

不確定ゲームにおけるエージェントの拡張ベイズ行動モデル

Extended Bayesian Model for Agent Behavior of Incomplete Information Games

笹井一人^{1*} 関根拓人¹
Kazuto Sasai¹ Takuto Sekine¹

¹ 茨城大学大学院理工学研究科

¹ Graduate School of Science and Engineering, Ibaraki University

Abstract: In contrast to Bayesian inference, which infers probabilistic models from observed data, extended Bayesian inference (EBI) combines inverse Bayesian inference that changes the hypotheses of the probabilistic models themselves. EBI shows the capability to adapt to rapid environmental changes by using changes in the model itself. In this study, we apply extended Bayesian reasoning to uncertain games to verify its effectiveness. The result of two cases of the games with the implemented behavioral model of EBI shows the Lévy distribution in the length of the successive win and lose sequence.

1 はじめに

ゲームにおけるコンピュータプレイヤーのアルゴリズムは近年飛躍的に進化し続け、NPC (Non-Player Character) などと呼ばれ、プレイヤーと同等の存在と認識されつつある。有名な Alpha Go や将棋の電王戦では、プロの人間が敗北することの意味が、単にコンピュータゲームで負ける以上の意味を持つと考えられている。また、ポーカーなどの不確定ゲームにおいて、ナッシュ均衡解を基にして戦略を決定し好成績を示す超人間 [1] と呼ばれる AI が提案され、近年話題となってきている。しかしながら、これらのモデルは、人間とかけ離れたアルゴリズムによってゲームをプレイするため、人間と相互作用する AI として果たして適切なものかという問題も提起されている [2]。

その一方で、認知科学がこれまで取り組んで来た人間の持つ (アルゴリズム的な意味での) 特徴とな何かという問題に対して、郡司らは「天然知能」モデル [3] を提案している。天然知能モデルとは、自らの信念が描く世界の外部、つまり想定外の出来事をノイズや外乱として扱うのではなく、それらに対して徹底した受動性を持つ (積極的に取り込む) モデルである。さらに郡司らは、天然知能モデルを表現する一例として、確率推論モデルであるベイズ推論を拡張した拡張ベイズ推論を提案している [4]。本研究では、超人間 AI では解決不能な、人間と共存共栄する AI のアルゴリズムにつ

いて検討するために、拡張ベイズ推論をエージェントの行動モデルとして応用し、人間との相互作用によって、どのような現象が生じるかを調査する。

2 拡張ベイズ行動モデル

本研究では、エージェントの行動モデルとして拡張ベイズ推論に基づく行動モデルを提案する。拡張ベイズ推論 (Extended Bayesian Inference, EBI) は、郡司ら [4] によって提案された、観測結果に基づく確率事象の推論モデルであるベイズ推論に、アレッキの提唱した人間の創造性をモデル化した逆ベイズ過程を組み合わせた推論方式である。これまでに、人間の行動選択予測や動物の群れ行動の分析などに応用され、成果を上げているモデルである。拡張ベイズ推論の特徴は、従来のベイズ推論では固定される必要のあった仮説自体を、逆ベイズ過程により変更することにある。仮説を変更することによって、観測対象の持つ真の確率分布が変化するような環境に対しても適応的に振る舞うことが可能となる。

図 1 に拡張ベイズ推論のアルゴリズムを示した。まず、ここでベイズ推論を、対象の確率モデルに適合する仮説を選択する過程とみなして、 $P^{t+1}(h) = P^t(h|d^t)$ と定義する。つまり、観測によって得られた事後分布を次の時間の事前知識 (分布) として用いる推論である。逆ベイズ推論 (過程) は、ベイズ推論に対向するように定義することができる、つまりベイズの定理において、ベイズ推論の更新式に出てこない要素に関す

*連絡先: 茨城大学大学院理工学研究科
〒316-8511 茨城県日立市中成沢町 4-12-1
E-mail: kazuto.sasai.z@vc.ibaraki.ac.jp

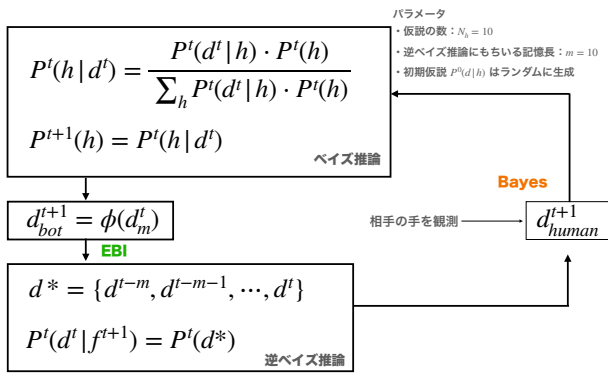


図 1: 拡張ベイズ推論の手順

る時間発展 $P^t(d|h^t) = P^{t-1}(d)$ を構成する。ベイズ推論において $P(d|h)$ は尤度と呼ばれ、確率モデルの選択肢に当たる部分であり、問題の定義そのものであるため、通常は推論過程においては不変の要素でなければならぬものであるが、逆ベイズ過程によってこれが変更されることで、想定外への適応過程が明示的に実装されている形式となっている。

3 Rock-Paper-Scissor (RPS)

Rock-Paper-Scissor (RPS) とは、いわゆる「じゃんけん」ゲームのことである。グー (Rock), チョキ (Scissor), パー (Paper) には, $r \rightarrow s \rightarrow p \rightarrow r$ の順で強弱が決まっており, cyclic game などとも呼ばれる。RPS は, マルチプレイヤーの条件で, 生態系モデルなどのシミュレーションに用いられることが多いが, ここでは, より日本的な「ジャンケンゲーム」として, 不確定ゲームの例に用いる [5].

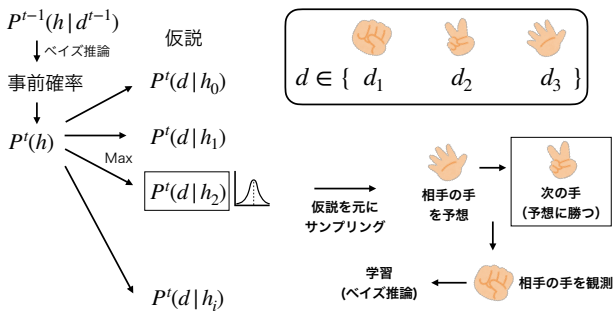


図 2: ベイズ推論モデルの実装

図 2 に RPS ゲームにおけるベイズ推論モデルを示す。グー・チョキ・パーのそれぞれを d_i として, 仮説 $d(d|h)$ をその確率分布とする。各回において相手の出した手を元にベイズ推論を行う。エージェントの行動モデルは, 仮説の確率 $P(h)$ の最も大きな h を選択し,

その確率分布によって生成された手を予想手として, それに勝つ手を次回出す。このようにして実装したジャンケンゲームを用いて, 実験を行うことにより, アルゴリズムの比較を行った。ジャンケンゲームはコマンドラインインタフェースを用いて作成した。

本実験では, (a) 人間 vs 人間 (この場合のみ通常のかけ声を用いたジャンケンゲーム), (b) 人間 vs ランダム, (c) 人間 vs ベイズのみ (Bayesian-Only Inference, BOI), そして (d) 人間 vs 拡張ベイズ推論 (EBI) の 4 種類のケースを比較した。本研究では, 勝ち負けのステップ長を分析した。人間同士の繰り返しジャンケンにおいては, しばしば勝ち負けが長く連続する場合が見られる。ジャンケンゲームのナッシュ均衡解は, 一様分布つまり等確率となることから, 勝ち負けは通常長くは続かないことが予想されるが, 特に動物の行動に見られるランダム・ウォークモデルであるレヴィ・ウォークは無制限長の連続行動を含むモデルであり, ここでもそのような性質が見られるのではないかという仮説の元にこれを調べたものである。

図 3 に, 勝ち負けの連続したステップ長の分布を示した。(a) の人間同士のゲームでは, 勝ち負けの連続ステップにおいて, べき分布が見られ, 仮説のようにレヴィ・ウォーク的な性質が存在することを示唆していることがわかる。これに対して, (b),(c) のランダム・BOI モデルではべき分布が見られないが, (d) の EBI モデルでは, (a) の場合に類似したべき分布形状がみられることがわかる。また, (b),(c) においても, 右側の長い部分で少し裾野が伸びているように見えるが, それ以外のほとんどの部分がきれいにカーブしていることから, 大部分の性質がポワソンの分布となっていることを示しているといえる。本結果から, EBI モデルが人間同士の結果に近い性質を示していることが分かった。また, 実験参加者の感想によれば, EBI モデルによるエージェントは, 「何か考えている」な, 「裏をかこうとしている」ように感じられると述べられてもいた。

4 Card-Chicken-Race (CCR)

本節では, 前節で示した RPS ゲームに加えて, 新たに定義した拡張形のゲームであるカードチキンレースゲームを用いた実験について説明する。図 4 にカードチキンレース (Card-Chicken-Race, CCR) ゲームの概要を示す。まず, 2 人のプレイヤーが衝立をはさんで向かい合い, それぞれが 10 枚のカードを持っているとする。互いにタイミングを合わせて, 「1, 2, ...」と声に出しながらカードを置いていく。このとき, プレイヤーは声だけでカードを出さないことができる。10 まで数えて, 最終的に出したカードが多い方が勝ち, そ

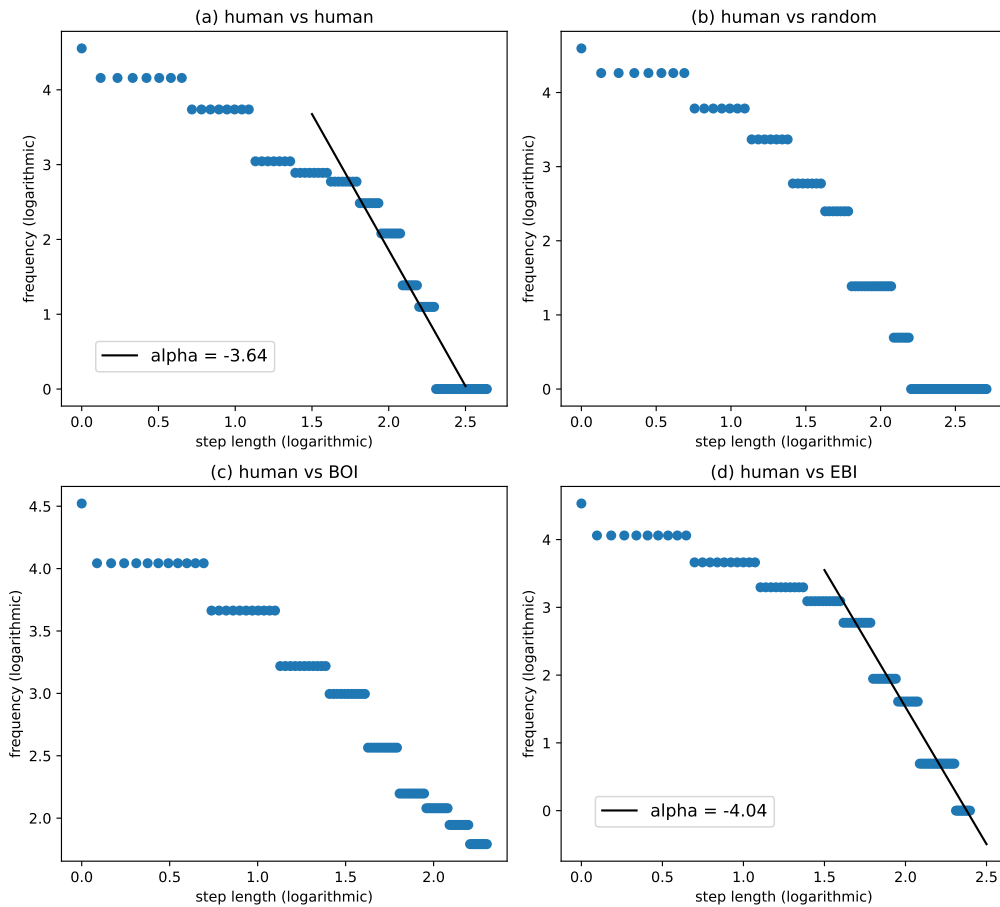


図 3: 勝ち負けの連続分布

の際に手元に出さずに残したカードの枚数分だけポイントがもらえる仕組みである。本研究では、このゲームをコンピュータプレイヤー（エージェント）と対戦するゲームを作成し、これをプレイしてもらう実験を行った [6]。

CCR ゲームは、RPS ゲームが $|d| = 3$ であったのに対して、 $|d| = 10$ のゲームであることがわかる。よって、RPS ゲームに用いたベイズ戦略を拡張することで、適応可能なゲームである。具体的には、手札の予想までは、ベイズ推論モデルをそのまま使用することができ、予想された相手の手札の数に対して、1 多い数の手札を出すことで、本稿における行動モデルとした。しかしながら、本モデルは完全に人間同士のプレイを模倣していない可能性がある。なぜならば、人間同士の場合は、声を出しながらカードを置く際に、相手が出たかどうかを予想しながら置くのに対して、エージェントは各回の最初に既に置く枚数を決めてしまっているからである。現状の行動モデル設計はシンプルさを重視して定義しているが、今後より詳細な検討が必要であると考えている。

図 5 に、勝ち負けの連続したステップ長の分布を示

した。ここでも (d) の拡張ベイズ推論モデル (EBI) の場合にべき分布が見られ、レヴィ・ウォーク的な戦況の変化が見られていると考えられる。べき分布は、戦略パターンの多様な表れを示しているともいえる。EBI の柔軟かつクイックな適応性が、レヴィ的な戦況変化に影響を及ぼしているのではないかと推測される。

また、本稿では誌面の都合で割愛したが、戦略変更を相関分析で解析したところ、人間同士のプレイに見られるのと同様な多彩な戦略変更のパターンが EBI エージェントとのプレイにも見られた [6] ことから、多様な戦略の存在を背景とした、勝ち負けのパターンが生成されているのではないかと考えることができる。よって、EBI 行動モデルは、人間らしさとその意図のからみ合いによる面白さの向上に関して一定の効果をもたらす可能性があることが見いだされた。

5 おわりに

本稿では、人間と相互作用するエージェントの行動モデルとして拡張ベイズ推論を用い、2つのゲームにお

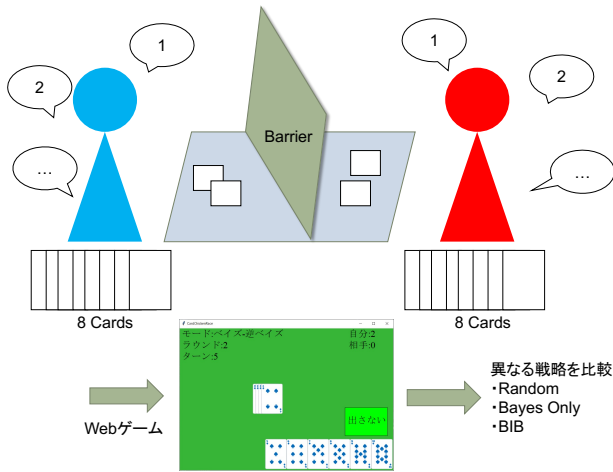


図 4: CCR ゲームの概要

ける結果を分析した。分析の結果、人間同士のプレイに見られるレヴィ・ウォークが拡張ベイズ推論の場合に類似してみられたことから、エージェントの行動モデルとして用いることに一定の意義があることが示唆された。

謝辞

本研究は JSPS 科研費 JP22K12143 の助成を受けたものです。

参考文献

- [1] Brown, N., Sandholm, T.: Superhuman AI for multiplayer poker, *Science*, Vol. 365, No. 6456, pp. 885–890 (2019)
- [2] Gupta, R. et al.: Aligning Superhuman AI with Human Behavior, *Proc 26th Acm Sigkdd Int Conf Knowl Discov Data Min*, pp. 1677–1687 (2020)
- [3] Gunji, Y.P., Nakamura, K.: Psychological origin of quantum logic: An orthomodular lattice derived from natural-born intelligence without Hilbert space, *Biosystems*, Vol. 215, pp. 104649 (2022)
- [4] Gunji, Y.-P., Shinohara, S., Haruna, T., Basios, V.: Inverse Bayesian inference as a key of consciousness featuring a macroscopic quantum logical structure, *Biosystems*, Vol 152, pp. 44–65 (2017)

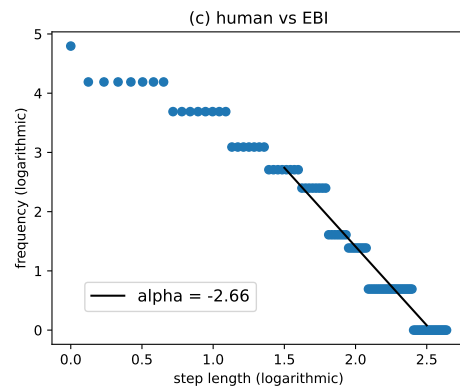
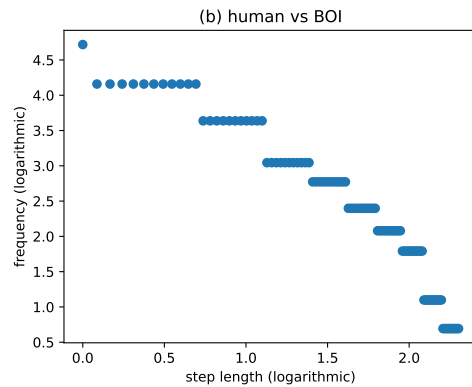
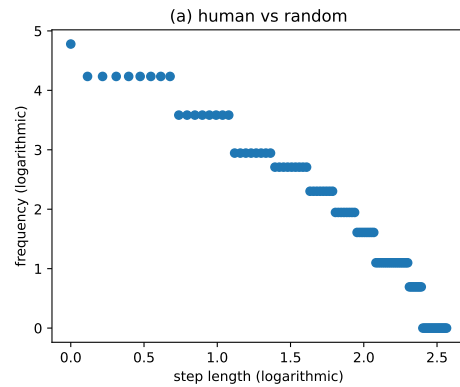


図 5: 勝ち負けの連続分布

- [5] Sasai, K., Hanazaki, R.: Affective Interaction Between Human and Agent via Inverse Bayesian Inference, *Proc the Joint Symposium of AROB 27th, ISBC 7th, and SWARM 5th*, pp. 1518–1522 (2022)
- [6] Sekine, T., Sasai, K.: Behavior Comparison in Game of Chicken with Cards Based on Bayesian Inference Models, *Proceedings of the Joint Symposium of AROB 28th, ISBC 8th, and SWARM 6th*, pp. 1349–1353 (2023)