

相互依存型決定過程 — 1 ステージモデル —

九州工業大学・大学院工学研究院 藤田 敏治

Toshiharu Fujita

Graduate School of Engineering, Kyushu Institute of Technology

1 はじめに

相互依存型決定過程 [3, 4, 6, 8] とは、複数の決定過程から構成され、各決定過程の利得関数が他の決定過程問題群の最適値の関数として定まるものである。このような依存構造が再帰的に与えられ、各決定過程における終端状態に対しては、通常の終端利得関数が割り当てられる。状態推移の立場からは、ノンシリアル動的計画 [1, 9] の一つの具体的モデルと考えられる。この種の相互依存構造により、多角形から凸多面体を構成する問題がうまく扱えることは [7] で示している。本稿では、2つの決定過程から構成される相互依存型決定過程で、特に1ステージの決定過程のみから構成されるモデルを考える。そして、この種のモデルを用いると、組合せゲーム [5] の必勝法を求める問題が、より容易に取り扱えることを示す。

2 定式化

本稿で扱う相互依存型決定は2つの決定過程からなるものである。ここでは、その2つをそれぞれ主決定過程、副決定過程と呼ぶこととし、ともに加法型評価の最小化問題である場合について考える。最初に一般の多段決定過程の場合について考え、その後、次節の応用例で用いる1ステージの場合について結果を与える。

2.1 主決定過程と副決定過程

まず、初期状態 $x_0 \in X \setminus T_X$ に対する主決定過程問題を考える。

$$\begin{aligned} \text{主決定過程 } P(x_0) \quad & \text{minimize} \quad \left[r(x_0, u_0) + r(x_1, u_1) + \cdots + r(x_{N-1}, u_{N-1}) + r_G(x_N) \right] \\ & \text{subject to} \quad x_{n+1} = f_{XX}(x_n, u_n) \quad n = 0, 1, \dots, N-1 \\ & \quad \quad \quad u_n \in U(x_n) \quad n = 0, 1, \dots, N-1 \\ & \quad \quad \quad N = N(x_0, u_0, x_1, u_1, \dots) = \max\{n : x_n \notin T_X\} + 1 \end{aligned}$$

ただし

- (1) X は有限状態空間をあらわし、 $T_X \subset X$ は終了状態集合とする。また、 $x_n \in X$ は時刻 n における状態をあらわし、 $x_N \in T_X$ となる時刻 N で推移は終了する。
- (2) U は有限決定空間をあらわし、 $u_n \in U$ は時刻 n における決定をあらわす。以下、一般に与えられた集合 S に対し、 S の部分集合全体を 2^S で表す。ここで、写像 $U : X \setminus T_X \rightarrow 2^U \setminus \{\emptyset\}$ は各非終了状態に対し実行可能な決定全体を与えるものとする。
- (3) $f_{XX} : G_r(U) \rightarrow X$ は確定的推移法則をあらわす。ただし、

$$G_r(U) = \{(x, u) \mid u \in U(x), x \in X \setminus T_X\}$$

である。ある時刻において状態 x に対し決定 u を選んだ際、次の時刻で状態 $f_{XX}(x, u) \in X$ へ推移する。

(4) $r: G_r(U) \rightarrow \mathbf{R}$ はコスト関数をあらわし、 $r_G: T_X \rightarrow \mathbf{R}$ は終端コスト関数をあらわす。

副決定過程は、初期状態 $y_0 \in Y \setminus T_Y$ に対し、次で与えられる。

$$\begin{aligned} \text{副決定過程 } Q(y_0) \quad & \text{minimize } [q(y_0, v_0) + q(y_1, v_1) + \cdots + q(y_{M-1}, v_{M-1}) + q_G(y_M)] \\ & \text{subject to } y_{n+1} = f_{YY}(y_n, v_n) \quad n = 0, 1, \dots, M-1 \\ & \quad v_n \in V(y_n) \quad n = 0, 1, \dots, M-1 \\ & \quad M = M(y_0, v_0, y_1, v_1, \dots) = \max\{n : y_n \notin T_Y\} + 1 \end{aligned}$$

ただし

- (1) Y は有限状態空間をあらわし、 $T_Y \subset Y$ は終了状態集合とする。また、 $y_n \in Y$ は時刻 n における状態をあらわし、 $y_M \in T_Y$ となる時刻 M で推移は終了する。
- (2) V は有限決定空間をあらわし、 $v_n \in V$ は時刻 n における決定をあらわす。ここで、写像 $V: Y \setminus T_Y \rightarrow 2^V \setminus \{\phi\}$ は各非終了状態に対し実行可能な決定全体を与えるものとする。
- (3) $f_{YY}: G_r(V) \rightarrow Y$ は確定的推移法則をあらわす。ある時刻において状態 y に対し決定 v を選んだ際、次の時刻で状態 $f_{YY}(y, v) \in Y$ へ推移する。
- (4) $q: G_r(V) \rightarrow \mathbf{R}$ はコスト関数をあらわし、 $q_G: T_Y \rightarrow \mathbf{R}$ は終端コスト関数をあらわす。

また、状態空間 X と Y の間に確定的推移：

$$f_{XY}: G_r(U) \rightarrow Y, \quad f_{YX}: G_r(V) \rightarrow X$$

を与える。ここで、 f_{XY} は、主決定過程の状態 $x \in X$ とその時の決定 $u \in U(x)$ に対し、副決定過程の初期状態 $f_{XY}(x, u) \in Y$ を定め、 f_{YX} は、副決定過程の状態 $y \in Y$ とその時の決定 $v \in V(y)$ に対し、主決定過程の初期状態 $f_{YX}(y, v) \in X$ を定める。さらに、主決定過程 $P(x_0)$ のコスト関数 $r(x, u)$ および副決定過程 $Q(y_0)$ のコスト関数 $q(y, v)$ は、副決定過程問題 $Q(f_{XY}(x, u))$ 、主決定過程問題 $P(f_{YX}(y, v))$ の最適値の関数としてそれぞれ次のように定まるものとする。なお、 $f_{XY}(x, u)$ 、 $f_{YX}(y, v)$ が終了状態のときは、終端利得関数値の関数となる。

$$\begin{aligned} r(x, u) &= \begin{cases} R(q_G(y_0)) & y_0 = f_{XY}(x, u) \in T_Y \\ R\left(\min_{\substack{v_n \in V(y_n) \\ n=0,1,\dots,M-1}} [q(y_0, v_0) + \cdots + q_G(y_M)]\right) & y_0 = f_{XY}(x, u) \notin T_Y \end{cases} \\ q(y, v) &= \begin{cases} Q(r_G(x_0)) & x_0 = f_{YX}(y, v) \in T_X \\ Q\left(\min_{\substack{u_n \in U(x_n) \\ n=0,1,\dots,N-1}} [r(x_0, u_0) + \cdots + r_G(x_N)]\right) & x_0 = f_{YX}(y, v) \notin T_X \end{cases} \end{aligned}$$

ただし、

$$R: \mathbf{R} \rightarrow \mathbf{R}, \quad Q: \mathbf{R} \rightarrow \mathbf{R}$$

は、あらかじめ与えられた関数とする。

このとき、解くべき問題は主決定過程側で与えられ、初期状態 $\bar{x}_0 \in X \setminus T_X$ に対し $(P, Q, \bar{x}_0)$ と表す。なお、再帰的に生じる両過程を通じて生成される全状態列に対し有限段での終了を仮定する。

2.2 再帰式

主決定過程および副決定過程の最適値関数 W, Z をそれぞれ次で定義する.

$$\begin{aligned} W(x_0) &= r_G(x_0) & x_0 \in T_X \\ W(x_0) &= \min_{\substack{u_n \in U(x_n) \\ n=0,1,\dots,N-1}} [r(x_0, u_0) + r(x_1, u_1) + \dots + r_G(x_N)] & x_0 \notin T_X \\ Z(y_0) &= q_G(y_0) & y_0 \in T_Y \\ Z(y_0) &= \min_{\substack{v_n \in V(y_n) \\ n=0,1,\dots,M-1}} [q(y_0, v_0) + q(y_1, v_1) + \dots + q_G(y_M)] & y_0 \notin T_Y \end{aligned}$$

このとき, W, Z に対し, それぞれ次の再帰式が成り立つ [2].

$$\begin{aligned} W(x) &= r_G(x) & x \in T_X \\ W(x) &= \min_{u \in U(x)} [r(x, u) + W(f_{XX}(x, u))] & x \notin T_X \\ Z(y) &= q_G(y) & y \in T_Y \\ Z(y) &= \min_{v \in V(y)} [q(y, v) + Z(f_{YY}(y, v))] & y \notin T_Y \end{aligned}$$

ここで,

$$r(x, u) = R\left(Z(f_{XY}(x, u))\right), \quad q(y, v) = Q\left(W(f_{YX}(y, v))\right)$$

と表されることに注意すると, 次の関係式を得る.

定理 2.1 最適値関数 W, Z に対し, 次の相互依存型再帰式が成り立つ.

$$\begin{aligned} W(x) &= r_G(x) & x \in T_X \\ W(x) &= \min_{u \in U(x)} \left[R\left(Z(f_{XY}(x, u))\right) + W(f_{XX}(x, u)) \right] & x \notin T_X \end{aligned} \quad (1)$$

$$\begin{aligned} Z(y) &= q_G(y) & y \in T_Y \\ Z(y) &= \min_{v \in V(y)} \left[Q\left(W(f_{YX}(y, v))\right) + Z(f_{YY}(y, v)) \right] & y \notin T_Y \end{aligned} \quad (2)$$

求めるべき最適値は $W(\bar{x}_0)$ である. また, (1) の最小値を与える決定 $u \in U$ を $\pi_X^*(x)$, (2) の最小値を与える決定 $v \in V$ を $\pi_Y^*(y)$ とおくと, (π_X^*, π_Y^*) が相互依存型決定過程 $(P, Q, \bar{x}_0)$ に対する最適政策を与える. 定理 2.1 により最適解を求めるための再帰的解法が与えられる.

2.3 1 ステージモデル

相互依存型決定過程において, 特に 1 ステージの場合について結果を整理しておく. この場合, 問題はそれぞれ

$$\begin{aligned} \text{主決定過程 } P(x_0) \quad & \text{minimize } [r(x_0, u_0) + r_G(x_1)] \\ & \text{subject to } x_1 = f_{XX}(x_0, u_0) \\ & \quad u_0 \in U(x_0), \end{aligned}$$

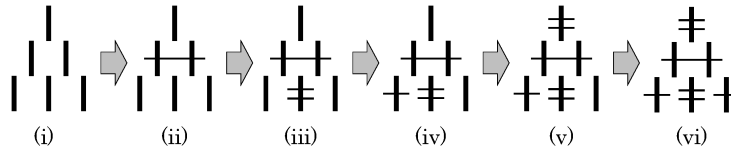


図 3.1: ゲームの流れ

$$\begin{aligned} \text{副決定過程 } Q(y_0) \quad & \text{minimize } [q(y_0, v_0) + q_G(y_1)] \\ & \text{subject to } y_1 = f_{YY}(y_0, v_0) \\ & v_0 \in V(y_0), \end{aligned}$$

ただし

$$\begin{aligned} r(x, u) &= \begin{cases} R(q_G(y_0)) & y_0 = f_{XY}(x, u) \in T_Y \\ R\left(\min_{v_0 \in V(y_0)} [q(y_0, v_0) + q_G(y_1)]\right) & y_0 = f_{XY}(x, u) \notin T_Y, \end{cases} \\ q(y, v) &= \begin{cases} Q(r_G(x_0)) & x_0 = f_{YX}(y, v) \in T_X \\ Q\left(\min_{u_0 \in U(x_0)} [r(x_0, u_0) + r_G(x_1)]\right) & x_0 = f_{YX}(y, v) \notin T_X \end{cases} \end{aligned}$$

となり、したがって再帰式は次で与えられる。

系 2.1

$$\begin{aligned} W(x) &= r_G(x) & x \in T_X \\ W(x) &= \min_{u \in U(x)} [R(Z(f_{XY}(x, u))) + r_G(f_{XX}(x, u))] & x \notin T_X \\ Z(y) &= q_G(y) & y \in T_Y \\ Z(y) &= \min_{v \in V(y)} [Q(W(f_{YX}(y, v))) + q_G(f_{YY}(y, v))] & y \notin T_Y \end{aligned}$$

3 ピラミッドゲーム必勝法への応用

3.1 ピラミッドゲームとは

ここで扱うピラミッドゲームとは、2人対戦のいわゆる組合せゲームの一種である。最初、縦棒が図 3.1(i) のように配置されており、2人のプレーヤーが次のルールに従って交互に棒を消していく。

1. 自分の順番では、必ず1本以上の棒を消さなければならない。
2. 複数の棒を消す場合、それらは一連の隣接した棒でなければならない。

そして、最後の1本を消したほうが負けとなる。図 3.1 は、ゲームの経過について表した例である。まず、先手が第2層の2本を消し、次いで後手が第3層の中央1本を消す。その後、先手が第3層の左端1本を消し、後手が第1層の1本を消す。そして、先手が最後に残った第3層の右端1本を消さなければならなくなり、後手の勝ちとなる。

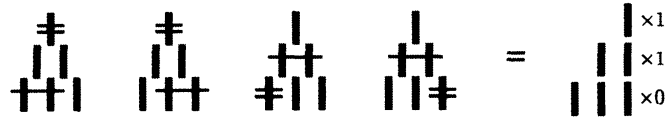


図 3.3: 同値な状況

一般に、棒の配置は何層でもよく、第 n 層には n 本の縦棒を並べる。この種のゲームが、筆者の所属する大学の学生の間でピラミッドゲームと呼ばれていたのも、ここでは、その名前と呼ぶこととする。

本稿での興味は、ピラミッドゲームに必勝法が存在するか否かを知ること、そして存在する場合は、その必勝法を求めることにある。そこで、相互依存型決定過程の 1 ステージモデルを用いて、先手の必勝法を求める問題を考える。

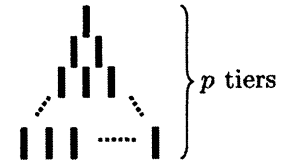
3.2 定式化

基本的な考え方は、主決定過程と副決定過程でそれぞれ先手と後手の手番を表すこととし、各時点でのゲームの状況を状態、プレイヤーの取る手を決定と考え、目的関数によって勝敗を判断させることである。なお、本節では状態および決定については主・副決定過程で共通に取られるため、

$$Y := X, \quad T_X = T_Y := T, \quad V := U, \quad f_{XX} = f_{YY} := f, \quad f_{XY} = f_{YX} := g$$

とおき、共通の記号 X, T, U, f, g を定める。

一般に p 層のピラミッドゲーム (図 3.2) を扱うこととする。最初に状態および状態空間について考える。図 3.3 に示される 4 つの状況は、一見異なるものであるが、その後のゲーム展開を考えれば実は同値な状況である。これらは、いずれも「1 本単独で存在する縦棒が 1 か所と 2 本の縦棒が並んで存在する所が 1 か所である」という点で同じなのである。すなわち、ゲームを考えるうえで必要な情報は、各時点で何本の縦棒が隣接して存在しているか、そして、それらが何か所あるか、である。縦棒が残っている場所についての情報は、本質的に不要なのである。したがって、状態は p 次元ベクトルで表現でき、状態 $(x_1, x_2, \dots, x_p)$ の x_n ($n = 1, 2, \dots, p$) が n 本並びの存在する個数を表すとする。状態空間は次で与えられる。

図 3.2: p 層ゲームの初期配置

$$X = \left\{ (x_1, x_2, \dots, x_p) \mid x_i = 0, 1, 2, \dots, \quad i = 1, 2, \dots, p \right\}.$$

初期状態は $(1, 1, \dots, 1)$ であり、図 3.3 の状況は $(1, 1, 0)$ と表される。なお、状態 $(1, 0, 0, \dots, 0)$ は、1 本単独で存在する縦棒が 1 か所のみとなり、この状況がプレイヤーの手番に回ってきたときは、最後の 1 本を消す以外に選択肢はない。すなわち、負けが決定した状態なのである。このことから、終了状態が定まり

$$T = \{(1, 0, 0, \dots, 0)\} \subset X.$$

とおくことができる。

次に決定について考える。図 3.4 は、ある手番における 1 つの縦棒の消し方であるが、このように各プレイヤーが取る手は 3 つの数値で定まる。すなわち、何本並びの縦棒を対象にするのか、そして、その何本目から何本目を消すのか、で定まる。

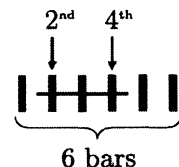


図 3.4: 棒の消し方 (例)

よって、決定を3次元ベクトル (u_1, u_2, u_3) で表すこととし、これが「 u_1 本並びの列において u_2 本目から u_3 本目までを消す」を意味するものとする。例えば、図 3.4 の決定は $(6, 2, 4)$ となる。以上より、決定空間は次で与えられる。

$$U = \bigcup_{n=1}^p U_n,$$

ただし

$$U_n = \left\{ (u_1, u_2, u_3) \left| \begin{array}{l} u_1 = n \\ u_2 < \frac{n}{2} + 1 \\ u_2 \leq u_3 \leq u_1 - u_2 + 1 \\ u_2, u_3 = 1, 2, \dots \end{array} \right. \right\}, \quad i = 1, 2, \dots, p$$

である。 U_n は n 本並びの列に対する決定全体を表している。また、 U_n の定義内には、左右対称な消し方を除く条件が含まれている。決定 $(6, 2, 4)$ と決定 $(6, 3, 5)$ は、ともに 6 本並びの列を 1 減少させ、1 本単独および 2 本並びの列を 1 ずつ増加させるという意味で同値であり、両方を考慮する必要はないのである。

さて、決定制約（実行可能な決定全体）についてであるが、基本的には

$$U(x) = \bigcup_{n; x_n \geq 1} U_n, \quad x = (x_1, x_2, \dots, x_p) \in X$$

で与えられることが容易にわかる。ここで注意したいのは、状態 $(0, 0, 1)$ に対する実行可能な決定 $(3, 1, 3)$ は、負ける必要がないときにわざと負ける手になっている点である。このような手を含めることは、問題の性質上、冗長になるため省略したい。よって、状態 $x = (x_1, x_2, \dots, x_p)$ が $x_1 = 0$ かつ $\sum_{i=2}^p x_i = 1$ を満たすとき、 $x_n = 1$ なる n に対しては U_n 次のように置き換えて考えるものとする。

$$U_n \leftarrow U_n \setminus \{(n, 1, n)\}.$$

各決定過程内の状態推移 f については、今回の定式化において終端利得の役割が不要であることから、ダミーの終了状態への推移を人為的に作っておく。

$$f(x, u) = (0, 0, \dots, 0), \quad x \in X, u \in U(x).$$

なお、決定制約の定め方により、必ず状態 $(1, 0, \dots, 0)$ へたどり着き、この状態で停止するため、決定過程間の推移 g によって状態 $(0, 0, \dots, 0)$ を生じることはない。これを終了状態に追加する。すなわち、改めて

$$T = \{(1, 0, \dots, 0), (0, 0, \dots, 0)\}$$

とおく。また、ゲームのルールにより、決定過程間の状態推移 $g(x, u) = (x'_1, x'_2, \dots, x'_p)$, $(x \in X, u \in U(x))$ は次のように定まる。

$$x'_i = \begin{cases} x_i - 1 & (i = u_1) \\ x_i + 1 & \left(\begin{array}{l} i = u_2 - 1 \neq u_1 - u_3, u_2 - 1 > 0 \\ \text{または} \\ i = u_1 - u_3 \neq u_2 - 1, u_1 - u_3 > 0 \end{array} \right) \\ x_i + 2 & (i = u_2 - 1 = u_1 - u_3) \\ x_i & (\text{その他}). \end{cases}$$

最後に利得関数を定める。ここでは、目的関数の値が、勝ち状態（必勝法あり）であれば 0 を、負け状態（必勝法なし）であれば 1 となるようにする。終端利得関数についてはダミー状態に対しては役割が不要なため 0 とし、

$$r_G(x) = q_G(x) = \begin{cases} 1 & x = (1, 0, \dots, 0) \\ 0 & x = (0, 0, \dots, 0) \end{cases}$$

とおく。利得関数は

$$R(\alpha(x, u)) = \begin{cases} 0 & \alpha(x, u) = 1 \\ 1 & \alpha(x, u) = 0 \end{cases} \quad \left(\begin{array}{l} = 1 - \alpha(x, u) \end{array} \right),$$

ただし

$$\begin{aligned} \alpha(x, u) &= W(g(x, u)) \quad \left(= Z(g(x, u)) \right) \\ &= \begin{cases} r_G(x_0) & x_0 = g(x, u) \in T \\ \min_{u_0 \in U(x_0)} [r(x_0, u_0) + r_G(x_1)] & x_0 = g(x, u) \notin T \end{cases} \end{aligned}$$

とし、 $r(x, u) = q(x, u) = R(x, u)$ とおく。

以上より、相互依存型決定過程を構成する 2 つの決定過程が全く同じものとなるため、再帰式は共通となり、かつダミー状態に対する終端利得が 0 であることから次で与えられる。

$$\begin{aligned} W(x) &= r_G(x) & x \in T \\ W(x) &= \min_{u \in U(x)} [R(\alpha(x, u))] & x \notin T. \end{aligned}$$

特に、 $x \notin T$ に対しては、 $R(\alpha(x, u)) = 1 - \alpha(x, u) = 1 - W(g(x, u))$ となるため

$$\begin{aligned} W(x) &= r_G(x) & x \in T \\ W(x) &= \min_{u \in U(x)} [1 - W(g(x, u))] & x \notin T \end{aligned}$$

である。この再帰式を用いて $W(1, 1, \dots, 1)$ を求め、それが 0 のとき、先手に必勝法が存在する。

3.3 数値例

例 3.1 (2 段ゲーム) $p = 2$ で、初期状態は $(1, 1)$ である。まず、終了状態に対しては $W(1, 0) = 1$ である。次に $U(1, 1) = \{(1, 1, 1), (2, 1, 1), (2, 1, 2)\}$ より

$$W(1, 1) = \min_{u \in U(1, 1)} [1 - W(g((1, 1), u))] = \min [(1 - W(0, 1)), (1 - W(2, 0)), (1 - W(1, 0))].$$

ここで $U(0, 1) = \{(2, 1, 1)\}$, $U(2, 0) = \{(1, 1, 1)\}$ より

$$W(0, 1) = \min_{u \in U(0, 1)} [1 - W(g((0, 1), u))] = 1 - W(1, 0) = 1 - 1 = 0, \quad \pi^*(0, 1) = (2, 1, 1)$$

$$W(2, 0) = \min_{u \in U(2, 0)} [1 - W(g((2, 0), u))] = 1 - W(1, 0) = 1 - 1 = 0, \quad \pi^*(2, 0) = (1, 1, 1).$$

よって

$$W(1, 1) = \min[(1 - 0), (1 - 0), (1 - 1)] = 0, \quad \pi^*(1, 1) = (2, 1, 2).$$

以上より、必勝法は存在し、先手は最初に 2 本並び列において 1 本目から 2 本目を消せばよい。□

例 3.2 (3 段ゲーム) $p = 3$ で, 初期状態は $(1, 1, 1)$ である. 一般に $k = 2, 3, \dots, p$ に対し

$$\begin{aligned} W(0, \dots, 0, \overset{k\text{th}}{\downarrow} 1, 0, \dots, 0) &= 0, & \pi^*(0, \dots, 0, \overset{k\text{th}}{\downarrow} 1, 0, \dots, 0) &= (k, 1, k-1) \\ W(1, 0, \dots, 0, \overset{k\text{th}}{\downarrow} 1, 0, \dots, 0) &= 0, & \pi^*(1, 0, \dots, 0, \overset{k\text{th}}{\downarrow} 1, 0, \dots, 0) &= (k, 1, k) \end{aligned}$$

が成り立ち, $l = 1, 2, \dots, p$ に対し

$$v(2l+1, 0, 0, \dots, 0) = 1, \quad v(2l, 0, 0, \dots, 0) = 0, \quad \pi^*(2l, 0, 0, \dots, 0) = (1, 1, 1)$$

が成り立つことを利用する. まず, 終了状態に対しては $W(1, 0, 0) = 1$ である. 次に $U(1, 1, 1) = \{(1, 1, 1), (2, 1, 1), (2, 1, 2), (3, 1, 1), (3, 1, 2), (3, 1, 3), (3, 2, 2)\}$ より

$$\begin{aligned} W(1, 1, 1) &= \min_{u \in U(1, 1, 1)} [1 - W(g((1, 1, 1), u))] \\ &= \min \left[(1 - W(0, 1, 1)), (1 - W(2, 0, 1)), (1 - W(1, 0, 1)), \right. \\ &\quad \left. (1 - W(1, 2, 0)), (1 - W(2, 1, 0)), (1 - W(1, 1, 0)), (1 - W(3, 1, 0)) \right]. \end{aligned}$$

ここで, $W(1, 1, 0) = W(1, 0, 1) = 0$ であり, $U(0, 1, 1) = \{(2, 1, 1), (2, 1, 2), (3, 1, 1), (3, 1, 2), (3, 1, 3), (3, 2, 2)\}$ より

$$\begin{aligned} W(0, 1, 1) &= \min_{u \in U(0, 1, 1)} [1 - W(g((0, 1, 1), u))] \\ &= \min \left[(1 - W(1, 0, 1)), (1 - W(0, 0, 1)), (1 - W(0, 2, 0)), \right. \\ &\quad \left. (1 - W(1, 1, 0)), (1 - W(0, 1, 0)), (1 - W(2, 1, 0)) \right]. \end{aligned}$$

$W(1, 1, 0) = W(1, 0, 1) = W(0, 1, 0) = W(0, 0, 1) = 0$ で, $U(0, 2, 0) = \{(2, 1, 1), (2, 1, 2)\}$ より

$$\begin{aligned} W(0, 2, 0) &= \min_{u \in U(0, 2, 0)} [1 - W(g((0, 2, 0), u))] = \min \left[(1 - W(1, 1, 0)), (1 - W(0, 1, 0)) \right] \\ &= \min \left[(1 - 0), (1 - 0) \right] = 1. \end{aligned}$$

さらに, $U(2, 1, 0) = \{(1, 1, 1), (2, 1, 1), (2, 1, 2)\}$ より

$$\begin{aligned} W(2, 1, 0) &= \min_{u \in U(2, 1, 0)} [1 - W(g((2, 1, 0), u))] \\ &= \min \left[(1 - W(1, 1, 0)), (1 - W(3, 0, 0)), (1 - W(2, 0, 0)) \right] \\ &= \min \left[(1 - 0), (1 - 1), (1 - 0) \right] = 0, \quad \pi^*(2, 1, 0) = (2, 1, 1). \end{aligned}$$

よって

$$\begin{aligned} W(0, 1, 1) &= \min \left[(1 - 0), (1 - 0), (1 - 1), (1 - 0), (1 - 0), (1 - 0) \right] = 0, \\ \pi^*(0, 1, 1) &= (3, 1, 1). \end{aligned}$$

同様に

$$W(2, 1, 0) = 0, \quad \pi^*(2, 1, 0) = (2, 1, 1), \quad W(0, 1, 1) = 0, \quad \pi^*(0, 1, 1) = (3, 1, 1),$$

$$W(2, 0, 1) = 0, \quad \pi^*(2, 0, 1) = (3, 1, 2), \quad W(1, 2, 0) = 0, \quad \pi^*(1, 2, 0) = (1, 1, 1), \\ W(3, 1, 0) = 0, \quad \pi^*(3, 1, 0) = (2, 1, 2).$$

したがって

$$W(1, 1, 1) = \min \left[(1-0), (1-0), (1-0), (1-0), (1-0), (1-0), (1-0) \right] = 1.$$

以上より, $p = 3$ のとき, 先手に必勝法は存在しない. $\square$

なお, 計算機による数値計算の結果, $p = 4, 5, 6, 8, 9, 10, 12, 13, 14, 16$ のとき, 先手に必勝法が存在することを確認した.

謝辞

本研究は科研費 15K05004 の助成を受けたものである.

参考文献

- [1] U. Bertelé and F. Brioschi, Nonserial Dynamic Programming, Academic Press, New York, 1972.
- [2] T. Fujita, Re-examination of Markov policies for additive decision process, Bulletin of Informatics Cybernetics, 29, No.1, 51-65, 1997.
- [3] T. Fujita, 結合型評価をもつ相互依存型決定過程, 京都大学数理解析研究所講究録 1802, 78-84, 2012.
- [4] T. Fujita, Associative Criteria in Mutually Dependent Markov Decision Processes, Proceedings of IIAI International Conference on Advanced Applied Informatics (IIAI AAI 2014), 147-150, 2014.
- [5] T. Fujita, Sure Ways to Win a Game — Nondeterministic Dynamic Programming Approach —, Bulletin of the Kyushu Institute of Technology. Pure and applied mathematics, 62, 1-14, 2015.
- [6] T. Fujita, 相互依存型マルコフ決定過程 — 結合型評価 —, 京都大学数理解析研究所講究録 1939, 189-195, 2015.
- [7] 藤田敏治, 長友健太郎, 折り紙ユニットを用いた凸多面体の構成 — 相互依存型決定過程によるアプローチ —, 京都大学数理解析研究所講究録, 1912, 17-25, 2014.
- [8] T. Fujita and A. Kira, Mutually Dependent Markov Decision Processes, Journal of Advanced Computational Intelligence and Intelligent Informatics, 18(6), 2014, 992-968.
- [9] G. L. Nemhauser, Introduction to Dynamic Programming, Wiley, New York, 1966.