# 人工知能によるボードゲームデザインについて

藤田　修[†1]

この論文は人工知能を利用したアブストラクト戦略ゲームのデザインについて論じる．ここで対象とするゲームは正方または六方格子状のボードとレーザー，ミラー，スプリッター等の光学部品を模した駒を使用する．プレイヤーは２人または３人以上で，各自が一個または複数個のレーザーを占有し，レーザービームを対戦相手のレーザーに照射して攻撃する．ゲームデザインとしては，ボードの形状や駒の機能と配置，プレイヤーの数，各種ルールなどについて様々なバリエーションが可能である．その中から，最適なデザインを人工知能によって探し出せるかどうかが課題である．

# Board Game Design by Artificial Intelligence

OSAMU FUJITA[†1]

This paper describes the design of an abstract strategy game by using artificial intelligence. The game uses a square or hexagonal grid board and pieces symbolizing optical components such as lasers, mirrors, splitters, etc. Two or more players each has one or more lasers, and controls them to shoot down opponents lasers. In game design, there can be a lot of variations in terms of playing fields framework, the function and configuration of pieces, the number of players, and playing rules. The problem is whether the artificial intelligence can find out the best design among these variations.

## 1.　はじめに

ボードゲームがマインドスポーツとしてエデュテインメントの分野で果たす役割は大きい．近年，注目を浴びつつあるプログラミング教育においても，主要な題材の一つになると思われる．既存の囲碁や将棋に限らず，多種多様なゲームが新たに生み出され，さらに発展する余地もある．それと同時に，急速な進歩を見せている人工知能が最強のプレイヤーとして台頭してきており，創造的なゲームデザインに活用されることも期待される．

今年（2016）は人工知能の歴史において画期的な年となった. Google の囲碁ソフト AlphaGo が囲碁の人類最強レベルのプロ棋士に勝利したという[1]．このニュースは人工知能の研究者にとっても衝撃的だったに違いない．また，時を同じくして，その陰にあってやや印象は薄いが，Boston Dynamics のヒューマノイド Atlas が自然原野の滑りやすい雪上を自律２足歩行で動き回る姿にも驚かされた[2]．デジタル計算機のソフトウェアが即応的な知的判断と運動制御の両方で人間を模倣する能力を具現的に示した点で，科学技術の急激な進歩には目を見張るものがある．とは言え，実行計算に膨大なハードウェアとエネルギーを使用している AlphaGo を，Atlas の頭部に個別に搭載することは困難で，まだまだ解決すべき課題は多い．

AlphaGo を解説した文献[1]によると，膨大な量の熟達者の棋譜を教師データとする教師あり学習と，自分の分身との対戦（セルフプレイ）において強化学習を行うディープニューラルネットワークを主体とし，それを用いてモンテカルロシミュレーションで先読みするという．熟達者の打ち方を模倣するだけではなく，過去に例のない新しい手筋を自ら探し出す革新性を兼ね備えていると推察される．明確なルールで条件が限定されたゲームにおいては，人工知能がプレイヤーとして最適な戦略を発見し，最強になりうることを示した．人工知能がこのレベルをクリアすると，次の新しいステージでは，ゲームデザインの最適化においても，航空機開発におけるテストパイロットのように活躍できる．テストしてみたい課題として，例えば，囲碁の場合，碁盤が正方格子ではなく，三角格子（または六方格子）にすると面白くなるのだろうか．プレイヤーが３人ではどうか．同等の複雑さや面白さを保ちながら，より短時間で勝敗が決まるルールに変更できないか．など，様々なバリエーションに対して，対戦シミュレーションを行い，ゲーム性の評価を自動化することはできないだろか.

この研究では，ボードゲームの題材として，囲碁や将棋などのルールが確立されているものではなく，人工知能プレイヤーの最適化も比較的容易と思われる新種のゲームを対象にする．具体的には，レーザー，ミラー，スプリッター等を模した駒を動かして，自分のレーザーから発したビームで対戦相手のレーザーを攻撃するゲームを題材とする．チェスや将棋に近いゲームである．類似のゲームがすでに”Khet”という商品名で販売されているが[3]，より単純化,

†1 大阪教育大学
Osaka Kyoiku University

抽象化したルールの典型的なゲームを第２章で紹介する．このゲームは Actionscript 3.0 でプログラミングした Flash ゲームとして，WEB 上で公開している[4]．第３章ではその様々なバリエーションを示す．第４章では人工知能による最適化について見通しを論じる．残念ながら，研究の初歩的段階にあり，結果を議論するまでには至っていない．最終目標までの道のりは遠く，困難も予想されるが，今後検討すべき課題について議論の材料としたい．

## 2. ゲームの原型

研究題材とするゲームの典型的な実例を紹介する．

### 2.1 ボードの形状

正方形のセルを縦横８×８に配置した正方格子のボードを図１に示す．便宜上，左下のセルを座標（0, 0），右下を（7, 0），左上を（0, 7），右上を（7, 7）と表示する．方向は（0, 0）から（7, 0）を 0 度，（0, 0）から（0, 7）を 90 度とする．

### 2.2 プレイヤーの数

２人．

### 2.3 プレイの順番

交互に一手ずつ，駒を操作できる．

### 2.4 駒の種類

駒はレーザー（砲）とミラーの２種類とする．レーザーは各プレイヤーが１個占有し，上下左右方向のいずれか１方向にレーザービーム（光線）を発射する．このビームは駒のない空セルは直進する．レーザーはその背面または側面にビームを受けると壊れる（ビーム発射できなくなる）．ミラーは極薄の両面鏡で±４５度に傾いており，縦方向の入射ビームは横方向へ反射し，横方向の入射ビームは縦方向へ反射する．

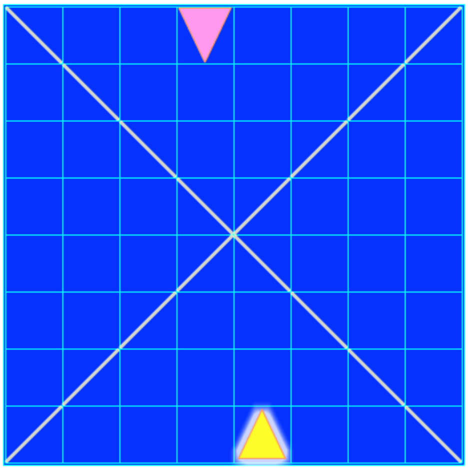

図 1　駒の初期配置

Figure 1　Initial configuration of lasers and mirrors.

### 2.5 駒の初期配置

初期配置の一例（図１）は下記の通り．

① 先手のレーザー：セル位置（4, 0），方向 90°．

② 後手のレーザー：セル位置（3, 7），方向 -90°．

③ ミラー：(0, 0) から (7, 7) の対角線上の８個のセル，鏡面 45 度．

④ ミラー：(0, 7) から (7, 0) の対角線上の８個のセル，鏡面 -45 度．

先手のレーザービームは（4, 0）から 90 度方向に直進後，（4, 3）のミラーで 180 度方向に反射され，さらに（3, 3）のミラーで反射されて 270 度方向に直進後，ボード外に出ている．

### 2.6 駒の動き

各駒は同じセル位置で１手当たり±90 度回転できる．あるいは，上下左右の隣接セルに他の駒がなければ，回転せずに１手当たり１マス平行移動できる．

### 2.7 情報表示

電子部品を使わないゲーム盤と駒を使用する場合，視覚的に明示される情報は駒の配置のみで，駒の可動範囲やレーザービームは基本的に不可視である．ビームの光路は各プレイヤーが想像しなければならない．

ビデオゲームとしてディスプレイに表示可能ならば，レーザービームの光路や次の一手の候補などを図示してもよい．

### 2.8 プレイヤーによる駒の操作

プレイヤーが操作できる駒は自分が占有するレーザーとそのビームが当たっているミラーのみ（図２）．

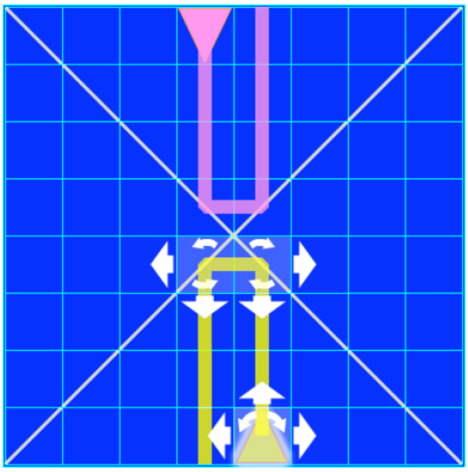

図 2　光路と初手の候補の表示

Figure 2　Choice of the first move.

### 2.9 勝敗

自分の手番の一手により，対戦相手のレーザーの背面または側面にレーザービームを当てて壊すと勝ち．ビームの発射源は問わず，対戦相手の発射したビームを利用してもよい．ただし，自分の手番で自分のビームと相手のビームを正面衝突させた場合は自分の負け（図３の自滅行為）．また，３個以上のレーザーがあって，他者のレーザー同士のビームを正面衝突させると双方相打ちで破壊されるとする．

**2.10 千日手**

相手の直前の一手を元に戻す手（一手前の局面を再現する手）は禁止．同一局面が再現された場合，前回の局面で打たれた手と同一の手は禁止．

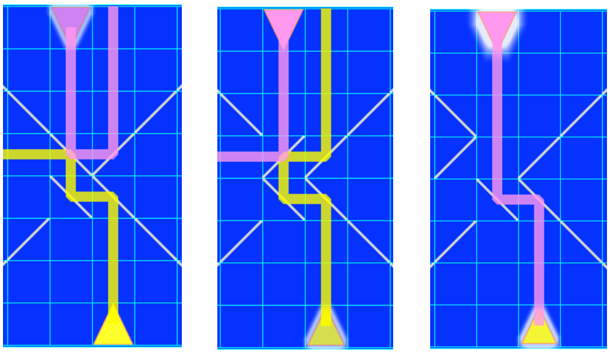

図３　進行例：左から初手（黄：ミラー回転），２手目（赤：ミラー回転），３手目（黄：ミラー移動で相打ちにより自滅，ちなみに，同じミラーの回転は千日手で反則）

Figure 3　Example of game play

## 3.　ゲームのバリエーション

各種の設定条件に対する変種を紹介する．

### 3.1　ボードの形状

外見上は六方格子（三角格子）が典型的である（図４）．ただし，セル配置は２次元座標で記述できる．その場合，正方格子上で光線の進行方向として縦横に加えて斜め 45 度方向が追加された場合と同等である．すると，正方格子において，さらに斜め 135 度方向を加えて縦横斜めの８方向の光路を認めるボードも可能である．また，２次元平面ボードを３次元立体格子に拡張することも考えられる．なお,一様な格子ではなく,光が曲線的に曲がって進んだり，界面で屈折したりするような，等角写像で描かれるような歪んだ空間格子も考えられる.ペンローズタイルも面白い．

なお，チェスにおいても，３人用のチェスのための，六角形盤（チェッカーボードの下半分を切出し，上辺の中点を引延ばして５角形（頂点の内角 120 度）に変形して，それを３面つなぎ合わせた形状[5]）や，円盤（チェッカーボード３枚分を左右の両端が重なるように円柱状に丸めて，それを平面に射影した形状[6]）などが商品化されている．

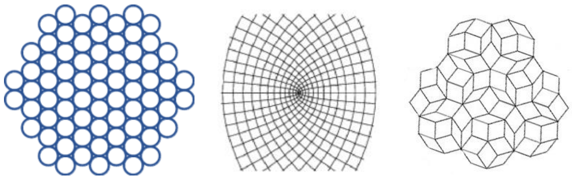

図４　各種ボード

Figure 4　Variation of grid boards.

### 3.2　プレイヤーの数

正方格子では２，４人，六方格子では２，３，４，６人．

### 3.3　プレイの順番

プレイヤーが２名ＡとＢの場合,通常は一手毎に交替する．

（1） A, B, A, B, A, B, A, B,・・・．（通常）

しかし，先手有利が明らかな場合は，何らかの緩和策を設けた方がよい．例えば,

（2） 初手に特別な制限を加える：{有利に働く手を禁じる，初期配置を工夫する}

（3） 先手の勝利の条件として、任意の手番で少なくとも１回パスする義務を課す．

（4） 後手に任意の手番で２手連続する権利を与え，権利を行使した後は，対戦相手にこの権利を移す．

具体的には,テニスのタイブレークのサービス順と同様に,

（5） A, B, B, A, A, B, B, A,・・・．

（6） 各自が２個のレーザーを所有する場合は，A のレーザーA1，A2，B のレーザー B1，B2 の操作順を A2, B1, B2, A1, A2, B1, B2, A1, A2, B1,・・・．

後手有利の場合には,

（7） 後手の勝利の条件として、任意の手番で少なくとも１回，２手連続して打つ義務を課す．

（8） 先手にパスする権利を与え，権利行使後はその権利を対戦相手に移す．ただし，連続のパスは禁止する．

複数回の勝負で競う場合は

（9） 互先で偶数回の勝負とする．

（10）敗者が次の試合の先手後手の選択権をもつ．

プレイヤーが A，B，C の３人の場合は，例えば，

（11）　A, B, C, A, B, C, A, B, C, A, B, C,・・・．（通常）

（12）　A, B, C, B, C, A, C, A, B, A, B, C,・・・．（先手の制限）

（13）　A, B, C, C, B, A, A, B, C, C, B, A,・・・．

しかし，２名が協力して１名を攻撃するような不公平が生じうる．この状況を正当な戦略とみなすゲームがあってもよいが，恣意的な連合を不正とみなす場合には不利な立場のプレイヤーに対する優遇策が必要である．例えば,

（14）盤上の駒の配置から各プレイヤーの形勢を評価して数値化する計算式を定義し，その値に基づいて最も劣勢なプレイヤーに優先的に手番を与える．ただし，２手を連続して打つ事は禁止する．

（15）各プレイヤーの手が各プレイヤーの形勢に与える影響を統計的に評価し，プレイヤー間の敵対関係の偏りを表す数値が一定値を超えた場合に，あるいは，特定のプレイヤーに攻撃が集中していると判定された場合に，劣勢なプレイヤーに特典を与える．特典としては，２手連続して打てる，駒の動きの制約を緩める，特殊な駒を使用できるなどが考えられる．

プレイヤーが４人以上になると手順による有利不利の補正はさらに難しくなる．確率的に公平化を図ることも検討課題となる．

(16) サイコロ等を利用して，次の手を打つプレイヤーをランダムに選択する．

### 3.4 駒の種類

各種の光学部品から類推すると，レーザーとミラー以外に,光路や光強度を変更する様々な駒があり得る.例えば,

(1) シャッター
(2) スプリッター
(3) 偏光板
(4) 減衰器／増幅器
(5) 波長変換器／位相変調器
(6) ミラーの数と配置

特殊な機能で意外な効果をもたらすことを期待したいが，あまり種類が多過ぎると面白さが失われるなどの逆効果もあり得る．

ゲームの進行状況に応じて，駒の機能や性質を変えることもあり得る，例えば,

(7) レーザーの破壊強度（耐性）を数値で表し，ビームを受けると数値が減少し，その値に応じて発射するビーム強度も低下する．
(8) レーザーがビームに打たれて壊れた後に，消滅する．
(9) レーザーがビームに打たれて壊れた後に，残骸として残り，ビームを遮断する障壁に変化する．

### 3.5 駒の初期配置

様々な変種を手軽に実現できるので，最初に取り組むべき検討課題として適している．例えば,

(1) 一人のプレイヤーが占有するレーザーの数と配置
(2) ミラーの数と配置
(3) ランダム配置
(4) 各プレイヤーの自由裁量配置
(5) ハンディキャップ

### 3.6 駒の動き

一手で動かせる駒の自由度を増やすとゲームの進展の速さや多様性が変化する可能性が高い．

(1) レーザー回転角：正方格子４方向 {0,±90, 180}, 六方格子６方向 {0, ±60, ±120, 180}, 正方格子８方向 {0, ±45, ±90, ±135, 180}
(2) ミラー回転角：正方格子４方向 {0, 90}, 六方格子６方向{0, ±30, ±60, 90}, 正方格子８方向{0, ±22.5, ±45, ±67.5, 90}
(3) 移動距離：隣接，縦横任意，斜め方向任意
(4) 移動方向：上，下，左，右，斜め
(5) 特定の局面における駒の生成／消滅
(6) 特定の局面における駒の機能変化
(7) 特定の局面における駒の種類の変化

### 3.7 情報表示

ゲームの進行状況の表示の仕方で面白さや難易度が変わる．情報表示を最小限にして，推理力や計算力，記憶力，忍耐力も競わせるべきか．しかし，勘違いやケアレスミスによる反則負けは不興の原因となるので，防止すべきかもしれない．

(1) レーザービームの表示：{常時，瞬時，不可視}
(2) 駒の可動範囲：{移動，回転}
(3) 千日手の警告
(4) 形勢評価（ゲームの進行規則に利用可能）

### 3.8 プレイヤーによる駒の操作

通常，自分が占有する駒を操作可能とすることは当然だが，ミラーなど，他の駒を非占有とした場合，操作可能な駒としては,

(1) 自分のレーザービームが当たっている駒のみ．
(2) 自分のレーザービームが当たっている駒と対戦相手のレーザービームが当たっていない駒．
(3) 対戦相手のレーザービームが当たっていない駒のみ．

### 3.9 勝敗

レーザービームの攻撃の有効性について，

(1) 照射の有効面（弱点）:{背面，側面，背面と側面}．
(2) 自分が発射したレーザーで自滅するか否か．
(3) １回の照射だけで決着させず，左右の側面と背面の３面をすべて照射するなどの条件の追加．
(4) レーザー強度の反映：複数回の照射の累積量が規定量に達する条件の追加．スプリッター等のビーム強度を変化させる駒を使用した場合に効果的である．
(5) レーザーが３個以上で，１つのレーザービームのミラー操作によって他のレーザー同士が正面で相打ちとなる場合，同士討ちで両者消滅とする．

### 3.10 千日手

一般的に，単純な繰り返しは禁じ手とし，複雑な手筋の繰り返しは引き分けとするのが妥当と思われる．しかし，引分けが多いと面白みに欠けるので，千日手が成立した局面における形勢判断やレーザービームの位置関係などを点数化して勝敗を決める方がよいかもしれない．

## 4. 人工知能の設計方針

いわゆる二人零和有限確定完全情報ゲームと分類されるボードゲームにおいて，プレイヤーとしての人工知能は勝負に勝つことを目的としており，その性能を表す最終的な評価関数は単純に勝率とすればよかった．将棋や囲碁で成功した人工知能は機械学習の手法を用いてそのような評価関数を最大化する行動選択関数（局面の情報を入力として次の一手を出力する関数）を獲得している．その性能が人間の能力を超えたという点で，AlphaGo はほぼ目的を達成したといえる．しかし，ここから先に，まだ達成困難な目標がある．人間の場合は，最強の競技者には指導者としての能力も求められる．今後，人工知能は一手毎に最良の手を示し，それが何故最良なのか，その理由を説明できるようになれるだろうか．様々な手筋のパターンから，抽象的な論理規則を見いだしてそれを格言にまとめたり，詰碁の

問題を作成したりできるだろうか．さらには，ゲームのルールについて価値ある修正案を提示できるだろうか．

本研究で題材とする光線ゲームは，従来の確立された人気の高いゲームとは異なり，特定の標準化されたルールはなく，ゲームとしての実績も乏しく，熟達者の棋譜など存在しない．もちろん，プレイヤーとしての人工知能もない．その状況で，いきなり，ゲームデザインのための人工知能をつくるのは無謀にも思われるが，単純で面白みに欠ける低レベルのゲームから，プレイヤーとしての人工知能をセルフプレイによる強化学習で徐々に高レベルに育てることを並行して行うことは可能であると思われる．ある意味で，「個体発生は系統発生を繰り返す」ように，これまでの人工知能の進化の過程を参考にしながら，ゲームの複雑さを高めていくことを考えたい．ゲームが比較的単純なレベルにおいて，強化学習で得られた人工知能の内部表現が論理的に解釈可能なものになっているか，それがどのような条件で得られるのか．厳密解が得られる状況でそれらを検証する価値もある．実はそちらの方がゲームデザインよりもはるかに興味深い研究課題かもしれない．

### 4.1 ゲームデザインの評価

よいボードゲームの評価基準とは何か，様々な観点で多くの議論がなされているようであるが，端的に言えば「面白い」あるいは「楽しい」というような，人間の脳に及ぼすある種の快感の程度ということになると思われる．しかし，残念ながら，心理学や神経生理学の分野で，この評価基準を客観的な数値として演繹的に決定できるような合理的知識が得られているとは思えない．最近のデータ科学の方法論で，すでに公開されたゲームに対するユーザー評価をインターネットで収集して，統計的にランク付けすることは可能であるが，公開前のゲームデザインには適用できない．そもそも，「面白い」「楽しい」というのは主観的で，「好き／嫌い」と同様，人によって感覚が大きく異なるので，単純な統計的平均で測ればよいというものではない．ほんの少数の人が面白いと思うだけでもそれなりの価値がある．

当面，「面白い」という評価基準はあきらめて，逆に「面白くない」条件を考えて許容範囲を絞ることにする．そうすると，

(1) 易し過ぎる：ゲーム木の探索空間が小さ過ぎる
(2) 難し過ぎる：ゲーム木の探索空間が大き過ぎる
(3) ルールが面倒：ゲームの制約条件が多過ぎる
(4) 不公正／不公平：先手後手の勝率差（または比）が大き過ぎる
(5) 引分けが多過ぎる

などのように，数量的に客観評価できる特性がある．これらを総合して一言でまとめれば「バランスのとれた適度な複雑性」といえるかもしれない．ただし，完全ゲーム木の単純なノード数で決まるものではなく，実現頻度等で重み付けしたものでなければならない．その判定基準となる具体的な数値は不明だが，厳密に決める必要性もない．人工知能によるセルフプレイのデータから統計的に推定すればよいと思われる．

本研究では，複雑性の数量的評価のため，集合間距離を利用する．まず，ゲーム木を有向グラフのノード集合とエッジ集合で表し，各集合の要素に関する要素間距離を定義して，その要素間距離をもとに集合間距離の統計的推定値を計算する．ただし，ノード集合の要素は単純なノードに限らず，部分木に対応するグラフを表す集合など，階層的な集合を一つの要素とみなしてもよい．ゲームの複雑性は単純なゲーム木との集合間距離で評価する．ゲームの多様性は異なるゲーム木間あるいは一つのゲーム木内の部分集合間の集合間距離で評価する．

集合を $A,B$ とした場合の集合間距離 $f(A,B)$ の計算式の一例を次式で示す．

$$(A,B)=\frac{|A/B|}{|A\cup B|}g(A/B\,,B)+\frac{|B/A|}{|A\cup B|}g(B/A\,,A)$$

ただし，$g(A,B)$ は要素間距離 $d(a,b)$ のいわゆる群平均距離

$$g(A,B)=\frac{1}{|A||B|}\sum_{a\in A}\sum_{b\in B}d(a,b)$$

で，$|A|$ は $A$ の濃度である．なお，集合 $A\cup B$ 内の要素間距離について工夫して冪平均をとる数式

$$u_{p,q}^{(00)}(A,B)=\frac{1}{p}\ln\left(\frac{1}{|A\cup B|}\sum_{\{x\in A\cup B\}}\left(\frac{I_{B\setminus A}(x)}{|A|}\sum_{\{y\in A\}}e^{qd(x,y)}+I_{A\cap B}(x)+\frac{I_{A\setminus B}(x)}{|B|}\sum_{\{y\in B\}}e^{qd(x,y)}\right)^{\frac{p}{q}}\right)$$

に書き換えると，$p\to+\infty$，$q\to-\infty$ の極限で有限集合に対するハウスドルフ距離に等価となる[7]．また，$p\to-\infty$，$q\to-\infty$ とすると最小距離になる．ただし，最小距離では三角不等式は成立しない．

集合間距離の統計的推定では，全要素を調べ上げる必要はなく，適切なサンプリングができればよい．従って，完全ゲーム木を求めず，ゲームプレイの実例を収集して，部分木を求める．その大部分は，人工知能をプレイヤーとするゲームのシミュレーションから得る．人工知能としては，ランダムな手を打つ初心者レベルから，強化学習などで勝率を高めた強いプレイヤーまでを用意する．

### 4.2 プレイヤーとしての人工知能の性能評価

ゲームプレイヤーとしての人工知能は，簡単なゲームに対しては完全ゲーム木の全探索を行うことで必勝法を得ることができる．ゲーム木の複雑性が増すと，探索範囲を限定してその範囲内をできるだけ効率よく先読みする．先読みが完全ゲーム木の葉ノードに届かない場合は，プレイヤーの勝率は先読みの正確さ，すなわち，探索範囲内の末端ノードに位置する局面の評価関数の値の正しさに依存する．

各局面の評価関数は盤面の駒の配置や位置関係に関する静的評価関数として定義することができる．将棋や囲碁の場合は，明確な根拠に基づいて演繹的に導かれた数式があ

るわけではないが，個々のゲームの特徴に適合する大雑把な近似式を仮定し，それに含まれる調整可能なパラメータを，熟達者の棋譜を教師データとする機械学習で決定している．例えば，将棋の Bonanza [8]では王を含む３個の駒の位置関係を表現する数式を用い，囲碁の AlphaGo [1]では石の配置の局所パターンを階層化された畳み込みニューラルネットで統合して全体像の特徴を認識する仕組みを用いるなど，ゲーム毎に異なる工夫が必要である．その一方で，ランダムな手を繰り返して終局まで先読みし，その結果を基に確率的に勝敗推定を行うようなモンテカルロ木探索[9]によるプレイアウトなど，比較的汎用性の高い手法が有効な場合もある．

本研究の場合は，適切な教師データがなく，また，モンテカルロ木探索によるプレイアウトがそれほど有効とも思えない．ゲームの特徴を分析して適切な評価関数を定め，その調整可能なパラメータをセルフプレイによる強化学習で最適化することを考える．

### 4.3 局面の評価

攻撃側の目的は相手のレーザーの弱点（背面または側面）にレーザービームを当てる事である．現在の局面で相手のレーザーが破壊されるビームの光路を想定し「当たりビーム」と呼ぶことにする．「当たりビーム」は相手のレーザーの背面または側面から垂直方向に逆進することで容易に光路を認識する事が出来る（図５）．ゲームの途中ではどのレーザービームも当たりビームに一致するものはないが，いずれかのレーザービームを敵レーザーの当たりビームに一致させることが駒を動かす目的となる．

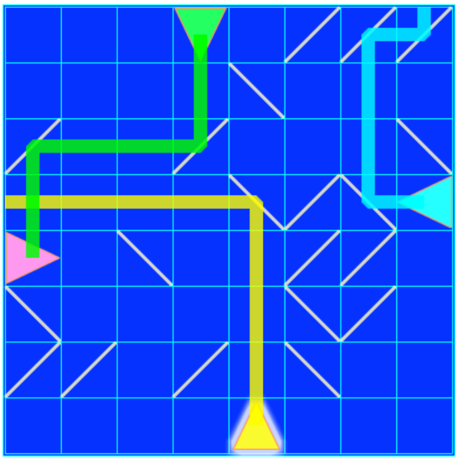

図５　当たりビーム（この図では，緑レーザーのビームが赤レーザーの当たりビームとなっている）

Figure 5　Hitting beam

局面の形勢は各プレイヤーのレーザーの当たりビームの集合と全レーザービームの集合との集合間距離で評価できる．ビームの経路は，始点となるレーザーのセル座標から，発射方向に空セルを直進して，駒に当たらなければ外縁セル座標まで到達してそのまま場外に出る．途中でミラーに当たるとミラーの角度に応じて±90 度反射する．この経路を記号で表現すると，経路に含まれる各セルの情報

（1）セルの位置座標（$x,y$）

（2）駒の種類と状態の識別記号（$k$）

（3）ビームの入出方向（$bi,bo$）

を始点から終点まで一次元配列状に並べた記号列で表される．ただし，ビームの入出方向は (1)の前後関係からわかるので，(3)を省略すると，セル情報は（$x,y,k$）でも十分．図５の例では，赤の当たりビームの記号列は

(3,8,133)(3,7,0)(3,6,20)(2,6,0)(1,6,0)(0,6,20)(0,5,0)(0,4,120)

ただし，$k = 0$：空セル，20：+45 度のミラー，120：右向の赤レーザー，133：下向の緑レーザー，を表す．なお，記号列でなく，素直に有向グラフとして表現しても良い．

ビーム間距離はこの記号列間の距離であるが，それは駒の操作の手数に比例するものでなければならない．文字列の編集距離も連想されるが，そのままでは適用できない．操作を加える位置が始点に近い程，ビームは大きく変化するので，始点をルートとして終点を末端の葉ノードとする木構造を想定し，一手の操作で枝を丸ごと交換する状況をイメージしなければならない．この状況を考えると，仮想的な一手によって変化したビームの集合がどうなるかを具体的に求めた方が良いと思われる．例えば，当たりビームの集合を $H0$ とし，その要素のビームに一手加えた（操作可能な駒に対して，それぞれ一手動かした）局面における当たりビームの集合を $H1$，同様に，当たりビームに n 手加えた場合の当たりビームの集合を $Hn$ とする．各レーザーが発射するビームの集合を $E0$ とし，同様に $n$ 手動かして出来るビームの集合を $En$ とする．集合 $Ei$ と $Hj$ の要素ビームの一部に重なりが見つかれば，その組み合わせの数と手数 $i$ と $j$ の適当な関数として，現在の局面から当たりビームを生成するまでの手数を距離とみなして計算する．この場合，最小の手数が重要なので，距離関数として三角不等式を満たす必要性はない．ただ，一手で終局する場合($i$+$j$=1)を除くと，対戦相手の手によって $Ei$ や $Hj$ とは異なる局面に遷移する可能性が高く，最小手数が実現する可能性は低い．従って，最小手数以外のビーム集合について平均的な距離を考慮すべき余地がある．その詳細につては今後の検討課題である．

時系列データのパターン認識では隠れマルコフモデルの有効性が認められている．ビームを表す記号列に対して，確率的状態遷移モデルから生成される記号列とみなして，それらの類似度を評価することを考えると，ビーム間距離の評価にあまり適しているとは思われないが，２手先，３手先の不確定性や，対戦相手の性格や癖などによる偏りを考慮する場合に有効な判断基準を示してくれるかもしれない．将来，モデル構築に必要な学習データが蓄積されれば，有効性を検証してみたい．

### 4.4 強化学習

前節の形勢評価関数の数式に調整可能なパラメータを組み込み，パラメータの異なる人工知能プレイヤー同士でゲームのシミュレーションを多数回行う．勝率に差があれば，勝率の高いプレイヤーのパラメータに近い値を採用する．パラメータの意図的変更に対して，システム性能の変化を測定し，それをフィードバックしてパラメータを最適化することができる．これはシステムの実行結果に対する報酬または罰を基に学習する強化学習の一種と言える．

これを計算式で表すと，パラメータ $p$ の人工知能の勝率が $v$ で，パラメータを $p+\Delta p$ に変化させた人工知能の勝率が $v+\Delta v$ だとすると，様々な条件で対戦させ，その平均的な相関 $\langle \Delta p \cdot \Delta v \rangle$ に比例する量だけ，次式のようにパラメータを修正する[10].

$$p_{new} \leftarrow p_{old} + \eta \langle \Delta p \cdot \Delta v \rangle$$

ただし，$\eta$ (>0) は学習係数である．この修正を繰り返すと，次第に勝率が向上する．複数のパラメータを同時に変化させ，勝率との相関から各パラメータの勝率への影響度を測定しつつ，学習過程を加速することもできる．ただし，この影響度は学習の進み具合で変化する．なお，システム性能の差分とパラメータの差分の比率 $\Delta v / \Delta p$ を，学習の目的関数のパラメータに関する勾配とみなして勾配法を適用しても良いが，パラメータの変化の与え方によっては収束性の悪化をもたらす場合があるので注意を要する．

ゲームデザインを行う人工知能においても，各種バリエーションを連続的数値または離散的数値のパラメータとして表現し，同様の手法でシステム評価関数が高くなるようにパラメータを調整できる．「面白くない」という性能をマイナスの数値で表す評価関数でどのように数式を組み立てるかについては，試行錯誤的な検討が必要と思われる．

## 5. おわりに

ゲームで遊び，さらに新種のゲームを考案することは楽しい．実生活でそれほど役に立つ訳ではないのに，生物の一形態である人間が，ゲームに何故没頭するのか，脳の発達とどう関係しているか，実に興味深い．その一方で，人工知能が驚異的な能力を示しつつある．現在の人工知能は論理的推論能力が乏しいまま，人間を外面的に模倣する能力のみ急速に発達したが，限定された具体的な目的が与えられない限り自発的な行動はできないものと思っていた．今やそれが自己組織化の能力を獲得しつつある．やがてゲームにのめり込んで時間とエネルギーを無駄に浪費する人工生命が誕生するのだろうか．その時は，おそらく，この研究も引き継いでくれるものと期待している．

**謝辞**　このゲームの試作段階でいろいろとご協力頂いた盛田宗昭さん，添本晃央さんを始めとする研究室の方々に，謹んで感謝の意を表する．

## 参考文献

1) Silver, D., et al.: Mastering the game of Go with deep neural networks and tree search, Nature 529, pp. 484–489, (2016). (http://www.nature.com/nature/journal/v529/n7587/full/nature16961.html)
2) Boston Dynamics: Atlas, The Next Generation (https://www.youtube.com/watch?v=rVlhMGQgDkY)
3) Khet (http://www.khet.com/)
4) Flash Board Game (http://www.osaka-kyoiku.ac.jp/~fuji/koushu/LBeam.html)
5) Three-player chess (https://en.wikipedia.org/wiki/Three-player_chess)
6) 3 Man Chess in The Round (http://www.3manchess.com)
7) Fujita, O.: Metrics based on average distance between sets, Japan J. Indust. Appl. Math., Vol. 30, No. 1, pp 1–19, (2013). (http://link.springer.com/article/10.1007%2Fs13160-012-0089-6)
8) Bonanza (https://ja.wikipedia.org/wiki/Bonanza) (http://www.geocities.jp/bonanza_shogi/)
9) Coulom, R.: Efficient selectivity and backup operators in Monte-Carlo tree search, In 5th Int. Conf. on Computers and Games, 72–83, (2006).
10) Fujita, O.: Trial-and-Error Correlation Learning, IEEE Trans. on Neural Networks, Vol. 4, No. 3, pp. 720-722, (1993).