# 研究レポート　No.16　 ～AI を制御する AI：忖度 AI～

2021 年 1 月 18 日　株式会社アイズファクトリー　https://bodais.com/company/

## 概要

人の教育や技能訓練において、学習者の技量や意向を忖度し、技量に応じた対応を行う AI を忖度 AI という。これらの機能を実装したゲーム AI を理解するうえで、関連技術の理解は有益である。本稿では、ゲーム情報の種類と構築方法の関係性、人の実力を示すレーティングに関する数理の概略を紹介する。

### 1.　忖度 AI の要件

「人間中心の AI 社会原則」では「AI は、人間の労働の一部を代替するのみならず、高度な道具として人間を補助することにより、人間の能力や創造性を拡大することができる」とされている [1]。人の能力や創造性を拡大させる活動には、教育や技能訓練などがある。忖度 AI（登録商標第 6282763 号）は、学習者の技量や意向を忖度し、その能力や創造性の拡大を支援すること、すなわち人間の成長支援を目的とした AI である。[2]

成長支援を行うには、学習者の技能水準がどのレベルにあるかを知り、レベルに応じた適切な対応を行う必要がある。初学者に対して上級者向けの訓練メニューを示しても意味がない。初学者に対しては興味を維持できるように平易なメニューを示し、上級者に対しては取り組み甲斐のある実力錬成のためのメニューを提示するという具合である。

ゲームを例にとると、既存のリバーシ AI は、すでに、人間のレベル以上の実力を備えているが、一般の人が楽しめる（創造性を刺激する）ゲームとして成立させるためには、リバーシ AI の手を制御する必要がある。従って忖度 AI は、予め人間のレベルを設定せずともリバーシ AI を自動制御し、人間と互角の対戦を実現するという特徴を備えており、実例が [2] に公開されている。

ゲーム忖度に上述の機能を実装するにあたって、関連する背景技術を整理しておくことは有益であるため、本稿では最初に、通常の AI に実装されている最適解探索機能に関連して、ゲーム展開の種類と AI ゲームの構築方法の関係性について概括する。後続の章では、レーティングに関する数理を紹介する。

### 2.　ゲームの分類

本章では、ゲームの分類軸として完備・完全・確定という概念を説明する (表 1)。完備とは、「経済学やゲーム理論において、参加プレイヤーが誰であるか、各プレイヤーの戦略、およびそれぞれの戦略の組み合わせが採られた場合の各プレイヤーの利得といったゲームの構造に関する情報を、全てのプレイヤーが保有している状況」を意味する [3]。よって、ゲームのルールが各プレイヤーで共有されてない場合や、相手のプレイヤーの打ち筋がわかない場合を不完備情報ゲームと呼び、ルールに基づき各プレイヤーが最善手を打つ場合は、完備情報ゲームと呼ぶ。一般的な AI ゲームでは完備情報を前提として構築されている。

完全とは、「展開型ゲームにおいて、すべての意思決定点において、それまでにとられた行動や実現した状態に関する情報がすべて与えられていること」を指す。完全情報ゲームとは、情報集合がすべて 1 点からなっており、どのノードにおいてもそこで手番をもつプレイヤーがそれまでの歴史を完全に把握できるようなゲームである [4]。即ち、自分以外の他のプレイヤーの手が公開されるゲームであり、チェス、将棋、囲碁やバックギャモンなどが挙げられる。ここで、バックギャモンのようにサイコロなどによるランダム要素が含まれてもよいことに注意する。他方、不完全情報ゲームは、軍人将棋、ポーカー、麻雀などのように、相手の手や選択するカードなどが伏せてあり、情報の一部が非公開のゲームが該当する。更に完備性をも考慮した分類も可能で、麻雀の例で言えば、何度も対戦したことのある相手とのゲームでは、相手の打ち筋もわかるため、完備・完全情報ゲームであり、初めて対戦する相手の場合、相手の打ち筋もわからないため、不完備・完全情報ゲームとなる。

確定情報とは、「サイコロを振った時の目の数に応じてコマを進める」や「シャフルしたトランプの束からカードを引く」などのランダム要素がある場合を不確定といい、ランダム要素がない場合を確定という。チェス、将棋、囲碁、軍人将棋は確定ゲームであり、バックギャモン、ポーカー、麻雀は不確定になる。

完備・完全・確定ゲームにおいて、一般的な AI ゲームでは、ゲーム木と呼ばれるツリーでゲームの展開を記載可能であり、このツリーに基づき AI ロジックを組んでいる。例えば、ミニマックス法 [5]によりオセロや将棋などの AI を作成できる。

**表 1: 完備・完全・確定性とゲームの事例**

| 完備性 | 完全性 | 確定性 | ゲーム名 |
|---|---|---|---|
| 完備 | 完全 | 確定 | チェス、将棋、囲碁 |
| 完備 | 完全 | 不確定 | バックギャモン |
| 完備 | 不完全 | 確定 | 軍人将棋 |
| 完備 | 不完全 | 不確定 | ポーカー、麻雀 |

### 3.　イロ・レーティング

イロ・レーティングとは、対戦型の競技（2 人のプレイヤーまたは 2 つのチームが対戦して勝敗を決めるタイプの競技）において、相対評価で実力を表すために使われる指標の一つである [6]。ここで、「イロ」は考案者の人名、アルパド・イロ(物理学者) に由来する。イロ・レーティングの利用は、チェスで始まり、将棋や囲碁、サッカー、ラグビーのような現実世界（オフライン）での試合だけでなく、インターネットを使用した対戦型のオンラインゲームでも使用されている。以下、イロ・レーティングという代わりに、簡便的にレーティングと称する。

あるプレイヤーのレーティングを$R$とすると、そのプレイヤーが平均的なプレイヤーと対戦した場合に予想される勝利確率を$W$, 敗北確率を$L\ (= 1 - W)$とすると、$R$は、

$$R = a\ \log_{10}\frac{W}{L} + b$$

で定義される。ここで、$a$と$b$は任意の定数である。プレイヤー$A$とプレイヤー$B$の対戦を考えた場合、プレイヤー$A$がプレイヤー$B$に勝利する確率を$W_{AB}$、敗北する確率を$L_{AB}(= 1 -$

$W_{AB} = W_{BA}$)とする。３人のプレイヤー$X, Y, Z$の対戦の組合せを考えた場合$W_{XY}, W_{YZ}, W_{XZ}$の関係に下記の式が成立することを仮定する。

$$\frac{W_{XZ}}{L_{XZ}} = \frac{W_{XY}}{L_{XY}} \cdot \frac{W_{YZ}}{L_{YZ}} \left(\Leftrightarrow \frac{W_{XZ}}{W_{ZX}} = \frac{W_{XY}}{W_{YX}} \cdot \frac{W_{YZ}}{W_{ZY}}\right)$$

ここで、プレイヤー$Y$を平均的なプレイヤーとすると、

$$\frac{W_{XZ}}{W_{ZX}} = \frac{W_{XZ}}{L_{XZ}} = \frac{W_{XY}}{L_{XY}} \cdot \frac{L_{ZY}}{W_{ZY}} = \frac{10^{\frac{R_X-b}{a}}}{10^{\frac{R_Z-b}{a}}} = 10^{\left(\frac{R_X-R_Z}{a}\right)}$$

となる。$W_{XZ} + W_{ZX} = 1$より、

$$W_{XZ} = 10^{\left(\frac{R_X-R_Z}{a}\right)} W_{ZX} = 10^{\left(\frac{R_X-R_Z}{a}\right)} \cdot (1 - W_{XZ})$$

$$W_{XZ} = \frac{10^{\left(\frac{R_X-R_Z}{a}\right)}}{1 + 10^{\left(\frac{R_X-R_Z}{a}\right)}} = \frac{1}{10^{\left(\frac{R_Z-R_X}{a}\right)} + 1}$$

となり、プレイヤー$X$のレーティング$R_X$とプレイヤー$Z$のレーティング$R_Z$の差$(R_z - R_X)$から勝利確率$W_{XZ}$を求めることができる。

次に、定数$a$と$b$の意味を考える。平均的なプレイヤーであれば、$W = L$が成立すると考えられるため、

$$R = a \log_{10} \frac{W}{L} + b = b$$

となる。つまり、定数$b$は、平均的なプレイヤーのレーティングを意味することがわかる。また、$a$ = 400、プレイヤー$X$とプレイヤー$Z$のレーティング差が 100 (=$R_z - R_X$) の場合、

$$W_{XZ} = \frac{1}{10^{\left(\frac{R_Z-R_X}{a}\right)} + 1} = \frac{1}{10^{\left(\frac{100}{400}\right)} + 1} \approx 0.35 \qquad (1)$$

であるから、3 回に 1 回程度は、レーティングが低いプレイヤー$X$が、レーティングの高いプレイヤー$Z$に勝てることを意味する。従って、$a$ はレーティング差と実力差（勝率）の関係を調整するパラメータと言える。

新たに参加したプレイヤーのレーティングの評価や、プレイヤーの実力が変化した場合のレーティングの評価のために、レーティングの値は一定ではなく、試合ごとに変化させる必要がある。新規プレイヤーのレーティングをある値（例えば、平均的なプレイヤーのレーティング $R = b$）を初期値としても、十分な回数の試合後は、この新規プレイヤーの実力に相当するレーティング $R(\neq b)$に収束するように調整しなければならない。

レーティング更新の仕組みを見るために、プレイヤー$X$とプレイヤー$Z$が、複数回($N$回)対戦することを考える。勝率は、式(1)で決まる$W_{XZ}$であり、$N$回の試合後の勝利回数の期待値は、$N \cdot W_{XZ}$回である。実際に勝った回数を$M$回とすると、期待値と実際に勝った回数の差 $M - N \cdot W_{XZ}$ に比例する値を現状のレーティング$R$に加えてレーティングを更新する。つまり、更新後のレーティングを$R'$とすると、

$$R' = R + K(M - N \cdot W_{XZ})$$

となる。ここで、$K$は一回の試合でレーティングを変化させる係数であり、$K$が大きい場合にはレーティングの収束が早くなるメリットがあるが、レーティングのぶれ（上げ・下げ）による不安定さが増す。一般的には$K$ = 32 をとることが多い [6]。

## 4. 段位とレーティングの関係

日本では、プレイヤーの実力評価方法として段位や級位による評価基準が有名である。一方、レーティングはチェスで始まり、今では、将棋においてもレーティングで評価することもある。本節では、レーティングと段位・級位という 2 つの異なる評価基準の関係を、実際の将棋試合結果を用いて、分析を行う。

参考文献[7]に記載されている将棋の段位とレーティングのデータ（サンプル数:172 人）を散布図で表示した(図 1)。横軸に段位、縦軸にレーティングを示す。段位と、レーティングには正の相関が期待されたが、興味深いことにその関係が顕著でない。実際、レーティングと段位の関係を線形近似すると図 1 の赤線のようにほぼレーティングと段位の関係は一定である。レーティングを$y$, 段位を$x$ で表現した式($y = a\,x + b$)で回帰分析をした結果、帰無仮説「$a = 0$である」は P 値 0.71 となるため、有意水準 5%で棄却できない。また、図 1 を見ると、段位があがるごとにレーティングの分散が大きくなっており、両者の関係性が薄いことがわかる。これは、段位の評価基準とレーティングの評価基準が異なるために生じた現象と考えられる。

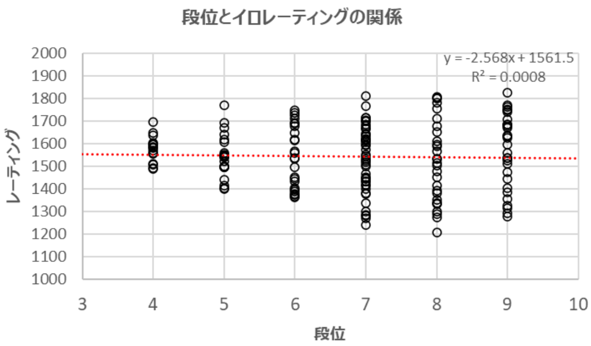


**図 1: 将棋での段位とレーティングの関係（サンプル数 172 人）**

## 5. まとめ

本稿では、忖度 AI の要件に関する理論的側面に焦点を絞り、関連技術（ゲーム情報の種類と構築方法の関係性、プレイヤーの実力を示すレーティングの数理）について解説を行った。人間中心の AI 社会原則に基づいた「AI を制御する AI、人に優しい AI」を構築するためには、高度な道具としての技術的側面と、人間の創造性を拡大していくソフトな側面を調和させていくことが大事である。

## 6. 参考文献

[1] 人間中心の AI 社会原則と AI 政策，須藤修，イノベーション政策強化推進のための有識者会議「AI 戦略」(AI 戦略実行会議)，第 3 回 資料 1，令和 2 年 https://www.kantei.go.jp/jp/singi/ai_senryaku/suuri_datascience_ai/dai3/siryou1.pdf
[2] 忖度 AI リバーシ，株式会社アイズファクトリー，https://bodais.com/rd/demo/sontaku-ai-reversi/
[3] 完備情報，フリー百科事典 ウィキペディア日本語版，https://ja.wikipedia.org/wiki/完備情報
[4] 完全情報ゲーム，フリー百科事典 ウィキペディア日本語版，https://ja.wikipedia.org/wiki/完全情報ゲーム
[5] ミニマックス法，フリー百科事典 ウィキペディア日本語版，https://ja.wikipedia.org/wiki/ミニマックス法
[6] イロレーティング，フリー百科事典 ウィキペディア日本語版，https://ja.wikipedia.org/wiki/イロレーティング
[7] 将棋棋士レーティングランキング，https://shogidata.info/list/rateranking.html