数理解析研究所講究録
947 巻 1996 年 110-120

二人確率ゲームの最適戦略に関する計算複雑性

The Computational Complexity of Optimal Strategies for Two-Person Games Under Uncertainty

山家 明男　　櫻井 幸一
Akio YANBE　Kouichi SAKURAI

九州大学 工学部 情報工学科
〒 812-81 福岡市東区箱崎 6-10-1
Department of Computer Science and Communication Engineering
Kyushu University
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-81, Japan
Phone.092-641-1101 Fax.092-632-5204
{yanbe,sakurai}@csce.kyushu-u.ac.jp

概要: On-line Metrical Task System [Borodin, A., Linial, N. and Saks, M.E., *J. Assoc. Comput. Mach.***39**] における最適なオンラインアルゴリズムは Mean Payoff Game [Ehrenfeucht, A. and Mycielski, J., *Int. J. Game Theory* **8**] という二人ゲームの最適戦略として考えることができる．このゲームの最適戦略の計算複雑性が，オンラインアルゴリズムの設計の難しさとなる．しかし，今のところ Mean Payoff Game の最適戦略を計算する多項式時間アルゴリズムは見つかっていない．

我々は On-line Metrical Task System で確率的な事象を扱うために，Mean Payoff Game の discounted version である Discounted Payoff Game に確率的な要素を入れた二人確率ゲームを考えた．そして，最適値および純粋最適戦略の存在を示すとともに，その計算複雑性は NP ∩ co-NP であるを示した．

さらに、二人のうち一人をランダムプレイヤーとした場合の Discounted Payoff Game における最適戦略は多項式時間で計算できることを示し，On-line Metrical Task System において要求がランダムにおこると仮定した場合には，最適なオンラインアルゴリズムを多項式時間で設計できることが示せた．

本研究の解析では，代表的な二人確率ゲームである Simple Stochastic Game [Shapley, L.S., *Proc. Nat. Acad. Sci. U.S.A.* **39**] が重要な役割をはたす．

Abstract: An optimal on-line algorithm for on-line metrical task systems [Borodin, A., Linial, N. and Saks, M.E., *J. Assoc. Comput. Mach.***39**] is regarded as an optimal strategy of Mean Payoff Game [Ehrenfeucht, A. and Mycielski, J., *Int. J. Game Theory* **8**]. Then, the complexity of finding an optimal strategy of Mean Payoff Game gives a measure of the hardness on the designing an on-line algorithm for task systems. Note that until today no polynomial-time algorithm is known for the mean payoff game.

For treating the probabilistic occurrences on task systems, we introduce Probabilistic Discounted Payoff Games as two-person games under uncertainty by adding a probabilistic move to Discounted Payoff Games(DPGs) that are the discounted version of Mean Payoff Games. We show that each player has an optimal pure strategy for Probabilistic DPGs and that the complexity class of Probabilistic DPGs belongs to NP ∩ co-NP.

Moreover, in DPGs where one of the two players acts as a random player, we show that an optimal strategy is computable in polynomial-time. This implies that, if requests are sent randomly in task system with discounted costs, an optimal on-line algorithm is computable in polynomial-time.

Simple Stochastic Game [Shapley, L.S., *Proc. Nat. Acad. Sci. U.S.A.* **39**] plays an important role in our analysis of Probabilistic DPGs.

1 はじめに

Mean Payoff Game は二人のプレイヤーによって行われるゲームである．枝に重みのついた有向グラフ上で，スタート地点から枝に沿って駒を交互に動かしていき，通った枝を $e_1, e_2, \cdots$ とする．通った枝の重みの和を通った枝の本数で割った値 (mean payoff) をこのゲームの利得とする．先手はこの利得を大きくすることが目的で，後手は小さくすることが目的である．このゲームには最適値 $\mathcal{V}$ が存在し，先手には $\mathcal{V}$ 以上の利得を保証する戦略，また後手には利得が $\mathcal{V}$ 以下になることを保証する戦略が存在する．

このゲームの最適値を求めたり最適戦略を計算したりする多項式時間アルゴリズムは見つかっておらず，最適値がある一定値より大きいかどうかを問う判定問題は NP $\cap$ co-NP であることが知られている [ZP95]．また，このゲームは metrical task system [BLS92] や string matching [CHPZ95] のオンラインアルゴリズムへの応用としても考えられている．

Mean Payoff game の変形版である Discounted Payoff Game(DPG) について，ランダムプレイヤーを追加したゲームを定義し，最適値および最適戦略が存在することを示すとともに，二人のプレイヤーのうち一人をランダムプレイヤーとした時の DPG では多項式時間で最適値を求められることを示した．これによって我々の定義したランダムプレイヤーを追加した DPG に関する判定問題も NP $\cap$ co-NP であることがわかった．

Mean Payoff Game および Discounted Payoff Game は代表的な二人確率ゲームである Simple Stochastic Game(単純確率ゲーム) [Con92] に変換することができる [ZP95]．単純確率ゲームは二人のプレイヤーによって争われるゲームであるが，二人以外にランダムプレイヤーの手番があり，ランダムプレイヤーの手番ではゲームの局面がランダムに変化するという確率的要素を含んでいる．

Shapley [Sha53] が確率ゲームを導入してから，確率ゲームの様々な変型モデルが考案され，その解法アルゴリズムが開発されてきた．単純確率ゲームは確率ゲームの最も簡単なモデルであり，P 完全なゲームとしてよく知られている TWO PLAYER GAME [GHR95] に，確率的要素を入れたものになっている．つまり，TWO PLAYER GAME は，頂点ごとに，駒を動かすことのできるプレイヤーが決められている有向グラフ上で行われるが，有向グラフにランダムプレイヤーが駒を動かす頂点も含めたものが単純確率ゲームである．この単純確率ゲームにも最適戦略を求めるたくさんのアルゴリズムが提案されているが，最悪の場合に指数時間かかることが示されたり [Con93]，アルゴリズム自体が間違いであることが証明されたりしている [Van78]．現在この単純確率ゲームが多項式時間で解けるかどうかは未解決である．Hoffman と Karp によって提案された戦略改良型アルゴリズムは正しいことは証明されているが，最悪の場合指数時間かかるかどうかわかっていない [HK66]．また，現在知られている正しくて最も速いアルゴリズムは期待値が $2^{O(\sqrt{n})}$ 時間の確率アルゴリズムである [Lud95]．

本研究では単純確率ゲームの全ての最適戦略の計算複雑性について考察を行なった．一般的に解の存在の判定問題は多項式時間で計算可能でも，解の個数を数え上げることはかなり難しくなる問題が知られている [Val79]．しかし，この単純確率ゲームでは最適戦略の数が実際に指数個になる場合があるにもかかわらず，データ構造をうまくとることによって，条件付きながら全ての最適戦略の記述を多項式サイズに収められることを証明した [YS95]．

また，ランダムプレイヤーを付け加えるのではなく，もとの P 完全な二人ゲームで二人のプレイヤーのうち一人をランダムプレイヤーとした場合は，SSG の判定問題は P 完全であることを示した．

2 On-line Metrical Task System

On-line Metrical Task System [BLS92] とは，n 状態を持ち k 種類の要求を入力とするシステムである．このシステムは 1 単位時間に次の三つの動作を行う．

step 1. システムが状態 i にあり，k 種類の要求のうち一つ t を受けとる．

step 2. 要求 t に対応する処理を行う．

step 3. システムは状態 i から状態 j へ遷移する．

このうち，step 2 において状態 i で処理 t を処理するには a_{it} のコストがかかり，step 3 において状態 i から状態 j に遷移するには b_{ij} のコストがかかる．つまり，1 単位時間で $a_{it}+b_{ij}$ のコストがかかることになる．このシステムが入力を受け続けたときに，最悪の場合，1 単位時間にかかる平均コストはどれだけ大きくなるかという問題が考えられている．

システム側はこの平均コストをなるべく低くおさえられるように，状態の遷移の仕方を考えなければならない．これはつまり，入力を読んで処理を行った後次に移る状態を決定するアルゴリズムを設計することであり，事前に入力がわからないことからオンラインアルゴリズムとして設計しなければならない．

このシステムは次に紹介する Mean Payoff Game として考えることができる．

3 Mean Payoff Game

Mean Payoff Game は有向グラフ $G=(V_1,V_2,E)$，枝の重み関数 $w:E\rightarrow R$ およびスタート頂点 $a_0\in V_1\cup V_2$，で与えられる．有向グラフではどの頂点も枝の次数は 1 以上である．

この有向グラフ上でのプレイは，まずスタート頂点に駒が置かれ，この駒を二人のプレイヤーが次のように動かしていく．V_1 頂点上に駒がある時にはプレイヤー I が，V_2 頂点上に駒がある時にはプレイヤー II が駒を動かす．この時通った枝を $e_1,e_2,\cdots$ とすると，プレイヤー I は利得 $\lim_{n\rightarrow\infty}\frac{1}{n}\sum_{i=1}^{n}w(e_i)$ を大きくすることを目的とし，プレイヤー II は小さくすることを目的とする．

この MPG では最適値 $\mathcal{V}$ が存在し，プレイヤー I には少なくとも $\mathcal{V}$ 以上の利得を保証する戦略があり，プレイヤー II には利得が多くて $\mathcal{V}$ 以下になることを保証する戦略がある [EM79]．この時の戦略を最適戦略という．なお，この最適戦略は，各頂点で枝を一本選び常にその枝に沿って駒を動かすという純粋戦略の形になっている．

MPG の最適値や最適戦略を求める多項式アルゴリズムは見つかっておらず，指数時間アルゴリズムが Gurvich らによって提案されている [GKK88].

4 Discounted Payoff Game

Discounted Payoff Game(DPG) は Mean Payoff Game の discounted version である．Discounted Payoff Game はそれ自体としても興味深いゲームであるとともに，Mean Payoff Game を Simple Stochastic Game(6章参照) に変換する時に中渡しとなるゲームである．

MPG と同様に二人のプレイヤーが駒を動かしていくが，利得の計算が MPG とは少し違っている．λ を $0<\lambda<1$ の実数とする．プレイヤーによって i 番目に選ばれた枝 e_i の重みには $(1-\lambda)\lambda^{i-1}$ がかけられ，ゲームの利得は $(1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}w(e_i)$ となる．プレイヤー I はこの利得を大きくすることが目的で，プレイヤー II は利得を小さくすることが目的である．λ は *discounting factor* と呼ばれる．

DPG は有向グラフ $G=(V_1,V_2,E)$，枝の重み関数 $w:E\rightarrow R$，スタート頂点 $i\in V_1\cup V_2$，および実数 λ で与えられる．有向グラフではどの頂点も枝の次数は 1 以上である．$V=V_1\cup V_2=\{1,2,\cdots,n\}$ とする．$x_i(=x_i(\lambda))$ は頂点 i からスタートする時の discounted game の利得とする．

この DPG の最適値や純粋最適戦略の存在は次の定理によって示されている．

定理 4.1 ([ZP95]) *discounted payoff game* の最適値 $\overline{x}=\{x_1,x_2,\cdots,x_n\}$ はつぎの方程式の唯一の解である

$$x_i=\begin{cases}\max_{(i,j)\in E}\{(1-\lambda)w_{ij}+\lambda x_j\} & \text{if } i\in V_1,\\ \min_{(i,j)\in E}\{(1-\lambda)w_{ij}+\lambda x_j\} & \text{if } i\in V_2.\end{cases}$$

discounting factor が λ である DPG の最適値を $\mathcal{V}(\lambda)$ とする．λ を 1 に近付けると $\mathcal{V}(\lambda)$ は Mean Payoff Game の最適値 V に近付くので，Mean Payoff Game は Discounted Payoff Game に帰着できることになる．

5 確率 discounted Payoff Games

確率 discounted payoff game(確率 DPG) は DPG にランダムプレイヤーを追加したものである．確率 DPG は有向グラフ $G = (V_1, V_2, V_3, E)$，枝の重み関数 $w : E \to R$，スタート頂点 $i \in V_1 \cup V_2 \cup V_3$，および実数 λ で与えられる．V_1 の頂点上に駒がある時にはプレイヤー I が駒を動かし，V_2 の頂点上に駒がある時にはプレイヤー II が駒を動かす．さらに，V_3 の頂点 i から出ている各枝 (i,j) には確率 r_{ij} がつけてあり，頂点 i に駒がある時にはこの確率にしたがって駒が動くことになる．各頂点 $i \in V_3$ について $\sum_{(i,j)\in E} r_{ij} = 1$ である．

DPG と同様に，プレイヤーによって i 番目に選ばれた枝 e_i の重みには $(1-\lambda)\lambda^{i-1}$ がかけられ，ゲームの利得は $(1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}w(e_i)$ となる．しかし，ランダムプレイヤーが存在するため，この利得の期待値を確率 DPG の利得とする．プレイヤー I はこの利得を大きくすることが目的で，プレイヤー II は利得を小さくすることが目的である．

5.1 確率 Discounted Payoff Game の最適値

定理 5.1 確率 *discounted payoff game* $G = (V_1, V_2, V_3, E)$ の最適値 $\overline{x} = \{x_1, x_2, \cdots, x_n\}$ はつぎの方程式の唯一の解である．

$$x_i = \begin{cases} \max_{(i,j)\in E}\{(1-\lambda)w_{ij} + \lambda x_j\} & \text{if } i \in V_1, \\ \min_{(i,j)\in E}\{(1-\lambda)w_{ij} + \lambda x_j\} & \text{if } i \in V_2, \\ \sum_{(i,j)\in E} r_{ij}\{(1-\lambda)w_{ij} + \lambda x_j\} & \text{if } i \in V_3. \end{cases}$$

証明： $\mathcal{F}$ を，任意のベクトル $\overline{x}$ を引数とし $\overline{y}$ を返す次のような関数とする．

$$y_i = \begin{cases} \max_{(i,j)\in E}\{(1-\lambda)w_{ij} + \lambda x_j\} & \text{if } i \in V_1, \\ \min_{(i,j)\in E}\{(1-\lambda)w_{ij} + \lambda x_j\} & \text{if } i \in V_2, \\ \sum_{(i,j)\in E} r_{ij}\{(1-\lambda)w_{ij} + \lambda x_j\} & \text{if } i \in V_3. \end{cases}$$

まず，$\overline{x} = \mathcal{F}(\overline{x})$ となるような $\overline{x}$ が存在することを示す．$||\overline{v}|| = max_i\{|v_i|\}$ とする (最大ノルム)．

$$\forall \overline{u}, \overline{v},\ ||\mathcal{F}(\overline{u}) - \mathcal{F}(\overline{v})|| \le \lambda||\overline{u} - \overline{v}||$$

よって，$0 < \lambda < 1$ より $\mathcal{F}$ は最大ノルムを小さくするような関数である．ゆえに，$\overline{x} = \lim_{n\to\infty}\mathcal{F}^n(\overline{0})$ が存在し，これが $\overline{x} = \mathcal{F}(\overline{x})$ の唯一の解となる．

$\overline{x}$ を $\overline{x} = \mathcal{F}(\overline{x})$ の唯一の解とする．プレイヤー I は戦略として各頂点 $i \in V_1$ で $(1-\lambda)w_{ij} + \lambda x_j$ を最大とするような頂点 j への枝 (i,j) を選べば，各頂点 i からゲームがスタートする時の利得の期待値を少なくとも x_i 以上に保てることは明らかである．同様に，プレイヤー II の戦略として各頂点 $i \in V_2$ で $(1-\lambda)w_{ij} + \lambda x_j$ を最小とするような頂点 j への枝 (i,j) を選べば，各頂点 i からゲームがスタートする時の利得の期待値を多くて x_i 以下に保てる．つまり，頂点 i からスタートするゲームの最適値は x_i となる． □

この定理から，二人のプレイヤーは各頂点で枝を一本選んで常にその枝に沿って駒を動かすような純粋戦略の形で最適戦略を持つことがわかる．

また，プレイヤー II をランダムプレイヤーとした DPG，プレイヤー I をランダムプレイヤーとした DPG においても最適値，最適戦略が存在することがわかる．

DPG は多項式時間で単純確率ゲーム (SSG) に変換できる [ZP95] が，この変換方法で，プレイヤー II をランダムプレイヤーとした DPG は MAX & AVE SSG に変換され，プレイヤー I をランダムプレイヤーとした DPG は MIN & AVE SSG に変換される．MAX & AVE SSG と MIN & AVE SSG は多項式時間で最適値および最適戦略が計算できるので，次の定理がいえる．

定理 5.2 プレイヤー *I*(または *II*) をランダムプレイヤーとした *DPG* の最適値および最適戦略は多項式時間で計算できる．

よって次の定理がいえる.

定理 5.3 確率 *Discounted Payoff Game* において相手の戦略を知っている場合に，その戦略に対して最適な戦略を多項式時間で計算することができる.

確率 DPG に関して自然に考えられる判定問題は次のようになる. 確率 DPG とある値 k が与えられたときに，確率 DPG の最適値が k より大きいか? という問題である.

定理 5.4 確率 *DPG* における判定問題は $NP \cap co\text{-}NP$ である.

6 単純確率ゲームの定義

6.1 単純確率ゲーム

単純確率ゲーム (Simple Stochastic Games,SSG) は，二人のプレイヤー (プレイヤー 0 とプレイヤー 1) によって行なわれるゲームであり，以下のような性質を持った有向グラフ $G = (V, E)$ とスタート頂点 $s \in V$ で表される. 頂点集合 V はそれぞれ MAX,MIN,AVE と呼ばれる頂点の集合 $V_{max}, V_{min}, V_{ave}$ の和集合であり，さらに 0-goal,1-goal と呼ばれる二つの特別な頂点 (この二つの頂点がそれぞれ二人のプレイヤーのゴールである) を含んでいる. 0-goal,1-goal 以外の各頂点からは 2 本の枝が出ている. G の中の一つの頂点をスタート頂点として決めてあり，最初に，スタート頂点に駒が置かれる. 駒はグラフの枝にそって動かされ，この時の動きは以下のルールによって決められている. 駒が MAX(MIN) にあるときはプレイヤー 1(プレイヤー 0) が駒の動く方向を選ぶ. 駒が AVE にあるときはその頂点から出ている 2 本の枝のうち 1 本を 2 分の 1 の確率でランダムに選ばれる. こうして 0-goal, または 1-goal に駒が到達した時にゲームは終了となる. 駒が 1-goal に到達した時にプレイヤー 1 の勝利とし，その他の場合はプレイヤー 0 の勝利とする. SSG 問題とはプレイヤー 1 の勝利確率が 1/2 より大きいかどうかを判定する問題である.

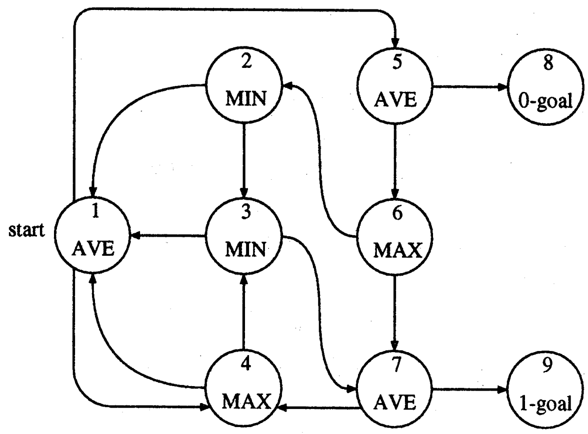

図 1: 単純確率ゲーム

6.2 戦略

戦略は，グラフ G の枝 E の部分集合である. プレイヤー 0(1) の戦略 $\tau(\sigma)$ とは，MIN(MAX) 頂点から出ている枝のうちの 1 本を選んだもので，ゲームの中でプレイヤー 0(1) は駒が MIN(MAX) 頂点にある時には，常にこの戦略にそって駒を動かすことにする.

ゲーム理論においては，次の条件を満たす戦略を純粋戦略 (pure strategy) という. (1) プレイヤーは駒の移動の選択に確率を用いない. (2) 各プレイヤーは，一つの頂点からの移動は，駒がその頂点上にある時は常に同じ移動を選択する. この論文では，この固定戦略のみについて考える. なぜなら，SSG ではプレイヤーはこの形の最適戦略を持つことが示されているからである [Con92].

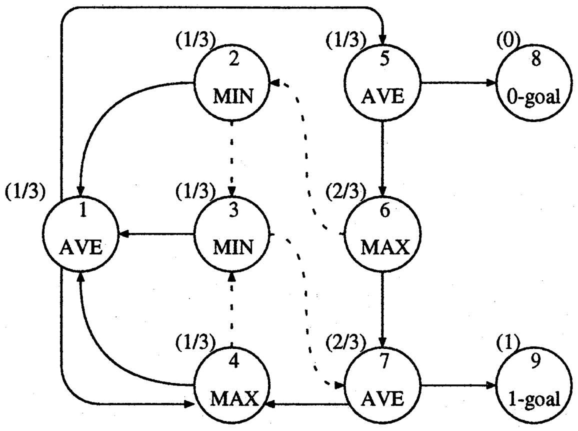

図 2: 戦略 $\sigma = \{(4,1),(6,7)\}, \tau = \{(2,1),(3,1)\}$ に対する $G_{\sigma,\tau}$ と各頂点の値

また，戦略 τ はグラフ G_τ と一致する．ここでグラフ G_τ とは G の部分グラフで，MIN 頂点から出ている枝のうち戦略 τ に入ってない方の枝を取り去ったものである．同様に戦略 σ は G_σ と一致する．グラフ G_σ は G の部分グラフで，MAX 頂点から出ている枝のうち戦略 σ に入ってない方の枝を取り去ったものである．グラフ $G_{\sigma,\tau}$ では全ての MAX,MIN 頂点から出ている枝は一本だけとなる．また，MAX,MIN 頂点の数を合わせて k 個とすると戦略の種類は 2^k ある．

さらにスタート頂点に依存する戦略を定義する．与えられた SSG のグラフにスタート頂点から到達不可能な頂点が存在する場合がある．その場合，到達不可能な頂点を除いた SSG' での戦略を考え，これをスタート頂点に依存する戦略とする．

6.3 値

戦略 σ,τ に関して G の頂点 i の値 $v_{\sigma,\tau}(i)$ を，駒が頂点 i にあるときにプレイヤーが戦略 σ,τ を使ってプレイヤー 1 が勝つ確率とする．つまり，頂点 i からグラフ $G_{\sigma,\tau}$ を通って 1-goal に到達する確率のことである．この値は G の全ての頂点について，次の線形方程式 (1) より多項式時間で計算できる．なお，1-goal,0-goal の値はそれぞれ 1,0 であり，$G_{\sigma,\tau}$ において 1-goal への道がない頂点の値は 0 である．

$$v_{\sigma,\tau}(i) = \begin{cases} \frac{1}{2}(v_{\sigma,\tau}(j) + v_{\sigma,\tau}(k)), \\ \qquad \text{if } i \text{ is an AVE vertex of } G_{\sigma,\tau} \\ \qquad \text{ with outgoing edges } (i,j),(i,k), \\ v_{\sigma,\tau}(j), \\ \qquad \text{if } i \text{ is a MAX or a MIN vertex of } G_{\sigma,\tau} \\ \qquad \text{ with an outgoing edge } (i,j). \end{cases} \tag{1}$$

6.4 停止型 SSG

以下の SSG の議論では停止型 SSG を考える．停止型 SSG とは，どんな $G_{\sigma,\tau}$ においても全ての頂点からゴール頂点に道が存在している SSG のことである．

与えられた SSG が停止型かどうか判定する問題を考えてみると，AVE 頂点を含まない SSG つまり，MAX & MIN SSG の停止性問題は与えられた SSG のグラフが Acyclic であるかどうかという問題と同値であり，多項式時間で判定可能である．AVE 頂点を含む SSG の場合は，図 3のようにグラフが acyclic でなくても停止型である SSG が存在する．よって AVE 頂点を含む SSG の停止性判定問題はグラフが acyclic かどうか調べるだけでは不十分である．

しかし，我々は SSG の停止性判定問題に対する多項式時間アルゴリズムを提案し以下の定理を示した．

定理 6.1 ([YS95c]) *SSG* の停止性判定問題は決定性多項式時間で計算可能である．

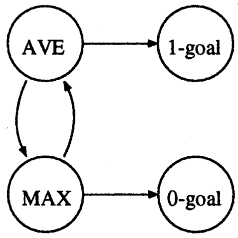

図 3: グラフが acyclic でなくても停止型である SSG の例

さらに，P 完全問題である Alternating Graph Accessibility Problem(AGAP) [GHR95] を帰着させることにより P 完全性も示した．

定理 6.2 ([YS95c]) *SSG* の停止性判定問題は *P* 完全である．

6.5 最適戦略と普遍最適戦略

停止型 SSG は有限ゼロ和ゲームであるので，ミニマックス定理より，最適な混合戦略が存在していることは明らかであるが，この停止型 SSG では，節 6.2で定義した純粋戦略だけを考えても鞍点が存在し，最適戦略が純粋戦略の形で存在することが以下の定理によって示されている [Con92].

定理 6.3 ([Con92]) 停止型 *SSG* では任意の頂点 i に関して次の式が成り立つ．

$$\max_{\sigma} \min_{\tau} v_{\sigma,\tau}(i) = \min_{\tau} \max_{\sigma} v_{\sigma,\tau}(i).$$

ここで，SSG の値を $max_\sigma\ min_\tau\ v_{\sigma,\tau}(start)$ と定義する．これはスタート頂点からお互いに最適戦略に従って駒を動かす時のプレイヤー 1 の勝つ確率である．SSG 問題とはこの SSG の値が 1/2 より大きいかどうかを問う問題である．

定理 6.3を満たす最適戦略は与えられたゲームのグラフに対して決まり，スタート頂点がどこであってもそのゲームの最適戦略となっている．この最適戦略を普遍最適戦略と呼ぶ．

停止型 SSG の各頂点 i(枝 $(i,j),(i,j)$ を持つ) の最適値 $v(i)$ は以下の方程式の唯一の解である [Con92].

$$v(i) = \begin{cases} max\{v(j), v(k)\}, & \text{if } i \text{ is a MAX vertex,} \\ min\{v(j), v(k)\}, & \text{if } i \text{ is a MIN vertex,} \\ \frac{1}{2}(v(j) + v(k)), & \text{if } i \text{ is an AVE vertex,} \\ 1, & \text{if } i \text{ is the 1-goal,} \\ 0. & \text{if } i \text{ is the 0-goal.} \end{cases} \tag{2}$$

また，与えられた SSG をスタート頂点に依存する SSG' に変換し，SSG' の全ての普遍最適戦略を SSG' のスタート頂点に依存する全ての最適戦略と定義する．

6.6 交換可能

ある戦略 (σ,τ) において G の頂点 i(枝 $(i,j),(i,k)$ が出ている) の値が次の状態のとき頂点 i が交換可能であるという．

$$\begin{cases} v_{\sigma,\tau}(i) < max\{v_{\sigma,\tau}(j), v_{\sigma,\tau}(k)\}, \\ \qquad \text{if } i \text{ is a MAX vertex,} \\ v_{\sigma,\tau}(i) > min\{v_{\sigma,\tau}(j), v_{\sigma,\tau}(k)\}, \\ \qquad \text{if } i \text{ is a MIN vertex.} \end{cases} \tag{3}$$

(i,j) を含む戦略 $\sigma(\tau)$ から頂点 i を交換することによって (i,k) を含む戦略 $\sigma'(\tau')$ に変わることを，$\sigma' = \sigma - \{(i,j)\} + \{(i,k)\}, (\tau' = \tau - \{(i,j)\} + \{(i,k)\})$ と定義する．

7 全ての普遍最適戦略を計算する複雑さ

ここではSSGを停止型SSG(どんな戦略に対しても，全ての頂点からgoal頂点への道が存在する)に制限して考える．一般のSSGは多項式時間で停止型SSGに変換可能である[Con92]．停止型SSGでは，ある戦略(σ,τ)が普遍最適戦略であることと，その戦略に対して値を計算して交換可能な頂点が存在しないことは同値である[Con92].

スタート頂点によらない普遍最適戦略とスタート頂点に依存する最適戦略には計算複雑性に関して次の補題7.1がいえる．

補題 7.1 停止型*SSG*の普遍最適戦略を求めるアルゴリズムBは，スタート頂点に依存する最適戦略を求めるアルゴリズムAの多項式回数繰り返しとして構成できる．

証明. 与えられたSSGの普遍最適戦略を求めるアルゴリズムをB，与えられたSSGのスタート頂点に依存する最適戦略を求めるアルゴリズムをAとすると，これらは次のように書ける．

$$\begin{aligned} B(G) &= (\sigma_{univ}, \tau_{univ}). \\ A(G, start) &= (\sigma_{start}, \tau_{start}). \end{aligned}$$

まず，全ての頂点$i(1 \le i \le n)$に対して$A(G,i)$を計算する．

$$\overline{v} = [v_{\sigma_1,\tau_1}(1), v_{\sigma_2,\tau_2}(2), \ldots, v_{\sigma_n,\tau_n}(n)]^T.$$

ここで，次の式が必ず成り立つことを利用すると，

$$v_{\sigma_{univ},\tau_{univ}}(start) = v_{\sigma_{start},\tau_{start}}(start).$$

上で得た$\overline{v}$は普遍最適戦略の各頂点の値と一致する．つまり，

$$\overline{v} = \overline{v}_{\sigma_{univ},\tau_{univ}}.$$

この$\overline{v}$が得られれば，式(3)を満たす頂点がないように戦略を構成することができ，これが普遍最適戦略となる． □

定理 7.2 ([Yan95, YS95, YS95a, YS95b]) 停止型*SSG*の全ての普遍最適戦略は非決定性多項式時間で計算可能である．

証明. 非決定性アルゴリズムを次のように構成する．

入力 = $\{G\}$
Gの普遍最適戦略σ,τを推測する;
σ,τにおける各頂点の値を式(1)を使って計算する;
交換可能な頂点が存在しないかどうか式(3)より調べる;
if 交換可能な頂点が存在しない
then { 全てのMAX,MIN頂点を調べて，2本の枝の先の頂点の値が等しくなっている頂点の頂点番号の集合を$U = (u_1, u_2, \ldots, u_t)$とする
(σ,τ,U)を出力する };
end.

ある普遍最適戦略σ,τにおいて，$\{u_1, u_2, \ldots, u_t\}$のうち任意個交換してできる戦略$\sigma',\tau'$も普遍最適戦略であることが，次の補題7.3よりいえる．

補題 7.3 ([Yan95, YS95, YS95a, YS95b]) ある戦略 (σ,τ) において，枝 $(i,j),(i,k)$ が出ており $v_{\sigma,\tau}(j) = v_{\sigma,\tau}(k)$ であるような頂点 i を任意個交換してもどの頂点の値も変わらない．

また，こうしてできる戦略が全ての普遍最適戦略を表していることが次の補題 7.4よりいえる．

補題 7.4 ([Yan95, YS95, YS95a, YS95b]) 普遍最適戦略が複数ある時，任意の *2* つの普遍最適戦略 (σ,τ), (σ',τ') において，(σ,τ) には (i,j) が含まれ (σ',τ') には (i,k) が含まれるような全ての頂点 i,j,k について $v_{\sigma,\tau}(j) = v_{\sigma,\tau}(k)$ である．

よって，この非決定性アルゴリズムは多項式時間で全ての普遍最適戦略を記述することができ，逆に $(\sigma,\tau,u_1, u_2,\ldots,u_t)$ が与えられれば，その正しさを多項式時間で検証することができる． □

なお，出力が $(\sigma,\tau,u_1,u_2,\ldots,u_t)$ のとき，最適戦略の数は 2^t で表されるので次の系 7.5が成り立つ．

系 7.5 ([Yan95, YS95, YS95a, YS95b]) 停止型 *SSG* の全ての普遍最適戦略の数は非決定性多項式時間で計算可能である．

例 7.6 図 1 の *SSG* が入力として与えられた時のアルゴリズムの動きは次のようになる．まず，戦略 $(\sigma = \{(4,1),(6,7)\}$, $\tau = \{(2,1),(3,1)\})$ を推測したとする (図 2). 線形方程式 (1) より値を計算する．

$$Q = \begin{bmatrix} 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \end{bmatrix}, \overline{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1/2 \end{bmatrix}.$$

$\overline{v}_{\sigma,\tau} = Q\overline{v}_{\sigma,\tau} + \overline{b}$ を解いて,

$$\overline{v}_{\sigma,\tau} = [\ 1/3,\ 1/3,\ 1/3,\ 1/3,\ 1/3,\ 2/3,\ 2/3\ ]^T.$$

次に条件式 (3) を満たす頂点が存在しないかどうか調べる．

$$\begin{aligned} v_{\sigma,\tau}(2) &= min\{v_{\sigma,\tau}(1), v_{\sigma,\tau}(3)\}, \\ v_{\sigma,\tau}(3) &= min\{v_{\sigma,\tau}(1), v_{\sigma,\tau}(7)\}, \\ v_{\sigma,\tau}(4) &= max\{v_{\sigma,\tau}(1), v_{\sigma,\tau}(3)\}, \\ v_{\sigma,\tau}(6) &= max\{v_{\sigma,\tau}(2), v_{\sigma,\tau}(7)\}. \end{aligned}$$

ここで頂点 2 の子である 1 と 3 の値を見てみると，$v_{\sigma,\tau}(1) = v_{\sigma,\tau}(3)(= 1/3)$ と一致している．頂点 4 の子である 1 と 3 の値も同様に一致している．よって $U = \{2,4\}$ となり，全ての普遍最適戦略の記述は (σ,τ,U) = $(\{(4,1),(6,7)\}$, $\{(2,1),(3,1)\}$,$\{2,4\})$ となる．これは次のような $2^2 = 4$ 通りの戦略を示していることになる．

$$\begin{aligned} &(\{(4,1),(6,7)\},\{(2,1),(3,1)\}), \\ &(\{(4,3),(6,7)\},\{(2,1),(3,1)\}), \\ &(\{(4,1),(6,7)\},\{(2,3),(3,1)\}), \\ &(\{(4,3),(6,7)\},\{(2,3),(3,1)\}). \end{aligned}$$

□

頂点数を 2 種類に制限した (1)MAX & MIN (2)MAX & AVE (3)MIN & AVE SSG の普遍最適戦略の 1 つは多項式時間で求めることができる [Con92] ので，次の系 7.7が成り立つ．

系 7.7 ([Yan95, YS95, YS95a, YS95b]) 頂点の種類を *2* 種類に制限した停止型 *SSG* の全ての普遍最適戦略は決定性多項式時間で計算可能である.

また, 回路値問題への帰着により P 完全性も示した.

定理 7.8 ([YS95a]) 頂点の種類を *2* 種類に制限した *SSG* 問題は *P* 完全である.

注意 7.9 一方が, ランダムに動作する二人ゲーム (Games against nature) に関しては, Papadimitriou [Pap85] が, 多項式時間限定ゲームを考察することで, その PSPACE 完全性を示している. Condon-Ladner [CL88] は, Papadimitriou [Pap85] の結果を, 領域限定の場合に一般化し, 領域 $s(n)$ の交代性 Turing 機械で計算可能な問題のクラス (ASPACE($s(n)$)) と, 領域 $s(n)$ 上で一方が, ランダムに動作する二人ゲーム (UC-SPACE($s(n)$)) とが一致するを示している. したがって, ASPACE($log(n)$)= P であることに注意すると, MAX & AVE SSG 問題と MIN & AVE 問題の P 完全性は Condon-Ladner [CL88] の結果としても導かれる.

8 Acyclic SSG の最適戦略の計算複雑性

有向グラフにおいて閉路を含まないようなものを acyclic なグラフであるという. SSG のグラフとして acyclic なグラフが与えられたとき, その SSG を Acyclic SSG と呼ぶ. Acyclic SSG は停止型 SSG に含まれる. 有向グラフが Acyclic であるかどうかは $O(max(|V|, |E|))$ で判定できる [AHU74] ので, SSG が Acyclic SSG であるかどうかは多項式時間で判定可能である.

SSG を Acyclic SSG に制限すると頂点の種類に関わらず以下の定理が成り立つ.

定理 8.1 ([YS95a]) *Acyclic SSG* の普遍最適戦略は決定性多項式時間で計算可能である.

よって, 全ての普遍最適戦略について次の定理 8.2がいえる.

定理 8.2 ([YS95a]) *Acyclic SSG* の全ての普遍最適戦略は決定性多項式時間で計算可能である.

定理 8.1より Acyclic SSG 問題が P であることは明らかである. また, 代表的な P 完全問題である回路値問題を帰着させることで次の定理 8.3が示せる.

定理 8.3 ([YS95a]) *Acyclic SSG* 問題は *P* 完全である.

9 まとめと今後の課題

今回, 確率 DPG について最適戦略および最適値の存在を示すことによって, On-line Metrical Task System の平均コストを discounted コストとして評価した場合には, ランダムな要求が System に送られると仮定したときの最適なオンラインアルゴリズムが多項式時間で計算可能であるということが得られたが, discounted factor λ をどこまで 1 に近付ければ平均コストとしての正確な最適値を求められるかという問題が残された.

また, 非停止型 SSG での全ての最適戦略の計算複雑性に関しては未解決であるが, 現在, 全ての最適戦略を多項式領域に記述することは困難だろうという予測に基づき理論的不可能性について検討中である.

参考文献

[AHU74] Aho, A.V., Hopcroft, J.E. and Ullman, J.D., *The Design and Analysis of Computer Algorithms*, Addison–Wesley, Reading MA (1974)

[BLS92] Borodin, A., Linial, N. and Saks, M.E., "An optimal on–line algorithm for metrical task system," *Journal of the Association for Computing Machinery* **39** (1992) pp. 745–763.

[CHPZ95] Cole, R., Hariharan, R., Paterson, M. and Zwick, U., "Tighter lower bounds on the exact complexity of string matching," *SIAM Journal on Computing* **24** (1995) pp. 30–45.

[CL88] Condon, A. and Ladner, R.E., "Probabilistic Game Automata," *JCSS* **36** (1988) pp. 452–489.

[Con92] Condon, A., "The Complexity of Stochastic Games," *Information and Computation* **96** (1992) pp. 203–224.

[Con93] Condon, A., "On Algorithms for Simple Stochastic Games," *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* **13** (1993) pp. 51–71.

[CT90] Chandra, A.K. and Tompa, M., "The complexity of short two–person games," *Discrete Applied Mathematics* **29(1)** (1990) pp. 21–33.

[EM79] Ehrenfeucht, A. and Mycielski, J., "Positional strategies for mean payoff games," *International Journal of Game Theory* **8** (1979) pp. 109–113.

[GHR95] Greenlaw, R., Hoover, H.J. and Ruzzo, W.L., *Limits to Parallel Computation: P–Completeness Theory*, Oxford University Press (1995).

[GKK88] Gurvich, V.A., Karzanov, A.V. and Khachivan, L.G., "Cyclic games and an algorithm to find minimax cycle means in directed graph," *USSR Comput. Maths. Math. Phys.* **28** (1988) pp. 85–91.

[HK66] Hoffman, A. and Karp, R., "On nondeterminating stochastic games," *Management Sci.* **12** (1966) pp. 359–370.

[Kar78] Karp, R.M., "A characterization of the minimum cycle mean in a digraph," *Discrete Mathematics* **23** (1978) pp. 309–311.

[KL93] Karzanov, A.V. and Lebedev, V.N., "Cyclical games with prohibitions," *Mathematical Programming* **60** (1993) pp. 277–293.

[Lud95] Ludwig, W., "A subexponential randomized algorithm for the simple stochastic game problem," *I & C* **117** (1995) pp. 151–155.

[Pap85] Papadimitriou, C.H., "Games against nature," *Proc. of FOCS'83*; also *JCSS* **31** (1985) pp. 288–301.

[Sch78] Schaefer, Thomas.J., "On the Complexity of Some Two–Person Perfect–Information Games," *Journal of Comput. and System Sci.* **16** (1978) pp. 185–225.

[Sha53] Shapley, L.S., "Stochastic games," *Proc. Nat. Acad. Sci. U.S.A* **39** (1953) pp. 1095–1100.

[Val79] Valiant, L.G., "The Complexity of Enumeration and Reliability Problems," *SIAM J. Comput.* **8** (1979) pp. 410–421.

[Van78] Van Der Wal, J., "Discounted Markov games: generalized policy iteration method," *J. Optim. Theory Appl.* **25** (1978) pp. 125–138.

[TA94] 遠山 宏明, 足立 暁生, "2 人ゲームにおける必勝手を数えあげる多項式時間アルゴリズム," 電子情報通信学会 コンピュテーション研究会 信学技報 COMP94–70 (1994) pp. 29–38.

[Yan95] 山家 明男, "単純確率ゲームの最適戦略に関する計算複雑性," 九州大学工学部情報工学科卒業論文 (1995).

[YS95] 山家 明男, 櫻井 幸一, "単純確率ゲームの普遍最適戦略に関する計算複雑性," 電子情報通信学会 1995 総合大会 情報・システム D–11 (1995) pp. 11.

[YS95a] 山家 明男, 櫻井 幸一, "二人確率ゲームの最適戦略数を数える計算複雑性," 電子情報通信学会 コンピュテーション研究会 (1995–07) **COMP95–38**.

[YS95b] Yanbe, A. and Sakurai, K., "A short certificate of the number of universal optimal strategies for stopping simple stochastic games," *Information Processing Letters* **57** (1996) pp. 17–24.

[YS95c] 山家 明男, 櫻井 幸一, "単純確率ゲームの停止性判定多項式時間アルゴリズム," 平成 7 年度電気関係学会九州支部連合大会 (1995) pp. 712.

[ZP95] Zwick, U. and Paterson, M, "The complexity of mean payoff games on graphs," *ECCC Reports Series 1995* **TR95–040**.