JAIST Repository
https://dspace.jaist.ac.jp/

Title	キャラクター設定に沿った振る舞いをするゲームAI
Author(s)	三上, 麟太郎
Citation	
Issue Date	2024-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/19037
Rights	
Description	Supervisor: 池田 心, 先端科学技術研究科, 修士(融合科学)

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

修士論文

キャラクター設定に沿った
振る舞いをするゲームAI

三上　麟太郎

主指導教員 池田　心

北陸先端科学技術大学院大学
融合科学共同専攻
（融合科学）

令和6年3月

キャラクター設定に沿った振る舞いをするゲーム AI
(Game AI that behaves according to character settings)

北陸先端科学技術大学院大学　学生番号 2250007
氏名　三上　麟太郎
主任研究指導教員氏名　池田　心

1. はじめに

近年，人工知能 (AI) は多岐にわたる分野で研究，活用されており，その進化も著しいものとなっている．ゲームにおいても，AI を用いたコンピュータプログラムが作成されている．それらは，チェスや囲碁，コンピューターゲームなど様々なゲームにおいて人間のトッププレイヤに勝利しており，その強さは人間を大きく超えるものであると言える．

このように，AI の「強さ」が十分な水準に達して以降，AI に「強さ」とは別の「人間らしさ」を表現させる研究も注目されるようになった．藤井ら [1] は，アクションゲームである「マリオブラザーズ」をプレイする AI に対して，学習時に生物学的制約を加えることで，AI に対し，入力に遅れや揺らぎがあることを前提とした，人間らしい安全な攻略法の学習を行わせた．ジェイコブら [2] は，FPS（一人称視点シューティングゲーム）の「Unreal Tournament 2004」にて，NeuroEvolution と模倣学習を用いることで，最も人間の対戦相手らしいと思われる AI の作成に成功した．

一方で，ゲームの AI に求められることは必ずしも「人間以上の強さを持つ対戦相手であること」か「人間プレイヤらしい振る舞いをすること」のみではない．様々なキャラクターが登場するゲームでは，AI が「ゲームの勝敗に関係が無くても，キャラクター設定に沿った振る舞いをすること」によって，プレイヤはリアルな世界観を体験することが可能となり，結果としてより楽しんでゲームをプレイできるようになると考える．

本研究では，キャラクターの作成に強化学習を用いる．ここで，シチュエーションと性格の組み合わせごとに報酬を用意しなくてもいいように，報酬に「精神症傾向」のような性質を関連付け，それとは独立に，性格ごとに各性質をどれくらい重要視するかの倍率を決定する．この倍率について，性格ごとに共通の値を用いることで,，ある程度報酬設計を使いまわしつつキャラクター表現ができるはずである．これにより，複雑なルール下であっても，機械が複数種類のキャラクターらしい行動を自動獲得することを目指す．

2. 研究方法

研究には Unity Technologies が公開しているゲームエンジンである，Unity ML-Agents[3] を用いる．Unity ML-Agents とは，同社が提供している Unity 上で強化学習ができるようになるオープンソースのライブラリである．

本研究の目的を達成するために，環境と報酬に応じて振舞いを学習する強化学習を用いる．学習は環境とキャラクターのペアごとに行う必要があるが，報酬は環境ごと，およびキャラクターごとに独立に行うことで，省力化を図る．

RPG をはじめとするゲームにはさまざまなシチュエーションが登場する．また，同じシチュエーションであっても登場するキャラクターの性格は異なっているかもしれない．逆に同じ性格のキャラクターが多くのシチュエーションに登場するかもしれない．これらのゲーム制作の要請に対し，提案手法では，シチュエーション（環境，状況）S ごとに，およびそれとは独立にキャラクター設定 C ごとに，報酬を設計する．そして，シチュエーションと登場キャラクターの組み合わせ $S \times C$ ごとに強化学習を行う．これにより，組み合わせごとに挙動を人手で定めるよりも簡単に，状況とキャラクタ設定に応じた挙動を獲得することを狙う．

3. 結果と考察

本研究では,「騎士 3 人が王様をドラゴン，暗殺者から守る」というシチュエーションを考える．ベース設定として「騎士 3 名のうち，勇敢な騎士が 2 名，臆病な騎士が 1 名，ドラゴンの HP が 4」という設定で騎士キャラクターを作成する実験を行った

その結果，勇敢な騎士は,「自分の命を犠牲にして，王様を守ろうとする」という行動が見られた．臆病な騎士は,「自分が死なないように，しかし王様も守ることができるように，攻撃への参加を最低限にとどめる」という行動が見られた．これらはどちらも，騎士であることをベースに，勇敢な性格らしい行動，臆病な性格らしい行動であるといえる．

また，シチュエーションを変更し,「勇敢な騎士 2 名，臆病な騎士 1 名，敵であるドラゴンの HP が 3」となった場合には，敵の HP が減っており，純粋に簡単な環境になったといえる．この環境において，臆病な騎士は「自分が攻撃しなくても王様を守れるので，自分がしないように逃げる」という行動を選択した．

「勇敢な騎士 2 名，普通の騎士 1 名」というシチュエーションでは，臆病な騎士の代わりに普通の騎士がチームに入っている．この環境において，普通の騎士は臆病な騎士よりも積極的に攻撃を行い，王様を守ることに成功することが多くなった．また，逃げ回ることのある臆病な騎士と比較して，普通の騎士は王様を守るために王様の周辺に向かうことが多かった．そのため，勇敢な騎士は王様の周辺の警戒を普通の騎士に任せ，ドラゴンの対処に向かうという明確な役割分担を行う様子も見られた．

本実験では，合計で 7 つのシチュエーションを取り扱い，シチュエーションが異なる場合であっても，そのシチュエーションに沿ったキャラクターらしさを表現することができた．しかし，すべてのシチュエーションで想定した通りの共同が実現できたわけではなく,「行動に他の性格との違いが表れていない」など，課題も残った．

4. まとめ

本研究では，ゲームに登場するキャラクターが，各自のキャラクター設定に沿った振る舞いを自動獲得すること目的にキャラクター AI の作成を行った．作成した AI は,「王様を騎士 3 人がドラゴンと暗殺者から守る」というシチュエーションに限定したものではあるが，自身の性格といった設定に応じて，キャラクターらしい行動ができるようになっていた．また，シチュエーションを変えた場合に，シチュエーションに合わせたキャラクターらしい行動の獲得に成功した．一方で，一部のシチュエーションではキャラクターらしさが十分に表現できてないという課題も残っている．

また，今回の研究環境は戦闘シーンにのみ注目したものであるが，日常生活シーンなど，それ以外の場面に対しても，この手法を発展させることで，キャラクターらしい行動の自動獲得が見込めると考えている．今後の課題として，キャラクター表現ができなかったキャラクターについて，報酬設計やシチュエーション設定の見直しを行っていくほか，戦闘以外のシチュエーションでキャラクター表現が可能になることについても目指していきたい．

参考文献（最大 5 件）

[1] 藤井叙人, et al. ”生物学的制約の導入による 「人間らしい」 振る舞いを伴うゲーム AI の自律的獲得.” ゲームプログラミングワークショップ 2013 論文集 (2013): 73-80.

[2] Schrum, Jacob, Igor V. Karpov, and Risto Miikkulainen. ”UT^2: Human-like behavior via neuroevolution of combat behavior and replay of human traces.” 2011 IEEE Conference on Computational Intelligence and Games (CIG’11). IEEE, 2011. [3] A. B and C. D, Journal name and so on.

[3] Juliani, Arthur, et al. ”Unity: A general platform for intelligent agents.” arXiv preprint arXiv:1809.02627 (2018).

目 次

第 1 章　はじめに　1

第 2 章　研究背景　3
2.1　ゲームにおける AI 3
2.1.1　デジタルゲームにおける敵 NPC 3
2.1.2　デジタルゲームにおける味方 NPC 4
2.1.3　デジタルゲームにおける中立 NPC 5
2.2　NPC とその個性 5

第 3 章　関連研究　7
3.1　人間らしい振る舞いをするゲーム AI 7
3.2　性格に沿った振る舞いをするゲーム AI 9
3.3　ビッグファイブ理論 11

第 4 章　提案手法　12
4.1　職業の表現 12
4.2　性格の表現 13

第 5 章　研究環境　15
5.1　Unity ML-Agents 15
5.1.1　DungeonEscape 15
5.2　実際に用いる環境 16
5.2.1　ゲーム全体のルールに関する変更点 16
5.2.2　ゲームに登場するキャラクターに関する変更点 17

第 6 章　騎士キャラクターの作成　19
6.1　騎士キャラクターの報酬設計 20
6.1.1　職業としての騎士 20
6.1.2　用意した性格 20
6.2　ベース設定での結果 22
6.2.1　学習曲線 23
6.2.2　テスト結果：チームの勝敗 23
6.2.3　テスト結果：チームの与えたダメージ 24

6.2.4　テスト結果：各騎士の序盤の行動 25
6.2.5　テスト結果：各騎士のゲーム終了時の状態 26
6.2.6　テスト結果：各騎士の攻撃状況 27
6.2.7　テスト結果：目視 27
6.3　バリエーションの設定と比較の全体像 27
6.3.1　キャラクター設定ごとの比較 28
6.3.2　敵の HP の値の変化 29
6.3.3　臆病な騎士と普通の騎士 30
6.3.4　勇敢な騎士と普通の騎士 31
6.3.5　臆病な騎士と英雄気取りな騎士 32
6.3.6　騎士の HP の値の変化 32
6.3.7　普通の騎士と英雄気取りな騎士 32
6.4　比較結果のまとめ 33

第 7 章　多様なシチュエーションでの利用　34

第 8 章　おわりに　36

付 録 A　その他統計データ　37
A.1　ベース設定でのテスト結果：各騎士の攻撃状況 37

図 目 次

3.1 生物学的制約の有無によるマリオプレイ AI の行動の差 8
3.2 坂田らの提案する大富豪プレイヤ AI の全体像 10

4.1 本研究で提案する手法の全体像 14

5.1 DungeonEscape の環境 [17] 16

6.1 騎士の獲得合計報酬に関する学習曲線 23
6.2 実験のバリエーションと比較関係 28

表 目 次

3.1 坂田らの提案する性格や能力，およびその対応する行動 10

5.1 DungeonEscape からのルール変更点 17
5.2 DungeonEscape からの構成キャラクター変更点 18

6.1 本研究における MA-POCA のハイパーパラメータ 19
6.2 本研究で作成したキャラクターの性格と報酬の倍率 21
6.3 実験における各キャラクターの HP 22
6.4 学習完了モデルのテスト時におけるチームの勝敗 24
6.5 学習完了モデルのテスト時におけるチームの与ダメージ 24
6.6 学習完了モデルのテスト時における騎士の序盤の行動 25
6.7 学習完了モデルのテスト時における各騎士の状態 26
6.8 各設定ごとの勝率 29
6.9 各設定ごとの騎士が敵に与えたダメージ（「接触した瞬間に与えたダメージ」＋「接触している間に与えたダメージ」） 30
6.10 各設定ごとの騎士の序盤の行動 31

A.1 学習完了モデルのテスト時における各騎士の攻撃状況 37

第1章　はじめに

近年，人工知能 (AI) は多岐にわたる分野で研究，活用されており，その進化も著しいものとなっている．ゲームにおいても，AI を用いた強いコンピュータプログラムが作成されている．チェスでは，1997 年に Deep Blue[1] が，当時の世界王者であるガルリ・カスパロフに勝利した．その後，戦略が複雑なことから AI が人間に勝利するのは難しいと言われていた囲碁では，2016 年に AlphaGo[2] がプロ棋士のイ・セドルに勝利した．大規模な探索空間を有し，より複雑なルールとリアルタイム性を持つことから，人間に勝利することが更に困難であると言われていたデジタルゲームの「StarCraft2」でも，2018 年に AlphaStar[3] がプロプレイヤである Grzegorz Kominc に勝利した．このように，様々なゲームにおいて AI は人間のトッププレイヤに勝利しており，その強さは人間を大きく超えるものであると言える．

このように，AI の「強さ」が十分な水準に達して以降，AI に「強さ」とは別の「人間らしさ」を表現させる研究も注目されるようになった．藤井らは，アクションゲームをプレイする AI に対して，学習時に生物学的制約を加えることで，AI に対し，入力に遅れや揺らぎがあることを前提とした安全な攻略法の学習を行わせた．その結果，非常に人間らしく見える AI の作成に成功している [4]．Schrum らは，ゲーム AI に関するチューリングテストである BotPrize において，AI に NeuroEvolution と模倣学習を取り入れることで，最も人間の対戦相手らしいと思われる AI の作成に成功した [5].

一方で，ゲームの AI に求められることは必ずしも「人間以上の強さを持つ対戦相手であること」か「人間プレイヤらしい振る舞いをすること」のみではない．カードゲームやボードゲームと比較して，デジタルゲームの多くには,「敵でも味方でもない一般市民」や,「非常に弱い敵」,「特殊な性格を持つ味方」など，様々なキャラクター設定を持つキャラクターが登場する．その中で，AI が「ゲームの勝敗に関係が無くても，キャラクター設定に沿った振る舞いをすること」によって，プレイヤはリアルな世界観を体験することが可能となり，結果としてより楽しんでゲームをプレイできるようになると考える．

現状では，このようなキャラクター表現は主にルールベースを用いて行われていると予想するが，ゲームが複雑になると，シチュエーションごとに取らせたい行動や，キャラクターの種類が膨大になり，ルールベースで網羅することが困難になってしまう．本研究では，強化学習を用いることで複雑なルール下であっても，シチュエーションごとの報酬と，キャラクターの性格による報酬の補正さえ

考えれば，以降は人が手を加えなくても機械が複数種類のキャラクターらしい行動を獲得することを目指す．

キャラクター設定の具体的な表現方法は以下のとおりである．まず，人間が普遍的に抱きうる「勝負ごとに勝ちたい」,「自分が死にたくない」といった感情を共通の報酬とする．そこにキャラクターの職業や，キャラクターが行動するシチュエーションに応じた報酬を加える．ここは，ゲームデザイナーが目的に応じて調整する必要がある．次に,「臆病」「忠誠心が高い」といったキャラクターごとの性格ごとに，各報酬に対して補正を加える．これらの補正については，できるだけシチュエーションに依存せず，キャラクターごとに一度決定すればよいものを目指す．これにより，騎士という同じ職業の中にも，自分の命を犠牲にして王様を守ろうとする洗練された者もいれば，戦って死ぬことを恐れる臆病者もいるというキャラクター設定が表現できるようになると考える．

本論文における構成は以下のようになっている．2 章では，ゲーム AI とそのキャラクター表現の背景について述べる．3 章では，ゲーム AI に人間らしい行動をさせる先行研究について述べる．4 章では，研究環境について述べる．5 章では，提案手法について説明する．6 章では，特定の性格らしい騎士キャラクターの作成実験についての結果と考察について述べる．7 章では，提案手法を用いて山賊を作成する方法と，それにより期待される結果について述べる．最後に，8 章で総括と今後の展望について述べる．

第2章　研究背景

2.1　ゲームにおける AI

本研究は，デジタルゲームにおける AI の行動に関するものである．デジタルゲームでは，AI はプレイヤが操作しないすべてのキャラクターを操作しており，NPC(Non-Player Character) と呼ばれることもある．

デジタルゲームには「敵キャラクター」がいれば,「味方キャラクター」もいて，さらには「敵でも味方でもない一般市民キャラクター」さえもいることがある．当然ながら，敵キャラクターに対して求められる役割と，味方たキャラクターに対して求められる役割は違うほか，同じ敵キャラクターの中でも，別の役割を持ってほしいということがありえる．

本章では，デジタルゲームに登場する各 NPC について，デジタルゲームにおける立ち位置と，求められる内容を述べる．

2.1.1　デジタルゲームにおける敵 NPC

デジタルゲームにおいて敵 NPC とは，スポーツゲームや格闘ゲーム，アクションゲームなど，プレイヤが対戦を楽しむゲームにおいて，競い合う相手であり，デジタルゲーム体験の中核を担っているともいえる．そのため，様々な表現が求められる．

敵 NPC に限らず，ゲーム内の人間や人間に近い存在を表現しようとすれば，そこには何らかの意味での人間らしさがあることが望ましい．分かりやすいものとしては，見た目，セリフ，物理的性質などである．そしてゲームに登場して何らかの意味でプレイヤがその活動を目にするのであれば，挙動が人間らしいことが求められる場合が多い．例えば町人がずっと同じところに立っていたり，特定の場所をいったりきたりすることは多いが，これは人間らしいとは言えない．勝負の相手をする敵 NPC が愚かすぎても，機械のように正確でも，楽しませることはできない．本節では以下，主に強さによって場合分けを行い，敵がプレイヤより強く設定されている場合，弱く設定されている場合，ほぼ同じに設定されている場合について，求められることを述べる．

プレイヤより強い敵であれば，例えば瞬間移動や連続攻撃などといった，強い行動を積極的に用いることで，プレイヤに対し乗り越える壁になることが求めら

れる．そのうえで，全体的な行動がワンパターンにならないよう，牽制や様子見などを織り交ぜたり，プレイヤ側の反撃の足掛かりとなるような，スキの大きな行動を行ったりすることも求められる．

プレイヤより弱い敵であれば，何もしないわけではないものの，プレイヤがNPCの行動への対処が容易であることが求められる．特にゲーム序盤は，プレイヤの操作するキャラクターの性能や，プレイヤ自身のゲームへの練度が低く，強敵に全く歯が立たないことも多い．そのような状況でも，プレイヤが戦って勝利する楽しさを感じられるよう，行動の幅が狭く，わかりやすい行動を多用することが望ましい．

プレイヤと同格の敵であれば，強い敵ほど勝つのが難しい相手ではないが，弱い敵ほど簡単に勝てる相手でもないというバランスが求められる．例えば，弱い敵であれば攻撃しかしてこなかったが，同格の敵はそれに加えて防御姿勢をとったり，プレイヤの攻撃を回避しようとしてくるなど，強い敵ほどではないが，行動選択の幅が少し広くなるということである．プレイヤに，行動が単調でなくなった相手に挑むスリルを与えたり，強い敵と戦うためのスキルアップを促したりすることが求められる．特に，人間であれば，戦うシチュエーションを考えたときに，相手が遠くにいるなら弓で攻撃をし，近づいてきたなら剣での攻撃に切り替えるといったような，道具をシチュエーションに最も合うように使うことが求められる．また，相手が遠距離攻撃をしてくるならば，遮蔽物を用いて回避することを試みたり，相手が剣で攻撃してくるならば，特に攻撃をためるモーションに注意して防御を行ったりするなど，多彩な道具を使われた場合の対処も求められる．

また，特にロールプレイングゲーム (RPG) においては，人間のみならず，野生動物や，ドラゴンをはじめとした架空の生物が敵として登場することもある．そのような，生物として異なることによる表現の差異もまた求められるものである．例えば，野生動物には，縄張り意識を持っているものも多い．そのため，目についたものを手当たり次第に攻撃するのではなく，縄張りに足を踏み入れようとした対象を優先して狙うといった表現が望ましい．一方で，ドラゴンは「大きな爪とうろこ状の皮膚，翼を有しており，口から火を吐くこともある」と表現されることも多い [6] 強力な存在であるため，その身体的特徴や，火を吐ける能力を生かして傍若無人にふるまうことが望ましい．

2.1.2 デジタルゲームにおける味方NPC

デジタルゲームにおいて味方NPCとは，プレイヤと実際に肩を並べて戦ったり，行動の補助を行ったりする存在である．味方NPCとの協力自体を楽しむこともあれば，味方NPCの存在によって強敵に勝利できたという喜びを感じられることもあり，ゲームの面白さと密接な関係があると考えている．

味方NPCに求められるものとして考えることができるのは，まず敵NPC同様に強さや．生物的な差異である．強い味方NPCには，プレイヤと協力する中で，

ゲームのクリアや，戦況を有利にすることにつながる行動をすることが望まれ，弱い味方 NPC には，何もしないわけではないが，プレイヤ側が，NPC をフォローしたいと思えるような弱さを演じることが望まれる．もちろん，動物や神話生物が味方 NPC としてゲームもあり，そのような場面では，自身の特徴を生かした行動をすることが望ましい．

また，味方との関係性によっても求められる表現が変わることがある．例えば，三つ巴の戦いの中で，1 つのチームだけが圧倒的に優勢な場合，残る 2 つのチームはルール上は敵であるが，協力して戦うことが予想される．そのような場面では，共闘や補助行動を行いつつも，どこかのタイミングで自チームだけが優勢に転じられるように，共闘をしている間から牽制を入れたり，最終的に裏切る行為といった表現の変化が求められる．また，ゲームによっては同士討ちシステムがあるものもあり，そのような環境では，基本的には味方だが，同士討ちを受けた場合には，怒って敵対的な行動をとるようになるといった表現の変化も求められる．

2.1.3 デジタルゲームにおける中立 NPC

デジタルゲームにおいて中立 NPC とは，敵でも味方でもない，例えば「町で歩いている一般市民のキャラクター」などが該当する．敵 NPC や味方 NPC とは異なり，中立 NPC の行動はプレイヤに直接働きかけるものではないことが多いため，中立 NPC の行動がゲームの面白さと直接つながるということは少ない．しかし，中立 NPC との会話や，中立 NPC が存在することそのものがゲームへの没入感を増加させ，結果としてより楽しいデジタルゲーム体験ができると考えている．

中立 NPC には，ゲームへの没入感の増加のために，一般市民役であるならば一般市民として自然なふるまいが求められる．例えば，現実世界で家を出て近くにある場所に向かうとき，多くの人は最短の直線距離を通るのではなく，なるべく舗装された道を歩いて向かうだろう．そのような，キャラクターが日常生活を送る中でしそうなことを，そのまま表現できることが望ましい．

2.2 NPC とその個性

デジタルゲームの NPC には，敵，味方，中立といった役割があり，その中でも強さによってさらに求められる表現が変化することを述べたが，それだけで十分ではない．NPC ひとりひとりに注目してみると，彼らはデジタルゲームの中の世界で実際に生きている，一個人であると考えることができる．そう考えると，彼らには職業があり，その職業に沿った振る舞いをするべきであると考えた．また，職業が同じキャラクターでも，一人一人性格は異なっており，その性格に沿った振る舞いをするはずであると考えた．しかし，デジタルゲームの中には，それらの個性による表現がされていないものも存在する．

Bethesda Softworks LLC が提供しているアクション PRG である The Elder Scrolls V Skyrim[7] では，山賊長という NPC と，山賊という NPC がいる．山賊長は山賊たちをまとめる，力のあるリーダーであり，戦闘中の行動も熟練していることが好ましいが，実際には山賊長は HP 等のステータスが高いだけで，山賊と戦闘中の行動は変わらない．

Bethesda Softworks LLC が提供しているアクション PRG である Fallout4[8] では，プレイヤと同行できる味方 NPC に対し，キャラクターごとに好きな行為と嫌いな行為が設定されている．例えば，物事を平和的に解決しようとする行動に対して，肯定的な反応を示すキャラクターと否定的な反応を示すキャラクターがいる．しかし，そのような好きな行為や嫌いな行為の設定は，味方 NPC のとる動作には反映されていない．平和的解決を例に挙げるならば，これが好きなキャラクターは，敵と戦うときに，敵を倒しきるのではなく，与えるダメージを調整して無力化にとどめることが好ましい．一方で，平和的解決が嫌いなキャラクターは，積極的に攻撃して相手を倒しきることが好ましい．実際には，すべての味方 NPC が敵を倒しきることを目的として戦闘を行ってしまっており，性格が行動に反映されていないと考えた．

CD Projekt S.A. が提供しているアクション RPG である Cyberpunk 2077[9] では，銃撃戦が始まると周辺の市民 NPC がその場から逃げ出すようになっている．しかしこのとき，ほかの逃走経路があるのにもかかわらず，戦っている主人公と敵の間を通って逃げようとする市民 NPC がいることがある．銃撃戦が行われているのであれば，主人公と敵の間は特に危険な場所であり，普通ならば近づくことはほとんどないため，市民 NPC の行動はキャラクター設定に則っていないと考えた．

そのため，デジタルゲームで，NPC に対してよりキャラクター設定に沿った振る舞いをさせたいと考えた．

第3章　関連研究

本章では，本研究の関連研究として，人間らしい振る舞いをするゲーム AI の研究について述べる．次に，性格に沿った振る舞いをするゲーム AI の研究について述べる．最後に，心理学領域における，人間の性格の分類に関する研究である，ビッグファイブ理論について述べる．

3.1　人間らしい振る舞いをするゲーム AI

本研究では,「職業や性格に沿った振る舞い」を AI プレイヤに行わせたい．一方で，そもそも AI プレイヤに「人間らしい振る舞い」を行わせること自体，ゲームによっては容易ではない．なぜなら，AI プレイヤの行動決定アルゴリズムと人間の行動決定法は大きく異なり，また，人間には遅れやゆらぎ，好みや感情など様々な人間らしさがあるためである．

人間らしさを演出するためのアプローチは，教師あり学習，制約付き強化学習，人間のプレイとの挙動差異の最小化など様々なものが存在するが，本節ではその一部を紹介する．

Schrum らは，FPS（一人称視点シューティングゲーム）において，人間らしい振る舞いをするビヘイビアベースのゲーム AI として，UT^2[5] を提案した．UT^2 は，対象のゲームにおける人間プレイヤの行動データを模倣することで，目的地へ移動するときの自然なルート作成や，ダッシュ，ジャンプ操作を獲得した．また，NeuroEvolution を用いることで，ビヘイビアツリーを最適化を行った．これにより,「自身の後ろに壁があるならば，後退しない」や,「相手が自分のロケットランチャーの射程内にいるならば,（自爆しないために）それ以上接近しない」といった戦闘中の行動を獲得した．これらを用いたことで，UT^2 は，2012 年に，人間らしい AI プレイヤについてのコンテストである 2K BotPrize にて，人間プレイヤよりも人間らしいと評価されることに成功している．

藤井らは，Q 学習と A*アルゴリズムを用いてマリオをプレイする AI を作成する際，学習時に追加で生物学的制約を加えた [4]．生物学的制約を加えると，制約のない状態でできていた「障害物ををぎりぎりの位置から飛び越える」「細かい足場の移動を連続で行う」といった高い精度を求められる行動がミスにつながりやすくなる．これにより，AI は制約が存在する前提で，安全な方策を学習し，人間らしいプレイが可能となった．生物学的制約とは，以下のようなものである．

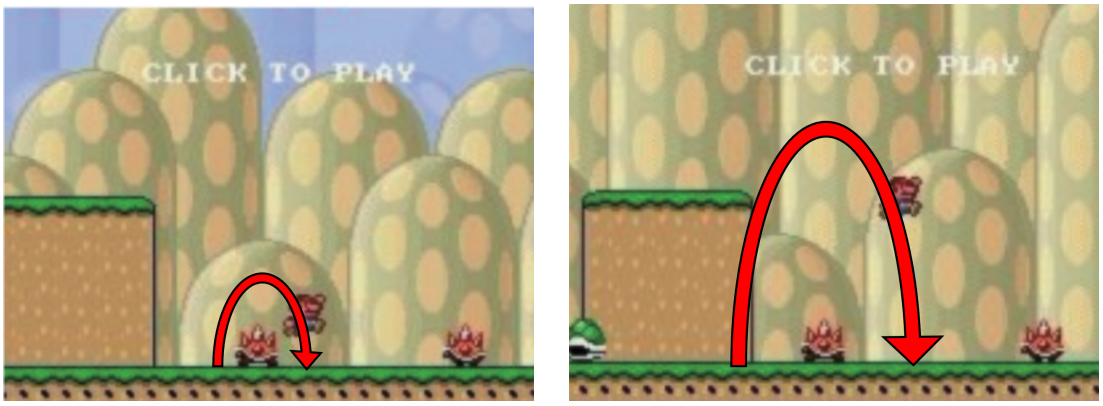
- 『ゆらぎ』: 人間プレイヤは画面内のオブジェクトの位置情報を，正確に把握し続けることは困難であり，見間違いが発生する．そのため，学習時の入力情報にガウスノイズを加えることで，この認識のゆらぎを再現する．

- 『遅れ』: 人間プレイヤは，画面に映る情報を取得してから，実際に操作を行うまでに遅れが発生する．そのため，学習時の入力とする情報を数百ミリ秒過去のものにすることで，この認識と操作の間の遅れを再現する．

- 『疲れ』: 人間プレイヤは,「ボタン連打」といった，短時間にキー操作が連続することや,「単調動作の繰り返し」といった，長時間操作を続けることで，疲れが生じる．そのため，学習時にキー操作を変更するたびに負の報酬を与えることで，この操作に対する疲れを再現する．

- 『訓練と挑戦のバランス』: 人間プレイヤは，成功している局面では，操作を安定させるために同じ行動を繰り返して訓練をする．反対に，失敗している局面では，失敗や飽きの解消のために，新しい行動に挑戦する．そのため，失敗が少ない場面ではこれまでと同じ行動をとりやすく，逆に失敗の多い場面ではこれまでと異なる動作を行うようにすることで，訓練と挑戦のバランスを再現する．

図 3.1 に生物学的制約を用いないマリオプレイ AI と，生物学的制約を用いたマリオプレイ AI の行動の差異を示す．



(a) 生物学的制約を用いない AI　　**(b)** 生物学的制約を用いた AI

図 3.1: 生物学的制約の有無によるマリオプレイ AI の行動の差

生物学的制約を用いていない，図 3.1a では，敵を最低限のジャンプで飛び越えているが，生物学的制約を用いた，図 3.1b では，遠くから大きなジャンプをして飛び越えることを選んでいる．これは，生物学的制約があることによって，より安全に障害物を超えることを学習することができたためである．そして，被験者実験では，これらを用いて作成したゲーム AI によるプレイ映像が，人間プレイヤのものよりも人間らしいという評価を受けることに成功している．

池田らは，マリオをプレイする AI に対し，特定の人間プレイヤらしい行動をさせる研究を行った [10]．池田らは，特定の人間プレイヤらしい行動とは，異なるステージでも共通して表れるものであるとし，関係する統計量を 6 つ用意した．そのうえで，それらの統計量のうち，人間の平均値と多く差がついているものが，模倣対象らしさを演出するための，強調すべき特徴であると決定した．学習には遺伝的アルゴリズムを用いており，特徴に関しては模倣対象の値よりもさらに平均値から離れた値を目指すように，特徴でない部分は人間の平均値を目指すようにボーナスとペナルティが調整されている．これらの結果として，模倣対象に各統計量の値が近い，特定の人間プレイヤらしいといえる行動の獲得に成功している．

McIlroy-Young らは，人間らしいふるまいをするチェス AI として，Maia を開発した [11]．Maia は AlphaZero[2] 的な深層学習のフレームワークを採用しているが，自己対局による棋譜の代わりに，人間の棋譜を学習している．McIlroy-Young らは人間の棋譜をレーティング（棋力を数値化したもの）ごとに 9 つのグループに分け，それぞれのレーティングに対応するように 9 つの異なる Maia を訓練した．例えば，Maia 1100 はプレイヤのレーティングが 1100 台の棋譜を 1200 万棋譜学習している．

学習の結果，低いレーティングで学習した Maia は低いレーティングのプレイヤの指し手をよく当てることができ，高いレーティングで学習した Maia は高いレーティンのプレイヤの指し手をよく当てることができた．このように棋力帯ごとのプレイヤの傾向をある程度捉えた Maia は，棋譜内の人間の指し手との一致率が 50 %を超えることができた．これは，α β 探索を浅く行う既存のチェス AI や，学習途中の AlphaZero 型のチェス AI の一致率約 40 %よりも高かった．

3.2　性格に沿った振る舞いをするゲーム AI

前節では生物学的制約を取り入れたり，人間のプレイを模倣したりして人間らしい AI を作成した先行研究について述べたが，人間をもっと細かく見ると，性格や能力などが 1 人 1 人異なっており，それらに基づいた個性がある．人間の代わりに AI が対戦相手になるゲームや，細かい設定を持つキャラクターが登場するゲームにおける AI に対しても，そのような個性があることが望ましい．

坂田ら [12] は，トランプゲームの大富豪をプレイする AI に対して，ルールベースを用いて能力や性格といった個性に応じて行動を変化させる研究を行った．坂田らは，AI プレイヤーに対し,「記憶力」といった能力や,「負けず嫌い」といった性格を合計で 5 つ設定した．表 3.1 に坂田らの研究で用いられた能力と性格，およびそれらに対応する行動を示す．

その 5 つの能力や性格をもとに，AI プレイヤーは手札のカードの選択確率を計算する．最終的に，選択確率を高い順にソートし，カードを場に出すか判断することで，AI のプレイに能力や性格が反映されるとしている．図 3.2 に坂田の提案手法の全体像を示す．

表 3.1: 坂田らの提案する性格や能力，およびその対応する行動

性格や能力	対応する行動．一部注釈をつけた
記憶力 （三上註：忘れっぽい）	場に出たカードを記憶する （三上註：場に出たカードを忘れやすい）
リスク志向	強いカードを出しやすい 相手が上がりそうなとき，阻止する
気まぐれ	適当にカードを出しやすい
負けず嫌い	一定以上の順位を目指す
焦りやすい	ピンチに強いカードを出しやすい

この手法によって作られた大富豪プレイヤ AI は人間の目から見て,「記憶力」による能力の差について，伝わりやすい表現ができていた．一方で,「焦りやすい」という性格に関しては，全体的に想定と異なる伝わり方をしており，課題の残る結果となっている．

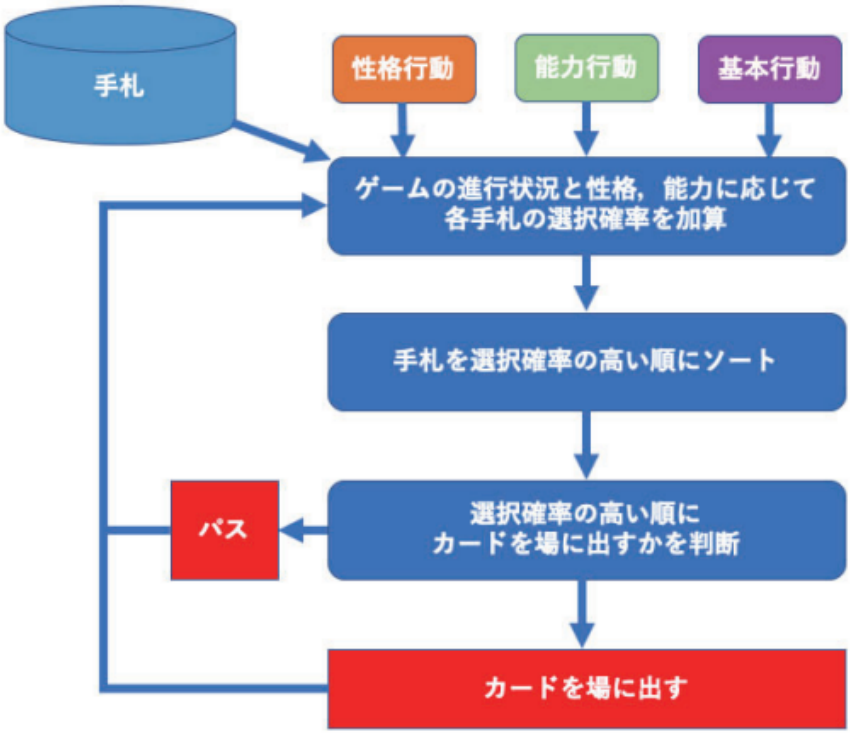


図 3.2: 坂田らの提案する大富豪プレイヤ AI の全体像

3.3　ビッグファイブ理論

本研究では，キャラクターの性格に基づいて，強化学習で与えられる報酬の量を調整する．それらの設定方法の一つの参考として，人間の性格を，5 つの因子とその強度で表現するビッグファイブ理論 [13] について述べる．

Goldberg らはビッグファイブ理論において，人間の性格は大きく分けて「開放性」「誠実性」「外向性」「協調性」「精神症傾向」の 5 つ因子および，それらの強度から定義できるとしている．なお，ビッグファイブは各因子の強度の値が高いこと，および低いことによる特性を述べているだけであり，それ自体の良し悪しを決定するものではないことに留意しなくてはならない．

「開放性」は，芸術性や探究心に関連する因子である．開放性の高い人物は良く言えば好奇心旺盛で，芸術的だが，悪く言えば思考が特異であるとされている．一方で，開放性の低い人物は，良く言えば伝統的なものを重視する人物であるが，悪く言えば鈍感で，多様性を受け入れづらい人物であるとされている．

「誠実性」は，まじめさや，勤勉さに関する因子である．誠実性の高い人物は良く言えば勤勉でまじめだが，悪く言えば厳格で高慢であるとされている．一方で，誠実性の低い人物は，良く言えば大胆で自立しており，悪く言えば非組織的で怠惰であるとされている．

「外向性」は，社交性や，明るさに関する因子である．外向性の高い人物は，良く言えば陽気で社交的な人物であるが，悪く言えば軽率で，うぬぼれることがある人物であるとされている．一方で，外向性の低い人物は，良く言えば謙虚で柔和な人物であり，悪く言えば無気力で，よそよそしい人物であるとされている．

「協調性」は，協力的であるかという点や，親切心に関する因子である．協調性の高い人物は，良く言えば寛容で親切な人物であるが，悪く言えば他人を疑わない，甘い人物であるとされている．一方で，協調性の低い人物は良く言えば強気な人物であり，悪く言えば融通の利かず，乱暴な人物であるとされている．

「精神症傾向」は，精神の不安定さや，ストレスに対する耐性に関する因子である．精神症傾向の高い人物は，良く言えば用心深い人物であるが，悪く言えば臆病で，自信のない人物であるとされている．一方で，精神症傾向の低い人は，良く言えば冷静で安定している人物であり，悪く言えば冷淡で，そっけない人物であるとされている．

これらのビッグファイブ理論は，今日では心理学において有力なモデルであるとされ，性格診断などにも多く用いられている．特に，性格診断では，最終的に診断者の性格と各因子との適合度合いを 1（適合していない）から 5（適合している）の数値で表現されることも多く,「性格の数値化」という点で本研究と強い関係性があると考えている．

第4章　提案手法

本章では，提案する手法の枠組みを述べる．本研究の目的は，性格や職業などさまざまな設定を持ったゲームキャラクターの，さまざまな状況におけるその設定らしい振舞いを，人手ではなく AI に自動獲得させることである．この目的達成のために，環境と報酬に応じて振舞いを学習する強化学習を用いることにする．学習は環境とキャラクターのペアごとに行う必要があるが，報酬は環境ごと，およびキャラクターごとに独立に行うことで，省力化を図る．5.1 節では環境ごとの報酬の設計法，5.2 節ではキャラクター設定ごとの報酬の調整法について述べる．以下に具体的な職業，報酬の設定方法を示す．

4.1　職業の表現

RPG をはじめとするゲームにはさまざまなシチュエーションが登場する．例えば「王の警護」「山賊が商人を襲う」といったものである．このとき,「王の警護」というシチュエーション自体は同じであっても登場するキャラクターの性格は異なっているかもしれない．逆に同じ性格のキャラクターが多くのシチュエーションに登場するかもしれない．これらのゲーム制作の要請に対し，提案手法では，シチュエーション（環境，状況）S ごとに，およびそれとは独立にキャラクター設定 C ごとに，報酬を設計する．そして，シチュエーションと登場キャラクターの組み合わせ $S \times C$ ごとに強化学習を行う．これにより，組み合わせごとに挙動を人手で定めるよりも簡単に，状況とキャラクタ設定に応じた挙動を獲得することを狙う．

シチュエーションごとに，いわゆる役割を持ったキャラクタが複数種類登場する．本論文ではこれを職業と呼称する．職業の異なるキャラクターは，他のキャラクターと異なる行動目的を持っているはずである．しかし，勝負に勝ちたい，自分が死にたくないなど，だれもが共通して持っている感情もあるはずである．それらを組み合わせることで，職業による表現が可能になるのではないかと考えた．

以下の手順で，キャラクターの職業に応じた報酬設計を決定することができる．

(1) 以下のような，ゲームにおいて登場しやすく，キャラクターだれもが共通して持っている感情を共通の報酬として用意する．

- 自分の命を惜しむ/自分の命を惜しまない．

- 仲間の命を惜しむ/仲間の命を惜しまない.
- 目標達成のために自分の力で貢献したい/仲間に貢献してもらいたい.

(2) それらに加えて以下のような，職業や，シチュエーションごとにとってほしい行動を，追加の報酬として用意する.

- 職業が騎士ならば，王様を守る.
- 職業が山賊ならば，自分が敵にとどめを刺して手柄を立てることを重視する.
- 職業が門番ならば，敵を門の奥に通さない.
- 因縁の敵が登場するシチュエーションなら，優先してその敵を対処する.

(3) 報酬の具体的な値を決定する．この時，次節で述べる性格による倍率の決定のため，普通のキャラクターとして報酬の値を設定することが望ましい.

4.2 性格の表現

同じ職業のキャラクターの中にも，異なる性格を持っているキャラクターが存在しうる．性格が異なれば，報酬設計において，どの部分を特に重視するのかという点が変化するはずである．これを，報酬設計に対する補正値として，性格ごとに決定すれば，性格による表現が可能になるのではないかと考えた.

以下の手順で，キャラクターの性格を用いて報酬の補正倍率を決定することができる.

(1) 共通の感情による報酬と，職業，シチュエーションごとに決定した報酬について，どのようなキャラクター設定の性質と関連するかを考える．また，この性質には，ビッグファイブ理論における5因子と関連性のあるものもある．以下に報酬と性質の関係の例を挙げる.

- 自分の命や，自分がリスクを負うことに関する報酬：自分の命重視性，精神症傾向（ビッグファイブ理論）
- 仲間の命や，仲間がリスクを負うことに関する報酬：仲間の命重視性，協調性（ビッグファイブ理論）
- 目的達成のための行動を自分が積極的に行うことによる報酬：自力貢献性
- 目的達成のための行動を仲間に任せることによる報酬：他力本願性
- 「王様を守る」など，職業に関係する報酬：任務成功重視性

(2) 性格ごとに，各性質とどれくらい強くつながりを持つかをもとに，報酬に加える倍率を決定する．ビッグファイブ理論と関連性のある性質については，その関連する因子の強度を用いて決定することもできる．以下に「臆病」な性格のキャラクターについて，性質とのつながりを考えた例を示す．

- 臆病なキャラクターは，自分が死んでしまうことをおそれるはずであり，「自分の命重要性」が高いはずである．そのため，「自分の命を惜しむ」という報酬に高い倍率をつける．
- 臆病なキャラクターは，自分の命が重要なので，仲間の命を考える余裕がないはずであり，「仲間の命重視性」が低いはずである．そのため，「仲間の命を惜しむ」という報酬に低い倍率をつける．
- 臆病なキャラクターは，自分が活躍するのではなく仲間が活躍してくれることを望むはずであり，「他力本願性」が高いはずである．そのため，「目的達成のために仲間に貢献してもらう」という報酬に高い倍率をつける．

(3) 別の職業についても，同じ性格を表現したいのであれば，ここで決定した性格と性質のつながりの倍率をそのまま用いて作成することができる．

本研究で提案する手法の全体像を，図 4.1 に示す．

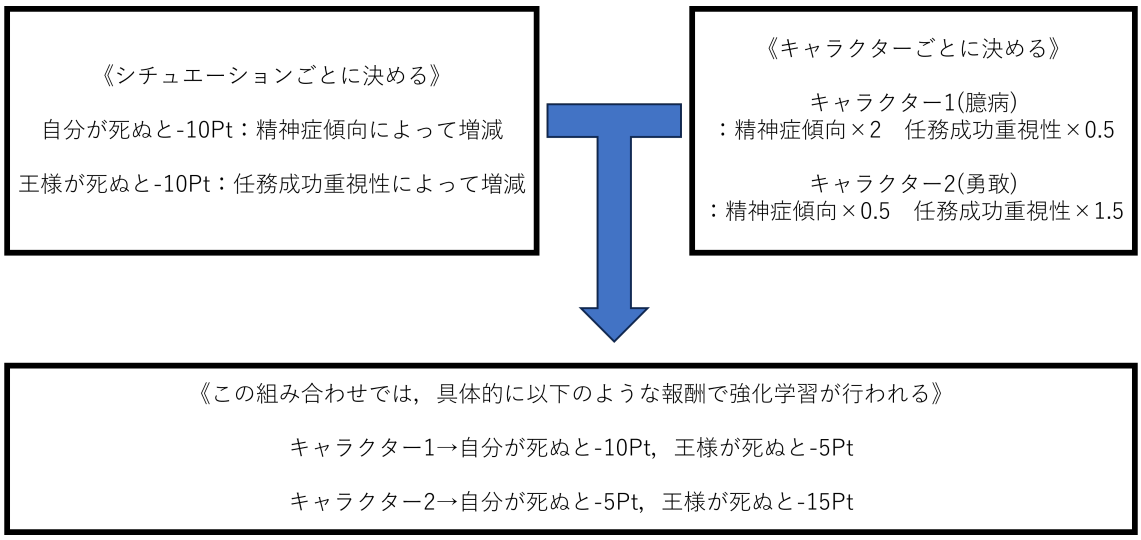


図 4.1: 本研究で提案する手法の全体像

第5章　研究環境

本章では，研究環境に用いたゲームと，その具体的なルールについて述べる．本研究では，キャラクターらしい行動の中でも，特に戦闘中に表れるものについて着目することとした．そのため，戦闘要素があり，かつ設定の異なる複数のキャラクターが登場可能な環境が必要となる．そこで本研究では，簡単に複数のキャラクターが登場する戦闘シーンを実装，学習可能である Unity ML-Agents[14] を研究環境として設定した．

5.1　Unity ML-Agents

Unity とは，Unity Technologies が公開しているゲームエンジンである．Unity は手軽にゲーム作成ができるというだけでなく，高品質な 3D の物理シミュレーションが可能であることから，建築設計に用いられたり，VR 開発の機能を用いて医療分野における手術シミュレーションソフトの作成に用いられたりすることもある．

Unity ML-Agents とは，同社が提供している，Unity 上で強化学習ができるようになるオープンソースのライブラリである．Unity 上で学習環境を構築し，Pytorch と連携することで学習を実行させることができる．また，学習が完了したモデルを Unity の環境に適用することで，シミュレーションを行うこともできる．Unity ML-Agents には，Proximal Policy Optimization(PPO)[15] や，Soft Actor-Critic(SAC)[16] など，強化学習用のアルゴリズムや，それを適用するための学習環境が複数用意されており，ユーザーは手軽にゲーム AI の学習を行うことができる．

5.1.1　DungeonEscape

本研究では，Unity ML-Agents に標準搭載されている環境の，DungeonEscape をベースにルールを変更したゲームを研究環境として用いている．標準の DungeonEscape（図 5.1）は，1 匹のドラゴン（緑）と 3 人のエージェント（青）から構成されている環境であり，エージェントは木製の扉から脱出することを目的とする．扉から脱出するためには，ポータル（紫）から逃走しようと単純な行動をとるドラゴンを倒し，鍵を入手し，扉の元まで持っていく必要がある．ドラゴンがポータルに到達すると，ゲームオーバーとなる．ドラゴンを倒すためには，エー

ジェントがドラゴンにぶつかって攻撃する必要があるが，ドラゴンにぶつかったエージェントは死亡するため，少なくともエージェント一人はクリアのために犠牲にならなくてはならない．

DungeonEscape では，戦闘の要素がドラゴンと 1 回ぶつかるという点だけであり，戦闘中に表れるキャラクターらしい行動を十分に表現できない可能性があると考えた．また，ドラゴンを倒した後に，鍵を扉まで持ち運ぶという点は，戦闘と関係のない点なので今回は必要がないと考えた．そのため，本研究ではこのゲームに対し，ルールの変更を適用したものを，研究環境として用いることにした．

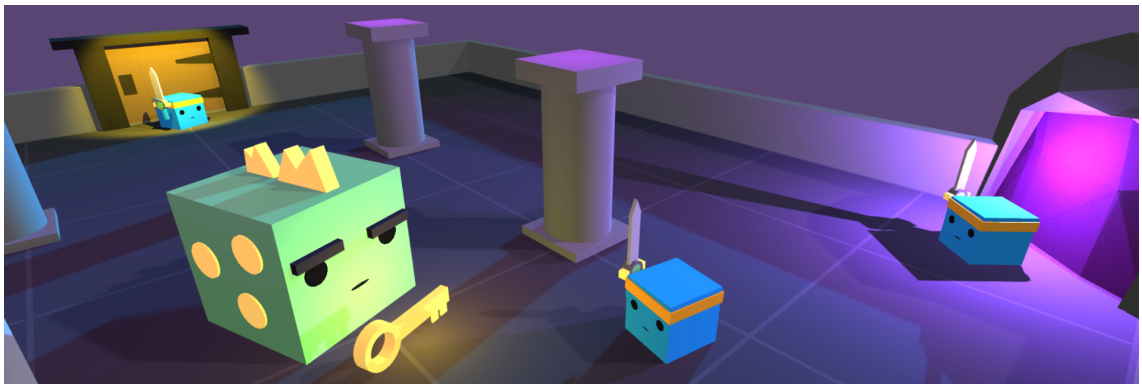


図 5.1: DungeonEscape の環境 [17]

5.2 実際に用いる環境

本節では，本研究で実際に用いた，ルール変更版 DungeonEscape について，ゲーム全体のルールと，登場キャラクターの 2 つの観点から説明する．

5.2.1 ゲーム全体のルールに関する変更点

表 5.1 に DungeonEscape からのゲーム全体のルールに関する変更点を示す．

前節で述べたように，DungeonEsacpe の環境では戦闘がエージェントとドラゴンの接触 1 回で終わってしまい，戦闘にかかる時間が非常に短いという問題があった．そのため，ゲームをお互いにぶつかり合って体力 (HP) を削りあう体力制に変更した．この体力制では，キャラクターごとに HP が割り当てられており，敵とぶつかることで，ぶつかった側とぶつかられた側の HP が 1 減少する．HP が 0 になったキャラクターは，死亡したという判定となる．また，体力制にした場合，1 度ぶつかっただけではドラゴンもエージェントも死亡せず，その後のドラゴンとエージェントが接触している間に何も発生しない時間が生まれてしまうことになりうる．そのため，キャラクターが敵と接触している間には HP を毎フレーム 0.1 ずつ減らすというルールに変更した．

DungeonEscape では，ドラゴンを倒した後に鍵を入手し，扉まで運ぶという操作が必要であった．しかし，本研究ではルールを体力制に変更し，その戦いの中で表れるキャラクターらしい行動の表現を目的としているため，ドラゴンが死亡した時をはじめとして，鍵に関する要素はすべて省くこととした．

DungeonEscape では，脱出に成功するか，ドラゴンが逃走するか，非常に長い時間をかけてもゲームのクリアができなかった場合に，ゲームが終了するようになっていた．本研究では，ルールが体力制に変更となったため，エージェントの脱出やドラゴンの逃走が存在しない．そのため，単純に 300 フレームが経過した時点で終了とし，次節で述べる王様の生死を主なゲームクリアの判定条件とした．

表 5.1: DungeonEscape からのルール変更点

要素	DungeonEscape	ルール変更版 DungeonEscape
ゲームの勝利条件	ドラゴンから鍵を入手し 扉から脱出	ゲーム終了時に 王様が生存
ゲームの敗北条件	ドラゴンがポータルに到達	王様が死亡
アイテム要素	鍵	（存在しない）
扉要素	ドラゴンのポータル ゴールの扉	（存在しない）
キャラクターが敵と ぶつかったときの挙動	お互いに死亡	お互いに HP が 1 減る
キャラクターが敵と ぶつかっている間の挙動	ぶつかったときにお互いに 死亡するので存在しない	お互いに HP が 0.1 減る
ドラゴン死亡時の挙動	ドラゴンがいた場所に 鍵を落として消滅	消滅
ゲームの終了条件	25000 フレームが経過	300 フレームが経過

5.2.2 ゲームに登場するキャラクターに関する変更点

表 5.2 に DungeonEscape からの構成キャラクターに関する変更点を示す．

DungeonEscape の環境は非常にシンプルであり，そのままでは戦闘中にキャラクターらしい行動が可能となるような設定を決定することが困難であった．そのため，エージェントが守る対象として，王様を 1 人作成した．王様はフィールドの下側に初めから出現する．ドラゴンはこの王様を倒そうと，単純な動きで王様に近づこうとする．一方で王様は，ドラゴンに殺されないように単純な動きでドラゴンから逃げようとする．王様の存在により,「自分が死ぬことを避けるため逃げ出す」キャラクターもいれば,「王様が死ぬことを避けるため命がけで王様をかばう」キャラクターも表現されうるのではないかと考えた．また,「ドラゴン対王

様を守る騎士」の形式だと，騎士は全員でドラゴンを囲んで攻撃すれば王様を守ることができてしまう．そのため，追加の敵キャラクターとして暗殺者を作成した．暗殺者は，毎フレーム 5%の確率でゲーム中最大 1 回だけ出現する．暗殺者はステージ左右にある壁付近のランダムな位置から出現する．暗殺者自体はドラゴンと同様に王様を倒そうと単純な動きで近づく挙動をとるが，いつ登場するかわからないため．これにより，王様を守りたいエージェントはドラゴンを倒しに行くのではなく，王様の周囲で警戒行動をとるのではないかと考えた．

また，DungeonEscape では，エージェントの観測は各エージェントの向いている前方のみであった．そのため，エージェントがドラゴンや暗殺者を見失うことがあった．エージェント自身が積極的に左右へ方向転換を行うことで，この問題を解決することはできるが，本研究ではそのような高度なキャラクター操作を学習させることが主な目的ではない．そのため，本研究では学習を容易にするため，エージェントの周囲すべてが観測可能となるように変更を行った．

表 5.2: DungeonEscape からの構成キャラクター変更点

要素	DungeonEscape	ルール変更版 DungeonEscape
構成キャラクター	ドラゴン 1 匹 エージェント 3 人	ドラゴン 1 匹 エージェント 3 人 王様 1 人 暗殺者 1 人
エージェントの目的	鍵を入手して脱出	性格によって異なる
ドラゴンの挙動	ポータルに向かう	王様に向かう
王様の挙動	（存在しない）	フィールド下側に登場 ドラゴンと反対方向に 向かう
暗殺者の挙動	（存在しない）	ランダムなタイミングで 左右から 1 度だけ登場 王様に向かう
エージェントの 観測方法	エージェントの前方のみ	エージェントの周囲すべて

第6章　騎士キャラクターの作成

本章では，4 章で示した提案手法を用いて，騎士キャラクターの作成を行った結果を示す．騎士キャラクターの作成，およびテストについては，5.2 節で述べた環境を用いた．学習アルゴリズムには，MA-POCA[18] を用いた．MA-POCA は，同じくマルチエージェント強化学習の手法である COMA[19] をもとに改良されたものである．これにより，マルチエージェント強化学習において，一部のエージェントが早期に終了してしまうことによって発生する課題である，死後信用割り当て問題に対処しようとした手法である．MA-POCA は方策勾配法を用いており，後述するハイパーパラメータにおける Beta の値に応じてランダムな行動を促している．

本研究の環境では，エージェントが敵と戦う中で，早期に死亡（終了）する者がいれば，最後まで死亡しない者もいるため，この手法が最も適していると考えた．以下に，本研究で用いた MA-POCA のハイパーパラメータを示す．これは，ML-Agents のデフォルト値である．

表 6.1: 本研究における MA-POCA のハイパーパラメータ

Hyperparameters	
Batch size	1024
Buffer size	10240
Learning late	0.03
Learning rate の減衰方法	減衰させず，常に一定に保つ
Hidden units	256
Number of layers	2
Beta	0.01

6.1 節では，騎士という職業の報酬設計や，性格による差異を表現するためにどのように考えたか示す．その後，6.2 節以降で，性格ごとに学習させた騎士キャラクターを用いてテストした際の統計結果や目視により判別した結果から，キャラクターらしい行動ができるようになったか考察する．

6.1　騎士キャラクターの報酬設計

本節では，本研究で騎士キャラクターや，それらの性格を表現可能な報酬設計を，どのような考えで設定するに至ったか述べる．

6.1.1　職業としての騎士

本研究で用いた環境は,「騎士，王様チーム対ドラゴン，暗殺者チーム」である．このシチュエーションでは，普通の騎士は「王様を敵から守ることを主な目的とし，余裕があれば自分や仲間の命も助かるように行動する」ことが望ましい．この点から，騎士キャラクターの職業に応じた報酬を以下のように設定した．

- 普通の騎士キャラクターは，守るべき対象（今回は王様）が死なないことを重要視するはずである．そのため，王様が死亡した時にマイナスの報酬を与える．これは騎士のすべき仕事であるため,「任務成功重視性」によって変動するものと定義する．

- 普通の騎士キャラクターは，王様を守るために命を懸けて戦うとはいえ，自分が死ぬことも少しだけ恐れているはずである．そのため，自分が死亡した時に，王様が死亡した時より少ないマイナスの報酬を与える．これは自身が負うリスクに関する報酬であるため,「自分の命重視性」や「精神症傾向」によって変動するものと定義する．

- 普通の騎士キャラクターは，王様を守るという目的が達成できるのならば，貢献するのは自分でも仲間でも構わないと考えるはずである．そのため，自分が攻撃した時と，仲間が攻撃した時とで，同じ量のプラスの報酬を与える．これは目的達成のために，だれが貢献するかという点に関する報酬であるため,「自力貢献性」や「他力本願性」によって変動するものと定義する．

- 普通の騎士キャラクターは，一番の目的こそ王様を守ることであるが，敵が弱かった時など，余裕があるときには仲間が多く生き残ってほしいと考えるはずである．そのため，王様を守ることに成功した時，生き残っている騎士の数に応じたプラスの報酬を与える．これは仲間の命に関する報酬であるため,「仲間の命重視性」や「協調性」によって変動するものと定義する．

6.1.2　用意した性格

前節で設定した報酬設計を「普通の騎士」に関する報酬設計とし，それをベースに「勇敢な性格」,「臆病な性格」,「英雄気取りで目立ちたがりな性格」を用意した．以下に具体的な報酬と性格ごとの倍率と，どのようにしてそれらの値を決

定したかを示す．これらの倍率の値は，5.2 節の環境のみに対応するものではなく，AI に対応させたいシチュエーションの報酬ごとに定められた,「任務成功重視性」「精神症傾向」などと対応した報酬の値に，同じように使えるものであると想定している．

表 6.2: 本研究で作成したキャラクターの性格と報酬の倍率

シチュエーションごとに決める		性格ごとに決める		
報酬	普通の騎士の報酬	勇敢な騎士	臆病な騎士	英雄気取りで 目立ちたがりな騎士
王様が死亡 （任務成功重視性）	-10	× 5	× 1	× 1
自分が死亡 （自分の命重視性, 精神症傾向）	-5	× 0	× 2	× 1
自分が敵に攻撃 （自力貢献性）	+2	× 1	× 1	× 2
仲間が敵に攻撃 （他力本願性）	+2	× 1	× 2	× 0.5
仲間が生存 （仲間の命重視性, 協調性）	1 人につき+1	× 4	× 0	× 2

- 勇敢な性格のキャラクターは，ビッグファイブにおける協調性が高い性格だと予想される．この性格のキャラクターは，仲間と共に任務を成功させることを重視し，そのためには協力や自己犠牲を惜しまない精神力の強さを持つ．6.1 節で定義された基準で言えば，任務成功重視性（× 5）や協調性（× 4）は高く，精神症傾向（× 0）は低いということである．結果として，このシチュエーションでは,「王様も仲間も死なせたくない，そのためには死を恐れない」という挙動を期待した報酬設計となる．

- 臆病な性格のキャラクターは，ビッグファイブにおける精神症傾向が高い性格だと予想される．この性格のキャラクターは，任務に失敗すること以外にも，自分が死亡することを恐れ，自分が死なずに任務をこなすことを重要視する利己的な考えを持つ．6.1 節で定義された基準で言えば，精神症傾向（× 2）や他力本願性（× 2）は高く，協調性（× 0）は低いということである．結果として，このシチュエーションでは,「王様を死なせないようにしたいが，自分も死なないようにしたいので，できるだけ戦いは味方に任せる」という挙動を期待した報酬設計となる．

- 英雄気取りで目立ちたがりなキャラクターは，自己顕示欲の高い性格だと予想される．この性格のキャラクターは，任務を遂行するなかで自分が積極的に貢献することや，仲間を守ることにより，英雄的だと思われたいという考えを持つ．6.1 節で定義された基準で言えば，自力貢献性（× 2）や協調性（× 2）は高く，他力本願性（× 0.5）は低いということである．結果として，このシチュエーションでは,「王様や仲間を守るために，かなり前のめりな体勢で敵に攻撃をしに行く．という挙動を期待した報酬設計となる．

6.2　ベース設定での結果

本節では，ベース設定として,「勇敢な性格の騎士 2 人，臆病な性格の騎士 1 人のチーム」で学習を行った結果について示す．各キャラクターの HP は以下の表 6.3 に示す．

表 6.3: 実験における各キャラクターの HP

	各騎士	王様	ドラゴン	暗殺者
HP	2	1	4	1

王様と暗殺者の HP は 1 であり，1 度敵と接触すると死亡する．また，王様のみ敵にダメージを与えることができない．騎士は 3 名いるため，2 名がドラゴンと 2 回ずつ接触，1 名が暗殺者と接触することで，1 名生き残りながら王様を守ることができる．騎士は以下のように，ぶつかっている間に与えられる 0.1 ずつのダメージを活用すれば，2 名生き残りながら王様を守ることができる．

- 騎士 A がドラゴンと暗殺者に 1 回ずつ接触する．ドラゴンは HP が 3 に減り，暗殺者と騎士 A は死亡する．
- 騎士 B と騎士 C がドラゴンに 1 回接触する．ドラゴンは HP が 1 に減り，騎士 B と騎士 C の残りの HP は 1 になる．
- その後騎士 B と騎士 C がドラゴンに，ぶつかっている間の 0.1 のダメージを 5 回ずつ与える．ドラゴンは死亡するが，騎士 B と騎士 C は HP が 0.5 残るため，生存することになる．

場合によっては，ドラゴンに複数名の騎士が同時に攻撃することで，その全員がドラゴンによる反撃を受けた結果，騎士 3 名全員が死亡するといったケースもありうる．

学習は 5.2.1 節で述べた 300 フレームの経過を 1 エピソードとして，200,000 エピソード行った．GPU を使わないデスクトップパソコン環境下で，実行完了までに 4 日ほど要した．ハイパーパラメータの Beta=0 として，テストは 500 エピソード行った．

6.2.1　学習曲線

図 6.1 に，訓練エピソードにおいて騎士の獲得した平均報酬に関する学習曲線を示す．学習曲線は，1000 エピソード分のデータを 1 区間として，移動平均を用いて出力した．

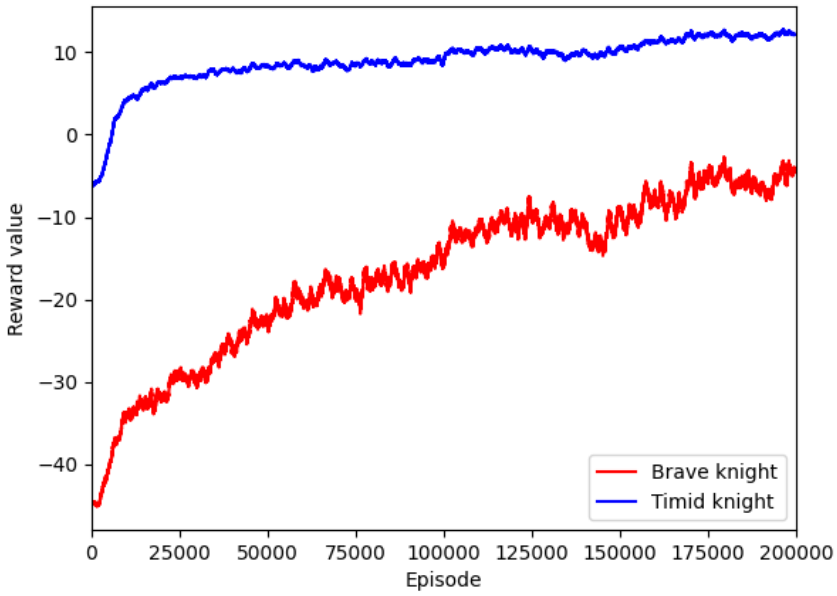


図 6.1: 騎士の獲得合計報酬に関する学習曲線

勇敢な騎士，臆病な騎士ともに，おおむね学習できているが，特に勇敢な騎士は，この後ももう少し改善することが見込めそうである．

6.2.2　テスト結果：チームの勝敗

表 6.4 にベース設定でテストを行った際の，全体の勝敗の内訳を示す．

ベース設定のチームは，約 58%の確率で王様を守ることに成功しており，チーム全体として見たときに，王様を守る騎士としての職業が表現できていることがわかる．また，王様を守ることの成功失敗を問わず，ゲームの中で騎士が 1 人以上死亡した割合は 94%である．これより，王様を守るために命を懸ける騎士らしさが表現できていることがわかる．王様を守れなかったケースとしては暗殺者を対処できず，王様が倒されてしまうということが多い．この原因として，騎士の王様を挟んだ反対側に暗殺者出現した時，王様にセンサーが阻まれて感知できなかったり，騎士が暗殺者を感知できる位置にいたとしても，スムーズに暗殺者の元まで移動できなかったりすることが挙げられる．これらの問題を解決するために，まず十分学習することが解決策の一つだと考えた．しかし，ほかにも「敵に

表 6.4: 学習完了モデルのテスト時におけるチームの勝敗

	騎士の 死亡数 0	騎士の 死亡数 1	騎士の 死亡数 2	騎士の 死亡数 3	合計
王様を 守った	24	123	141	1	289
王様がドラゴンに 殺された	0	0	64	1	65
王様が暗殺者に 殺された	6	74	65	1	146
合計	30	197	270	3	500

ダメージを与えたときの報酬は，王様が生きていないと与えられない」や,「王様の近くにいたにもかかわらず，守れなかった場合さらにマイナスの報酬を与える」など，新たな報酬設計を用意することが有効なのではないかと考えた.

6.2.3 テスト結果：チームの与えたダメージ

表 6.5 にベース設定でテストを行った際の，騎士全体の与えたダメージを示す. 表の数値は,「接触した瞬間に与えたダメージ」+「接触している間に与えたダメージ」の形式で表現されている．本研究の環境では，敵の HP は 1，もしくは 0.1 ずつ減少するため，残りの HP が 0.3 である敵が 1 ダメージを受けた場合，0.7 点分の余剰ダメージを受けたことになる．この余剰ダメージは，敵を倒すために必要なダメージではない．そのため,「接触した瞬間に与えたダメージ」には，余剰ダメージが減算された後の数値が記載されており，本来 1 点ずつ与えられる接触した瞬間ダメージの値が小数点を含んでいることがある.

表 6.5: 学習完了モデルのテスト時におけるチームの与ダメージ

	勇敢な騎士の 与えたダメージ（2 人の平均）	臆病な騎士の 与えたダメージ	合計
ダメージ量	653.85 + 261.9	363.5 + 8.2	1671.2 + 532.0

「勇敢な騎士の与えたダメージ」の項目は，勇敢な騎士 2 人が与えたダメージの合計値を示している．つまり，勇敢な騎士は，平均 650 点の「接触した瞬間のダメージ」と，平均 260 点の「接触している間のダメージ」を与えているということである．これらは臆病な騎士の与えたそれぞれのダメージよりも明らかに高い．本研究のゲームルールでは，攻撃をするということは，自分も同じだけのダメージを受けるということである．そのため，勇敢な騎士たちはダメージを恐れずに積極的に攻撃していることがわかる．一方で，臆病な騎士の与えた「接触し

た瞬間のダメージ」は約 360 点である．テストは 500 回エピソード行ったため，1 エピソードにつき 1 回も攻撃しないことさえあるという，低い攻撃頻度であることがわかる．これは，臆病な性格であるがゆえに死を恐れ，攻撃に踏み出せないというキャラクターらしい行動であるといえる．

また，表 6.4 にて，騎士の死亡数が 2 名で王様を守った回数が 141 回に対し．騎士の死亡数が 1 名で王様を守った回数は 123 回と，ほとんど変わらない．6.2 節で述べたように，騎士の死亡数 1 名で敵をすべて倒すためには,「接触している間ダメージ」を活用しなくてはならない．表 6.5 を見ると，0.1 点ずつ与えられる「接触している間ダメージ」をチーム全員の合計で 532.0 点分与えており，シチュエーションに合わせた勝ち方を学習できているといえる．

6.2.4　テスト結果：各騎士の序盤の行動

表 6.6 にベース設定でテストを行った際の，各騎士の序盤の行動を示す．ここでいう序盤とは，ゲーム開始後 20 フレームから 60 フレームの間のことを指す．この期間に,「自分と王様が一定距離以内にいる．かつその距離が自分とドラゴンの間の距離より短い」，または「自分とドラゴンが一定距離以内にいる．かつその距離が自分と王様の間の距離より短い」の条件を満たしているフレームを数え，それの量に応じて「王様守護志向」,「ドラゴン討伐志向」,「王様・ドラゴン両方志向」,「どちらでもない」に分類した.「どちらでもない」とは序盤に王様の元にもドラゴンの元にも向かわなかったということであり，主に「逃げ出している状態」や「どちらのもとに行くか悩み，動き出しが遅れている」状態などが該当する．

表 6.6: 学習完了モデルのテスト時における騎士の序盤の行動

	勇敢な騎士 1 の行動	勇敢な騎士 2 の行動	臆病な騎士の行動
ドラゴン討伐志向	274	280	79
王様守護志向	137	138	194
王様・ドラゴン両方志向	7	5	10
どちらでもない	82	77	217

勇敢な騎士はドラゴンに向かうことが多く，臆病な騎士はドラゴン，王様のどちらにも近づかない，もしくは王様に近づくという行動をとることが多い．表 6.3 より，本実験のドラゴンの HP は 4 であり，騎士の HP は 2 である．そのため，対処のためには基本的に 2 人の騎士が犠牲にならなくてはならない．そのため，勇敢な騎士が積極的にドラゴンに向かい，その役割を引き受けているのだと考えることができる．王を守るため，自分が犠牲になることをいとわず行動できている点は,“ 勇敢な騎士らしい ” 行動であるといえる．また，勇敢な騎士たちの序盤の行動として,「王様に向かうこと」も多い．これは，後述の臆病な騎士の行動の影

響であると考えている．臆病な騎士は「自身の生存」「王様の守護」どちらも達成できる王様守護志向を選ぶことがあるが，時折王様守護志向ですらない，逃げ出す行動をとってしまうこともある．そのような場合，暗殺者から王様を守るために，勇敢な騎士がカバーに入っているのだと考える．この点から，仲間の振る舞いに合わせて，最適なキャラクター表現の自動獲得ができていることがわかる．

臆病な騎士の序盤の行動として，最も多いものは「どちらでもない」であった．これは，臆病な騎士が「王様を守れなさそうなので，無理しないで逃げた方がよい」と考えた結果行われたものであると考えた．この序盤の行動は,「王様の命より自分の命を優先して逃げ出してしまう」「王様やドラゴンに向かわなくてはいけないが，戦って死ぬことを恐れて動けなくなってしまう」という表現であったと考えることができる．これは，非常に臆病な性格らしい行動であるといえる．また，臆病な騎士の行動として次いで多いのは「王様の元へ向かう」であった．これは，臆病な騎士が「ほかのキャラクターがドラゴンを対処してくれるため，自分は暗殺者対処をすればよい」と考えた結果行われたものであると考えた．表 6.3 より，暗殺者の HP は 1 である．そのため，騎士は暗殺者の対処をするだけであれば，死亡せずに王様の守護に貢献することができる．そのため，王様に向かうことは「自分が死なないようにしつつも，王様を守るために暗殺者対処をする」ことであるといえる．これは,“ 臆病な騎士らしい ” 行動であるといえる．

6.2.5　テスト結果：各騎士のゲーム終了時の状態

表 6.7 にベース設定でテストを行った際の，各騎士のエピソード終了時の状態を示す．

表 6.7: 学習完了モデルのテスト時における各騎士の状態

		勇敢な騎士 1 の結果	勇敢な騎士 2 の結果	臆病な騎士の結果
王を守った	自分は無傷	28	19	56
	自分は負傷	151	166	224
	自分は死亡	110	104	10
王がドラゴンに殺された	自分は無傷	0	1	44
	自分は負傷	21	21	16
	自分は死亡	44	43	16
王が暗殺者に殺された	自分は無傷	11	6	56
	自分は負傷	90	92	85
	自分は死亡	45	48	5

王様の守護に成功したケース，失敗したケース問わず，勇敢な騎士はエピソード終了時に死亡していることが多いのに対し，臆病な騎士はほとんど死亡するこ

とがない．その一方で，勇敢な騎士は無傷，すなわち攻撃せずエピソード終了を迎えることがほぼなかったのに対し，臆病な騎士は無傷でエピソード終了を迎えることが 150 回ほどあった．ここからも，勇敢な騎士は自分の死を恐れず攻撃できている反面，臆病な騎士は死を恐れ，自分が死なないように立ち回っていることがわかる．また，臆病な騎士はエピソード終了時の結果において，負傷しているケースが最も多い．ここから，臆病な騎士はただ逃げ回っていただけではなく，騎士らしく王様を守るために戦っていることがわかる．

6.2.6　テスト結果：各騎士の攻撃状況

ベース設定でテストを行った際の，各騎士のエピソード終了時の攻撃状況について，付録 A.1 にまとめた．

重要な点として，勇敢な騎士はドラゴンに対して 1.1 点以上，特に 2 点のダメージを与えることが多く，臆病な騎士は暗殺者に攻撃する，もしくは一度も攻撃しないことが多いことが観測された．これは，6.2.4 節で述べた，勇敢な騎士と臆病な騎士の序盤の行動を裏付けるような結果である．これらからも，勇敢な騎士は自分の命を犠牲にして王を守るというキャラクターらしさが，臆病な騎士は王様を守りながらも自分が死なないように立ち回るというキャラクターらしさが表現できていることがわかる．

6.2.7　テスト結果：目視

テスト映像を目視で観測した結果，以下のような行動を確認することができた．

- 勇敢な騎士は，積極的にドラゴンに向かって攻撃しに行く．自分とドラゴンの出現位置が遠い時は，ドラゴンと王様の間に立ち，迎え撃つような体制をとることもある．
- 臆病な騎士は，ドラゴンからも王様からも逃げ出すことが多い．王様の元へ向かって，暗殺者の対処を行うことも多いが，ドラゴンが迫ってきたときはなるべく攻撃しないように後退することが多い．

これらは，6.1.2 節で「勇敢な性格のキャラクター」,「臆病な性格のキャラクター」に対して想定した通りの行動であり，キャラクター設定に沿った振る舞いの獲得ができているといえる．

6.3　バリエーションの設定と比較の全体像

前節では，ベース設定（勇敢 2，臆病 1，ドラゴン HP4）について騎士キャラクター AI の統計量と挙動をまとめた．

本節では，騎士の性格の組み合わせやキャラクターの HP など，状況を様々に変化させ，いずれの場合でも想定する挙動（あるいは，納得できる挙動）が獲得できているかを比較する．これにより，「特定シチュエーションでの性格」ごとに報酬を毎回調整するのではなく，それぞれを独立に決めることでもキャラクターらしい挙動が獲得できるか確認する

図 6.2 に，本研究で行った騎士作成実験における，性格の組み合わせや，キャラクターの HP といったバリエーションの設定を示す．また，矢印でつながれている設定同士について，次節で結果の比較を行い，キャラクターらしい行動が表現されているか考察する．

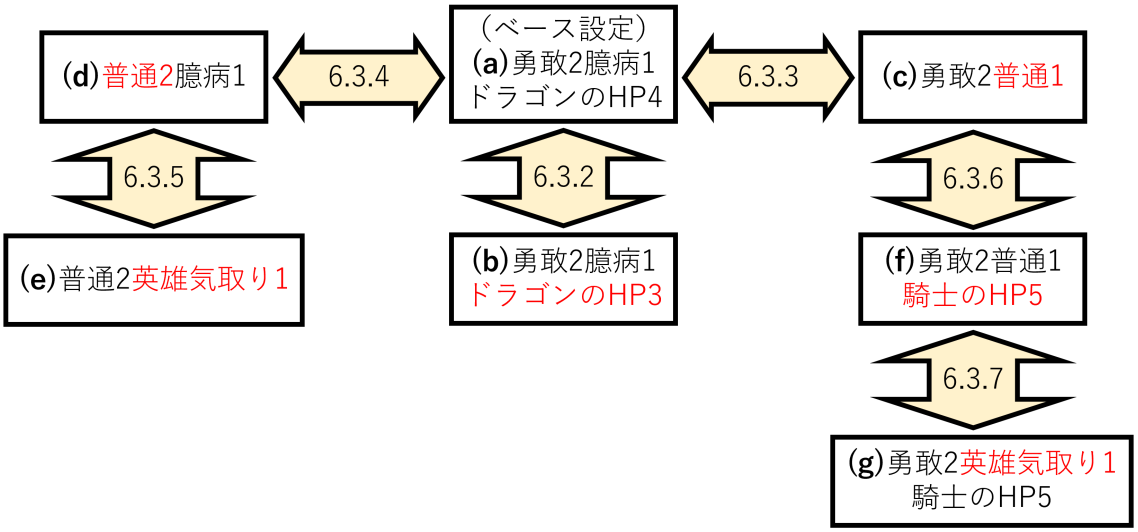


図 6.2: 実験のバリエーションと比較関係

6.3.1 キャラクター設定ごとの比較

本節以降では，6.3 節で示した各環境においても，キャラクターがそれぞれの設定に沿った行動をしていることを確認する．

各バリエーション設定ごとの結果として，表 6.8 に勝率を，表 6.9 に騎士が与えたダメージを，表 6.10 に騎士の序盤の行動を示す．なお，表 6.9 の数値は，「接触した瞬間に与えたダメージ」＋「接触している間に与えたダメージ」の形式で表現されている．

表 6.8 より，ドラゴンの HP が低い，騎士の HP が高いといった純粋にゲームが簡単になる場合には，王を守る確率が上がることがわかった．また，性格という面では，勇敢な騎士がいることで王を守る確率が上がり，臆病な騎士がいることで王を守る確率が下がることがわかった．普通の騎士と英雄気取りな騎士に関しては，勝率という面ではあまり変わらないこともわかった．

表 6.8: 各設定ごとの勝率

	王を守った	王がドラゴンに殺された	王が暗殺者に殺された
(a) 勇敢 2 臆病 1 ドラゴン HP4	57.8%	13%	29.2%
(b) 勇敢 2 臆病 1 ドラゴン HP3	69.2%	11.4%	19.4%
(c) 勇敢 2 普通 1	70.8%	3.4%	25.8%
(d) 普通 2 臆病 1	51.8%	35.0%	13.2%
(e) 普通 2 英雄気取り 1	67.8%	6.6%	25.6%
(f) 勇敢 2 普通 1 騎士 HP5	83.4%	1.2%	15.4%
(g) 勇敢 2 英雄気取り 1 騎士 HP5	80.6%	1.4%	18.0%

表 6.9 より，騎士の HP が 2 の環境（a）-（e）では，勇敢な騎士，英雄気取りな騎士，普通の騎士，臆病な騎士の順でダメージを与えることがわかる．英雄気取りな騎士と普通の騎士は，環境の差によるダメージの差こそあるが,（c）と（e）を比較すると，あまり与えるダメージに差がなさそうであることが予想される．臆病な騎士は与えたダメージの最大値が 360 点，最小値が 66 点と差が大きく，状況に応じて仕方なく攻撃する，もしくは攻撃せず逃げていることが予想される．一方で，騎士の HP が 5 になると，勇敢な騎士の与えたダメージは少なくなり，普通の騎士，英雄気取りな騎士の与ダメージが上がっている．

表 6.10 より，勇敢な騎士，普通の騎士，臆病な騎士の順でドラゴンに向かう行動が多いことがわかる．ドラゴンの HP が低かったり，騎士の HP が高かったりするシチュエーションでは，ドラゴンよりも暗殺者の対処の方が優先され，王様に向かう行動が多くなっていることもわかる．一方で,（e）の英雄気取りな騎士の行動や,（f）における普通の騎士の行動など，想定と異なる動きをすることもあった．

6.3.2 敵の HP の値の変化

ベース設定である「（a）勇敢 2, 臆病 1, ドラゴン HP4」と比較して,「（b）勇敢 2, 臆病 1, ドラゴン HP3」は純粋に敵の HP が低く，簡単な環境になったといえる．そのため，王を守ることに成功することも多く（57.8% → 69.2%）なっている．

騎士の動きをみると，臆病な騎士がドラゴンにも，王様にも向かわないことが増え（217 → 343），また，臆病な騎士の与えたダメージが大きく減っている（363.5+8.2 → 66.6+4.9）．これは，敵の HP が 3 のとき，暗殺者とドラゴンの両方を 2 人の勇敢な騎士で対処可能であるからである．そのため，臆病な騎士は攻撃を勇敢な騎

表 6.9: 各設定ごとの騎士が敵に与えたダメージ（「接触した瞬間に与えたダメージ」+「接触している間に与えたダメージ」）

		与えたダメージ
(a) 勇敢 2 臆病 1 ドラゴン HP4	勇敢（平均）	635.8+261.9
	臆病	363.5+8.2
(b) 勇敢 2 臆病 1 ドラゴン HP3	勇敢（平均）	694.0+155.4
	臆病	66.6+4.9
(c) 勇敢 2 普通 1	勇敢（平均）	650.2+273.7
	普通	433.1+12.2
(d) 普通 2 臆病 1	普通（平均）	643.5+197.2
	臆病	288.3+42.6
(e) 普通 2 英雄気取り 1	普通（平均）	572.3+288.6
	英雄気取り	459.9+17.1
(f) 勇敢 2 普通 1 騎士 HP5	勇敢（平均）	301.6+63.1
	普通	739.6+889.8
(g) 勇敢 2 英雄気取り 1 騎士 HP5	勇敢（平均）	321.0+62.0
	英雄気取り	1077.0+590.6

士に任せ，逃げる行動を選んだのである．

逆に言えば，ドラゴンの HP が 4 の時は，王様を守るために臆病な騎士も攻撃しなくてはならない．そのため，攻撃を行っているのだといえる．このように，敵の HP に合わせて，キャラクターの行動が適切に選択されていることがわかる．

6.3.3 臆病な騎士と普通の騎士

ベース設定である「(a) 勇敢 2, 臆病 1, ドラゴン HP4」と比較して,「(c) 勇敢 2, 普通 1」は，臆病な騎士の代わりに普通の騎士がチームに入っている．普通の騎士は，報酬的に臆病な騎士ほど自分が死ぬことを恐れないため，臆病な騎士と比較して戦いに参加することが多い．そのため，より多くのダメージを与える（363.5+8.2 → 433.1+12.2）ことができている．その結果，王を守ることに成功することも多く（57.8% → 70.8%）なっている．

また，臆病な騎士はゲームの序盤で王にもドラゴンにも向かわない,「どちらでもない」行動をとることが多い（217 回）が，普通の騎士は「王様守護」に向かうことが多い（268 回）．これは，勇敢な騎士が基本的にドラゴンに攻撃しに行くためであると考えている．これにより，普通の騎士は王様の下で暗殺者対処を行えば王様を守ることができる．とはいえ，暗殺者対処を行うと，自分はダメージを受けることになる.「このダメージのリスクを恐れず行動する普通の騎士」と,「ダメージのリスクを恐れて行動できないことのある臆病な騎士」という対比は，キャ

表 6.10: 各設定ごとの騎士の序盤の行動

		ドラゴン討伐	王様守護	どちらでもない
(a) 勇敢 2 臆病 1 ドラゴン HP4	勇敢（平均）	277	137	79
	臆病	79	194	217
(b) 勇敢 2 臆病 1 ドラゴン HP3	勇敢（平均）	123	258	115
	臆病	122	30	343
(c) 勇敢 2 普通 1	勇敢（平均）	331	80	86
	普通	60	268	163
(d) 普通 2 臆病 1	普通（平均）	123	270	100
	臆病	116	143	233
(e) 普通 2 英雄気取り 1	普通（平均）	290	87	120
	英雄気取り	60	301	121
(f) 勇敢 2 普通 1 騎士 HP5	勇敢（平均）	80	244	142
	普通	337	50	102
(g) 勇敢 2 英雄気取り 1 騎士 HP5	勇敢（平均）	84	265	142
	英雄気取り	353	39	102

ラクター設定に沿った行動であるといえる．

また，勇敢な騎士の行動に目を向けると，チームメイトが臆病な騎士であるときより，普通の騎士であるときの方が，王様に向かうことが少なく，ドラゴンに向かうことが多く（277 → 331）なっている．臆病な騎士は逃げ出すことが多く，王様に向かってくれないことがあることを考えると，勇敢な騎士と普通の騎士のチームでは役割分担が可能になったといえる．このような，チームメイトの変化による行動の変化という点も，キャラクターらしいと考えた．

6.3.4 勇敢な騎士と普通の騎士

ベース設定である「(a) 勇敢 2, 臆病 1, ドラゴン HP4」と比較して,「(d) 普通 2, 臆病 1」は，勇敢な騎士の代わりに普通の騎士がチームに入っている．普通の騎士は勇敢な騎士とは違い，少しは自分が死ぬことを恐れるため，序盤にドラゴンに向かうのではなく，王様に向かうことで迎え撃つ態勢をとることが多い（137 → 270）．そのため,「勇敢 2, 臆病 1 チーム」で臆病な騎士が行っていた，王様の近くで暗殺者対処をすることができない．また，勇敢な騎士と比較して，普通の騎士は死を恐れ，王様でもドラゴンでもない序盤の行動をすることが多く発生する（79 → 100）．この結果，王を守ることに成功することが少なくなっている（57.8% → 51.8%）．このような，チームメンバーが変わることで，チームとして弱くなってしまうという点は，キャラクターらしいと考えた．

6.3.5　臆病な騎士と英雄気取りな騎士

「(d) 普通 2, 臆病 1」と比較して,「(e) 普通 2, 英雄気取り 1」は，臆病な騎士の代わりに英雄気取りな騎士がチームに入っている．英雄気取りな騎士は，王様を守ることや，自分もできれば死なないということに加え，活躍して目立ちたいと考えており，敵に積極的に攻撃することを望んでいる．そのため,「普通 2, 臆病 1 チーム」における臆病な騎士と比較して明らかに与えたダメージが多い（288.3+42.6 → 459.9+17.1）.

しかし，その攻撃の積極性から，英雄気取りな騎士はドラゴンに向かうのではないかと予想していたが，予想に反して，序盤の行動より，英雄気取りな騎士は王様に向かうことが多い（301 回）ことがわかった．この原因として，環境と報酬設計のバランスが悪く，英雄気取りな騎士にとって「ドラゴンに 2 回攻撃し，死亡する」ことと「暗殺者に 1 回攻撃し，生存する」ことの価値が同じになっていたことを考えた．そのため，積極的に攻撃している点こそキャラクターらしいが，もう少し英雄気取りな騎士らしさの発揮できる環境の設定が必要となる．

また,「英雄気取り」は，本来「勇敢」や「普通」と比較すべき性格であり,「臆病」と比較することになってしまった点は，実験計画における反省点である．

6.3.6　騎士の HP の値の変化

「(c) 勇敢 2, 普通 1」と比較して,「(f) 勇敢 2, 普通 1, 騎士 HP5」は，騎士の HP が高いため，簡単な環境になったといえる．そのため，王様を守ることに成功することが多く（70.8% → 83.4%）なっている．

また，これまでは，騎士の HP が 2 であるのに対し，ドラゴンの HP が 4 であるため，ドラゴンを倒すために騎士が 2 人犠牲にならなくてはならなかったが，この環境では，騎士の犠牲なしにドラゴンを倒すことができる．そのため,「勇敢 2, 普通 1」では自分が死なないように暗殺者対処を行っていた，普通の騎士がドラゴンに向かうことが多くなったのではないかと考える．この点から，6.3.2 節と同様に，自身の HP の値よっても，キャラクターらしい行動が適切に選択されていることがわかる

6.3.7　普通の騎士と英雄気取りな騎士

「(f) 勇敢 2, 普通 1, 騎士 HP5」と比較して,「(g) 勇敢 2, 英雄気取り 1, 騎士 HP5」は，普通の騎士の代わりに英雄気取りな騎士がチームに入っている．英雄気取りな騎士はドラゴンに向かうことが多いという点は普通の騎士と同じだが，与えたダメージの中でも，接触した瞬間ダメージによるものが多い（739.6 → 1077.0）という違いがある．これは，ドラゴンのみならず，暗殺者への攻撃も狙えるのであれば暗殺者も狙っていたからではないかと考えた．

つまり，普通の騎士は「自分が死なずに王様を守ることができるため，ドラゴンの対処を行った．その結果ドラゴン接触している間ダメージを多く与えていた」のだと考える．一方で，英雄気取りな騎士は「ドラゴンの対処を行うが，敵に攻撃して目立ちたいので，暗殺者も攻撃する」のだと考えた．これは，キャラクターらしい行動の差であると考えた．

6.4 比較結果のまとめ

本章では，ベース設定として用意した「勇敢 2，臆病 1，ドラゴン HP4」での行動統計の確認や，合計 7 つのシチュエーションを変更した環境での結果を比較した．

特に「勇敢な騎士」や「臆病な騎士」は想定していた挙動を獲得させることができたほか，シチュエーションを変更した実験では，チームメンバーの性格やキャラクターの HP に応じてその状況に適した行動を獲得することができていた．一方で,「英雄気取りな騎士」は「普通の騎士」とのわかりやすい行動の違いを獲得することができなかった．この原因としては,「勇敢な騎士」と「臆病な騎士」は「王様を守ること」や「自分が死なないこと」といった結果に直接関係する報酬に倍率がかかっていたのに対し,「英雄気取りな騎士」は「敵に自分が攻撃する」という結果に間接的に関係する報酬に倍率がかかっていたことがあると考えている．

第7章　多様なシチュエーションでの利用

本研究では,「ドラゴンと暗殺者から王様を守る騎士」というシチュエーションに限定して実験を行っており，もっと多様なシチュエーションを用意し，実験するという点に至らなかった．しかし，本研究の提案手法はもっと様々なシチュエーションに対して適用可能なものであり，今回取り扱わなかったシチュエーションにおいても，キャラクターらしさを獲得できると期待している．本章では，いくつか想定した新たなシチュエーションをもとに，本研究の提案手法が「ドラゴンと暗殺者から王様を守る騎士」以外にも適用可能であることを述べる．

本研究では，表現する職業として「騎士」のみに焦点を当てていた．しかし，本来 RPG にはもっと様々な設定のキャラクターが登場し，それらすべてに対して，キャラクターらしさを獲得できるようにすることが理想である．例として，その中の一つである,「山賊」について表現することを考える.「山賊」とは，主に山を拠点として，通行人から略奪行為を行う盗賊であるとされている [20]．ここから，山賊であるキャラクターが職業として求めていることは，相手を倒した先にある，略奪行為を行って戦利品を獲ることだと考えた．そのため,「敵にとどめを刺す」ことや，アイテムが配置されている環境であれば,「戦利品を入手する」ことで報酬が与えられるように設定すると，山賊らしいキャラクターが作成できるのではないかと考えた．ここでいう,「褒美をもらうために敵を倒す」や「戦利品を入手する」といった報酬は，4.2 節で述べた，任務成功重視性に関連する報酬としてとらえることができる．

このような「山賊」が登場しうるシチュエーションとして,「山賊が町を襲っている途中で，金貨の入った袋が落ちているのと，兵士に守られている貴族を発見した」というものを考える．この時，金貨の入った袋を略奪するのは簡単だが，少し頑張って兵士と戦えば，貴族から略奪行為をすることができ，より多くの戦利品が入手できるかもしれない．ここで,「金品を略奪すること」による報酬を「任務成功重視性」,「自分が死亡すること」による報酬を「精神症傾向」と結びつけて考える．山賊の中でも勇敢な者は，任務成功重視性が高く，精神症傾向が低いので，少しでも多くの戦利品を獲得するために積極的に兵士に戦いを挑みに行くことが予想される．一方で，山賊の中でも臆病な者は，精神症傾向が高いので，兵士に負けて死んでしまうリスクを恐れ，金貨の入った袋を略奪するにとどまることが予想される．この，勇敢さや臆病さを表現するために報酬に加える倍率は，騎

士を作成するときに，勇敢な者と臆病な者に対して用意した倍率をそのまま使うことができると考えている．

前述の通り，本研究で取り扱わなかった職業に対しても，本研究の提案手法を適用することで，キャラクターらしさを獲得させることができそうである．一方で，本研究で取り扱った職業である「騎士」が，全く異なるシチュエーションで登場することがあり得る．さらに，そのシチュエーションは戦闘要素が一切ない，日常生活シーンであるかもしれない．そのため，本研究の提案手法を用いて「騎士の日常生活シーン」における表現を考える．騎士は，戦いのない日には，町で見回りをしたり，城の門番役になったりすることが多いことが予想される．このとき，騎士は戦い出なくても，誠実に仕事をしているようにふるまうのが望ましい．具体的には,「決められた持ち場を離れない」「怪しい人物を見つけたら，その人物を警戒するようにふるまう」といったものである．このとき，普通の騎士はこの通りにふるまうのに対し，英雄気取りの騎士は，仕事の地味さを嫌がって，持ち場を離れてしまうことがあるかもしれない．そのような表現も，キャラクターらしいといえる．このような，自分の仕事を完遂することの報酬は，ビッグファイブ理論における「誠実性」とのつながりがあることが予想され，それにより報酬の倍率が決定できると考えている．

また，日常生活でのみ表現できる性格表現もあると考えている．例えば,「世間話好き」や「すれ違ったら挨拶をしてくれる」という点である．これらを報酬に直すとしたら,「1 日につき特定の回数まで，ほかのキャラクターと軽い雑談をしたら+1」などがある．これらの報酬は，キャラクターの「積極交流性」やビッグファイブ理論における「外向性」とのつながりの強さから，性格表現のための倍率を決定することができるのではないかと考えている．

第8章　おわりに

本研究では，ゲーム AI がキャラクター設定に沿った振る舞い，特に戦闘シーンに表れるキャラクターらしい行動を自動獲得することを目的とした．本研究では，キャラクター設定を「職業」「性格」の 2 点に分け，これらをもとに報酬設計を適用した強化学習を行うことで，目標の達成が可能であると考えた．

提案手法を用いて，様々な性格の「騎士」キャラクターの作成を行った.「ドラゴンと暗殺者から，王様を守る騎士」というシンプルな環境の元ではあるが,「勇敢 2, 臆病 1, ドラゴン HP4」というベース設定では，勇敢な騎士は「自分の命を犠牲にしてドラゴンに攻撃しに行く」のに対し，臆病な騎士は「攻撃は自分が死なない程度にとどめる」という，勇敢な性格らしい行動と臆病な性格らしい行動の差を確認することができた．

さらに，騎士チームの構成や，敵の HP など，シチュエーションを変更した環境を，ベース設定以外に 6 つ用意し，それぞれの騎士がキャラクターらしい行動を獲得できているかの確認を行った．その結果,「攻撃しなくてもクリアできるなら，自分が死なないように逃げる臆病な騎士」や「普通の騎士と勇敢な騎士がそれぞれ敵の対処の役割分担をする」といったキャラクターらしい表現の獲得に成功した．一方で，英雄気取りな騎士に関する実験では，報酬設計や，環境設定に問題があり，予想に反して「普通の騎士との違いがわからなくなってしまう」という課題もあった．

本研究で提案した手法を発展させることで，他の職業，性格を表現することはもちろん，日常生活シーンなど，戦闘シーン以外のシチュエーションでもキャラクターらしさの表現をすることも考えていきたい．しかし，本研究では「騎士」という職業の戦闘シーンにおける表現にとどまり，さらに広い表現に至ることができなかった．また，用意した AI の評価に関して，統計データと自分自身の目視による確認にとどまり，被験者実験による評価に至らなかった．そのため，戦闘以外のシチュエーションに対しても本研究の手法を適用すること，より多くのキャラクター表現を可能とすること，被験者実験を行い，実際にプレイヤーが見たときに伝わるようなキャラクター表現ができているか確認することの 3 点を今後の課題としたい．

付 録A　その他統計データ

本付録では，論文本編中に示したもの以外で，収集したキャラクター AI の統計データを掲載する．

A.1　ベース設定でのテスト結果：各騎士の攻撃状況

表 A.1 に，ベース設定である「勇敢 2 臆病 1 ドラゴン HP4」においてテストを行った際の，各騎士のエピソード終了時の攻撃状況を示す．

表 A.1: 学習完了モデルのテスト時における各騎士の攻撃状況

	勇敢な騎士 1 の攻撃状況	勇敢な騎士 2 の攻撃状況	臆病な騎士の攻撃状況
攻撃しなかった	40	27	156
ドラゴンに 1 点攻撃した	9	5	102
ドラゴンに 1.1~1.9 点攻撃した	80	73	7
ドラゴンに 2 点攻撃した	241	281	3
暗殺者に 1 点攻撃した	24	23	213
ドラゴンに 1 点暗殺者に 1 点攻撃した	97	87	16
ドラゴンに 1.1~1.9 点暗殺者に 1 点攻撃した	10	5	4

表 A.1 からは，勇敢な騎士はドラゴンに対して 1.1 点以上，特に 2 点のダメージを与えることが多く，臆病な騎士は暗殺者に攻撃する，もしくは一度も攻撃しないことが多いことがわかる．論文本編の 6.2.4 節では，勇敢な騎士は，序盤の行動でドラゴンに向かうことが多く，臆病な騎士は序盤の行動で王様に向かうか，ド

ラゴン，王様のどちらにも向かわないことが多いことがわかった．これらから，勇敢な騎士は「自分の命を犠牲にしても王を守るために戦う」というキャラクターらしさが，臆病な騎士は「王様を守りながらも自分が死なないように立ち回る」というキャラクターらしさが表現できていることがわかる．また，臆病な騎士は，ドラゴンに 1 点だけ攻撃することも多いことがわかる．これは，6.2.4 節で述べたように，臆病な騎士が序盤に逃げ出してしまった場合，代わりに勇敢な騎士が暗殺者対処を行うことがあるためである．実際に，表 A.1 からは，勇敢な騎士がドラゴンと暗殺者にそれぞれ攻撃するケースも多いことがわかる．そのような状況では，ドラゴンの HP は勇敢な騎士 2 人の攻撃の結果，1 だけ残っていることが多い．HP が 1 しかないのであれば，臆病な騎士はドラゴンに攻撃しても死亡することはない．そのため，臆病な騎士がドラゴンに 1 点だけ攻撃することが多かったのだと考える．この点からも,「王を守りたいが，自分が犠牲になるような方法はとりたくない」という臆病な騎士らしさが表現できていることがわかる．

参考文献

[1] Campbell, Murray, A. Joseph Hoane Jr, and Feng-hsiung Hsu. ”Deep blue.” Artificial intelligence 134.1-2 (2002): 57-83.

[2] Silver, David, et al. ”Mastering the game of Go with deep neural networks and tree search.” nature 529.7587 (2016): 484-489.

[3] Vinyals, Oriol, et al. ”Grandmaster level in StarCraft II using multi-agent reinforcement learning.” Nature 575.7782 (2019): 350-354.

[4] 藤井叙人, et al. ”生物学的制約の導入による 「人間らしい」 振る舞いを伴うゲーム AI の自律的獲得.” ゲームプログラミングワークショップ 2013 論文集 (2013): 73-80.

[5] Schrum, Jacob, Igor V. Karpov, and Risto Miikkulainen. ”UT^2: Human-like behavior via neuroevolution of combat behavior and replay of human traces.” 2011 IEEE Conference on Computational Intelligence and Games (CIG’11). IEEE, 2011.

[6] Oxford English Dictionary, s.v. “dragon (n.1), sense 2.a,” December 2023, https://doi.org/10.1093/OED/1117915186.

[7] The Elder Scrolls, Bethesda Softworks LLC, https://elderscrolls.bethesda.net/ja (access: 2023-01-12).

[8] Fallout 4: Game of the Year Edition - Fallout 76 - Bethesda.net, Bethesda Softworks LLC,https://fallout.bethesda.net/ja/games/fallout-4 (access: 2023-01-15).

[9] 究極の『サイバーパンク 2077』が今ここに - Cyberpunk 2077, CD Projekt S.A., https://www.cyberpunk.net/us/ja/ (access: 2023-01-12).

[10] 池田裕太郎, and 池田心. ”アクションゲームにおける特定のプレイヤの特徴を模倣する AI プレイヤの作成.” 情報処理学会研究報告 (2020).

[11] McIlroy-Young, Reid, et al. ”Aligning superhuman ai with human behavior: Chess as a model system.” Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020.

[12] 坂田駿允, et al. ”対戦型ゲームにおける個人の性格や能力を反映した AI プレイヤーの作成.” 人工知能学会全国大会論文集 第 33 回 (2019). 一般社団法人人工知能学会, 2019.

[13] Goldberg, Lewis R. ”An alternative “description of personality”: The Big-Five factor structure.” Personality and Personality Disorders. Routledge, 2013. 34-47.

[14] Juliani, Arthur, et al. ”Unity: A general platform for intelligent agents.” arXiv preprint arXiv:1809.02627 (2018).

[15] Schulman, John, et al. ”Proximal policy optimization algorithms.” arXiv preprint arXiv:1707.06347 (2017).

[16] Haarnoja, Tuomas, et al. ”Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor.” International conference on machine learning. PMLR, 2018.

[17] ML-Agents 2.0 がリリースされました — 複雑な協調動作のサポートを追加, Unity Technologies, https://blog.unity.com/ja/engine-platform/ml-agents-v20-release-now-supports-training-complex-cooperative-behaviors (access: 2023-01-15).

[18] Cohen, Andrew, et al. ”On the use and misuse of absorbing states in multi-agent reinforcement learning.” arXiv preprint arXiv:2111.05992 (2021).

[19] Foerster, Jakob, et al. ”Counterfactual multi-agent policy gradients.” Proceedings of the AAAI conference on artificial intelligence. Vol. 32. No. 1. 2018.

[20] デジタル大辞泉, “山賊”, https://www.weblio.jp/content/山賊?dictCode=SGKDJ (access: 2023-01-24).