MC758A/MO758A

Teoria dos Jogos Algorítmica

Professor: Rafael C. S. Schouery

# Large Richer Representations: Beyond the Normal and Extensive Forms

Lucas Ismaily B. Freitas RA:142696

24 de junho de 2018

**Resumo**

Este texto apresenta um resumo do Capítulo 6 do livro [1]. Especificamente, jogos que são jogados uma quantidade finita e infinita de vezes pelo mesmo conjunto de jogadores e jogos de máquinas, que são jogos cuja as estratégias dos jogadores são representadas por autômatos. Estes tipos de jogos são importantes porque modelam problemas que não são possíveis de serem modelados com jogos na forma normal e extensiva.

## 1 Introdução

Os jogos representandos na forma normal e extensiva nem sempre são capazes de modelar jogos grandes (em termos de número de jogadores ou estratégia) ou mais realísticos. Eles assumem que os jogadores tem conhecimento de todas as utilidades e que tanto o número de jogadores quanto o conjunto de estratégias são finitos. Aqui vamos considerar cenários como: o que acontece quando um jogo é jogado infinitas vezes pelo mesmo conjunto de jogadores? Ou o que ocorre quando um jogador tem um número infintos de ações (estratégias) possíveis?

Assim, considere o jogo do dilema dos prisioneiros conforme ilustrado na Figura 1. Nas próximas seções usaremos esse jogo como exemplo para ilustrar jogos de repetição (in)finita e jogos de máquina.

| | $C$ | $D$ |
|---|---|---|
| $C$ | $-1, -1$ | $-4, 0$ |
| $D$ | $0, -4$ | $-3, -3$ |

Figura 1: Dilema dos primeiros na forma normal.

## 2 Jogos finita e infinitamente repetidos

Em jogos repetidos, um jogo representado na forma normal é jogado (in)finitamente pelo mesmo conjunto de jogadores. Cada repetição do jogo é chamado de *jogo de estágio*. Esses tipos de jogos são modelados como jogos de informação imperfeita na forma extensiva. Essa representação deixa claro que $(i)$ A utilidade dos jogadores é definida em função dos jogos de estágios, $(ii)$ Em cada jogo de estágio um jogador não sabe o que os outros estão jogando, e $(iii)$ Os jogadores sabem o que fizeram e o que os outros jogadores fizeram nos jogos de estágios anteriores. A Figura 2 ilustra o jogo do dilema dos prisioneiros repetido duas vezes.

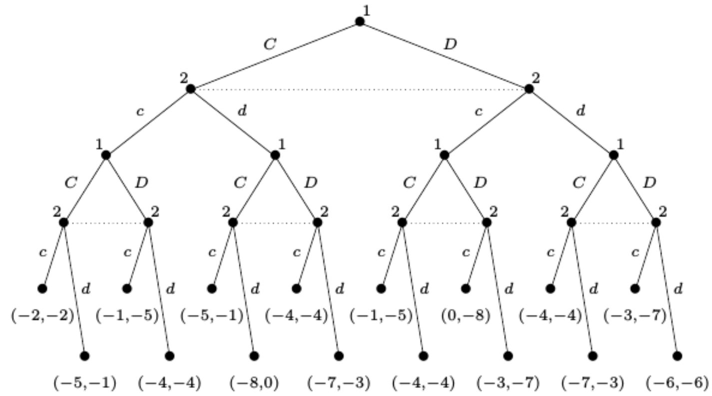



Figura 2: Dilema dos primeiros repetido duas vezes.

Quando o jogo é repetido infinitamente a árvore da forma extensiva tem altura infinita, e desse modo a utilidade precisar ser definida. Considerando $r_i^k$ como a utilidade do jogador $i$ no jogo de estágio $k$. Duas definições conhecidas são: *utilidade média* que consistem em $\lim_{k\to\infty} \frac{\sum_{j=1}^{k} r_i^j}{k}$, e *utilidade descontada* que é $\sum_{j=1}^{\infty} \beta^j r_i^j$, para $0 \leq \beta \leq 1$.

As estratégias nos jogos repetidos são mais ricas do que as do jogo de estágio. A luz da repetição do jogo do dilema dos prisioneiros, podemos definir a estratégia *estacionária* como aquela onde o jogador escolhe a mesma estratégia em cada jogo de estágio. Outras duas são a *Tit-for-Tat (TfT)* que é a estratégia no qual o jogador começa cooperando e se houver algum desvio na rodada $k$, ele escolhe, na rodada $k+1$, o que seu adversário escolheu na rodada $k$, e a *estratégia gatilho*, na estratégia gatilho o jogador começa cooperando, se houver desvio, ele nunca mais coopera. O Teorema 2.1 mostra que as estratégias TfT e gatilho produzem equilíbrios na repetição infinita do jogo do dilema dos prisioneiros.

**Teorema 2.1.** *Seja $G$ o jogo do dilemas dos prisioneiros repetido infinitamente. Então, (1) Ambos os jogadores jogando TfT é um equilíbrio em $G$; (2) Ambos os jogadores jogando a estratégia gatilho é um equilíbrio em $G$; e (3) Um jogador jogando a estratégia gatilho e outro jogando TfT é um equilíbrio em $G$.* □

Considerando a repetição de um jogo qualquer, o Teorema 2.2 apresenta uma caracterização de um equilíbrio, porém antes de enunciar e provar o Teorema 2.2, é preciso introduzir algumas definições. Sejam $G$ um jogo $(N, A, u)$ com $n$ jogadores e um perfil de utilidades $r = (r_1, r_2, \ldots, r_n)$ para a repetição infinita de $G$. E seja $v_i = \min_{s_{-i}\in S_{-i}} \max_{s_i \in S_i} u_i(s_{-i}, s_i)$, isto é, $v_i$ é a utilidade de $i$ quando os outros jogadores estão jogando a estratégia minimax contra ele. O perfil $r$ é *executável* se $\forall_i \in N, r_i \geq v_i$. Dizemos que $r$ é *viável* se existem valores $\alpha_a$ racionais e não negativos, tal que para todo $i \in N$, $r_i$ pode ser expressado como $\sum_{a\in A} \alpha_a u_i(a)$, com $\sum_{a\in A} \alpha_a = 1$.

**Teorema 2.2.** *Considere um jogo $G$ infinitamente repetido com utilidade média $r$. Então, (1) Se $r$ é um perfil para algum equilíbrio de $G$, então para cada jogador $i$, $r_i$ é executável; (2) Se $r$ é um perfil viável e executável, então $r$ é um perfil de utilidade média para algum equilíbrio em $G$.*

*Demonstração.* Parte 1. Suponha que $r$ não seja executável para algum jogador $i$, isto é, $r_i < v_i$. Considere uma estratégia alternativa para $i$: jogar $BR(s_{-i}(h))$, onde $s_{-i}(h)$ é a estratégia dos outros jogadores na rodada $h$, e $BR(s_{-i}(h))$ é uma função que retorna a melhor resposta para $i$, dado o perfil de estratégia $s_{-i}$. Pela definição da estratégia minimax, o jogador $i$ recebe utilidade de pelo menos $v_i$ em cada jogo de estágio $h$, se ele joga $BR(s_{-i}(h))$. Portanto, uma contradição, uma vez que por hipótese $r$ é um perfil de um equilíbrio.

Parte 2. Como $r$ é viável e executável, para todo $i$ podemos escrever $r_i = \sum_{a \in A} \left(\frac{\beta_a}{\gamma}\right) u_i(a)$, com $\beta_a$ e $\gamma$ não negativos e $\gamma = \sum_{a \in A} \beta_a$. Seja $i$ um jogador qualquer. Vamos construir uma estratégia para $i$ que resulta na utilidade acima e é um equilíbrio: a estratégia $s_i$ passará por todo $a \in A$ em $G$, em blocos de tamanho $\gamma$, sendo que em cada bloco $a$ será repetindo exatamente $\beta_a$ vezes. Se em algum momento $t$, um jogador $j \neq i$ desviar, então $i$ passará a jogar a estratégia minimax contra $j$. Assim, se todos jogarem $s_i$, obtém utilidade $r_i$. Ademais, nenhum jogador $j$ deseja desviar, pois pela definição de minimax, receberá $v_j$, porém o perfil é executável, por hipótese. $\square$

Estamos assumindo que os jogadores conseguem raciocinar sobre toda a profundidade da árvore e na prática isso não ocorre, em torneios de xadrez ocorrem vitórias, por exemplo. Desse modo, é interessante limitar essa racionalidade. Uma primeira abordagem foi a noção de $\epsilon$-equilíbrio, e uma outra abordagem é através de jogos de máquinas que modelam as estratégias como autômatos e permitem impor um limite no número de estados. A intuição é que autômatos com com poucos estados representam estratégias mais simples. A seguir mais detalhes e formalismos sobre jogos de máquina.

## 3 Jogos de máquina

Antes de definir jogo de máquina, é preciso definir um autômato no contexto de teoria dos jogos. Desse modo, dado um jogo $G = (N, A, u)$ que será jogado repetidamente, um *autômato* $M_i$ para um jogador $i$ é definido pela quádrupla $(Q_i, q_i^0, \delta_i, f_i)$, sendo:

- $Q_i$ é um conjunto de estados;
- $q_i^0$ é o estado inicial;
- $\delta_i : Q_i \times A \to Q_i$ é uma função de transição;
- $f_i : Q_i \to A$ é uma função de estratégia que associa cada estado a uma ação do jogador $i$.

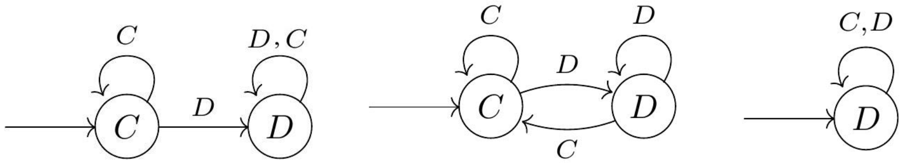
O autômato é usado pra representar cada estratégia de um jogador $i$ em cada jogo de estágio. A sua execução começa no estado $q_i^0$ com $i$ jogando $f_i(q_i^0)$, e usando a função de transição $\delta_i$ com as ações dos outros jogadores, $i$ muda de estado $\delta_i(q_i^0, a_1, ..., a_n)$ e joga $f_i(\delta_i(q_i^0, a_1, ..., a_n))$ na rodada 2. De maneira geral, em um jogo de estado $t$, temos que:

- $a_i^t = f_i(q_i^t)$;
- $q_i^{t+1} = \delta_i(q_i^t, a_1^t, \ldots, a_n^t)$.

A Figura 3 ilustra os autômatos referentes as estratégias TfT, gatilho e estacionária $D$ no jogo do dilema dos prisioneiros.

Vamos definir jogo de máquina com dois jogadores, porém é possível generalizar para $n$. Um *jogo de máquina* $G^M$ com dois jogadores é definido como $G^M = (\{1, 2\}, \mathcal{M}, G)$, onde:

- $\{1, 2\}$ é um conjunto de jogadores;
- $\mathcal{M} = \{M_1, M_2\}$, sendo $M_i$ os autômatos para o jogador $i$;
- $G = (\{1, 2\}, A, u)$ um jogo na forma normal que será repetido (in)finitamente.




(a) Estratégia gatilho.
(b) Estratégia TfT.
(c) Estratégia estacionária.

Figura 3: Autômatos representando estratégias.

Se há uma $k$ repetição de $G$. Então, podemos transformar um jogo de máquina para um jogo na forma normal. Assim, sabendo que o par $M_1, M_2 \in \mathcal{M}$ gera uma saída em cada jogo de estágio $t$, denote $o^t(M_1, M_2)$ a função que obtém a saída. O jogo $G^*$ na forma normal a partir de $G^M$ é definido por $G^* = (\{1,2\}, \mathcal{M}, U)$, sendo

- $\{1,2\}$ o conjunto de jogadores;
- $\mathcal{M}$ o conjunto de estratégias (cada jogador $i$ escolhe um $M_i \in \mathcal{M}$);
- A utilidade de $i$ definida como $U_i(M_1, M_2) = \sum_{t=1}^{k} u_i(o^t(M_1, M_2))$.

Como dito anteriormente, podemos utilizar autômatos para limitar a racionalidade dos jogadores. Desse modo, dado um jogo de máquina $G^M$, considere

- $s : M \to \mathbb{N}$ uma função que recebe um autômato e retorna o número de estados;
- $S(\mathcal{M}_i) = \max_{M \in \mathcal{M}_i} s(M)$ uma função que retorna o maior autômato de $\mathcal{M}_i$.

No jogo de máquina do dilema dos prisioneiros repetido $k$ vezes, se $2 \leq \max(S(\mathcal{M}_1), S(\mathcal{M}_2)) \leq k$, a estratégia estacionária não é mais um equilíbrio enquanto que a estratégia TfT continua sendo. Além disso, no jogo do dilema dos prisioneiros da Figura 4 é possível provar que a utilidade de um jogador é próximo de 3, conforme o Teorema 3.1.

| | $C$ | $D$ |
|---|---|---|
| $C$ | $3,3$ | $0,4$ |
| $D$ | $4,0$ | $1,1$ |

Figura 4: Dilema dos primeiros na forma normal.

**Teorema 3.1.** *Para qualquer inteiro $x$, existe um inteiro $k_0$ tal que para todo $k > k_0$, qualquer jogo de máquina $G^M = (\{1,2\}, \mathcal{M}, G)$ do dilema dos prisioneiros repetido $k$ vezes, no qual*

$$k^{1/x} \leq \min\{S(\mathcal{M}_1, \mathcal{M}_2)\} \leq \max\{S(\mathcal{M}_1, \mathcal{M}_2)\} \leq k^x$$

*tem um equilíbrio com utilidade média para cada jogador de pelo menos $3 - 1/x$.* □

Note que quanto maior o valor de $x$, mais a utilidade se aproxima de três.

# Referências

[1] Shoham, Y., and Leyton-Brown, K. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations.* Cambridge University Press, 2008.