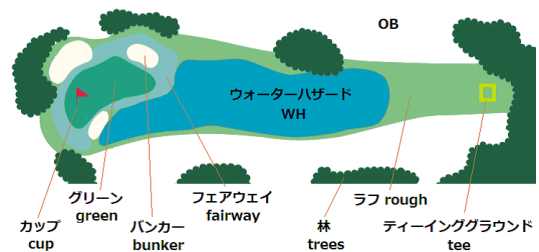


図2 コース区分

$$L = \{tee, fairway, rough, bunker, WH, trees, green, OB, cup\} \quad (1)$$

ゴルフの競技は通常18ホールをプレイし、18ホールを一巡することをラウンドと呼ぶ。ホール h の集合を H とし、 h は状態 s の集合である S とある状態 s でのコース区分を導く関数 l を要素として持つ。

$$H = \{h_1, h_2, h_3, \dots, h_{18}\} \quad (2)$$

$$h = \{S, l\} \quad (3)$$

$$s = (x, y) \quad (s \in S, S \in \mathbb{R}^2) \quad (4)$$

$$l: S \rightarrow L \quad (5)$$

2.2 ゴルフショット

ショットとはゴルフクラブを用いてボールを打つことである。競技中使用できるクラブは14本で、通常はウッド(W)3本、アイアン(I)10本、パター1本の組み合わせが多い。クラブには番号がついており、番号が大きくなるほど飛距離が短くなる。最も飛距離の長い1番ウッドは特別にドライバー(DR)と呼ばれる。本稿ではウッドとアイアンを使うことをショット、パターを使うことをパッティングとする。飛距離の単位はヤードを用い、ショットされたボールの空中における飛距離をキャリー、落下後ボールが停止するまでの距離をランと呼ぶ。

2.2.1 ショット

ボールを打ったとき狙った位置と実際のボール落下位置の間にずれが生じる．ずれの大きさは確率的であるとし，ボールの落下位置がある確率分布で表わされると仮定する．つまりスキルレベルの高低が狙った位置からのずれの大きさを表す．

ショットの際，使用するクラブ，ボールを打ち出す水平方向と垂直角度，ボールを打ち出す力加減，ボールを曲げるかどうかの5つを決める．本稿では垂直角度（インテンショナルハイ・ローボール）とボールの曲がり（フック，スライスなど）は定義せず，ゴルファーの行動 a をクラブ c ，ボールを打ち出す水平方向 θ ，力加減 $pw \in \mathbb{R}$ から定義する．また，行動の集合は A である．

$$a = \{c, \theta, pw\} \quad (c \in C, 0 \leq \theta < 2\pi, 0 \leq pw \leq 1) \quad (6)$$

$$C = \{DR, 3W, 5W, 2I, 3I, 4I, 5I, 6I, 7I, 8I, 9I, PW, SW\} \quad (7)$$

本稿では単純にキャリーと方向がそれぞれ正規分布に従うものとし，そのパラメータをスキルとして定義する．各クラブで飛距離が異なり，一般的に番手が1つ上がると飛距離が10ヤード伸びる．平均キャリー d の集合を D ，飛距離の標準偏差 σ_d の集合を Sig_d ，分布の中心の左右のずれを表す方向の平均値 μ_θ の集合を M ，方向の標準偏差 σ_θ の集合を Sig_θ とする．

$$D = \{d_c | c \in C\} \quad (8)$$

$$Sig_d = \{\sigma_{d_c} | c \in C\} \quad (9)$$

$$Sig_\theta = \{\sigma_{\theta_c} | c \in C\} \quad (10)$$

$$M = \{\mu_{\theta_c} | c \in C\} \quad (11)$$

b_d と b_θ はコース区分から決まる値でショットの正確性に影響し，フェアウェイとティーインググラウンドが基準で，ラフやバンカーだと値が大きくなる．

$$b_d : L \rightarrow \mathbb{R} \quad (12)$$

$$b_\theta : L \rightarrow \mathbb{R} \quad (13)$$

これらにより，ボールを打った後の落下位置 $s' = (x', y')$ はボールを打つ前の位置 $s = (x, y)$ から以下のように表わされる．

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d' \cos \theta' \\ d' \sin \theta' \end{pmatrix} \quad (14)$$

$$d' \sim N\left(d_c \cdot pw, (b_d(l) \cdot \sigma_{d_c})^2\right) \quad (15)$$

$$\theta' \sim N\left(\theta + \mu_\theta, (b_\theta(l) \cdot \sigma_{\theta_c})^2\right) \quad (16)$$

また，ミスショットする確率 p_{miss} を定義し，ミスショッ

ト（ダフリ，チョロ）の場合は使用クラブ c よりキャリーの短いクラブ c' を使うと考え，飛距離が短くなることを表す．つまりショットの分布は2つの正規分布による混合分布として表現される．

ランについては，地面の起伏など物理シミュレーションのように扱えば精度良く計算できるが，簡易かつ高速に処理するため「carry fraction」を計算する．ランとキャリーを含めた全体の飛距離 r ，キャリー d ，carry fraction を cf とすると， $cf = d/r$ ，ランは $d(1 - cf)/cf$ となる．carry fraction はクラブデザインについての文献 [6] から推定している．図3は，ハンディキャップの異なる4人の

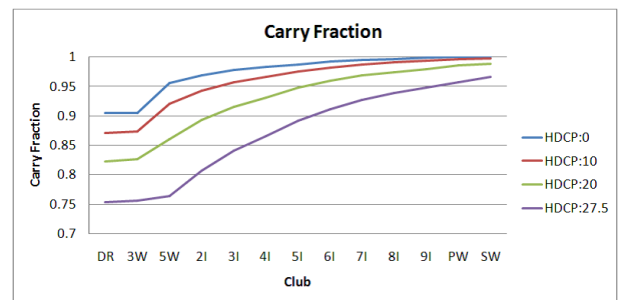


図3 Carry Fraction

ゴルファーの carry fraction である．ハンディキャップとは，技量の異なるプレイヤーが公平な基準で競えるように定められたゴルファーの技量を示す尺度で，値が小さいゴルファーほど技量が高いと扱われる．この論文では carry fraction を，クラブ番号を変数とする対数関数として定義し，そのパラメータは各ゴルファーのドライバーの飛距離によって決める．carry fraction はフェアウェイを基準としており，ラフを通るときランはフェアウェイの半分の距離となり，バンカーではボールは転がらない．

最後に， p_{miss} の集合を P_{miss} とし，スキル $skill$ を以下のように定義する．

$$skill = \{D, Sig_d, Sig_\theta, M_\theta, P_{miss}\} \quad (17)$$

$$P_{miss} = \{p_{miss_c} | c \in C\} \quad (18)$$

また，戦略 st を状態空間 S から行動空間 A への写像とし，集合を St とする．

$$st : S \rightarrow A \quad (19)$$

2.2.2 パッティング

パッティングでは左右と飛距離のずれが少ない正確性の高さ，グリーン上で起伏や芝が生えている向きを分析するかも必要となる．パッティングにおいて重要な起伏についてモデル化していないため，パッティングについてはボールの位置を確率的に決めるのではなく，Bansal らの研究 [7] を参考にカップインまでにかかるパット数を確率的に決める．図4，5はプロゴルファーについてのグラフである．

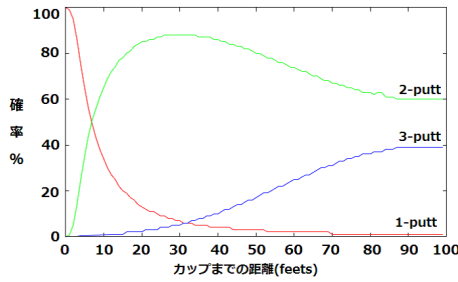


図 4 カップまでの距離に対するパット数別のカップイン確率

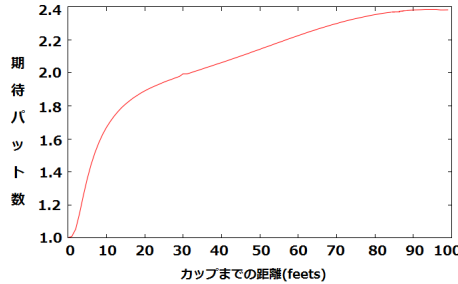


図 5 カップまでの距離に対する期待パット数

2.3 期待スコアと最適化問題

ここでは、期待スコアを求める計算式を定義し、それを用いて最適化の目的関数を定義する。スコアはグリーンオンするまでの打数、パット数、ペナルティ数（罰打数）の合計である。また、ショットとパッティングの期待スコアの2つに分けて定式化する。1 ホールの期待スコアは戦略 st とスキル $skill$ とホール h から決まる条件付き期待値とし、1 ホールのスコアを N 、1 ホールのショットのスコアを N_1 、1 ホールのパッティングのスコアを N_2 とする。

$$E[N|st, skill, h] = E[N_1|st, skill, h] + E[N_2|st, skill, h] \quad (20)$$

目的は、あるホールでスキルが不変の場合に、戦略を変化させて期待スコアを最小化することである。よって目的関数は以下の式となる。

$$\min_{st \in St} E[N|st, skill, h] \quad (21)$$

まず、グリーンオンしたときにスコアが n である確率 $P_1(n|st, skill, h)$ を定義し、ショットの期待スコア $E[N_1|st, skill, h]$ を以下の式で表す。

$$E[N_1|st, skill, h] = \sum_{n=0}^{\infty} P_1(n|st, skill, h) \cdot n \quad (22)$$

スコア $n \in \mathbb{N}$ は打数と罰打数の合計である。また、ゴルファーがペナルティを受けるコース区分にボールを打ち込んだことを示すため、ペナルティ u を定義する。 u は 0 か 1 をとり、1 のときペナルティを受けるものとする。ショット前のボールの位置が s_n 、ペナルティが u_n で、ショット後のボールの位置が s_{n+1} 、ペナルティが u_{n+1} となる確率を $p(s_{n+1}, u_{n+1}|s_n, u_n, st, skill, h)$ とし、ホール h の中でコース区分がグリーン領域を D 、ティーインググラウ

ンドで 1 打目を打つボールの位置を st_{ee} とする。よって $P_1(n|st, skill, h)$ は以下のように定義される。

$$P_1(n|st, skill, h) = \sum_{u_n=0}^1 \left(\int_D p(s_n, u_n|st, skill, h) ds_n \right) \quad (23)$$

$$p(s_n, u_n|st, skill, h) = \sum_{u_{n-1}=0}^1 \left(\int p(s_n, u_n|s_{n-1}, u_{n-1}, st, skill, h) \cdot p(s_{n-1}, u_{n-1}|st, skill, h) ds_{n-1} \right) \quad (24)$$

$$p(s_0, u_0|st, skill, h) = \delta(s_0 - st_{ee}) \cdot \delta_{u_0 0} \quad (25)$$

$\delta(x)$ はディラックのデルタ関数と呼ばれる超関数、 δ_{ij} はクロネッカーのデルタと呼ばれる記号で、以下の性質を持つ。

$$\delta(x) = \begin{cases} \infty & (x = 0) \\ 0 & (x \neq 0) \end{cases} \quad (26)$$

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad (x \in \mathbb{R}) \quad (27)$$

$$\delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \quad (28)$$

$$\sum_{n=0}^N \delta_{n0} = 1 \quad (n \in \mathbb{N}) \quad (29)$$

$p(s_{n+1}, u_{n+1}|s_n, u_n, st, skill, h)$ の定義は後述する。

最後に、グリーン上で m パットしてカップインする確率 $P_2(m|st, skill, h)$ を定義することにより、期待スコア $E[N_1|st, skill, h]$ でボールがグリーンオンした後のパッティングの期待スコア $E[N_2|st, skill, h]$ を以下の式で表す。

$$E[N_2|st, skill, h] = \sum_{m=0}^{\infty} P_2(m|st, skill, h) \cdot m \quad (30)$$

ボールが位置 s にある確率を $p(s|st, skill, h)$ とする。

$$p(s|st, skill, h) = \sum_{m=0}^{\infty} \left(\sum_{u_m=0}^1 p(s_m, u_m|st, skill, h) \right) \quad (31)$$

また、 $g(s, m)$ は、グリーン上の s の位置から m 回パッティングしてカップインする確率を表す関数である。関数 g は 2.2.2 章の図 5 から定義している。 $p(s|st, skill, h)$ と $g(s, m)$ から、 $P_2(m|st, skill, h)$ は以下のように定義される。

$$P_2(m|st, skill, h) = \int_D p(s|st, skill, h) \cdot g(s, m) ds \quad (32)$$

2.3.1 $p(s_{n+1}, u_{n+1} | s_n, u_n, st, skill, h)$ の定式化

$p(s_{n+1}, u_{n+1} | s_n, u_n, st, skill, h)$ を関数 f で表すと、以下のようになる。

$$p(s_{n+1}, u_{n+1} | s_n, u_n, st, skill, h) = f(s_{n+1}, u_{n+1}; s_n, u_n, st, skill, h) \quad (33)$$

関数 f はショット後のボール位置の分布であり、以下の複数の分布から成る。

- ショットによるボール落下位置の分布 $f_1(s_{n+1}^{(1)}; s_n, a, skill, h)$: ボール位置 s_n からスキル $skill$ のゴルファーが行動 a をとった後のボール落下位置 $s_{n+1}^{(1)}$ の分布。木の衝突とランは考えない。
- 林に関わるボール位置の分布 $f_2(s_{n+1}^{(2)}; s_{n+1}^{(1)}, s_n, h)$: ボールが林を通過する場合、一定距離ごとに木と衝突するか判定する反復試行を考える。木の高さについては簡単のため、軌道を3つに分けた真ん中の部分ではボールが木より高い位置にあるみなして衝突確率を0とする。ショットによるボール落下位置 $s_{n+1}^{(1)}$ 、関数 f_2 によって新たに決まる位置を $s_{n+1}^{(2)}$ とする。
- ランについての分布 $f_3(s_{n+1}^{(3)}; s_{n+1}^{(2)}, s_n, skill, h)$: ショットによるボール落下位置を $s_{n+1}^{(2)}$ 、関数 f_3 によって新たに決まるランを含めたボール位置を $s_{n+1}^{(3)}$ とする。
- 次状態のペナルティの分布 $f_4(u_{n+1}; s_{n+1}^{(3)}, h)$: 位置 $s_{n+1}^{(3)}$ のコース区分が OB か WH なら u_{n+1} は1、それ以外なら0。
- 現在状態のペナルティによるボール位置の分布 $f_5(s_{n+1}^{(4)}; s_{n+1}^{(3)}, s_n, u_n, h)$: ペナルティ u_n が1なら1打罰を受ける。つまりスコアは1増えるがボール位置が変わらない。
- 池に関わるボール位置の分布 $f_6(s_{n+1}^{(5)}; s_{n+1}^{(4)}, s_n, h)$: ショット後のボール位置を $s_{n+1}^{(4)}$ とし、コース区分が WH の場合、最後にボールが WH の区域を横切った位置が新たなボール位置 $s_{n+1}^{(5)}$ となる。
- OB のボール位置の分布 $f_7(s_{n+1}; s_{n+1}^{(5)}, s_n, h)$: ショット後のボール位置を $s_{n+1}^{(5)}$ とし、コース区分が OB の場合、新たなボール位置 s_{n+1} はショット前のボール位置 s_n と等しくなる。

つまり関数 f は以下の式で求められる。

$$f(s_{n+1}, u_{n+1}; s_n, u_n, st, skill, h) = \int \int \int \int \int f_1(s_{n+1}^{(1)}; s_n, st(s_n), skill, h) \cdot f_2(s_{n+1}^{(2)}; s_{n+1}^{(1)}, s_n, h) \cdot f_3(s_{n+1}^{(3)}; s_{n+1}^{(2)}, s_n, st(s_n), skill, h) \cdot f_4(u_{n+1}; s_{n+1}^{(3)}, h) \cdot f_5(s_{n+1}^{(4)}; s_{n+1}^{(3)}, s_n, u_n, h) \cdot f_6(s_{n+1}^{(5)}; s_{n+1}^{(4)}, s_n, h) \cdot f_7(s_{n+1}; s_{n+1}^{(5)}, s_n, h) ds_{n+1}^{(1)} ds_{n+1}^{(2)} ds_{n+1}^{(3)} ds_{n+1}^{(4)} ds_{n+1}^{(5)} \quad (34)$$

関数 $f_1 \sim f_7$ の詳細な説明は紙面の都合上省略する。

3. Q 学習の設定

前章で期待スコアを最小化する最適化問題を定義したが、複数の確率分布の組み合わせにより定義されているこの問題を解析的に解くことは困難であるため、定義したゴルフモデルが、未来の状態が過去の状態によらず現在の状態から決まるマルコフ性を有しているという特徴から、マルコフ決定過程のフレームワークで有効な手法である Q 学習 [8] によって数値的に解く。この章では Q 学習をどのようにゴルフに適用するかを説明する。

(1) 行動: モデルではクラブ、力加減、方向の3つから定義したが、Q 学習を適用するため離散化と行動空間の縮小を行う。まず力加減は常にフルショットとし、一番飛距離の短い SW より短い距離が狙えなくなないように、SW だけ飛距離 10 ヤード刻みで力加減できるものとする。そしてそれぞれを1つのクラブとして扱い、クラブ集合 C に SW を力加減した飛距離のクラブを足したクラブ集合を C' とする。また、方向 θ は $\theta \sim [0, 360)$ の範囲で整数値をとる。

(2) 状態: 状態 s はコースの座標 (x, y) で連続値のため、ある大きさのグリッドで区切り離散化する。計算量と精度を見て、グリッドの大きさは約3ヤード四方とする。コースレイアウトは図6のようなコース画像を考え、 s はコース画像上にあり、状態空間の大きさは画像の大きさと等しく、 s がコース画像上からはみ出した場合は OB とする。

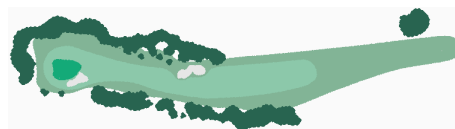


図6 学習で用いるコース画像

(3) 報酬: 報酬 r の与え方はスコアをもとにし、次状態が OB か WH のときは2、グリーンなら1+期待パット数、その他は1である。期待パット数は図5の期待パット数を用いる。この研究ではパッティングの戦略の最適化は目的に含まれていないため、条件をそろえるために全てのスキルのゴルファーで同じ値を用いる。

(4) Q の初期化: グリーン上にボールがのった時点でエピソードを終了する。グリーン上の Q 値は更新されないため、グリーン上の Q 値の初期値は全て0とする。グリーン以外の Q 値については、学習に用いるスキルから見積もる予想の期待打数より大きい値に定める。

(5) 行動選択: 行動選択は ϵ -greedy を用いる。ここでは学習時間と学習結果を見比べた際に、複数のコース、スキルに渡って効率よく学習が行われた $\epsilon = 0.9$ という値を用いる。

(6) 学習アルゴリズム: 学習の開始位置 (状態の初期位

Q(s, a)を初期化
各エピソードごとに以下を繰り返す
状態sの初期化(ティーインググラウンドから開始)
以下を繰り返す
ε-greedyから、状態sでの行動aを選択
・確率εでランダムな行動選択
・確率1-εでsに対するmin_a Q(s, a)となるaを選択
aをとり、報酬rと次状態s'を観測する
・s'がグリーンなら r ← 1 + 期待パット数
・s'がOBかWHなら r ← 2
・それ以外 r ← 1
Q(s, a) ← Q(s, a) + α[r + γ min_a Q(s', a') - Q(s, a)]
s ← s'
sがグリーンになるまで

図 7 本研究の Q 学習アルゴリズム

置)はティーインググラウンドとし、ボールがグリーンオンした時点で1つのエピソードを終了する。学習率 α, 割引率 γ は、学習時間と学習結果を見比べた際に、複数のコース、スキルに渡って効率よく学習が行われた値 α = 0.1, γ = 0.9 という値を用いる。最後に、図 7 にアルゴリズムをまとめる。

4. 実験

この章では、スキルとコースレイアウトに応じた学習が行われたかを確認する。学習のエピソード数を2千万回とし、学習過程でティーインググラウンドの位置のQ値が収束しかつ最小の値となったときのQテーブルを戦略とし

	平均スコア	スコアの標準偏差	フェアウェイキープ率(%)	パーオン率(%)	サンドセーブ率(%)	平均パット数
simulation	73.70	12.43	90.25	77.33	21.03	1.86
masters (2012)	73.5		71.92	58.72	46.65	1.66
masters (2011)	72.43		69.07	62.84	40.29	1.67

図 8 大会とシミュレーション結果の比較

Hole No.	simulation	masters (2013)	simulationとmastersの差
1	4.045	4.292	-0.247
2	5.129	4.714	0.415
3	3.825	4.010	-0.184
4	3.226	3.390	-0.164
5	4.058	4.227	-0.169
6	3.024	3.107	-0.083
7	4.312	4.240	0.071
8	4.700	4.679	0.021
9	4.258	4.107	0.151
10	4.595	4.201	0.393
11	4.361	4.276	0.085
12	2.712	3.218	-0.505
13	4.522	4.753	-0.232
14	4.006	4.101	-0.095
15	5.111	4.640	0.471
16	2.981	3.055	-0.074
17	4.392	4.221	0.171
18	4.351	4.182	0.169
Total	73.606	73.412	0.194

図 9 各ホールの大会とシミュレーションの平均スコア

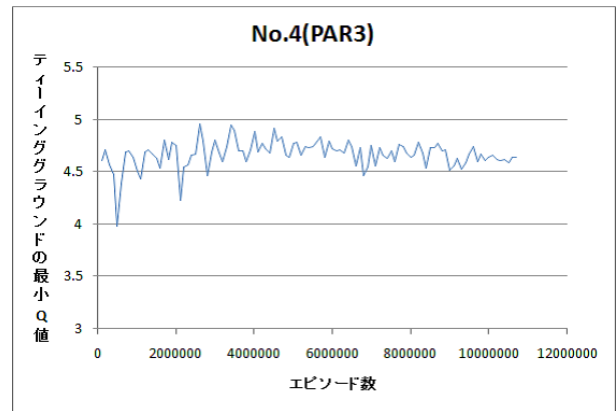


図 10 パー 3 ホールの学習曲線

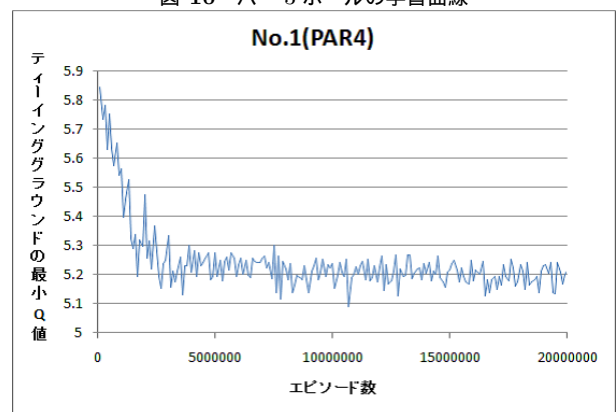


図 11 パー 4 ホールの学習曲線

て採用する。そして、Qテーブルからある位置で最小のQ値となる行動を選択するティーショットからカップインまでを1万回繰り返すシミュレーションを行った。Qテーブルを用いた行動選択はグリーンオンまで行われ、パッティングについては図4を用いてカップまでの残り距離に応じて確率的にパット数を決定する。

実験に用いる基準となるスキルとして、比較するデータの多いプロゴルファーの平均的なスキルを設定する。コースはアメリカにあるオーガスタ・ナショナル・ゴルフクラブのコース画像を用い、そこで毎年開かれるマスターズ・トーナメント [9] の大会結果と近い結果となるようにスキルパラメータを調節する。

フェアウェイキープ率は、パー3以外のホールにおいてティーショットで打ったボールがフェアウェイにとどまる割合、パーオン率は、1ラウンド中でパーオンする割合(各ホールのパーから2打少ない打数でグリーンオンすることをパーオンと呼ぶ)、サンドセーブ率は、グリーン周り

		DR	3W	5W	2I	3I	4I	5I	6I	7I	8I	9I	PW	SW
1σ	σ _r	7.72	7.12	6.63	5.73	5.25	4.77	4.34	4.04	3.5	3.08	2.59	2.29	1.99
	σ _e	1.81	1.69	1.57	1.51	1.39	1.33	1.27	1.21	1.15	1.03	0.91	0.78	0.66
3σ	σ _r	23.16	21.36	19.99	17.19	15.75	14.31	13.02	12.12	10.5	9.24	7.77	6.87	5.97
	σ _e	5.43	5.07	4.71	4.53	4.17	3.99	3.81	3.63	3.45	3.09	2.73	2.34	1.98

図 12 飛距離と方向の標準偏差

dir	c	DR	3W	5W	2I	3I	4I	5I	6I	7I	8I	9I	PW	SW
270Y	c	270	243	230	225	212	203	194	183	172	160	148	136	124
230Y	c	230	215	200	195	180	170	160	150	140	130	115	105	95

図 13 飛距離の平均値

のバンカーにボールを打ち込んだ後 2 打以内にカップインする割合を表す。図 8, 9 から、スコアについてはシミュレーションと大会でほとんど差がないことが確認できた。また、図 10, 11 のように、エピソード数に対してティーインググラウンドの位置で最小となる Q 値をプロットすることにより、Q 値が収束に向かっていくことが確認できる。

プロゴルファーの飛距離と方向の正確性スキルをまとめて 1σ と表記し、そのパラメータを定数倍することで複数の正確性を設定し（図 12）、飛距離についてもプロゴルファーの飛距離のスキルを 270Y と表記して、それをもとに複数の飛距離を設定する（図 13）。ここでは方向の平均は全てのクラブで 0、ミスショットの確率は Sig_d をもとに与え、 1σ では全てのクラブで 0、 3.0σ では全てのクラブで 0.1 とする。

最後に、スキルとコースレイアウトに応じた戦略が獲得できたかを視覚的に確認する。図 14, 15 はパー 4 ホール

で、ティーインググラウンドからカップまで 290 ヤードで、グリーン手前の池を越えるには 265 ヤード以上の飛距離が必要となる。Q テーブルにより各位置で指示される行動による狙いどころを、狙われる回数が多い位置の温度が高くなるようなヒートマップで表し、ティーインググラウンドと代表的な狙いどころを赤い点で示した。

図 14 は同じ 270Y の飛距離で、正確性が異なる場合である。正確性が高いと 1 打目で池越えのショットを狙い、正確性が低いと池より手前のフェアウェイセンターを狙っていることがわかる。正確性が低いゴルファーが池越えを狙うと、池に打ち込んでしまう可能性が高いため、スキルとコースレイアウトから池越えを狙わない戦略を学習したことを確認できた。図 15 は同じ 230Y の飛距離で、正確性が異なる場合である。265 ヤード以上の飛距離を出すことが出来ないため、このスキルでも池の手前のフェアウェイを狙う戦略を獲得した。

このように提案したフレームワークによって、スキルパラメータとコースレイアウトに応じて適切な戦略を獲得できることを示した。

5. バーチャルゴルフコーチ

攻略ルートとはグリーンオンまでの戦略の一部で、自身のスキルとコースレイアウトからグリーンオンまでの目標打数を定め、現在位置からその打数分の狙いどころを結んでいったルートのことである。ゴルフでは視界の範囲内で狙いどころを決めて打つと、次のショットが打ちづらくなる場面があり、コースを全体から広く見て狙いどころを決めることが重要と言われている。だが、グリーンオンまでに取り得る行動のパターンは数多く存在し、さらに行動の結果としてボール位置が一意に決まらず確率的に考える必要があるため、初心者であるほど攻略ルートをたてることは難しい。

そこで、今回提案したフレームワークを用いて、ゴルファーが攻略ルートをたてることの支援を行うバーチャルゴルフコーチを開発することを考えた。バーチャルゴルフコーチは以下のような機能を持つ。

- データ収集：プレー中のボール位置、コース区分、行動などのデータを蓄え、それによってスキルを推定したり、統計データを算出したりする。
- 戦略の計算：推定したスキルとコースレイアウトをもとに、戦略を計算する。
- 攻略ルート支援情報の提示：攻略ルートを考えるのに役立つ情報を提示する。

攻略ルート支援情報として、ゴルファーのショットの分布、目標打数として最適戦略をとるときの平均スコア、ある位置である行動をとることの価値、危険領域として周囲と比べて期待スコアが低い領域の図示を実装する。

実際の場面から適用例を示す。図 16 は、筆者が大札幌

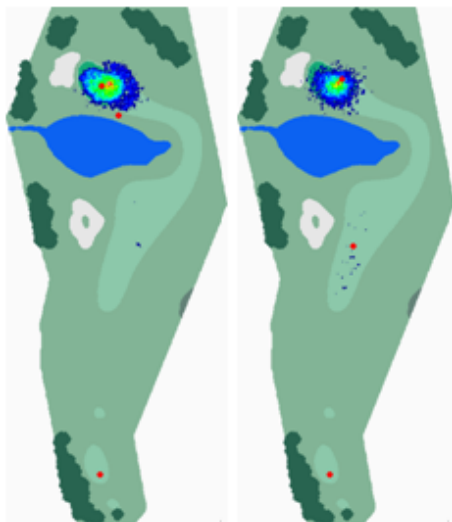


図 14 Q テーブルによる狙いどころのヒートマップ
(左: 270Y, 1σ 右: 270Y, 3.0σ)

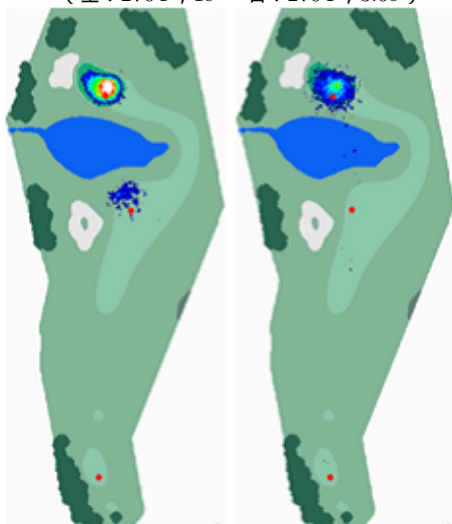


図 15 Q テーブルによる狙いどころのヒートマップ
(左: 230Y, 1σ 右: 270Y, 3.0σ)

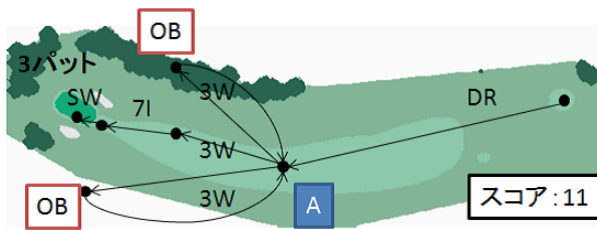


図 16 パー 4 ホールのプレイデータ

カントリークラブの 15 番ホール (パー 4) をプレイしたデータであり, 規定の 4 打を大きく超えてしまい, その原因として A 地点で 3W を選んだ行動選択の失敗を考えた.

	DR	3W	5W	7I	9I	11I	13I	15I	17I	19I	21I	23I	25I	27I	31I	35I	39I	43I	47I	51I	55I	59I	63I	67I	71I	75I	79I	83I	87I	91I	PW	SW	
σ_r	11.11	9.669	10.77	11.87	12.97	14.07	15.17	16.27	17.37	18.47	19.57	20.67	21.77	22.87	23.97	25.07	26.17	27.27	28.37	29.47	30.57	31.67	32.77	33.87	34.97	36.07	37.17	38.27	39.37	40.47	41.57	42.67	
σ_θ	12.5	18.09	16.31	14.53	12.75	15.72	18.69	16.2	13.71	16.2	16.7	13.54	8.985																				
d	1.77	1.61	1.55	1.48	1.42	1.43	1.45	1.38	1.31	1.18	1.06	0.95	0.83																				
μ_θ	2.723	0.454	4.202	7.95	11.7	4.34	-3.02	-1.85	-0.69	-0.96	-1.23	1.072	3.375																				

図 17 筆者のスキルパラメータ

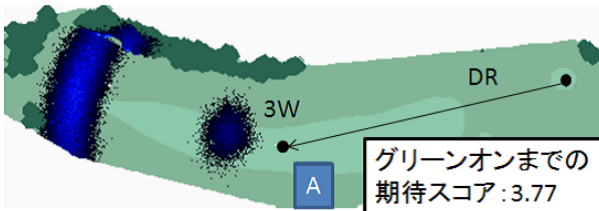


図 18 3W によるショット分布

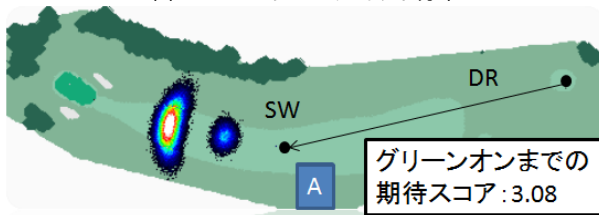


図 19 SW によるショット分布

筆者のスキルを推定すると図 17 のようになり, 図 18 は 3W のショット分布と 3W を使った場合のグリーンオンまでの期待スコア, 図 19 は一番期待スコアの低い行動である SW を使った場合のショット分布と期待スコアを示している. 3W は 1 打でグリーンオンすることを狙った選択なのだが, 期待スコアを見るとプラスして約 3 打多くかかってしまうことがわかり, SW を使って安全をとった方がスコアがよくなることが理解できた.

ゴルフの上達を考えると, 苦手なクラブをコースで使うことにより苦手なクラブが克服される場合もあり, 提示された通りに行動するのがそのまま上達につながるわけではない. だが, 何も情報がないときと比べ, バーチャルゴルフコーチにより行動を決定するための方針をはっきりさせることができ, それによってよりよい攻略ルートがたてられるのではないかと考えている.

6. まとめ

本研究は, スコアアップにつながる戦略を獲得するため

のフレームワークを開発することを目的とし, ゴルフのシミュレーションモデルを開発して, マルコフ決定過程において用いられる最適化の手法である Q 学習を適用することで, 期待スコアが最小となるような戦略を獲得できるフレームワークを新たに提案した. 実験から, そのフレームワークによりどの程度戦略が最適化されたかを示すとともに, フレームワークの応用例として, 攻略ルートの支援ツールであるバーチャルゴルフコーチの機能と支援例を提示した.

今後の課題として, シミュレーションモデルの詳細化と Q 学習の高速化が挙げられる. Q 学習は即時的に戦略を獲得できる手法ではないため, ツールとして考えたとき予め学習をしておかなければコースで情報を見ることが出来ない. そのため, 学習のたび Q 値を初期化するのではなく, 似たコースで学習した Q テーブルを再利用することで, 学習の効率化が出来るか考えている. また, ゴルファーのスキルパラメータを少ないサンプル数でも精度よく推定できる手法も必要である.

参考文献

- [1] Harold O. Fried, James Lambrinos and James Tyner. Evaluating the performance of professional golfers on the PGA, LPGA and SPGA tours. *European Journal of Operational Research*, 154: 548-561, April 2004.
- [2] Soichiro Suzuki, Yohei Hoshino and Yukinori Kobayashi. Skill Analysis of the Wrist Release in the Golf Swings Utilizing Shaft Elasticity. *Journal of System Design and Dynamics*, 3(1): 47-58, 2009.
- [3] Mark Broadie and Soonmin Ko. A simulation model to analyze the impact of distance and direction on golf scores. In *Winter Simulation Conference, WSC '09*, pages 3109-3120, 2009.
- [4] Christopher Archibald, Alon Altman and Yoav Shoham. Success, strategy and skill: an experimental study. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1, AAMAS '10*, pages 1089-1096, 2010.
- [5] Soonmin Ko. A Simulation Model to Analyze the Impact of Golf Skills and a Scenario-based Approach to Options Portfolio Optimization. In *Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences, Columbia University*, 2012.
- [6] Frank D. Werner and Richard C. Greig. *How Golf Clubs Really Work and How to Optimize Their Designs*. Origin Incorporated, 2000.
- [7] Matulya Bansal and Mark Broadie. A simulation model to analyze the impact of hole size on putting in golf. In *Winter Simulation Conference, WSC '08*, pages 2826-2834, 2008.
- [8] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning. A Bradford Book*, MIT Press Cambridge, MA, 1998.
- [9] <http://www.masters.com/>
マスターズ・トーナメント公式サイト