

コスト関数をもつ有限システム：逐次決定過程

京大・工学部 茨木 俊秀

1. まごがき. ここでいう逐次決定過程とは、簡単にい
えば、有限オートマトンにコスト関数を付与したものである。
このモデルは、いわゆる組合せ最適化問題を一般的に表現す
る手段として研究されてきた。しかし、対象を組合せ最適化
問題に限定せず、一つのシステムモデルと考えることも可能
であろう。

以下、逐次決定過程に関するこれまでの成果を簡単に紹介
するとともに、システムモデルとしての種々の可能性を考察
する。

2. 諸定義. $M = (Q, \Sigma, q_0, \lambda, Q_F)$ を有限オートマ
トン (fa) とする。ただし、

Q : 状態の空でない有限集合

Σ : 有限アルファベット. $a \in \Sigma$ を 決定 という。

$q_0 \in Q$: 初期状態

$\lambda : Q \times \Sigma \rightarrow Q$, 状態遷移関数

$Q_F \subset Q$: 最終状態の集合.

λ は通常の方法で $Q \times \Sigma^* \rightarrow Q$ に拡張可能. Σ^* は Σ から生成される有限系列 (方策という) の集合.

逐次決定過程 (sdp, Sequential Decision Process) とはシステム $\Pi = (M, r, \xi_0)$ である. ただし,

$M : fa$

$r : E \times Q \times \Sigma \rightarrow E$ (E は実数の集合)

$\xi_0 \in E$: 初期状態 q_0 のもつ初期コスト値.

r は $r(\xi, q, \lambda(q, x), a) = r(r(\xi, q, x), \lambda(q, x), a)$ によって $E \times Q \times \Sigma^* \rightarrow E$ に拡張できる.

$r(\xi, q, a)$ が ξ に関して単調ならば Π は msdp (Monotone sdp) であるという.

また, $F(\Pi) = \{x \in \Sigma^* \mid \lambda(q_0, x) \in Q_F\}$ を Π が 受理する方策 の集合, $O(\Pi) = \{x \in F(\Pi) \mid (\forall y \in F(\Pi)) (r(\xi_0, q_0, x) \leq r(\xi_0, q_0, y))\}$ を 最適方策 の集合という.

3. 組合せ最適化問題の sdp による表現. 簡単のため例を用いて説明する.

行商人問題: n 個の節点を一巡する最短閉路を求めよ. E

ただし、節集 i から節集 j への距離を c_{ij} とする。

この問題に対し、 λ の sdp $\Pi = (M(Q, \Sigma, \delta_0, \lambda, Q_F), k, \xi_0)$ を考へる (出発集を節集 1 とする)。

$$Q = \{ [\mathcal{S}, i] \mid \mathcal{S} \subset \{2, 3, \dots, n\}, i \in \mathcal{S} \} \cup \{ [\emptyset, 1], \delta_d, \delta_F \}$$

$$\Sigma = \{ a_1, a_2, \dots, a_n \} \quad (a_i \text{ は " } \lambda \text{ に } i \text{ へ進む" 決定})$$

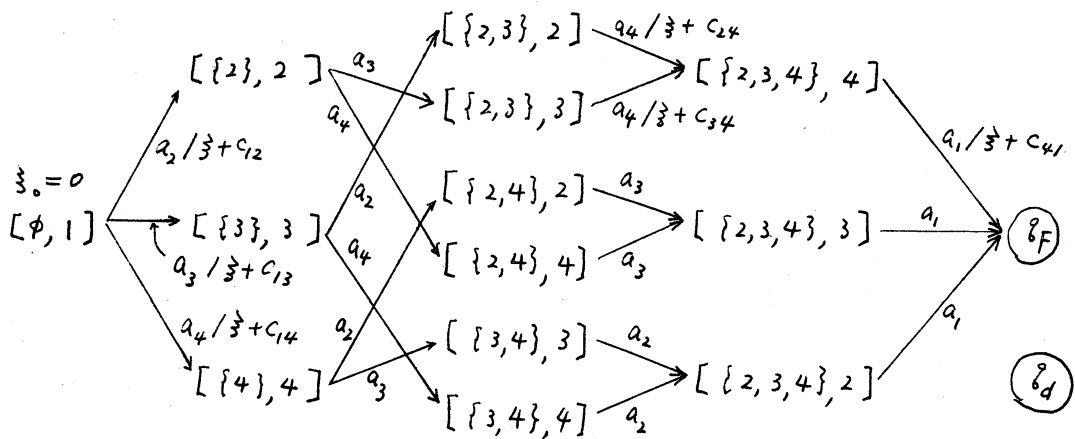
$$\delta_0 = [\emptyset, 1], \quad Q_F = \{ \delta_F \}$$

$$\lambda([\mathcal{S}, i], a_j) = \begin{cases} [\mathcal{S} \cup \{j\}, j] & \text{if } j \notin \mathcal{S} \wedge j \neq 1 \\ \delta_d & \text{if } j \in \mathcal{S} \text{ or } \mathcal{S} \neq \{2, \dots, n\} \wedge j = 1 \\ \delta_F & \text{if } \mathcal{S} = \{2, \dots, n\} \wedge j = 1 \end{cases}$$

$$\lambda(\delta_d, a_j) = \lambda(\delta_F, a_j) = \delta_d$$

$$k(\xi, [\mathcal{S}, i], a_j) = \xi + c_{ij}, \quad \xi_0 = 0.$$

$n = 4$ の場合の例を下图に示す



(各枝に $a / k(\xi, \delta, a)$ を付す (k は一部省略). $\delta_d \wedge a$)
 状態遷移も省略。

Π はつぎの意味で行商人問題を表現している. 1) $x \in F(\Pi)$
 $\Leftrightarrow x$ は n 節実を一巡する閉路を表す. 2) $x \in F(\Pi)$ に対
して $l(s_0, s_0, x)$ は閉路の長さを与える.

$sdp \Pi$ は, M, l, s_0 をうまく定めることによつて, 種々の
組合せ最適化問題を表現できる [1]. 表現問題に関しては
[1][2] に詳しい議論がある. [3] はより一般的な前提で論じ
ている. また, (状態数) 最小表現が [4][5] で扱われている.

4. 最適方策. sdp に対しその最適方策を求めるアルゴ
リズムを与えることができれば, sdp に表現可能なすべての
組合せ最適化問題のアルゴリズムとして使用できる. この観
点から $msdp$ は重要である. なぜなら, $msdp$ に表現可能なこと
と, 動的計画法 (DP) における最適性の原理が成立することを
同一視できるからである [1].

しかし, 任意の $msdp$ に対して, その最適方策の集合 $O(\Pi)$
あるいは最適方策 $x \in O(\Pi)$ を求める一般的なアルゴリズムは
存在しない* [6]. アルゴリズムの存在を保障するにはさらに
制約を加える必要がある. そのようなクラスとしてつぎのも

* アルゴリズムの存在を論じる際には, $l: \mathbb{Z} \times Q \times \Sigma \rightarrow \mathbb{Z}$ (\mathbb{Z}
は整数の集合) とし, l を帰納関数と仮定する等の変更を要
するが, 詳細は略し区別せずを用いる.

のが代表的である。

pmsdp (Positively msdp): $\text{msdp} \wedge \ell(\xi, \delta, a) \geq \xi$

smsdp (Strictly msdp): $\xi_1 > \xi_2 \Rightarrow \ell(\xi_1, \delta, a) > \ell(\xi_2, \delta, a)$

lmsdp (Loop-Free msdp): $\text{msdp} \wedge |F(\Pi)| < \infty$.

具体的なアルゴリズムは [7] 参照。

また、各クラスに対する $O(\Pi) \subset \Sigma^*$ の性質を明らかにすることはいずれ自身興味ある問題であり、ある程度の結果が得られている [2][6][7]。閉包性に関する結果を次表にまとめる。(なお、dfs 以下は §6.1 参照)

	$\sigma \cap V$	$\sigma \cup V$	$\bar{\sigma}$	$\sigma \setminus V$	$\min \sigma$		$\sigma \cap V$	$\sigma \cup V$	$\bar{\sigma}$	$\sigma \setminus V$	$\min \sigma$
sdp	○	✓	✓	✓	○	dfs	○	✓	✓	✓	○
msdp	✓	✓	✓	✓	✓	mdfs	✓	✓	✓	✓	✓
smsdp	○	○	○	○	○	smdfs	○	✓	✓	✓	○
pmsdp	○	○	○	○	○	pmdfs	✓	✓	✓	✓	○
lmsdp	○	○	✓	○	○	lmdfs	○	○	✓	○	○

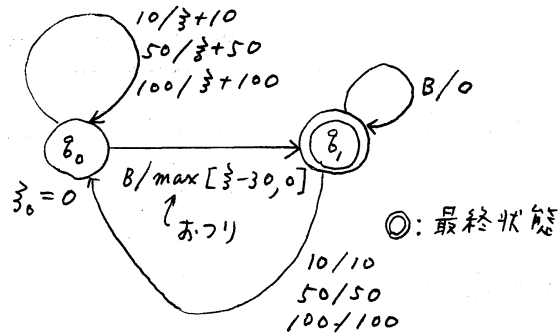
✓ : 閉じていない, ○ : 閉じている。

smsdp , pmsdp においては, $\sigma = O(\Pi)$ をみたす Π が存在する必要十分条件は σ が正規集合になることである。

5. システムモデルと LZ の sdp : この場合、関数 ℓ をシステム動作のコスト、何らかの Figure of Merit、あるいは記憶内容を表わす一つのパラメータと解釈できよう。

例. 自動入場券販売機:

$\Sigma = \{10 \text{ (10円入れる)}, 50 \text{ (50円)}, 100 \text{ (100円)}, B \text{ (ボタンを押す)}\}$



この立場に立つとき、 $O(\pi)$ 以外の性質にも注目する必要がある。たとえば、ハードウェアで実現するための理論、合成・分解、あるコスト値をもつ方策集合の性質、同定、可制御・可観測性 etc. である。成果次第では sdp をデジタルシステムの有力な表現手段とできよう。

6. 概念の拡張.

6.1 時間依存コスト関数. 一つの可能性は、時変システムへの拡張である。一例として、 ν を $E \times Q \times \Sigma \times T \rightarrow E$ とした 離散的有限システム (dfs, Discrete Finite System) がある。ただし、 $T = \{0, 1, \dots\}$ は離散時刻の集合である。[9] [10]の結果によれば、表現問題に関しては sdp の結果をそのまま拡張できるが、 $O(\pi)$ の性質、たとえばその閉包性については前表のように相当変化する (dfs の部分クラスも sdp と同様に定義する)。また、このモデルではほとんどの決定問題が時変性のために決定不能となり、あまり意味のある結果は得られないようである。

6.2 コストに依存した状態遷移. 状態遷移関数を $\lambda: \mathcal{Q} \times \Sigma \times E \rightarrow \mathcal{Q}$, すなわちその時々のコスト値きによ、遷移先が変動することを許せば、システムの表現能力が体系的に増大する.

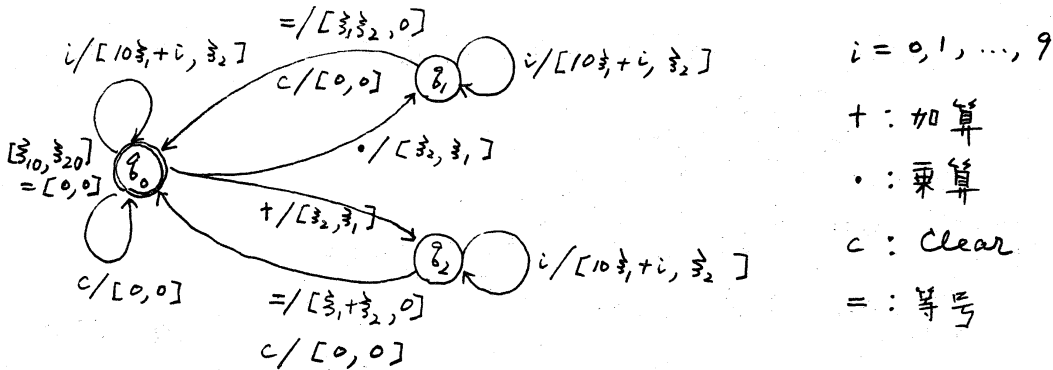
しかし、このモデルは、適当な仮定の下で、Turing 機械と同じ能力をもつことになり、理論的に目新しい結果は得られないようである。すなわち、 M として Turing 機械の制御部、コスト値きとしてその時々のテープの内容を数値化したものを想定すればよい。

なお、このモデルはプログラム理論との関連においても興味深い。

6.3 多重コスト関数. コスト関数として、 c_1, c_2, \dots, c_R の長個を考慮することもできる。一般には、 c_1, c_2, \dots, c_R を Γ の元へ数値化 (e.g., Gödel Numbering) できるから、 sdp としての能力は変化しないことに注意する必要がある。しかし、表現の簡潔性、コスト関数の単調性などについては本質的な影響がある。また、 $O(\Pi)$ の定義も Γ のポイントになる。

なお、このモデルは、 c_1, c_2, \dots, c_R をレジスタの内容と見ると理解しやすく、デジタルシステムの便利な表現手段となる可能性がある。

例. 電卓 (整数の加算・乗算のみを行なう) :



(たとえば, $123 \cdot 231 = + 13 =$ と押すと $123 \times 231 + 13$ が z_1 に入る.) 各枝には $a/[r_1(z_1, z_2, a), r_2(z_1, z_2, a)]$ が付されている.

6.4 非決定性状態遷移. 非決定性有限オートマトンに

たうとて, $\lambda : Q \times \Sigma \rightarrow 2^Q$ とする. すなわち, $\lambda(q, a) = \{q_1, q_2, \dots, q_k\}$ の形になる. コスト関数は遷移先に応じて $r_1(z_1, z_2, a), \dots, r_k(z_1, z_2, a)$ とするのが自然であろう. $x \in \Sigma^*$ による遷移 $\lambda(q_0, x)$ はいくつかの経路 π_1, \dots, π_s で表わされ, そのそれぞれに対しコスト値 $r_{\pi_1}(z_0, z_0, x), \dots, r_{\pi_s}(z_0, z_0, x)$ が定まる. r および $O(\pi)$ をつぎのように定義する.

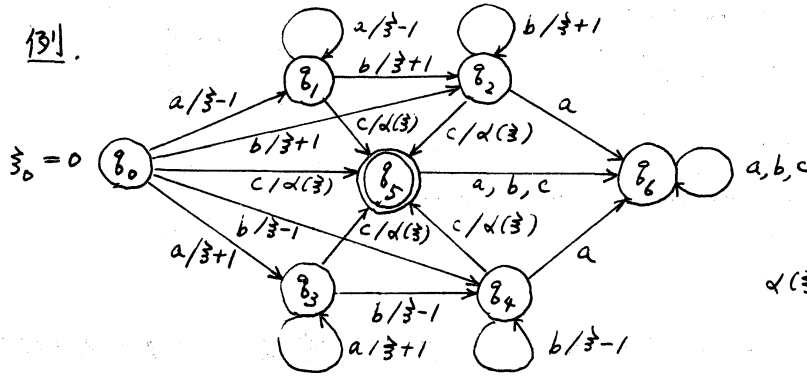
$$r(z_0, z_0, x) = \begin{cases} \infty & \text{if } \lambda(q_0, x) \cap Q_F = \emptyset \\ \min \{ r_{\pi_i}(z_0, z_0, x) \mid \pi_i \text{ は } q \in Q_F \text{ に至る} \}, & \text{その他} \end{cases}$$

$$O(\pi) = \{ x \in \Sigma^* \mid (\forall y \in \Sigma^*) (r(z_0, z_0, x) \leq r(z_0, z_0, y)) \}.$$

以上のモデルを非決定性 sdp と呼ぶ. このとき, 任意の非

決定性 sdp Π に対し, $O(\Pi) = O(\Pi)$ をみたす (決定性) sdp が存在する. しかし, この性質を msdp に拡張することはできない.

例.



$$\alpha(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

明らかに $O(\Pi) = \{a^r b^s c \mid r \neq s\}$ かつ各コスト関数は単調であるが, この $O(\Pi)$ は msdp の最適方針の集合とはならない. なお, 本モデルは阿部のペナルティオートマトン [11], 磯道の Parallel Machine [12] などと特殊な場合とを含むという意味でも興味がある.

6.5 確率的状態遷移. 状態遷移を確率的にした確率 sdp

を考之る. すなわち, 各 $a \in \Sigma$ に対し $n \times n$ 確率行列 $P^a = [p_{ij}^a]$ およびコスト行列 $H^a = [h_{ij}^a]$ が定められ, 初期確率ベクトル p_0 および最終状態ベクトル $\gamma = [\gamma_j]$ ($\gamma_j = 1$ if $s_j \in Q_F$, 0 if $s_j \notin Q_F$) に対し, $x = a_1 a_2 \dots a_t$ が

$$p_0 P^{a_1} P^{a_2} \dots P^{a_t} \gamma \geq c \text{ (定数)}$$

を満足せば Π に受理されるとする. x による遷移は n^t 個の経路で表わされ, 各経路 π に対し π をとる確率 p_π が計算でき, $h_\pi(s_0, s_0, x)$ も非決定性 sdp と同様に定まる. よって,

$$h(s_0, s_0, x) = \sum_{\pi} p_\pi h_\pi(s_0, s_0, x)$$

$$O(\Pi) = \{x \in F(\Pi) \mid (\forall y \in F(\Pi)) (h(s_0, s_0, x) \leq h(s_0, s_0, y))\}$$

と定義する。

このモデルはマルコフ決定過程 Π' と類似しているが、つぎの点で異なる。1) Π' は Q_F をもたない、つまり $F(\Pi) = \Sigma^*$ 。
 2) Π' は通常、定常状態での平均コストを問題にする。3) Π' では状態観測可能であり、各状態に対する決定を考える。4) Π' のコスト関数は、通常、 $g(z, \alpha) = \xi + c_2 \alpha$ に限定される。

確率 sdp における $O(\Pi)$ の性質、 $O(\Pi)$ を求めるアルゴリズム等に関しては今後の研究に待たねばならない。

6.6 Fuzzy 状態遷移。状態遷移に Fuzzy 性をもたせると Fuzzy sdp が得られる。定義自体は確率 sdp にたよってできるが、その性質についてはほとんど未知である。

文献

- [1] R. M. Karp and M. Held, Finite-state processes and dynamic programming, SIAM J. on Applied Math. 15(1967) 693-718.
- [2] T. Ibaraki, Representation theorems for equivalent optimization problems, Information and Control 21(1972), 397-435; 茨木, 有限状態逐次決定過程の基礎理論と動的計画法への応用, 信学論 55-D (1972), 419-426.
- [3] -----, Finite state representations of discrete optimization problems, SIAM J. Computing 2(1973), 193-822.
- [4] -----, Minimal representations of some classes of dynamic programming, to appear in Information and Control.

- [5] 田中, 茨木, 長谷川, ある種の有限状態最適化問題の最小表現, 信学論 57-D (1974), 637-644.
- [6] T. Ibaraki, Classes of discrete optimization problems and their decision problems, J. Computer and System Sciences 8(1974), 84-116.
- [7] -----, Solvable classes of discrete dynamic programming, J. Mathematical Analysis and Applications 43(1973), 642-693; 茨木, 動的計画法の可解なクラス, 信学論 55-D (1972), 738-745; 茨木, ある種の離散的最適化問題の有限状態表現について, 信学論 55-D (1972), 815-822.
- [8] -----, On the optimality of algorithms for finite state sequential decision processes, to appear in J. Math. Analysis and Applications.
- [9] -----, 離散的有限システムの最適制御, 信学論 57-D (1974), 85-92.
- [10] -----, 離散的有限システムのいくつかの部分クラス, 信学論 58-D (1975), 掲載予定.
- [11] 阿部, Penalty automatonによる記号系列パターンの識別, 本講究録.
- [12] 磯道, 時系列パターンのセグメンテーションフリーな認識, 本講究録.