北海道大学工学部研究報告
第167号（平成6年）

Bulletin of the Faculty of Engineering
Hokkaido University, No. 167 (1994)

アナロジーを利用した問題解決に関する研究

松浦　賢一　　嘉数　侑昇

（平成5年8月31日受理）

A Study on Problem Solving Using Analogy

Kenichi MATSUURA and Yukinori KAKAZU

(Received August 31, 1993)

Abstract

To solve a given problem, the strategy that is suitable for solving that problem is used in general. This strategies for solving a problem can be treated as a mapping from the problem into a solution. Besides, there will exist some kinds of analogical relationship among these strategies. Thus, we can apply the method, that takes analogy of strategies with the problem that has already been solved and that is similar to the unsolved problem, to find unsolved problem. In this paper, we attempt to perform problem solving using such analogical method. To this end, first, heuristics are generated by generalizing a set of problem solving strategies that are similar to each other. Second, these heuristics are applied to the targeted problem and a solution is found. We have performed some computer experiments to examine the role of the heuristics and the effectiveness of the analogy.

1．はじめに

問題を解決するためには，対象とする問題に対して有効な問題解決戦略を用いるのが一般的である。この問題解決戦略は，ある問題とそれに対する解との関係であると考えることができ，さらに個々の問題解決戦略間にも，順序関係，包含関係などの様々な関係が存在する[3]。それらの関係の中で類似関係に着目すると，これを利用した問題解決を考えることができる。つまり，ある問題に対してある問題解決戦略が有効であるなら，その問題解決戦略に類似する別の問題解決戦略もまた有効であろうと期待できる。またこれとは逆に，ある問題解決戦略がある問題に対して有効であるなら，その問題と類似する別の問題に対しても有効であろうと期待することも考えられる[5]。

ここでは，後者に着目した問題解決を考える。つまり，過去に解決した問題の問題解決戦略を，未解決の問題の解決に利用することに着目するが，これにはさらに2通りの立場がある。一つは，過去に解決した問題における問題解決戦略が，直接的に未解決の問題に適用されるもので，もう一つは直接的ではなく，問題解決戦略が一度一般化されてから，未解決の問題に適用されるというものである[2]。

精密工学科　精密器機学第一講座

これらのうち，本論文ではさらに後者の立場をとる。すなわち，問題解決戦略間のアナロジーの利用を，一般化によって行なわれるものとして，まずは個々の問題解決戦略を要素とし，それらの類似関係を距離とした空間を設定する。次に，この空間の距離を基準としてカテゴリを形成し,それぞれのカテゴリ毎に一般化を行なうことでいくつかのヒューリスティックスを生成する。そして，このヒューリスティックスを未解決の問題に適用することで，問題の解決が行なわれる。さらに，これらについて計算機実験を行ない，生成されたヒューリスティックスの機能とアナロジーの有効性について検討する。

2．問題解決戦略とその一般化

アナロジーを利用した問題解決を試みるにあたって，問題解決戦略とその一般化で生成するヒューリスティックスについて，ここでの基本的な考え方を以下で述べる。

問題解決における問題とは，(1)式の4項で定義されている[1]。

$$Problem = (IS, \{GS\}, Op, Co) \tag{1}$$

where　IS：初期状態
GS：目標状態
Op：演算子集合
Co：制約条件

そして問題を解決することは，IS から GS へ至る演算子の順列を求めることであり，これが解 *Solution* である。Co は解の満たすべき性質である。

このような定義から問題解決戦略 *Strategy* を考えると，(2)式のようになる。

$$Strategy:\ Problem \rightarrow Solution \tag{2}$$

ここで，(1)式の4項を考えると，これら4項のうち，GS は IS と *Solution* より一意に決定される。また問題領域を考えた場合，同一の問題領域では Op，Co が共通であると考えられる。したがって，ある特定の問題領域を対象とした場合，問題解決戦略は初期状態と解との関係を表すのみで十分であると考えられる。

$$Strategy:\ IS \rightarrow Solution \tag{3}$$

(3)式のように表現されるいくつかの問題解決戦略を一般化することでヒューリスティックスを考えるが，その問題解決戦略が表現される空間と，初期状態が表現される空間を次のようにする。

$$\Gamma = (\{Strategy\},\ d_{St}) \tag{4}$$

where　d_{St}：Γ における距離関数

$$\Lambda = (\{IS\},\ d_{IS}) \tag{5}$$

where　d_{IS}：Λ における距離関数

また，一般化には“∩”(交わり)の演算子が有効であるといわれている[4)6)]。これにしたがうと，ヒューリスティックス H は(6)式のように表現することができる。

$$H = \bigcap_{i} Strategy_i \tag{6}$$

3．アナロジーを利用した問題解決

アナロジーを利用した問題解決として，ここでは問題解決戦略間のアナロジーに着目するが，まずは問題解決戦略にいくつかの前提を設ける。次に，この前提のもとでのヒューリスティックスの生成法と，これを用いた問題解決法について述べる。

3.1 問題解決戦略の前提

ここでのヒューリスティックスは問題解決戦略を一般化することによって生成していくが，その一般化と関連して，ここでの問題解決戦略はいくつかの部分戦略から構成されるものとする。

$$Strategy=\{ps_i|i=1,\ \ldots,\ n\} \quad (7)$$
$$\text{where} \quad ps_i：\text{部分戦略}$$
$$n：\text{定数}$$

前述のように，一般化には“∩”（交わり）の演算子が有効であるが，この演算子は，共通部分の抽出・強調を行なうものである。そこで問題解決戦略の一般化には，問題解決戦略を集合とみた場合の部分集合，さらにはその構成要素を考慮する必要があると考えられる。また，問題解決戦略を構成する部分戦略の構成要素数 n は，問題設定に依存するものとする。

また(7)式より，問題解決戦略の定義域 IS と値域 $Solution$ もそれぞれの構成要素 ppr_i，pso_iから成る集合とする。

$$ps_i:\ ppr_i \rightarrow pso_i \quad (8)$$
$$\text{where} \quad ppr_i \in IS$$
$$pso_i \in Solution$$

さらに，このような構成要素間には，一般に距離が存在するものと考えられるため，それぞれの集合は要素間の距離の定義可能な集合，すなわち距離空間であるとする。

3.2 一般化によるヒューリスティックスの生成

(7)(8)式より，ひとつの問題解決戦略は $IS \times Solution$ の空間上に点群として表現できる。すなわち，(9)式のような部分集合で表現される。

$$Strategy \subset IS \times Solution \quad (9)$$
$$\because \quad ps_i \in IS \times Solution$$
$$Strategy=\{ps_i|i=1,\ \ldots,\ n\}$$

ヒューリスティックスの生成に先だって，戦略間の距離に基づいた Γ のカテゴライズを行なうが，その距離として，ここでは点群の第一主成分の係数ベクトルをとって(10)式のような距離 d_{St} を考える。

$$d_{St}(Strategy_i,\ Strategy_j)=\|pe(Strategy_i)-pe(Strategy_j)\| \quad (10)$$
$$\text{where} \quad pe(Strategy_i)：Strategy_i \text{ の第一主成分の係数ベクトル}$$

このような距離 d_{St} にしたがって，問題解決戦略の集合からカテゴリ C_k^{St} を形成していく。カテゴライズの方法として，ここでは非階層的クラスタ分析法を用いる[7)8)]。

次に，得られたカテゴリからヒューリスティックスを生成する。これは，部分戦略の概念より次のように考えることができる。すなわち，あるカテゴリにおける全ての部分戦略 ps_i の集合を次

の手順によって分類し，分類されたそれぞれのクラスをその濃度で代表させることによって一般化と考えることができる。

(1) 次の初期設定をする。

$$Classes=\{class\},\ |class|=0$$

(2) PS の各要素 ps_i に対して，以下を行なう。

(2.1) $Classes$ の各要素 $class$ に対して，以下を行なう。

$$\exists\, mem,\ ps_i \cong mem \Rightarrow \text{add } ps_i \text{ to } class \quad (mem \in class)$$

(2.2) ps_i がどの $class$ にも追加されなかったとき，以下を行なう。

$$\text{add } Classes \text{ to } new.class \quad (new.class=\{ps_i\})$$

(3) $Classes$ をもって分類を完了する。

ここで，クラスの濃度の表現として，部分戦略の頻度 Q を導入する。つまり，この頻度が大きいほど，その部分戦略がカテゴリ中に多く存在することを意味する。このように，頻度を導入することによって，共通部分の強調の割合が表現できる。従って，$Classes\times Q$ の直積空間上に(11)式のヒューリスティックス H が生成される。

$$H=\left\{(class^i,\ q^i)\ \middle|\ class^i\in Classes,\ q^i=\frac{|class^i|}{\sum_j |class^j|},\ i=1,\ \ldots,\ |Classes|\right\} \tag{11}$$
$$\text{where}\quad H\subset Strategy\times Q$$

3.3 ヒューリスティックスを用いた問題解決

以上のようにして生成されたヒューリスティックスを用いることで，戦略間のアナロジーによる問題解決を行なうが，これは次の手順で行なわれる。

(1) 与えられた未解決の問題が属する，カテゴリを選択する。

(2) 選択されたカテゴリのヒューリスティックスを用いて，問題を解決する。

ここで，(1)のカテゴリ選択であるが，未解決の問題は Λ 上で表現されるため，この空間でのカテゴリの選択となる。IS の空間でのカテゴリは，距離 d_{St} をもとに形成した Γ でのカテゴリの位相と対応するようなものが望ましい。そこで，$Strategy$ の空間でのカテゴリをもとに，IS の空間でのカテゴリ形成を次のように考える。

(1) Λ において，過去に解決した問題に対するボロノイの多角形分割を考える。これは，(12)式で示される領域 S_k への分割である。

$$S_k=\{x|x\in\Lambda,\ d_{IS}(x,\ X_i)\leqq d_{IS}(x,\ X_j),\ i\neq j,\ i=1,\ \ldots,\ kn\} \tag{12}$$

where　X_i：過去に解決した問題 $(i=1,\ \ldots,\ kn)$

d_{IS}：距離関数

kn：過去に解決した問題の総数

(2) 次に，分割された S_i を Γ でのカテゴリにしたがって，(13)式のように統合し，IS の空間におけるカテゴリ C_k^{IS} とする。

$$C_k^{IS}=\bigcup_{X_i\in C_k^{SI}} S_i \tag{13}$$

IS の空間におけるカテゴリが決定されると，そのカテゴリに対するヒューリスティックスを未解決の問題に適用することで問題解決を行なうが，(8)式のように，未解決の問題における初期状態は，いくつかの構成要素 ppr_i から成るため，それぞれの要素に対してヒューリスティックスを適用する。すなわち，初期状態の各要素 ppr_i に対して，頻度が最大となるような解の構成要素 pso_i を求めていく。そして，pso_i の集合が解となる。

4．計算機実験

以上で展開した手法の検証を目的として，ここではナップザック問題を対象例として計算機実験を行なった。

4.2 問題設定

ナップザック問題は，どの荷物をナップザックに入れていくのかという問題である。そこで，(7)式の問題解決戦略を次のように設定する。

$$Strategy=\{ps_i | i=1, \ldots, ln\} \tag{14}$$

$$\text{where} \quad ps_i=(size'_i,\ cost'_i,\ ap_i)$$

$$size'_i=\frac{size_i}{\max_j (size_j)}$$

$$cost'_i=\frac{cost_i}{\max_j (cost_j)}$$

$$ap_i=\begin{cases} 1 & i\text{ 番目の荷物を入れる} \\ 0 & i\text{ 番目の荷物を入れない} \end{cases}$$

ただし，$size_i$，$cost_i$ はそれぞれ i 番目の荷物の大きさ，コストを，ln は荷物数を表すものとする。

また，このような戦略の設定より，Λ を(15)式のように設定する。

$$\Lambda=\left(V(size),\ V(cost),\ \frac{knapsack}{\sum_i size_i}\right) \tag{15}$$

$$\text{where} \quad V(size)\text{：荷物の大きさの分散}$$

$$V(cost)\text{：荷物のコストの分散}$$

4.2 実験結果と考察

過去に解決した問題として，ナップザック問題のいくつかの例を与え，それぞれの問題解決戦略をカテゴライズした結果，**Fig. 1** のような結果が得られた。ただし，軸 a_1，a_2，a_3 は，(16)式で表される第一主成分の係数である。

$$a_1 size'_i + a_2 cost'_i + a_3 ap_i \tag{16}$$

Fig. 1 において，5個のカテゴリが形成されているが，これらのうち複数の戦略から構成されているカテゴリについて，戦略の一般化によってヒューリスティックスを生成した。この結果に

ついて，**Fig. 2** に示す。ただし，これらは頻度 Q が最大の部分について再構成して表現してある。一方，初期状態 IS の空間におけるカテゴリについて，**Fig. 3** に示す。これより，未解決の問題の属するカテゴリを選択し，そのヒューリスティックスを用いることで，問題解決が可能となる。

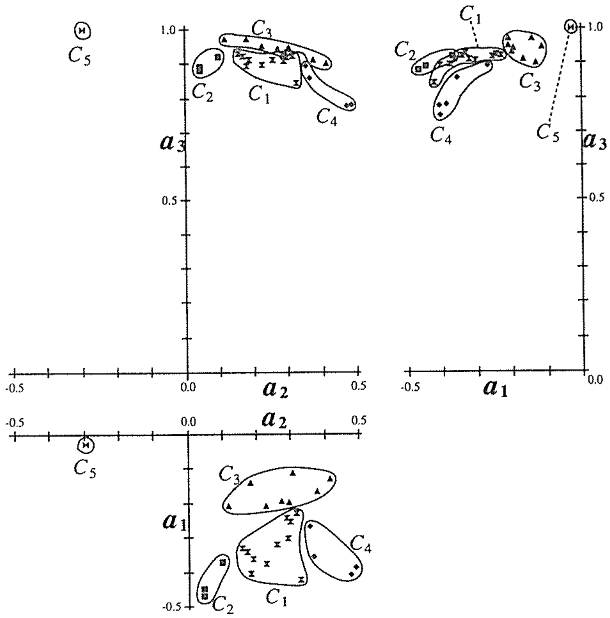

Fig. 1 Result of Categorizing Problem Solving Strategies

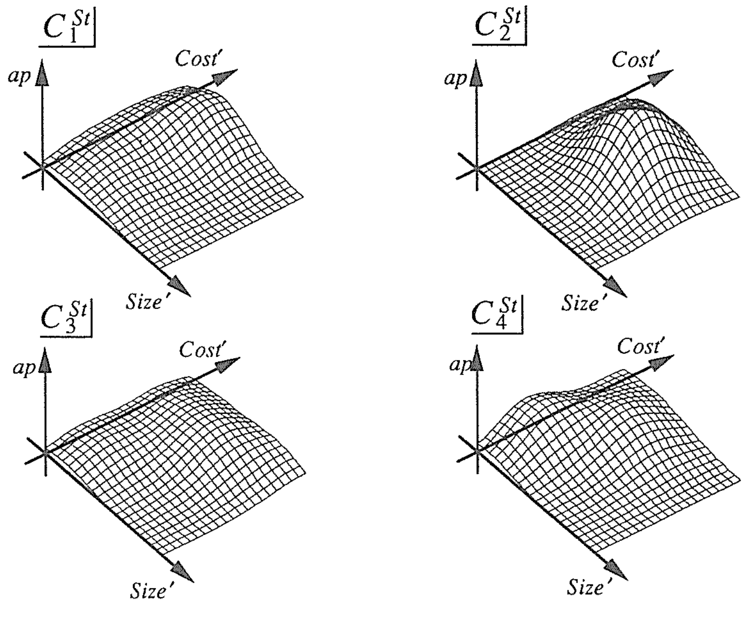

Fig. 2 Generated Heuristics of each Category

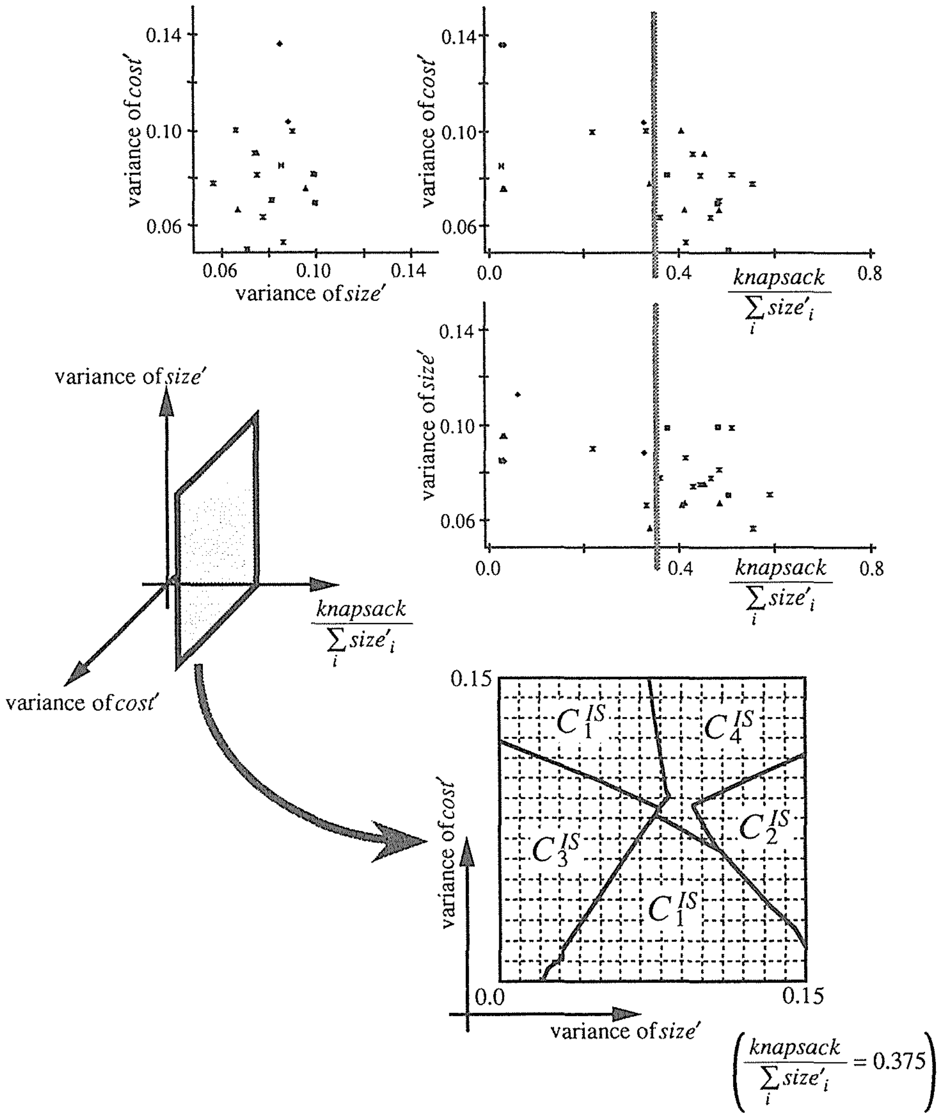

Fig. 3 Category of Initial State

これらを用いた問題解決の結果について，**Fig. 4** に示す。これより，得られた解の指標は最適解に近く，生成されたヒューリスティックスが問題に対して適切であるといえる。また，これは問題解決戦略の一般化としてのアナロジーが有効に機能しているということであり，問題解決戦略間の類似関係を基準とした，問題領域の位相構造の同定が可能になるものと考えられる。

problem:	knapsack		387
	loads	size	19 86 61 78 55 12 75 14 71 55 99 44 35 67 86 87 19 30 24 49
		cost	46 67 63 13 87 18 84 60 48 59 27 97 43 22 44 55 78 27 38 24
solution:			1 0 1 0 1 0 1 1 0 1 0 1 0 0 0 0 1 0 1 0
	performance index / optimal index		612 / 630
strategy:	category		C_3

Fig. 4 Result of Problem Solving Using Heuristics

5. おわりに

問題解決戦略間のアナロジーを利用した問題解決として，次のことを試みた。すなわち，互いに類似関係にある問題解決戦略のカテゴリを形成し，それぞれのカテゴリを一般化することによってヒューリスティックスを生成した。そして，これを未解決の問題に適用することで問題解決を行なった。

ここで，問題解決戦略の一般化を考えるにあたって，問題解決戦略の構成要素として部分戦略の概念を導入し，また生成されたヒューリスティックスの表現のため，頻度を導入した。これらにより，ヒューリスティックスは頻度の最も大きな部分戦略の集合で表現される。

さらに，このようなヒューリスティックスの生成と，これを利用した問題解決について，ナップザック問題を対象例題とした計算機実験を行ない，ヒューリスティックスの機能と問題解決戦略間のアナロジーの有効性を確認した。これより，問題領域における位相構造の同定について考察した。

参考文献

1) Banerji, R., B.: Theory of Problem Solving-An Approach to Artificial Intelligence, American Elsevier Publishing, (1969).
2) Gick, M., L. and Holyoak, K., J.: Schema Induction and Analogical Transfer, Cognitive Psychology, 15, pp. 1-38, (1983)
3) Holland, J., H., Holyoak, K., F., Nisbett, E. and Thagard, P., R.: Induction-Process of Inference, Learning and Discovery, The MIT Press, (1986).
4) Laird, P., D.: Learning from Good and Bad Data, Kluwer Academic Publishers, (1988).
5) Lenat, D.: The Role of Heuristics in Learning by Discovery: Three Case Studies, Machine Learning, Toiga Press, (1983).
6) Winston, P. H.: Learning New Principles from Precedents and Exercises, Artificial Intelligence. 19, pp. 321-350, (1982).
7) 奥野，ほか：多変量解析法，日科技連出版社，(1971).
8) 河口　至商：多変量解析入門Ⅰ，Ⅱ，森北出版，(1973).