

招待論文

機械の学習と人間の学習——ゲーム情報学を題材として

松原 仁¹

受付日 2016年11月2日, 採録日 2016年11月3日

概要: ゲームは人工知能研究のいい題材である。コンピュータチェスの膨大な研究がなされた結果、1997年に Deep Blue が世界チャンピオンの Kasparov に勝利した。将棋と囲碁はコンピュータにとってチェスよりはるかに複雑なゲームであり、将棋と囲碁の強いプログラムを作ることはとてもむずかしかった。最近機械学習の手法を用いることによって将棋と囲碁のプログラムは人間よりも強くなった。その機械学習と比較することで人間の学習に関する知見を得ることができる。

キーワード: 将棋, 囲碁, 機械学習, 創造性

Machine Learning and Human Learning: Case of Game Informatics

HITOSHI MATSUBARA¹

Received: November 2, 2016, Accepted: November 3, 2016

Abstract: Games are good examples for artificial intelligence research. There were a lot of computer chess researches and computer chess Deep Blue won against human world champion Kasparov in 1997. Shogi and Go are more and more complex than chess so to make strong Shogi and Go programs was very difficult. Recently by using machine learning Shogi and Go programs became stronger than human professional players. We can learn about human learning by comparing with the machine learning.

Keywords: Shogi, Go, machine learning, creativity

1. はじめに

最近コンピュータ将棋とコンピュータ囲碁はとても強くなり、プロ棋士に勝つまでになっている。その技術的な基盤は機械学習である。この機械学習を人間の学習と比較することは人間の学習を理解するための手助けになると考えられる。ゲームは、ルールが明確である、勝ち負けによって手法の良し悪しが明確に評価できる、目標にできる強い人間が存在する、など人工知能の良い例題として取り上げられてきた。以下ではまずコンピュータチェスについて簡単に振り返った後で、コンピュータ将棋とコンピュータ囲碁でどのように機械学習が貢献したか、それによってどのようなことが起きたか、人間の学習と比較するとどういった特徴があるかについて述べる。

2. コンピュータチェス

人工知能で最初に例題として取り上げられたゲームはチェスであった。Shannon [1] と Turing が 1950 年前後に相次いで独立にコンピュータチェスの論文を書いて、コンピュータにもチェスが指せること（ゲーム理論で用いられているミニマックス法を適用することを提案した）、原理的にはコンピュータが人間を超えられること（世界チャンピオンに勝つコンピュータチェスが開発可能であることを主張した。Simon, Newell, McCarthy など人工知能のパイオニアのほとんどはコンピュータチェスの研究を手がけた（手がけなかったパイオニアは Minsky ぐらいである）。McCarthy はチェスのことを人工知能の「ハエ」と称した。遺伝学が「ハエ」を題材として大きな進歩をしたように、人工知能はチェスを題材として大きな進歩をしたという意味である。チェスの場合の数はほぼ 10^{120} である。ある局面でルール上指せる合法手の数を分岐数というが、チェスの平均分岐数は約 35 である。チェスは平均 80 手で勝負がつ

¹ 公立はこだて未来大学
Future University Hakodate, Hakodate, Hokkaido 041-8655,
Japan

くので、35の80乗すなわちほぼ 10^{120} が場合の数となる。

ShannonとTuringが提案したミニマックス法およびその改良版であるアルファベータ法がその後のゲームの探索の基本となった。コンピュータチェスでも機械学習の試みは何度もなされたが、結果的には機械学習を使わずに強くなった。チェスは何度も人間との対戦を経たのちに1997年にDeep Blueが世界チャンピオンのKasparovに勝利した。6回戦で5回戦が終わった時点では1勝1敗3引き分けのイーブンであったが、最終戦でKasparovが緊張のあまり序盤で大悪手を指して負けてしまった。これは番狂わせによるコンピュータチェスの勝ちで、この時点ではまだKasparovの実力がDeep Blueに勝っていたと思われる。その後の進歩によりコンピュータは人間より明らかに強くなっている。Deep Blueはスーパーコンピュータにチェス専用マシンを数百台並べた構成であったが、現在ではパソコン1台はおろかスマートフォン1台でも人間が敵わないまでになっている。

3. コンピュータ将棋

チェスの 10^{120} よりも将棋は 10^{220} と場合の数がはるかに大きい。将棋はチェスと同じ敵の重要な駒（キングあるいは玉）を捕まえるゲームであるが、チェスは敵から取った駒が使えないのに対して将棋では敵から取った駒が再利用できる（「持ち駒」制度と呼ばれる）ため、終盤は序盤より分岐数が大きくなる。チェスは終盤になるに従って分岐数が小さくなる収束型ゲームであるが、将棋は発散型ゲームなのである。この違いによってコンピュータ将棋を強くするためにコンピュータチェスとは異なる工夫が必要になる。将棋は日本固有のゲームなので、将棋を対象とした研究は当然のこととして日本が中心になった[2]。

将棋のプログラムの開発は1970年代に始まったが、当時のコンピュータの能力では将棋はチェスのような全数探索は無理だったので、前向き枝刈りの探索手法がさかんに研究された。ミニマックス法（アルファベータ法）などチェスで有効だった手法で将棋でも使える手法はおよそすべて使った。評価関数はチェス同様に手作業で作成と改良を行っていた。チェスの評価関数は駒の損得という明快な基準があったものの、将棋の評価関数は駒の損得以外にも玉の堅さや攻め駒の効率などの要素が複雑に絡み合っているため、優れたものを手作業で作るのは非常にむずかかった。そのためにコンピュータ将棋はなかなか強くならず、2000年代の半ばまでアマの高段者レベルの実力にとどまっていた。

2005年に登場した保木邦仁（当時東北大、現在電通大）のボナンザによってコンピュータ将棋は革命的な進歩を果たした[3]。保木の工夫は、これまで前向き枝刈りで候補手を絞って探索していたのをチェスのように全数探索をするようにしたこと、およびこれまで手作業で作っていた評価

関数を数万局の大量の棋譜からの機械学習で作るようにしたこと、の2点である。この工夫をしたボナンザが圧倒的な強さでコンピュータ将棋のトップに立ったので、他の研究者開発者もこぞってこれらの方法を取り入れた。特に後半の機械学習はその後の強豪のプログラムすべてが取り入れており、「ボナンザ・メソッド」と呼ばれている。保木はボナンザのアルゴリズムをすぐに公開し、またプログラムのソースコードも無償で公開した。研究成果を公開するという習慣がゲーム情報学の研究領域の発展を支えているものと思われる（最近も技巧という強豪のコンピュータ将棋がソースコードを公開している）。

人間が手作業でコンピュータ将棋の評価関数を作成していたときは評価の要素は駒の損得、玉の堅さ、攻め駒の効率など人間が思いつくものであった。要素の数もたかだか数十から百程度であった。人間が制御できる範囲に限られていたのである。機械学習で評価関数を作成することになれば、膨大な数の評価の要素を扱うことができるようになる。保木は3個の駒のあらゆる位置関係を評価の要素とした。将棋には40個8種類（成り駒も含めれば14種類）の駒があり、そのあらゆる3駒関係というのは膨大な数になる。ボナンザはプロ棋士などの大量の棋譜からこの3駒関係を評価関数の要素として学習した。膨大な3駒関係のうちほとんどは将棋の形勢判断には関係ない（要素の重みが実質的にゼロ）であるが、一部は要素の重みに一定の値がある。その3駒がそういう位置関係にあるのはプラスであったりマイナスであったりするという意味を意味する。その中にはこれまでプロ棋士を含む人間が気付かなかったものもある。それらを取り入れることによって、それまでより優れた評価関数ができたのである。コンピュータ将棋の評価関数が優れたものになったといっても、トップレベルのプロ棋士のような形勢判断がいつもできるようになったものではない。形勢判断はコンピュータ将棋よりもプロ棋士の方が優れている部分も多い。コンピュータ将棋が人間を圧倒しているのは探索の深さと幅であり、評価関数が以前より良くなったことによりその探索能力を生かせるようになったのである。

2010年代になってコンピュータとプロ棋士が対戦するようになった。2010年には情報処理学会の50周年記念のイベントとして女流プロ棋士の清水市代女流王将（当時）と合議制のコンピュータ将棋「あから2010」が対戦して「あから2010」が勝利した（図1）。2013年、2014年と電王戦と称してプロ棋士5人とプログラム5つが対戦したが、3勝1敗1分け、4勝1敗とともにコンピュータが優勝した。その時点でコンピュータはトップクラスのプロ棋士（竜王、名人）のレベルに達したと思われる。いまトッププロ棋士とコンピュータが対戦すれば勝率的にはコンピュータが勝ち越す可能性が高い。現在のトップ棋士である羽生善治氏との対戦はすぐに実現しないと思われるので、情報処



図 1 2010 年 清水市代女流王将対あから 2000
 Fig. 1 Michiyo Shimizu vs. Akara 2000 Shogi match.

理学会は 2015 年 10 月に将棋で人間とコンピュータの強さを問うことは学問的には結論が出たという終了宣言を行った [4].

将棋（と囲碁）は段位で強さを表現する．東洋の伝統的な尺度であり，柔道や剣道なども段位を用いている．西洋のチェスでは 4 桁のレーティングという数字によって強さを表現している．段位は一度上がると弱くなっても下がらないため，現在の強さを表す指標としては欠点を有している．レーティングは勝てば上がって負ければ下がるので，現在の強さを示す指標としては段位よりも適切といえる．将棋でも非公式にプロ棋士とコンピュータの強さをレーティングによって評価する試みが行われている．それによれば 2016 年現在最強のコンピュータ将棋である ponanza はプロ棋士最強レベル（羽生善治氏，渡辺明氏など）よりも 400 点以上レーティングが上である [5]．レーティングの差が 200 点あれば上位は下位に 75% の確率で勝ち，400 点差があれば 90% の確率で勝つ．学問的には結論が出たという宣言はこのような状況をふまえてのことである．

2005 年のボナンザ以降しばらくの間コンピュータ将棋は人間（プロ棋士）の棋譜からの機械学習によって強くなった．当時はプロ棋士の方がコンピュータよりも強かったのでプロ棋士の棋譜から学ぶことが有効であったのである．ここ数年はコンピュータ将棋の方が事実上人間よりも強く

なったので，もはや人間の棋譜から学ぶことが有効ではない状態になっている．そこで現在ではコンピュータ将棋どうしがたくさん対局することによってそこから機械学習によって強くするという強化学習 [6] の手法がとられている．人間の棋譜から学んでいたときのコンピュータ将棋は序盤などで人間の指し手に近い手（定跡）を選んでしたが，コンピュータどうしの棋譜から学ぶようになって序盤からコンピュータ独自の手を指すようになってきている．人間にはそのコンピュータの指し手の意味が理解しにくいので，そのことも人間がコンピュータに勝ちにくくなっている理由となっている．

4. コンピュータ囲碁

囲碁は中国発祥のゲームであるが，中国ではいったん廃れて日本に入ってきてさかんになった（いま中国でさかんになったのはいわば日本からの逆輸入である）．囲碁は，他に似たルールのゲームが存在しない，漢字を使っていないので親しみやすい，などの理由で世界的に普及している．最初にコンピュータ囲碁の研究がなされたのは 1960 年代である（チェスよりは遅いが将棋より早い）．囲碁も最初はチェスのようにミニマックス法の探索によって次の手を決めようとしたが，囲碁の場合の数は 10^{360} とチェスや将棋よりはるかに大きく，ふつうの探索によって良い手を見

つけるには候補手が多すぎて強くならなかった。2000年代になってもまだとても弱い状態であった（初心者レベルよりはましでもせいぜい初級者レベルであった）。

囲碁も将棋のボナンザ・メソッドのような革命的な手法が現れた。それがモンテカルロ木探索である。この元となったモンテカルロ法は von Neumann の命名といわれるシミュレーションによって解を求める方法である。1993年にこれを囲碁に適用するというアイデアが発表されたものの、そのときは成功しなかった。2006年になって Coulom が Crazy Stone という囲碁プログラムの中でモンテカルロ法を応用した（UCT というアイデアを付け加えた）モンテカルロ木探索を採用し、この Crazy Stone が圧倒的な強さを示した。囲碁にモンテカルロ法を適用するということは、ある局面から白と黒が交互にランダムに終局まで打ち進めるというシミュレーションを多数行って勝つ確率が一番高い手を選ぶということである。そこには囲碁の知識はほとんど何もはっていない。この一見単純な方法（実は UCT など細かい工夫が重要ではある）で強くなることに驚きその後の囲碁プログラムはみんなこの方法を取り入れている。それで囲碁プログラムは一気にアマチュアの6段程度の実力に達した [7]。モンテカルロ法自体は機械学習とは異なる統計的な方法である。

最近までは日本の ZEN（これもモンテカルロ木探索を用いている）が Crazy Stone を抜いて最も強い囲碁プログラムであった。これらのプログラムはまだ互先（ハンディなし）で戦うのは無理であったが、トッププロ棋士と4子（初期局面に4個の石をあらかじめ置く）程度のハンディで勝つまでになっていた。2015年の時点ではコンピュータ囲碁がトッププロ棋士に勝つのはまだ10年はかかると思われた（コンピュータ将棋の10年前の強さに相当すると思われていた）。

2016年1月に Google の AlphaGo というプログラムが2段のプロ棋士に互先で5戦5勝の成績をあげたと Nature という雑誌に発表して大ニュースになった [8]。AlphaGo は、深層学習（ディープ・ラーニング）、モンテカルロ木探索、強化学習、という3つの手法をうまく組み合わせている。大量のプロ棋士の棋譜（約3,000万局面とのことである）をデータとして深層学習によってある程度の強さのプログラムを作り、そのプログラムどうしの強化学習によってさらに強くした。これまでコンピュータ囲碁では優れた評価関数を作ることに失敗してきたが、初めて優れた評価関数を実質的に作ったことが AlphaGo の大きな特徴である。手を決める部分では従来手法であるモンテカルロ木探索を使っている。これらの3つの手法それぞれはよくある技術であるが、それらをうまく使ったことに AlphaGo の成功の理由がある。また深層学習は複雑なニューラルネットワークなので、計算にコンピュータパワーを必要とする。AlphaGo は13層のネットワークで、莫大なコンピュータ

パワーを使っている。

その後2016年3月に韓国のイ・セドル9段というプロ棋士と AlphaGo が対戦した。1月の Nature に載った論文では AlphaGo はプロ2段の人間に勝ったものの、トップレベルのプロ棋士とはまだ実力差があった（AlphaGo とプロ2段との対局の棋譜を分析したプロ棋士がそう判断していた）。イ・セドルは直前まで世界一のランキングに位置付けられてこの時点でも世界ランキング4位のトップのプロ棋士である（日本で最強の井山6冠よりもランキングは上である）。プロ棋士もコンピュータ囲碁の関係者も AlphaGo はイ・セドルにはまだ敵わないという見方が強かった。しかし予想に反して AlphaGo はイ・セドルに4勝1敗と圧勝した。AlphaGo は機械学習によって強くなったプログラムであり、2段に勝った Nature の論文の時点からイ・セドルと対戦するまでにさらなる機械学習でさらに強くなっていたのである。

イ・セドルと対戦した AlphaGo の強さは衝撃的であった [9]。AlphaGo は序盤から対戦相手のイ・セドルを含めたプロ棋士に理解不能の手を打った。イ・セドルの対局後の感想によれば、イ・セドルはその手が打たれたときには意味が分からず AlphaGo が悪手を打ったと思ったそうである。解説をしていたプロ棋士も悪手と断定してイ・セドルが有利と判断していた。イ・セドル自身の判断も同様であった。イ・セドルは中盤で自分が不利であることに気付いた。なぜ不利になってしまったかを考える中で、序盤の AlphaGo の意味不明な手がとても良い手であることが分かったとのことである（解説のプロ棋士は終盤になるまでそのことに気付くのが遅れた）。囲碁（や将棋）で局面を直感的に把握する能力のことを大局観という。これまで大局観は強い人間だけが有していてコンピュータは有していなかった。AlphaGo は機械学習によってトッププロ棋士を超える大局観を獲得したのである。コンピュータ将棋は人間にまさる探索能力で強くなったが、コンピュータ囲碁は人間にまさる大局観で強くなった。大局観でコンピュータが人間にまさることはないと思われていたので衝撃であったのである。

5. 機械学習の特徴

ここではコンピュータ将棋とコンピュータ囲碁の機械学習の特徴を整理して人間の学習と比較する。

(1) 学習に必要なデータ量が膨大である

機械学習は一般に膨大なデータを必要とする。コンピュータ将棋で用いられている機械学習では一般に数万局から数十万局の棋譜を使っている（コンピュータ将棋では深層学習ではない従来の機械学習の手法が用いられることが多い）。コンピュータ囲碁では深層学習に AlphaGo は3000万局面程度の棋譜を使っている。ゲームではないが深層学習で猫の概念を学んだときも数万匹から十数万匹の猫の画

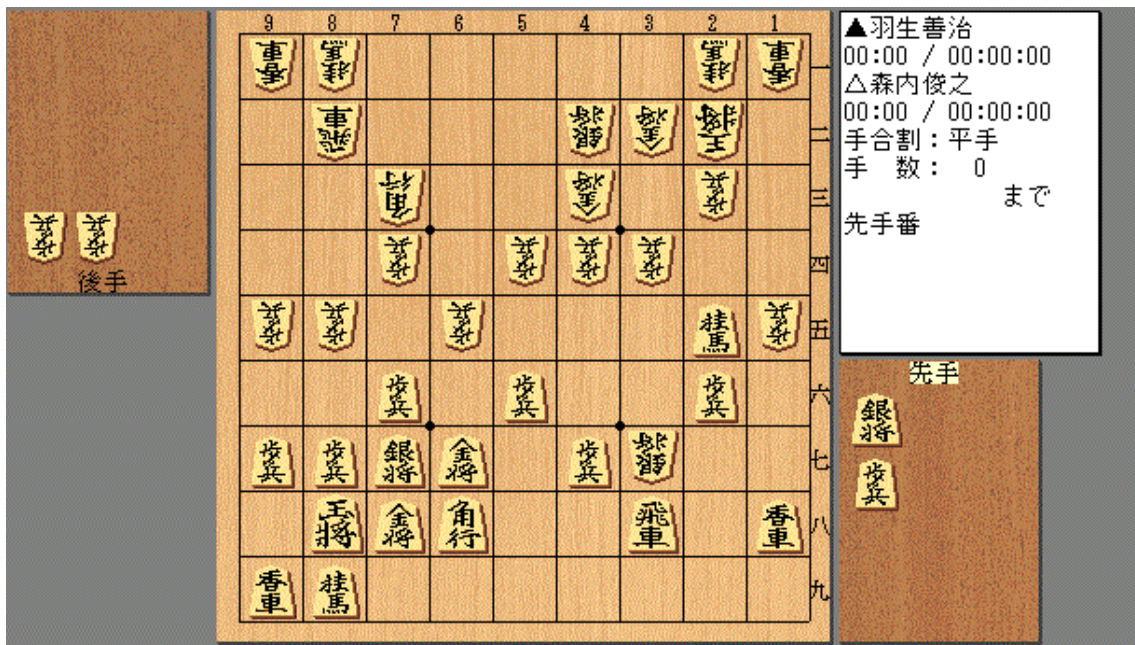


図 2 2013 年 71 期名人戦第 5 局後手 3 七銀の局面

Fig. 2 37 Gin move in a professional match.

像を使っている。以前に比べると深層学習を含めた機械学習の能力は格段に進歩をとげたが、能力を発揮するためには膨大なデータがなければならない。

一方で人間の学習はこれほどのデータは必要としていない。人間の学習に必要なデータを量的に見積もるのはむずかしいが、将棋にしても囲碁にしてもプロ棋士は強くなるのにコンピュータほどには棋譜は勉強していない。我々はみな猫の概念を子供のときに学習しているが、学習する前に見た猫はせいぜい数十匹であって（テレビや本でも見ることはあるとはいえ）数万匹の猫は見えていない。

機械学習は学習に必要なデータ量において人間の学習にはるかに非効率的である。この点において人間の学習の能力はコンピュータよりもはるかにまさっている。

(2) 学習によって創造性を発揮する

コンピュータは人間に命じられたことしかできないので、コンピュータは原理的に創造性を発揮できないという考え方があ。しかし機械学習で強くなったコンピュータ将棋とコンピュータ囲碁は創造性を発揮し始めている。何をもって創造性を発揮しているかとの定義はむずかしいが、ここでは誰（何）かが新しく見つけたものが第三者から優れたものと評価されたときに創造性を発揮したと見なすことにする。

最近のコンピュータ将棋はいくつもの「新手」を見つけている。「新手」とはプロ棋士の間でそれまでは評価の低かった指し手に誰（何）かが良い手であることに気付いてその手が優れたものと評価されたものである。図 2 は 2013 年の名人戦（プロ棋士で最も権威のあるタイトル戦）で先手が羽生善治氏、後手が森内俊之氏で後手の森内氏が

3 七銀と打った局面である。3 七銀を打つ前の局面はプロ棋士の間でとても有名でこれまで長い間検討が重ねられてきたが、この局面で 3 七銀と打ったプロ棋士は森内氏が初めてであった。これが良い手でこの対局は後手の森内氏が勝利した。感想戦で森内氏が語ったところによると 3 七銀は森内氏自身が気づいたものではなく、この対局の少し前にコンピュータ将棋の ponanza が同じ局面で指したのを真似したとのことである。コンピュータ将棋が見つけた手をトッププロ棋士が最も権威のあるタイトル戦で採用したのである。この手は ponanza 新手と呼ばれている。その後プロ棋士同士の対局でコンピュータが見つけた新手が何度も採用されている。コンピュータ将棋が創造性を発揮しているのである。コンピュータ囲碁でも同様の手が観察されている。コンピュータは人間に命じられたままに機械学習をしている（だけ）であるが、その結果として創造性を発揮できているのである。

コンピュータが機械学習によって創造性を発揮しているということは不思議に感じられるかもしれないが、人間の学習を考えればまったく不思議なことではない。人間も既存のデータから学習しているが、その結果として新しいことを思いつくという創造性を発揮できている。結果的に、うまくいった学習は人間の場合であれコンピュータの場合であれ創造性を発揮できる段階までいっているといえよう。

6. おわりに

チェスは機械学習を使わずにコンピュータが人間よりも強くなり、将棋と囲碁は機械学習を使うことによってコンピュータが人間よりも強くなった。これはその時点（チェ

スは1997年で将棋は2015年、囲碁は2016年)でのコンピュータの性能の違いもあるが、ゲームの性質(特に場合の数の大きさ)の違いもあると思われる。人間の学習を理解するために機械学習は反面教師の面も含めて参考になるはずである。機械学習も人間の学習の効率性を参考にしてもっと効率的なものにしていかなければならない。

参考文献

- [1] Shannon, C.: Programming a Computer for Playing Chess, *Philosophical Magazine*, Ser.7, Vol.41 (1950).
- [2] 松原 仁(編著):コンピュータ将棋の進歩1-6, 共立出版(1996-2012).
- [3] 保木邦仁:局面評価の学習を目指した探索結果の最適制御, 第11回ゲームプログラミングワークショップ招待講演資料(2006).
- [4] 松原 仁:コンピュータ将棋プロジェクトの終了宣言, 情報処理学会誌, Vol.56, No.11, pp.1054-1055 (2015).
- [5] 伊藤毅志:コンピュータ将棋と不正疑惑, 情報処理学会誌, Vol.57, No.12, pp.1184-1185 (2016).
- [6] Sutton, R. and Barto, A. (著), 三上貞芳, 皆川雅章(訳):強化学習, 森北出版(2000).
- [7] 松原 仁(監修), 美添一樹, 山下 宏(著):コンピュータ囲碁—モンテカルロ法の理論と実践, 共立出版(2012).
- [8] Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D.: Mastering the Game of Go with Deep Neural Networks and Tree Search, *Nature*, Vol.529, pp.484-489 (2016).
- [9] 松原 仁:速報AlphaGoの勝利, 情報処理学会誌, Vol.57, No.6, pp.502-503 (2016).



松原 仁 (正会員)

1986年東京大学大学院工学系研究科情報工学専攻博士課程修了。工学博士。同年通産省工技院電子技術総合研究所(現,産業技術総合研究所)入所。2000年公立はこだて未来大学教授。

2016年同副理事長。人工知能,ゲーム

情報学,観光情報学等に興味を持つ。著書に『コンピュータ将棋の進歩』,『鉄腕アトムは実現できるか』,『先を読む頭脳』,『観光情報学入門』等。本会理事,前人工知能学会会長,観光情報学会理事。