バイオメディカル・ファジィ・システム学会誌　Vol.24，No.1，pp.13－26（2022）

[Original article]

（2022 年 5 月 20 日　Accepted）

敵が適応的に行動選択するロールプレイングゲームの攻略法

前田　康成[1]

1）北見工業大学・地域未来デザイン工学科

要約：従来からロールプレイングゲーム（RPG）の攻略法に関する研究は多い．攻略法は，初心者に適切な行動を推薦する初心者支援機能に利用可能である．また，攻略法と被験者のプレイ履歴の比較による，RPG の楽しさに相当するゲーム設定の発見が期待される．敵の戦略の高度化による，習熟した遊び手の満足度向上が可能と考えられる．しかし，従来研究の敵の戦略は単純である．そこで，本研究では敵が適応的に行動選択する RPG の攻略法を検討する．提案方法では行動は max-min 基準に基づいて選択される．敵はプレイヤーの期待総利得を最小化する．プレイヤーは最小期待総利得を最大化する．モデル化にはマルコフ決定過程を利用し，提案方法では動的計画法を利用する．提案方法の有効性を数値計算例で示す．提案方法の RPG の難易度は敵の戦略が単純な比較対象の RPG よりも高いため，提案方法の期待総利得は比較対象の期待総利得よりも小さいことを確認した．

キーワード：RPG，マルコフ決定過程，動的計画法，max-min 基準，難易度

Strategies for Role-Playing Games in which enemies select actions adaptively

Yasunari MAEDA[1]

1) School of Regional Innovation and Social Design Engineering, Kitami Institute of Technology

Abstract: *There is a lot of previous research on strategies in role-playing games(RPG). The strategies can be used in a beginner support function that recommends appropriate actions to beginners. By comparing the strategy and subjects' playing histories, it is expected that the game setting that creates the enjoyment of RPG will be discovered. The sophistication of the enemies' strategies is expected to increase the satisfaction of improved players. But in the previous research, enemies have only simple strategies. In this research, a strategy in RPG with adaptive actions of enemies is studied. In the proposed method, actions are selected based on the max-min criterion. The enemy minimizes the player's expected total rewards. The player maximizes the minimum expected total rewards. Markov decision processes is used in modeling. In the proposed method, dynamic programming is used. The effectiveness of the proposed method is shown by some computational examples. Since the RPG of the proposed method is more difficult than the RPG of a comparison target with an enemy's simple strategy, the expected total rewards of the proposed method is smaller than that of the comparison target.*

Keywords: *RPG, Markov decision processes, dynamic programming, max-min criterion, difficulty*

Yasunari MAEDA
165 Koen-cho, Kitami-shi, Hokkaido, 090-8507, Japan
Phone:　+81-157-26-9328,　Fax: +81-157-26-9344,　E-mail: maedaya@mail.kitami-it.ac.jp

1. はじめに

2020年の新型コロナウィルス感染症の世界的流行の影響により，多くの産業が売上を下げている．しかし，ゲーム産業には新型コロナウィルス感染症の世界的流行下における生活様式の変化がプラスに作用し，売上を伸ばしている[1]．2020年の日本国内のゲーム市場規模は2兆円を超え，世界市場は20兆円を超えた[2]．

数理工学分野では，従来からゲーム機／パーソナルコンピュータ（PC）／スマートホンなどで遊ぶゲームの攻略法が検討されている．例えば，従来研究[3]ではビデオトレーディングカードゲームの攻略法，従来研究[4]～[8]では遊び手が操作するプレイヤーとコンピュータが操作する敵が戦うタイプのロールプレイングゲーム（RPG）の攻略法が検討されている．本研究でも従来研究同様にRPGの攻略法を対象とする．

攻略法の従来研究では，攻略法を検討する目的として，ゲーム初心者に対する支援機能，ゲーム開発の省力化などが挙げられている．しかし，どの従来研究でも将来的な目的としており，実際に初心者に対する支援機能や開発の省力化に貢献できるレベルには達しておらず，基礎研究の段階である．本研究も従来研究同様に基礎研究であり，実際に初心者に対する支援機能や開発の省力化に貢献するためには，今後の拡張研究が必要である．

従来研究や本研究では，数理工学の視点から何らかの基準（本研究ではmax-min基準）のもとで最適な（あるいは近似の）攻略法を提案／算出するが，最適な攻略法の検討と初心者支援機能，あるいは開発の省力化との関係がわかりにくいと思われるので，その関係を解説する．

最初に初心者支援機能との関係についてである．近年のRPGでは，当該RPGを初めて遊ぶ（あるいは不慣れな）遊び手に向けて，チュートリアルが用意されていることがある．内容はさまざまであるが，基本的にはゲーム中の何らかの状況を想定して，適切な行動（コマンド）を遊び手に提示し，実際に当該行動を遊び手に選択させて，適切なゲーム進行例を遊び手に体験させる機能である．現状では，このようなチュートリアルを提供するためには，チュートリアルで使用するゲーム中の状況や適切な行動を開発者がゲーム中から人手で選んで準備する必要がある．よって，開発コストの視点から，チュートリアルで提供できる状況（ゲームシーン）の数は少数に限られる．しかし，攻略法を自動的に算出できれば，従来のチュートリアルよりもより多くのゲームシーン（あるいはゲーム全体）で初心者に適切な行動を提示すること（初心者支援機能）が可能になる[7]．

次に，攻略法と開発の省力化との関係について解説する[4]～[8]．従来のチュートリアル開発において開発者が適切な行動を人手で選んでいる作業工程を，攻略法を自動的に算出するシステムで代替するのであれば，上記の初心者支援機能に関しても開発の省力化の一種と考えられる．

また，RPGの開発の際に既に多くの愛好者がいるような原作（アニメ等）に基づくこともある．原作の中には，登場キャラクタが必殺技など特定の戦闘の仕方（勝ち方）をすることもある．原作の愛好者にもRPGに満足してもらうには，原作の必殺技がRPGの当該状況における最適な行動と一致するゲーム設定（遊び手が操作する登場キャラクタや敵の能力値などの設定）が好ましい．現状では，あるゲーム設定のもとで被験者にプレイしてもらって，被験者が習熟した際に必殺技を選択するかどうか，あるいは被験者の代わりに何らかのシミュレータを使用して，シミュレータ上のエージェント（被験者の代わり）が必殺技を選択するかどうかを確認して，当該設定が妥当かどうかを確認する必要がある．他方，攻略法の自動算出が可能になれば，当該設定のもとで算出した攻略法での選択行動と原作の必殺技が一致するかどうか確認することで，当該設定の妥当性を容易に確認できる．上記のような設定の妥当性の確認（あるいは，妥当な設定の探索）に要する時間や稼働等の開発コストは攻略法の自動算出によって省力化が可能と考える．

また，RPGを被験者がプレイした際の行動選択の履歴データと自動算出した攻略法を比較することにより，人間が最適な攻略法と同様の行動選択をするのかどうか比較検証できる．仮にRPGに習熟した被験者が攻略法と異なる行動を選択する場合，その最適ではない選択の中に人間が感じる楽しさなどが存在する可能性がある．被験者の選択と攻略法の比較検証から人間の感じる楽しさをみつけ，RPG上の各種設定との関係（例えば，人間が楽しいと感じる設定）を発見できれば，人間が楽しいと感じる要素を多く含むRPGの効率的な開発が期待される．現状では，従来研究も本研究も実用化（実際のRPGへの適用）レベルには到達しておらず，検討する問題設定を実際のRPGに徐々に近づけている基礎研究段階である．

RPGの従来研究では，いろいろな問題設定のもとで遊び手のプレイヤーに関する最適な攻略法が検討されている．しかし，敵の攻撃については常に確定的にプレイヤー側にダメージが発生すると仮定する場合や，攻撃の結果が確率的な場合でも攻撃の種類が1種類のみなど，基本的に単純な攻撃／戦略に限られている．敵の攻撃を

単純なものに限定していると，敵の強さ（難易度設定）と遊び手の満足度との関係などの検討が十分にできない．初心者の遊び手に対しては低い難易度（単純な敵の戦略）が求められるが，慣れた（習熟した）遊び手に対しては高い難易度（高度な敵の戦略）が求められると考える．例えば，RPGではないが，従来研究[3]ではビデオトレーディングカードゲームに関して，上達した遊び手に対する難易度の高い強い敵の必要性が示唆されている．

そこで，本研究では敵の攻撃についても，複数種類の行動の中から適応的に敵が選択するような問題設定を考える．敵の戦略は敵の目的に依存するが，本研究ではプレイヤーの目的と逆の目的を敵の目的とする．プレイヤーが敵を倒すことによって得られる値を本研究では利得と呼ぶが，プレイヤーの目的が得られる総利得の期待値の最大化であるのに対して，敵の目的はプレイヤーの期待総利得の最小化とする．具体的には従来研究[4]～[8]同様に意思決定を伴う確率モデルであるマルコフ決定過程[9]を利用してモデル化し，動的計画法[9]によって逐次的に行動を選択する．

なお，本研究は基礎研究であり，研究内容そのものの実用化は考えていない．議論を簡便にするために対象とするRPGは簡易な小規模のRPGにしている．特にプレイヤーと敵の戦闘に焦点を当てるため，RPG中でプレイヤーと敵が戦う戦闘場面（戦闘モード）のみを対象としている．実際のRPGでは敵と遭遇する前に地図（マップ）上を移動する移動場面（マップモード）も存在する．また，本研究ではプレイヤーと敵の攻撃等が成功／失敗する確率を既知と仮定する．実際には，これらの情報は開発者のみ知る情報である．よって，実際の遊び手を想定するためには，これらの情報を未知とした問題設定を考える必要がある．これらの検討は今後の課題である．

ただし，敵の攻撃が単純なRPGに関しては従来研究[6]～[8]において，マップモードを含むRPGや，各種確率が未知の実際の遊び手の立場での攻略法も検討されており，本研究の拡張の際にはこれらの従来研究の知見の多くが活用可能である．また，近年のRPGでは複数の遊び手が参加するタイプのRPGもある．類似のRPGには，遊び手は1人で，遊び手が操作するプレイヤー（プレイヤーキャラクタ）と，コンピュータが操作するノンプレイヤーキャラクタがグループ（パーティ）になって協力するタイプのRPGがある．後者の遊び手のプレイヤーとノンプレイヤーキャラクタが協力するRPGの攻略法については従来研究[7]で検討されている．本研究をノンプレイヤーキャラクタとの協力を含むRPGや複数の遊び手の場合に拡張する際には，従来研究[7]が参考になると考える．

2. 準備

従来研究[6]～[8]では，マップ上をプレイヤーが動いて敵に遭遇すると戦闘する，マップモードと戦闘モードで構成されるRPGについて検討している．従来研究[4][5]では，敵と戦闘する戦闘モードのみのRPGについて検討している．本研究では，従来研究[4][5]同様に戦闘モードのみのRPGを対象とする．

本研究で使用する記号などを説明する．記号やモデルの基本部分の多くは従来研究[6]～[8]と同様である．プレイヤーはヒットポイント（HP）と呼ばれる体力に相当する数値を持ち，開始時は最大値$M(pl)$とする．HPが0以下になると倒されたことを意味し，次の期には最大値まで復活する．$e_i \in E$はi番目の敵を示し，$E=\left\{e_1, e_2, \dots, e_{|E|}\right\}$は敵集合である．敵$e_i$の出現時のHPは最大値$M(e_i)$とする．プレイヤーが同時に戦う敵の数は1体とし，敵のHPが0以下になるとプレイヤーの勝利である．プレイヤーが敵を倒すと，敵e_iの発生確率$\Pr(e_i)$によって次の敵が最大HPで発生する．なお，プレイヤーと敵のどちらが勝った場合でも，勝った側の勝った時点でのHPの値が次の期（次の戦闘）に引き継がれる．

プレイヤーの行動には敵に攻撃する攻撃行動$a_{pl,1,i}$とプレイヤーのHPを回復する回復行動$a_{pl,2,i}$がある．$a_{pl,1,i}$はプレイヤーのi番目の攻撃行動，$a_{pl,2,i}$はプレイヤーのi番目の回復行動を示し，$A_{pl,1}=\left\{a_{pl,1,1}, a_{pl,1,2}, \dots, a_{pl,1,\left|A_{pl,1}\right|}\right\}$はプレイヤーの攻撃行動集合，$A_{pl,2}=\left\{a_{pl,2,1}, a_{pl,2,2}, \dots, a_{pl,2,\left|A_{pl,2}\right|}\right\}$はプレイヤーの回復行動集合である．$A_{pl}=A_{pl,1} \cup A_{pl,2}$はプレイヤーの行動集合で，攻撃行動集合と回復行動集合の和集合である．

プレイヤーの攻撃行動$a_{pl,1,i}$は確率$\Pr\left(B_1\left(a_{pl,1,i}\right) \middle| a_{pl,1,i}\right)$で成功し，敵のHPを$B_1\left(a_{pl,1,i}\right)$だけ減少させる．攻撃確率は二項分布であり，確率$1-\Pr\left(B_1\left(a_{pl,1,i}\right) \middle| a_{pl,1,i}\right)$で失敗する．回復行動$a_{pl,2,i}$は確率$\Pr\left(B_2\left(a_{pl,2,i}\right) \middle| a_{pl,2,i}\right)$で成功し，プレイヤーのHPを

$B_2\left(a_{pl,2,i}\right)$だけ回復させる．回復後のHPが最大値$M(pl)$を超える場合は最大値$M(pl)$まで回復させる．回復確率は二項分布であり，確率$1-\Pr\left(B_2\left(a_{pl,2,i}\right)\middle|a_{pl,2,i}\right)$で失敗する．

敵e_iの行動にはプレイヤーに攻撃する攻撃行動$a_{e_i,1,j}$と敵e_iのHPを回復する回復行動$a_{e_i,2,j}$がある．$a_{e_i,1,j}$は敵e_iのj番目の攻撃行動，$a_{e_i,2,j}$は敵e_iのj番目の回復行動を示し，$A_{e_i,1}=\left\{a_{e_i,1,1},a_{e_i,1,2},\dots,a_{e_i,1,\left|A_{e_i,1}\right|}\right\}$は敵$e_i$の攻撃行動集合，$A_{e_i,2}=\left\{a_{e_i,2,1},a_{e_i,2,2},\dots,a_{e_i,2,\left|A_{e_i,2}\right|}\right\}$は敵$e_i$の回復行動集合である．$A_{e_i}=A_{e_i,1}\cup A_{e_i,2}$は敵$e_i$の行動集合で，攻撃行動集合と回復行動集合の和集合である．

攻撃行動$a_{e_i,1,j}$は確率$\Pr\left(B_1\left(a_{e_i,1,j}\right)\middle|a_{e_i,1,j}\right)$で成功し，プレイヤーのHPを$B_1\left(a_{e_i,1,j}\right)$だけ減少させる．攻撃確率は二項分布であり，確率$1-\Pr\left(B_1\left(a_{e_i,1,j}\right)\middle|a_{e_i,1,j}\right)$で失敗する．回復行動$a_{e_i,2,j}$は確率$\Pr\left(B_2\left(a_{e_i,2,j}\right)\middle|a_{e_i,2,j}\right)$で成功し，敵$e_i$のHPを$B_2\left(a_{e_i,2,j}\right)$だけ回復させる．回復後のHPが最大値$M(e_i)$を超える場合は最大値$M(e_i)$まで回復させる．回復確率は二項分布であり，確率$1-\Pr\left(B_2\left(a_{e_i,2,j}\right)\middle|a_{e_i,2,j}\right)$で失敗する．

選択した行動が攻撃でも回復でも常にプレイヤーが先攻，敵が後攻で，プレイヤーと敵の1回ずつの行動実施を1期（1ターン）とし，有限のT期間の問題設定を考える．プレイヤーを主体とした問題設定とし，プレイヤーが敵を倒した際の利得を$G(e_i)$（正の利得），敵がプレイヤーを倒した際の利得を$G(pl)$（負の利得）とする．利得はすべてプレイヤーが得るものとし，プレイヤーはT期間での期待総利得の最大化，敵はプレイヤーの期待総利得の最小化を目的とする．

T期間の問題設定におけるt期の状態を$X_t=\left(X_{t,1},X_{t,2},X_{t,3}\right)$とする．ただし，$X_{t,1}$は$t$期のプレイヤーのHPを示す変数，$X_{t,2}\in E$はプレイヤーと戦闘中の敵の種類を示す変数，$X_{t,3}$は敵$X_{t,2}$のHPを示す変数である．$Y_t=\left(Y_{t,1},Y_{t,2}\right)$は$t$期の行動で，$Y_{t,1}\in A_{pl}$はプレイヤーの行動を示す変数，$Y_{t,2}\in A_{X_{t,2}}$は敵$X_{t,2}$の行動を示す変数である．状態$X_t$で行動$Y_t$を選択（実施）し状態$X_{t+1}$に遷移した場合の利得を$r\left(X_t,Y_t,X_{t+1}\right)$とする．利得の値は当該遷移においてプレイヤーが倒された場合は$G(pl)$，敵$X_{t,2}$が倒された場合は$G\left(X_{t,2}\right)$，その他の場合は0である．

3.　状態遷移と状態遷移確率

4章で解説する提案方法（プレイヤーと敵の行動選択の計算方法）では，状態X_tで行動Y_tを選択（実施）し状態X_{t+1}に遷移する確率である状態遷移確率$\Pr\left(X_{t+1}\middle|X_t,Y_t\right)$を計算に使用する．状態遷移確率は具体的には，プレイヤーと敵の攻撃確率や回復確率の同時確率になる．また，プレイヤーと敵の行動の成功／失敗の組合せが同じでも，その際のプレイヤーや敵のHP次第で倒される場合と倒されない場合がある．以下で，4章の提案方法の解説の前に，状態遷移の各種パターンとその際の状態遷移確率，状態の更新，利得を解説する．

3.1　プレイヤーと敵がともに攻撃の場合

・プレイヤーの攻撃が成功かつ$B_1\left(Y_{t,1}\right)<X_{t,3}$，敵の攻撃が成功かつ$B_1\left(Y_{t,2}\right)<X_{t,1}$の場合

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right)=$$

$$\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\Pr\left(B_1\left(Y_{t,2}\right)\middle|Y_{t,2}\right) \tag{1}$$

$$X_{t+1,1}=X_{t,1}-B_1\left(Y_{t,2}\right) \tag{2}$$

$$X_{t+1,2}=X_{t,2} \tag{3}$$

$$X_{t+1,3}=X_{t,3}-B_1\left(Y_{t,1}\right) \tag{4}$$

$$r\left(X_t,Y_t,X_{t+1}\right)=0 \tag{5}$$

・プレイヤーの攻撃が成功かつ$B_1\left(Y_{t,1}\right)<X_{t,3}$，敵の攻

撃が成功かつ$B_1\left(Y_{t,2}\right) \geq X_{t,1}$の場合（プレイヤーが倒される）

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right) =$$

$$\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\Pr\left(B_1\left(Y_{t,2}\right)\middle|Y_{t,2}\right) \quad (6)$$

$$X_{t+1,1} = M(pl) \quad (7)$$

$$X_{t+1,2} = X_{t,2} \quad (8)$$

$$X_{t+1,3} = X_{t,3} - B_1\left(Y_{t,1}\right) \quad (9)$$

$$r\left(X_t,Y_t,X_{t+1}\right) = G(pl) \quad (10)$$

・プレイヤーの攻撃が成功かつ$B_1\left(Y_{t,1}\right) \geq X_{t,3}$の場合（敵が倒される）

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right) =$$

$$\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\Pr\left(X_{t+1,2} = e_i\right) \quad (11)$$

$$X_{t+1,1} = X_{t,1} \quad (12)$$

$$X_{t+1,2} = e_i \quad (13)$$

$$X_{t+1,3} = M(e_i) \quad (14)$$

$$r\left(X_t,Y_t,X_{t+1}\right) = G\left(X_{t,2}\right) \quad (15)$$

・プレイヤーの攻撃が成功かつ$B_1\left(Y_{t,1}\right) < X_{t,3}$，敵の攻撃が失敗の場合

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right) =$$

$$\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\left(1-\Pr\left(B_1\left(Y_{t,2}\right)\middle|Y_{t,2}\right)\right) \quad (16)$$

$$X_{t+1,1} = X_{t,1} \quad (17)$$

$$X_{t+1,2} = X_{t,2} \quad (18)$$

$$X_{t+1,3} = X_{t,3} - B_1\left(Y_{t,1}\right) \quad (19)$$

$$r\left(X_t,Y_t,X_{t+1}\right) = 0 \quad (20)$$

・プレイヤーの攻撃が失敗，敵の攻撃が成功かつ$B_1\left(Y_{t,2}\right) < X_{t,1}$の場合

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right) =$$

$$\left(1-\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\right)\Pr\left(B_1\left(Y_{t,2}\right)\middle|Y_{t,2}\right) \quad (21)$$

$$X_{t+1,1} = X_{t,1} - B_1\left(Y_{t,2}\right) \quad (22)$$

$$X_{t+1,2} = X_{t,2} \quad (23)$$

$$X_{t+1,3} = X_{t,3} \quad (24)$$

$$r\left(X_t,Y_t,X_{t+1}\right) = 0 \quad (25)$$

・プレイヤーの攻撃が失敗，敵の攻撃が成功かつ$B_1\left(Y_{t,2}\right) \geq X_{t,1}$の場合（プレイヤーが倒される）

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right) =$$

$$\left(1-\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\right)\Pr\left(B_1\left(Y_{t,2}\right)\middle|Y_{t,2}\right) \quad (26)$$

$$X_{t+1,1} = M(pl) \quad (27)$$

$$X_{t+1,2} = X_{t,2} \quad (28)$$

$$X_{t+1,3} = X_{t,3} \quad (29)$$

$$r\left(X_t,Y_t,X_{t+1}\right) = G(pl) \quad (30)$$

・プレイヤーの攻撃が失敗，敵の攻撃が失敗の場合

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right) =$$

$$\left(1-\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\right)\left(1-\Pr\left(B_1\left(Y_{t,2}\right)\middle|Y_{t,2}\right)\right) \quad (31)$$

$$X_{t+1,1} = X_{t,1} \quad (32)$$

$$X_{t+1,2} = X_{t,2} \quad (33)$$

$$X_{t+1,3} = X_{t,3} \quad (34)$$

$$r\left(X_t,Y_t,X_{t+1}\right) = 0 \quad (35)$$

3.2　プレイヤーが攻撃，敵が回復の場合

・プレイヤーの攻撃が成功かつ$B_1\left(Y_{t,1}\right) < X_{t,3}$，敵の回復が成功の場合

$$\Pr\left(X_{t+1}\middle|X_t,Y_t\right) =$$

$$\Pr\left(B_1\left(Y_{t,1}\right)\middle|Y_{t,1}\right)\Pr\left(B_2\left(Y_{t,2}\right)\middle|Y_{t,2}\right) \quad (36)$$

$$X_{t+1,1} = X_{t,1} \quad (37)$$

$$X_{t+1,2} = X_{t,2} \quad (38)$$

$$X_{t+1,3} = \min\left\{X_{t,3} - B_1\left(Y_{t,1}\right) + B_2\left(Y_{t,2}\right), M\left(X_{t,2}\right)\right\}$$

(39)

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{40}$$

・プレイヤーの攻撃が成功かつ$B_1\left(Y_{t,1}\right) \geq X_{t,3}$の場合（敵が倒される）

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\Pr\left(B_1\left(Y_{t,1}\right) \middle| Y_{t,1}\right) \Pr\left(X_{t+1,2} = e_i\right) \tag{41}$$

$$X_{t+1,1} = X_{t,1} \tag{42}$$

$$X_{t+1,2} = e_i \tag{43}$$

$$X_{t+1,3} = M(e_i) \tag{44}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = G\left(X_{t,2}\right) \tag{45}$$

・プレイヤーの攻撃が成功かつ$B_1\left(Y_{t,1}\right) < X_{t,3}$，敵の回復が失敗の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\Pr\left(B_1\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\left(1 - \Pr\left(B_2\left(Y_{t,2}\right) \middle| Y_{t,2}\right)\right) \tag{46}$$

$$X_{t+1,1} = X_{t,1} \tag{47}$$

$$X_{t+1,2} = X_{t,2} \tag{48}$$

$$X_{t+1,3} = X_{t,3} - B_1\left(Y_{t,1}\right) \tag{49}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{50}$$

・プレイヤーの攻撃が失敗，敵の回復が成功の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\left(1 - \Pr\left(B_1\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\right) \Pr\left(B_2\left(Y_{t,2}\right) \middle| Y_{t,2}\right) \tag{51}$$

$$X_{t+1,1} = X_{t,1} \tag{52}$$

$$X_{t+1,2} = X_{t,2} \tag{53}$$

$$X_{t+1,3} = \min\left\{X_{t,3} + B_2\left(Y_{t,2}\right), M\left(X_{t,2}\right)\right\} \tag{54}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{55}$$

・プレイヤーの攻撃が失敗，敵の回復が失敗の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\left(1 - \Pr\left(B_1\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\right)\left(1 - \Pr\left(B_2\left(Y_{t,2}\right) \middle| Y_{t,2}\right)\right)$$

(56)

$$X_{t+1,1} = X_{t,1} \tag{57}$$

$$X_{t+1,2} = X_{t,2} \tag{58}$$

$$X_{t+1,3} = X_{t,3} \tag{59}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{60}$$

3.3　プレイヤーが回復，敵が攻撃の場合

・プレイヤーの回復が成功，敵の攻撃が成功かつ$B_1\left(Y_{t,2}\right) < \min\left\{X_{t,1} + B_2\left(Y_{t,1}\right), M(pl)\right\}$の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right) \Pr\left(B_1\left(Y_{t,2}\right) \middle| Y_{t,2}\right) \tag{61}$$

$$X_{t+1,1} = \min\left\{X_{t,1} + B_2\left(Y_{t,1}\right), M(pl)\right\} - B_1\left(Y_{t,2}\right)$$

(62)

$$X_{t+1,2} = X_{t,2} \tag{63}$$

$$X_{t+1,3} = X_{t,3} \tag{64}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{65}$$

・プレイヤーの回復が成功，敵の攻撃が成功かつ$B_1\left(Y_{t,2}\right) \geq \min\left\{X_{t,1} + B_2\left(Y_{t,1}\right), M(pl)\right\}$の場合（プレイヤーが倒される）

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right) \Pr\left(B_1\left(Y_{t,2}\right) \middle| Y_{t,2}\right) \tag{66}$$

$$X_{t+1,1} = M(pl) \tag{67}$$

$$X_{t+1,2} = X_{t,2} \tag{68}$$

$$X_{t+1,3} = X_{t,3} \tag{69}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = G(pl) \tag{70}$$

・プレイヤーの回復が成功，敵の攻撃が失敗の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\left(1 - \Pr\left(B_1\left(Y_{t,2}\right) \middle| Y_{t,2}\right)\right) \tag{71}$$

$$X_{t+1,1} = \min\left\{X_{t,1} + B_2\left(Y_{t,1}\right), M(pl)\right\} \tag{72}$$

$$X_{t+1,2} = X_{t,2} \tag{73}$$

$$X_{t+1,3} = X_{t,3} \tag{74}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{75}$$

・プレイヤーの回復が失敗，敵の攻撃が成功かつ$B_1\left(Y_{t,2}\right) < X_{t,1}$の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\left(1 - \Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\right) \Pr\left(B_1\left(Y_{t,2}\right) \middle| Y_{t,2}\right) \tag{76}$$

$$X_{t+1,1} = X_{t,1} - B_1\left(Y_{t,2}\right) \tag{77}$$

$$X_{t+1,2} = X_{t,2} \tag{78}$$

$$X_{t+1,3} = X_{t,3} \tag{79}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{80}$$

・プレイヤーの回復が失敗，敵の攻撃が成功かつ$B_1\left(Y_{t,2}\right) \geq X_{t,1}$の場合（プレイヤーが倒される）

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\left(1 - \Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\right) \Pr\left(B_1\left(Y_{t,2}\right) \middle| Y_{t,2}\right) \tag{81}$$

$$X_{t+1,1} = M(pl) \tag{82}$$

$$X_{t+1,2} = X_{t,2} \tag{83}$$

$$X_{t+1,3} = X_{t,3} \tag{84}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = G(pl) \tag{85}$$

・プレイヤーの回復が失敗，敵の攻撃が失敗の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\left(1 - \Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\right)\left(1 - \Pr\left(B_1\left(Y_{t,2}\right) \middle| Y_{t,2}\right)\right) \tag{86}$$

$$X_{t+1,1} = X_{t,1} \tag{87}$$

$$X_{t+1,2} = X_{t,2} \tag{88}$$

$$X_{t+1,3} = X_{t,3} \tag{89}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{90}$$

3.4　プレイヤーと敵がともに回復の場合

・プレイヤーの回復が成功，敵の回復が成功の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right) \Pr\left(B_2\left(Y_{t,2}\right) \middle| Y_{t,2}\right) \tag{91}$$

$$X_{t+1,1} = \min\left\{X_{t,1} + B_2\left(Y_{t,1}\right), M(pl)\right\} \tag{92}$$

$$X_{t+1,2} = X_{t,2} \tag{93}$$

$$X_{t+1,3} = \min\left\{X_{t,3} + B_2\left(Y_{t,2}\right), M\left(X_{t,2}\right)\right\} \tag{94}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{95}$$

・プレイヤーの回復が成功，敵の回復が失敗の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\left(1 - \Pr\left(B_2\left(Y_{t,2}\right) \middle| Y_{t,2}\right)\right) \tag{96}$$

$$X_{t+1,1} = \min\left\{X_{t,1} + B_2\left(Y_{t,1}\right), M(pl)\right\} \tag{97}$$

$$X_{t+1,2} = X_{t,2} \tag{98}$$

$$X_{t+1,3} = X_{t,3} \tag{99}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{100}$$

・プレイヤーの回復が失敗，敵の回復が成功の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\left(1 - \Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\right) \Pr\left(B_2\left(Y_{t,2}\right) \middle| Y_{t,2}\right) \tag{101}$$

$$X_{t+1,1} = X_{t,1} \tag{102}$$

$$X_{t+1,2} = X_{t,2} \tag{103}$$

$$X_{t+1,3} = \min\left\{X_{t,3} + B_2\left(Y_{t,2}\right), M\left(X_{t,2}\right)\right\} \tag{104}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{105}$$

・プレイヤーの回復が失敗，敵の回復が失敗の場合

$$\Pr\left(X_{t+1} \middle| X_t, Y_t\right) =$$

$$\left(1 - \Pr\left(B_2\left(Y_{t,1}\right) \middle| Y_{t,1}\right)\right)\left(1 - \Pr\left(B_2\left(Y_{t,2}\right) \middle| Y_{t,2}\right)\right) \tag{106}$$

$$X_{t+1,1} = X_{t,1} \tag{107}$$

$$X_{t+1,2} = X_{t,2} \tag{108}$$

$$X_{t+1,3} = X_{t,3} \tag{109}$$

$$r\left(X_t, Y_t, X_{t+1}\right) = 0 \tag{110}$$

4.　提案方法

動的計画法による提案方法では，T期から1期まで遡りながら処理を実施する．T期の処理を以下に示す．

$$V(X_T,T)=\max_{Y_{T,1}\in A_{pl}}\min_{Y_{T,2}\in A_{X_{T,2}}}$$

$$\sum_{X_{T+1}}\Pr(X_{T+1}|X_T,Y_T)r(X_T,Y_T,X_{T+1}) \quad (111)$$

$V(X_T,T)$はT期のプレイヤーの行動$Y_{T,1}$に対して敵が最小化（式(111)の内側のmin）するプレイヤーの期待利得の最小値を，プレイヤーが行動$Y_{T,1}$の選択によって最大化（式(111)の外側のmax）する期待利得の値である．式(111)のmax-minに対応する行動の組がT期のプレイヤーの行動$Y_{T,1}$と敵の行動$Y_{T,2}$である．

t期（$1\le t<T$）の処理を以下に示す．

$$V(X_t,t)=\max_{Y_{t,1}\in A_{pl}}\min_{Y_{t,2}\in A_{X_{t,2}}}\sum_{X_{t+1}}\Pr\left(X_{t+1}\middle|X_t,Y_t\right)$$

$$\left(r(X_t,Y_t,X_{t+1})+V(X_{t+1},t+1)\right) \quad (112)$$

$V(X_t,t)$はt期のプレイヤーの行動$Y_{t,1}$に対して敵が最小化（式(112)の内側のmin）するプレイヤーのt期以降の期待総利得の最小値を，プレイヤーが行動$Y_{t,1}$の選択によって最大化（式(112)の外側のmax）する期待総利得の値である．式(112)のmax-minに対応する行動の組がt期のプレイヤーの行動$Y_{t,1}$と敵の行動$Y_{t,2}$である．

5.　数値計算例

提案方法について検証するために，以下に数値計算例を示す．設定が異なる3タイプの数値計算例を紹介する．

5.1　数値計算例タイプ1

敵の種類数$|E|=3$，プレイヤーの攻撃行動数$\left|A_{pl,1}\right|=2$，プレイヤーの回復行動数$\left|A_{pl,2}\right|=2$，敵e_iの攻撃行動数$\left|A_{e_1,1}\right|=\left|A_{e_2,1}\right|=\left|A_{e_3,1}\right|=2$，敵$e_i$の回復行動数$\left|A_{e_1,2}\right|=\left|A_{e_2,2}\right|=\left|A_{e_3,2}\right|=2$，プレイヤーの最大HPが$M(pl)=10$，敵$e_1$の最大HPが$M\left(e_1\right)=4$，敵$e_2$の最大HPが$M\left(e_2\right)=7$，敵$e_3$の最大HPが$M\left(e_3\right)=10$，敵の発生確率$\Pr(e_i)$は等確率，期間の長さ$T=100$，プレイヤーが敵に倒された場合の利得$G(pl)=-1$，敵$e_i$を倒した場合の利得$G\left(e_1\right)=G\left(e_2\right)=G\left(e_3\right)=1$，プレイヤーと敵の攻撃行動のダメージ量$B_1\left(a_{\cdot,1,i}\right)$と回復行動の回復量$B_2\left(a_{\cdot,2,i}\right)$を表1，プレイヤーと敵の攻撃行動の成功確率$\Pr\left(B_1\left(a_{\cdot,1,i}\right)\middle|a_{\cdot,1,i}\right)$と回復行動の成功確率$\Pr\left(B_2\left(a_{\cdot,2,i}\right)\middle|a_{\cdot,2,i}\right)$を表2に示す．表1および表2では，末尾の添え字（番号）の大きな攻撃，回復の方が効果（ダメージ量，回復量）が大きくて，成功確率が小さい設定である．初期（1期）のプレイヤーと敵のHPは最大HPとする．

表1.　ダメージ量$B_1\left(a_{\cdot,1,i}\right)$と回復量$B_2\left(a_{\cdot,2,i}\right)$

	$B_1\left(a_{\cdot,1,1}\right)$	$B_1\left(a_{\cdot,1,2}\right)$	$B_2\left(a_{\cdot,2,1}\right)$	$B_2\left(a_{\cdot,2,2}\right)$
pl	4	8	4	8
e_1	1	2	1	2
e_2	2	4	2	4
e_3	4	8	4	8

$B_1\left(a_{\cdot,1,i}\right)$：プレイヤーと敵の攻撃行動のダメージ量

$B_2\left(a_{\cdot,2,i}\right)$：プレイヤーと敵の回復行動の回復量

表2.　確率$\Pr\left(B_1\left(a_{\cdot,1,i}\right)\middle|a_{\cdot,1,i}\right)$と$\Pr\left(B_2\left(a_{\cdot,2,i}\right)\middle|a_{\cdot,2,i}\right)$

	$a_{\cdot,1,1}$	$a_{\cdot,1,2}$	$a_{\cdot,2,1}$	$a_{\cdot,2,2}$
pl	0.8	0.4	0.8	0.4
e_1	0.8	0.4	0.8	0.4
e_2	0.8	0.4	0.8	0.4
e_3	0.8	0.4	0.8	0.4

$\Pr\left(B_1\left(a_{\cdot,1,i}\right)\middle|a_{\cdot,1,i}\right)$：プレイヤーと敵の攻撃成功確率

$\Pr\left(B_2\left(a_{\cdot,2,i}\right)\middle|a_{\cdot,2,i}\right)$：プレイヤーと敵の回復成功確率

提案方法の比較対象として比較対象COMの結果も示す．比較対象COMでは，敵の行動選択は等確率のランダム選択とし，プレイヤーの行動選択のみ最適化する．比較対象COMは従来研究相当の単純な敵戦略の一例である．比較対象COMの行動選択（T期とt期（$1\le t<T$））は以下

のとおりである．

$$V_{COM}(X_T,T)=\max_{Y_{T,1}\in A_{pl}}\sum_{Y_{T,2}\in A_{X_{T,2}}}\Pr\left(Y_{T,2}\right)$$

$$\sum_{X_{T+1}}\Pr(X_{T+1}|X_T,Y_T)r(X_T,Y_T,X_{T+1}) \qquad (113)$$

ただし，敵の行動選択確率$\Pr\left(Y_{T,2}\right)$は等確率である．

$$V_{COM}(X_t,t)$$

$$=\max_{Y_{t,1}\in A_{pl}}\sum_{Y_{t,2}\in A_{X_{t,2}}}\Pr\left(Y_{t,2}\right)\sum_{X_{t+1}}\Pr\left(X_{t+1}\middle|X_t,Y_t\right)$$

$$\left(r(X_t,Y_t,X_{t+1})+V(X_{t+1},t+1)\right) \qquad (114)$$

比較対象COMは行動選択がランダムな敵に対してプレイヤーが期待総利得を最大化する場合である．なお，提案方法の敵がプレイヤーの期待総利得を最小化しようと行動選択するのに対し，比較対象COMの敵はランダムな行動選択なので，提案方法での期待総利得が比較対象COMでの期待総利得よりも小さくなることは最適化問題の性質より自明である．しかし，本章では本研究における敵戦略の高度化の効果を具体的に確認するために，数値計算例を紹介する．

提案方法と比較対象COMの期待総利得等の数値計算結果を紹介するが，その前にプレイヤーと敵の間での具体的な戦闘例を紹介する．ここで紹介する戦闘例（図1～図3）は，すべて数値計算例タイプ1に関して提案方法の式(111)と式(112)での行動選択による戦闘例である．

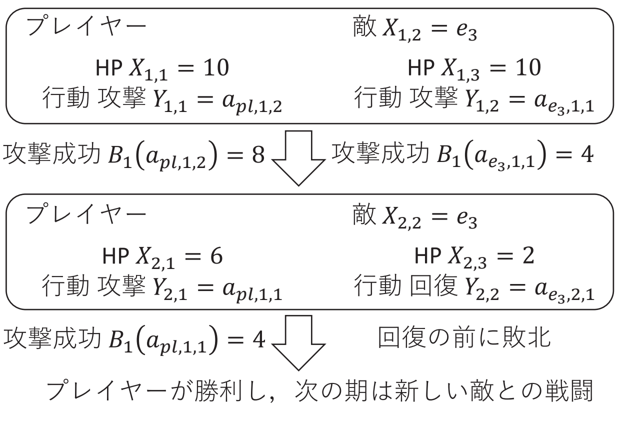

図1.　プレイヤーが敵e_3に最小回数で勝つ例

図1は1期（初期）の敵が敵e_3の場合の例である．プレイヤーに注目して解説すると，プレイヤーの攻撃は2種あるが，1期にはダメージ量のより大きな攻撃$a_{pl,1,2}$を選択して成功し，敵e_3のHPを10から2に減少させている．2期では，敵e_3に勝つためには成功する確率が攻撃$a_{pl,1,2}$よりも大きな攻撃$a_{pl,1,1}$のダメージ量4で十分なため，攻撃$a_{pl,1,1}$を選択して成功し，敵e_3に勝っている．次の期（3期）には，プレイヤーはHPが6のまま，新しい敵（敵は最大HP）との戦闘である．

図1の例ではプレイヤーが敵e_3に2回の行動選択で勝っているが，2回はタイプ1においてプレイヤーが敵e_3に勝つ最小回数である．他方，敵e_1の最大HPは4，敵e_2の最大HPは7なので，プレイヤーがそれぞれの敵に勝つ最小回数はともに1回である．

なお，プレイヤー，敵ともに攻撃等の成功／失敗は確率的なので，図1の戦闘例も確率的に発生する．1期の敵が敵e_3の条件のもとで，1期のプレイヤーの攻撃$a_{pl,1,2}$，敵e_3の攻撃$a_{e_3,1,1}$，2期のプレイヤーの攻撃$a_{pl,1,1}$が成功しているので，発生する確率はその同時確率で$0.4\times0.8\times0.8=0.256$である．

次に，図2と図3はともに，敵e_3が最大HP10の状況のプレイヤーに勝つ戦闘例である．

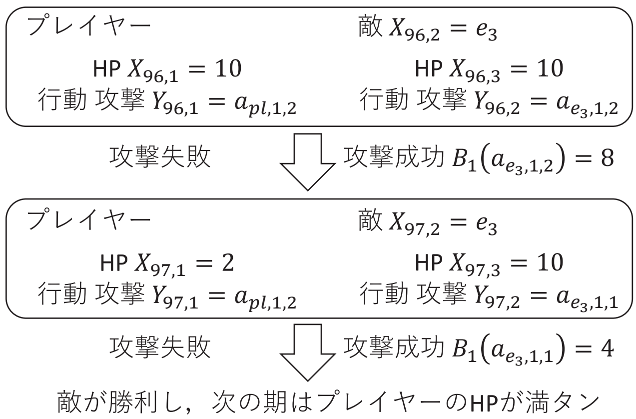

図2.　敵e_3がプレイヤーに最小回数で勝つ例

図2は96期，97期の例であるが，図1のプレイヤー同様に敵e_3が，96期にダメージ量8の攻撃$a_{e_3,1,2}$を選択して成功し，97期にダメージ量4の攻撃$a_{e_3,1,1}$を選択して成功し，プレイヤーに勝っている．2回の行動選択で敵e_3がプレイヤーに勝っているが，2回は敵e_3がHPが最大HPの場合のプレイヤーに勝つ最小回数である．ただし，プレイヤーが先攻で敵は後攻のため，このように敵e_3がプレイヤーに最小回数で勝つためには，図2のようにプレイヤーの攻

撃が失敗する等の敵側の運も必要である．なお，敵e_1がプレイヤーに勝つ最小回数は5回，敵e_2がプレイヤーに勝つ最小回数は3回である．

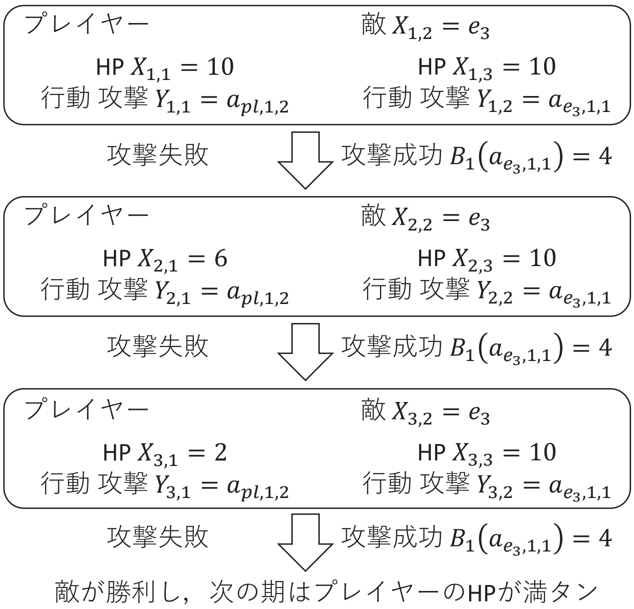

図3.　敵e_3がプレイヤーに勝つ例

図2は96期，97期という対象期間である$T = 100$期間の終わり付近であるが，もっと早期の敵e_3の戦略は異なる．図3が1期から3期の攻撃で敵e_3がプレイヤーに勝つ例である．早期の段階では，後攻の不利な状況ではダメージ量は小さくても成功確率が大きな攻撃で，確実に少しずつプレイヤーにダメージを与える戦略である．本研究では，プレイヤーか敵のどちらかが負けた場合に，勝った側のHPは勝った時点での値のまま次の期（次の戦闘）に引き継がれる設定のため，上記のような敵e_3の戦略の効果があると考えられる．なお，全ての敵で戦略が同じではない．HPや攻撃等の能力設定が敵3種で異なるので，戦略も全ての敵で同じとは限らない．

次に，提案方法と比較対象COMの期待総利得等の数値計算結果を表3に示す．表3では，$T = 100$期間の初期（1期）の敵$X_{1,2}$がそれぞれe_1，e_2，e_3の場合の提案方法における1期のプレイヤーの行動と敵の行動，比較対象COMのプレイヤーの行動を紹介している．$\bar{V}(\cdot,1)$は$T = 100$期間の期待総利得の初期の敵3種に関する平均値である．提案方法では，敵の戦略はプレイヤーの期待総利得の最小化のため，プレイヤーにとっては戦略に関して強敵である．他方，比較対象COMでは，敵の戦略は各行動を等確率で選択する単純なランダム選択なので，プレイヤーにとって提案方法の敵よりも戦略に関して弱い敵である．

表3.　タイプ1の数値計算結果

	提案方法			比較対象COM	
$X_{1,2}$	$\bar{V}(\cdot,1)$	$Y_{1,1}$	$Y_{1,2}$	$\bar{V}(\cdot,1)$	$Y_{1,1}$
e_1	25.33	$a_{pl,1,1}$	$a_{e_1,1,2}$	29.57	$a_{pl,1,1}$
e_2		$a_{pl,1,2}$	$a_{e_2,1,1}$		$a_{pl,1,2}$
e_3		$a_{pl,1,2}$	$a_{e_3,1,1}$		$a_{pl,1,2}$

$X_{1,2}$：初期（1期）の敵

$\bar{V}(\cdot,1)$：100期間の期待総利得の初期の敵に関する平均

$Y_{1,1}$：初期（1期）のプレイヤーの行動

$Y_{1,2}$：初期（1期）の敵の行動

例えば，提案方法に関する図3の3期に敵e_3はプレイヤーに勝つのに十分なダメージ量4で成功確率が0.8と大きな攻撃$a_{e_3,1,1}$を選択する．図3の3期のこの状況では，プレイヤーの攻撃の成否に関係なく，敵e_3は確率0.8でプレイヤーに勝つことができる．しかし，比較対象COMでは等確率で行動を選択するので，同じ状況でも敵e_3がこの期にプレイヤーに勝つ確率は提案方法よりも小さい．このような意味で，提案方法の敵は戦略に関して比較対象COMの敵よりも強いと考えられる．

そのため，期待総利得の平均値$\bar{V}(\cdot,1)$は提案方法よりも比較対象COMの方が大きい．この差分は敵の戦略を提案方法のように高度化することによるゲームの難易度の上昇を示唆していると解釈できる．なお，敵を倒した場合の利得がすべての敵の種類について$G(e_i) = 1$で，プレイヤーが倒された場合の利得が$G(pl) = -1$なので，期待総利得は$T = 100$期間において倒した敵の数からプレイヤーが倒された数を引いた数の期待値に相当する．

表3には1期の行動選択しか掲載していないが，提案方法において，プレイヤーが敵の種類に応じて適応的に攻撃行動を選択していること（プレイヤーの行動$Y_{1,1}$が敵e_1に対しては$a_{pl,1,1}$であるのに対して，敵e_2と敵e_3に対しては$a_{pl,1,2}$であること），敵の種類によってプレイヤーに対する敵の攻撃行動も異なること（敵e_1の行動$Y_{1,2}$が

成功確率の小さな$a_{e_1,1,2}$であるのに対して，敵e_2と敵e_3の行動$Y_{1,2}$は成功確率の大きな$a_{e_2,1,1}$，$a_{e_3,1,1}$であること）が確認できる．

この数値計算例タイプ1では，HP，攻撃のダメージ量，回復量に関して敵e_1が弱い敵，敵e_2が中程度の敵，敵e_3が強い敵の設定で，ちょうど敵e_3の各種設定（HP，攻撃のダメージ量，回復量等）がプレイヤーと同じである．敵の強さについては，前述のHPが最大HPの場合のプレイヤーに勝つ最小回数（敵e_1が5回，敵e_2が3回，敵e_3が2回）から判断することも出来る．次の数値計算例タイプ2ではプレイヤーよりも強い敵を含む場合を紹介する．表2のとおり，攻撃行動と回復行動の成功確率はプレイヤーと敵3種で添え字の末尾の番号が同じ場合には同じ確率とした．よって，プレイヤーや敵の強さはHP，攻撃のダメージ量，回復量に依存する．

5.2　数値計算例タイプ2

数値計算例タイプ2では敵の設定の一部分以外はタイプ1と同様である．タイプ1の敵e_2の設定をプレイヤーよりも強い，攻撃のダメージ量が$B_1\left(a_{e_2,1,1}\right)=6$と$B_1\left(a_{e_2,1,2}\right)=12$，回復量が$B_2\left(a_{e_2,2,1}\right)=6$と$B_2\left(a_{e_2,2,2}\right)=12$，最大HPが$M\left(e_2\right)=12$という設定に変更した．この新しい敵$e_2$がHPが最大HPの場合のプレイヤーに勝つ最小回数は1回である．他方，プレイヤーがHPが最大HPの場合の敵e_2に勝つ最小回数は2回である．このようにHPが最大HPの場合の相手に勝つための最小回数がプレイヤーの方が大きいということからも，敵e_2の設定がプレイヤーの設定よりも強い設定であることが確認できる．タイプ2の数値計算結果を表4に示す．

タイプ1と比較してタイプ2では強力な敵e_2が含まれるため，提案方法と比較対象COMともに期待総利得の平均値$\bar{V}(\cdot,1)$がタイプ1よりも小さくなっており，ゲームの設定としてタイプ2の難易度がタイプ1よりも高くなったと考えられる．敵の戦略がより高度な提案方法におけるプレイヤーの期待総利得の平均値が，敵の戦略が単純なランダム選択である比較対象COMよりも小さいことはタイプ1と同様である．また，提案方法における適応的な行動選択もタイプ1と同様に確認できる．

表4.　タイプ2の数値計算結果

	提案方法			比較対象COM	
$X_{1,2}$	$\bar{V}(\cdot,1)$	$Y_{1,1}$	$Y_{1,2}$	$\bar{V}(\cdot,1)$	$Y_{1,1}$
e_1	11.18	$a_{pl,1,1}$	$a_{e_1,1,2}$	17.29	$a_{pl,1,1}$
e_2		$a_{pl,1,2}$	$a_{e_2,1,2}$		$a_{pl,1,2}$
e_3		$a_{pl,1,2}$	$a_{e_3,1,1}$		$a_{pl,1,2}$

$X_{1,2}$：初期（1期）の敵

$\bar{V}(\cdot,1)$：100期間の期待総利得の初期の敵に関する平均

$Y_{1,1}$：初期（1期）のプレイヤーの行動

$Y_{1,2}$：初期（1期）の敵の行動

5.3　数値計算例タイプ3

数値計算例のタイプ3として，敵がプレイヤーと能力が同じタイプ1及びタイプ2の敵e_3の1種のみの場合を紹介する．タイプ3の数値計算結果を表5に示す．

表5.　タイプ3の数値計算結果

	提案方法			比較対象COM	
$X_{1,2}$	$\bar{V}(\cdot,1)$	$Y_{1,1}$	$Y_{1,2}$	$\bar{V}(\cdot,1)$	$Y_{1,1}$
e_3	6.54	$a_{pl,1,2}$	$a_{e_1,1,2}$	12.48	$a_{pl,1,2}$

$X_{1,2}$：初期（1期）の敵

$\bar{V}(\cdot,1)$：100期間の期待総利得の初期の敵に関する平均

$Y_{1,1}$：初期（1期）のプレイヤーの行動

$Y_{1,2}$：初期（1期）の敵の行動

タイプ1の設定では敵3種の内，2種がプレイヤーよりも弱く，1種がプレイヤーと同能力だった．タイプ2の設定では敵3種の内，1種がプレイヤーよりも強く，1種がプレイヤーと同能力で，残り1種がプレイヤーよりも弱かった．タイプ3の設定ではプレイヤーと同能力の1種のみである．タイプ1，タイプ2，タイプ3を比較しやすいように，各タイプの特徴や数値計算結果のまとめを表6に示す．

表6には，プレイヤーとタイプ1，タイプ2，タイプ3の各敵（表6のタi敵e_jはタイプiの敵e_jを示す）の能力が書かれている．タイプ1の敵e_2とタイプ2の敵e_2の能力設定は異なる．また，タイプ3では敵は1種のみで，タイプ1及びタイプ2の敵e_3と同じ能力設定のため，1種のみであるが敵e_3と表記している．HPは最大HP，攻iはプレイヤーと敵の攻撃$a_{\cdot,1,i}$のダメージ量$B_1\left(a_{\cdot,1,i}\right)$を示す．プ最はプレイヤーがHP満タン（最大HP）の当該敵に勝つ最小回数，敵最は当該敵がHP満タン（最大HP）のプレイヤーに勝つ最小回数，提V，比Vは提案方法，比較対象COMの$\bar{V}(\cdot,1)$（表3，表4，表5掲載の100期間の期待総利得$\bar{V}(\cdot,1)$の再掲）である．

表6.　各タイプの特徴と数値計算結果

	HP	攻1	攻2	プ最	敵最	提V	比V
プレイヤー	10	4	8	−	−	−	−
タ1敵e_1	4	1	2	1	5	25.33	29.57
タ1敵e_2	7	2	4	1	3		
タ1敵e_3	10	4	8	2	2		
タ2敵e_1	4	1	2	1	5	11.18	17.29
タ2敵e_2	12	6	12	2	1		
タ2敵e_3	10	4	8	2	2		
タ3敵e_3	10	4	8	2	2	6.54	12.48

タi敵e_j：タイプiの敵e_j

攻i：攻撃行動$a_{\cdot,1,i}$のダメージ$B_1\left(a_{\cdot,1,i}\right)$

プ最：プレイヤーがHP満タンの敵に勝つ最小回数

敵最：敵がHP満タンのプレイヤーに勝つ最小回数

提V，比V：提案方法と比較対象COMの$\bar{V}(\cdot,1)$

期待総利得が提案方法と比較対象COMともにタイプ3の方がタイプ2よりも小さかったことから，今回の設定のもとでは，同能力の敵1種のみのタイプ3の方がタイプ2よりも難易度が高い設定に相当すると解釈できる．タイプ1，タイプ2，タイプ3の難易度の順番は，表6の相手に勝つ最小回数（プ最，敵最）でも確認できる．例えば，プレイヤーがHPが最大HPの場合の当該敵に勝つ最小回数（プ最）に関する各タイプでの平均値は，各敵の発生確率が等確率なので，タイプ1：$\frac{1}{3}\times 1+\frac{1}{3}\times 1+\frac{1}{3}\times 2=1.33$，タイプ2：$\frac{1}{3}\times 1+\frac{1}{3}\times 2+\frac{1}{3}\times 2=1.67$，タイプ3（敵は1種のみ）：2である．プレイヤーが敵に勝つ最小回数の平均値が小さければ，小さいほど，$T=100$期間中にプレイヤーが勝つ回数が増える可能性があるので，タイプ1，タイプ2，タイプ3の順で難易度が高くなる（タイプ3が最も難易度が高い）ことが確認できる．

同様に，各敵がプレイヤーに勝つ最小回数（敵最）の平均は，タイプ1：$\frac{1}{3}\times 5+\frac{1}{3}\times 3+\frac{1}{3}\times 2=3.33$，タイプ2：$\frac{1}{3}\times 5+\frac{1}{3}\times 1+\frac{1}{3}\times 2=2.67$，タイプ3（敵は1種のみ）：2である．敵がプレイヤーに勝つ最小回数の平均値が大きければ，大きいほど，$T=100$期間中に敵が勝つ回数が減る可能性があるので，タイプ1，タイプ2，タイプ3の順で難易度が高くなる（タイプ3が最も難易度が高い）ことが確認できる．

6.　考察と今後の課題

6.1　考察

従来からロールプレイングゲームにおいて人間の遊び手が操作するプレイヤーと敵が戦う戦闘モードの攻略法に関して検討されている．しかし，従来研究では敵の攻撃（行動）に関して毎回確定的なダメージ，あるいは敵の攻撃は1種のみでダメージ（攻撃の成否）は確率的など，プレイヤーの戦略（行動選択の仕方）を最適化するのに対して敵の戦略は単純な場合がほとんどである．そこで，本研究では敵の戦略にも高度な要素を加味したもとでの攻略法を検討した．

提案方法では，敵はプレイヤーが得る期待総利得の最小化を目的とし，プレイヤーはそのもとで期待総利得を最大化する．具体的には動的計画法を用いてmax-min戦略を算出する．

数値計算例では，提案方法と，敵の行動選択は単純に等確率のランダム選択のもとでプレイヤーの期待総利得を最大化する比較対象を比較した．提案方法と比較対象における敵の戦略を比較すると，提案方法の敵の戦略の方がより高度である．数値計算例では，プレイヤーが得る期待総利得に関して，提案方法が比較対象よりも小さくなることが具体的に確認できた．

また，体力に相当するHPの最大値，攻撃の際に相手に与えるダメージ量や回復の際の回復量がプレイヤーの最大HP，攻撃行動や回復行動よりも大きい敵（プレイヤーよりも強い敵）の有無，プレイヤーよりも弱い敵の有無等の設定が異なる複数タイプの数値計算例を紹介した．提案方法，比較対象ともにプレイヤーよりも弱い

敵が存在すれば期待総利得が大きくなる傾向，強い敵が存在すれば小さくなる傾向を数値計算例において具体的に確認できた．

6.2　今後の課題

本研究では従来研究における問題設定よりも，より高度な戦略を敵が有するRPGの攻略法に関して検討した．数値計算例では，各種設定（最大HP，ダメージ量，回復量）によって弱い敵／強い敵の有無による，プレイヤーが得る期待総利得に関する違いを確認した．

数値計算例を見ると，ゲームの難易度を高くするには，本研究のように敵の戦略を高度化する場合と，敵の各種設定を強く設定する場合の2つのケースが考えられる．ゲームの難易度の調整として，敵戦略の調整と敵の各種設定のどちらがよいのかは数値計算例の期待総利得を見るのみではわからない．

本研究によって，従来研究でも可能だった敵の各種設定によるゲームの難易度調整以外に，敵の戦略の高度化によるゲームの難易度調整の可能性を確認できた．本研究はRPGの攻略法に関する従来研究同様に基礎研究であり，現状ではゲームの難易度調整は実用化のレベルではない．しかし，本研究や今後の本研究の拡張研究の成果などに基づいて，どのような難易度が実際の遊び手（被験者）に好まれるのか，ゲーム進行に従ってどのような難易度の組合せが実際の遊び手（被験者）に好まれるのか等の検討を積み重ねることによって，実用化に近づけることが可能と考える．

例えば，ゲームに慣れていて難易度が高めのゲームを好む人間の遊び手を考える．この遊び手に対して，本研究の提案方法のように敵の戦略が高度化されて難易度が高めのゲームと，敵の戦略は本研究の比較対象COMのような単純な戦略で敵の各種設定を強く設定することによる難易度高めのゲームを提供する．このとき，提案方法と比較対象COMによる期待総利得が同程度であれば，期待総利得に基づく難易度評価では同程度の難易度と考えられる．しかし，実際の人間の遊び手の満足度がどのようになるかは現時点では不明であり，満足度の比較は人間の満足度に関する考察として興味深いと考える．このような被験者（遊び手）の満足度に関する評価実験は今後の課題である．

また，遊び手が何らかのRPGを購入した際（遊び始めた際），当該RPGについて最初は全員が不慣れである．RPGに慣れている／いない人，得意／不得意な人での習熟の個人差はあるが，基本的にゲーム進行に従って全員が徐々に当該RPGに慣れていくと考えられる．この点を考慮すると，本研究のような習熟した遊び手向けの敵の戦略はRPGの後半で採用し，前半は従来どおりの単純な戦略を採用することによって，RPGの後半の難易度を前半よりも高めにするような設定も考えられる．

7. まとめ

RPGの従来研究では，プレイヤーに対する敵の攻撃はダメージ量が確定的，あるいは攻撃の種類が1種類のみなど，敵の戦略は単純である．そこで，本研究ではプレイヤーも敵もともに複数の攻撃行動，回復行動から行動選択可能で，かつ敵もプレイヤーが得る期待総利得を最小化することを目的とする高度な戦略を有する問題設定のもとで，プレイヤーが得る期待総利得を最大化する提案方法を検討した．

従来研究でも，敵の攻撃力などの数値設定次第でゲームの難易度を調整することは可能だったが，本研究では提案方法による敵の戦略の高度化によってゲームの難易度を高くできることを数値計算例によって確認した．

本研究では，敵の戦略としてプレイヤーと同様に動的計画法で期間全体の先読みを実施する高度な戦略を提案した．数値計算例では比較対象として，敵の行動選択が単純な等確率でのランダム選択の場合と比較した．しかし，敵の戦略として目先の短期間の間にプレイヤーを倒す確率の最大化等，本研究の提案方法と比較対象以外の戦略候補もいろいろと考えられる．

本研究は基礎研究であり，議論を簡便にするために各種確率を既知と仮定するなど簡易な問題設定を対象とした．より実際のRPGに近い問題設定への拡張や，いろいろな敵の戦略候補の検討，6.2節で述べた被験者による満足度に関する評価実験等は今後の課題である．なお，確率が未知のマルコフ決定過程に関する機械学習[10]等，従来研究の中には本研究の今後の拡張に活用可能な知見も多いと考える．

また，確率が未知の場合への拡張や，実際のRPGに近づける今後の拡張に際しては，モデルの大規模化・複雑化に伴う計算量の膨大化が予想される．対象とする最適化問題が最適解の算出が難しい規模にまで大きくなった際には，何らかの近似解法の検討が必要になる．近年，囲碁や各種テレビゲームをコンピュータにプレイさせる検討の中で，深層学習の一種である深層強化学習が利用されることが多い．深層強化学習は，未知情報を伴うマルコフ決定過程に対する機械学習の一種である強化学習（Q学習）を深層学習によって近似する機械学習であ

る．よって，確率モデルとしてマルコフ決定過程を採用する本研究の拡張研究においても，近似解法を検討する際には深層強化学習の適用が有力な候補の1つである．

謝辞

本研究の一部はJSPS科研費JP21K04543の助成による．

参考文献

[1] 経済産業省：活況を呈する国内ゲームソフト産業, https://www.meti.go.jp/statistics/toppage/report/minikaisetsu/hitokoto_kako/20210721hitokoto.html, 参照 (2021.12.16)

[2] gamebiz：世界ゲーム市場は初の20兆円の大台突破　国内も2兆円突破、アプリが3分の2占める 『あつ森』が記録的ヒット 『ファミ通ゲーム白書 2021』発刊, https://gamebiz.jp/news/300622, 参照(2021.12.16)

[3] 藤井叙人, 片寄晴弘：戦略型トレーディングカードゲームのための戦略獲得手法, 情報処理学会論文誌, Vol.50, No.12, pp.2796-2806, 2009.

[4] 高木幸一郎, 雨宮真人：ロールプレイングゲーム (RPG) の戦闘におけるバランス自動調整システム開発のための基礎的考察, 情報処理学会研究報告GI, Vol.2001, No.28, pp.31-38, 2001.

[5] 高木幸一郎, 雨宮真人：ロールプレイングゲーム (RPG) のバランスとは何か：分析およびその調整に関する提案, 情報処理学会研究報告GI, Vol.2001, No.58, pp.67-74, 2001.

[6] 前田康成, 後藤文太朗, 升井洋志, 桝井文人, 鈴木正清：マルコフ決定過程のロールプレイングゲームへの適用, 情報処理学会論文誌, Vol.53, No.6, pp.1608-1616, 2012.

[7] 前田康成, 後藤文太朗, 升井洋志, 桝井文人, 鈴木正清, 松嶋敏泰：ノンプレイヤキャラクタを伴うロールプレイングゲームの攻略法に関する一考察, 電子情報通信学会論文誌A, Vol.96, No.8, pp.572-581, 2013.

[8] 前田康成, 後藤文太朗, 升井洋志, 桝井文人, 鈴木正清, 松嶋敏泰：マルコフ決定過程で表現されたロールプレイングゲームにおける攻略法の能動学習, バイオメディカル・ファジィ・システム学会誌, Vol.15, No.1, pp.69-81, 2013.

[9] 森村英典, 高橋幸雄：マルコフ解析, 日科技連, 東京, 1979.

[10] J.J. Martin : Bayesian Decision Problems and Markov Chains, John Wiley & Sons, 1967.

前田康成（まえだやすなり）

平成7年早大・理工卒．平成9年同大学院理工学研究科修士課程修了．日本電信電話（株），東日本電信電話（株），北見工大助手，助教，准教授を経て平成28年同大学教授，現在に至る．博士（工学）．統計的決定理論の学習問題への応用に関する研究に従事．電子情報通信学会等各会員．