Modelo de Melhor Defesa em Jogos com Informação Imperfeita

Fábio Takashi Hirano[1], Flávio Soares Corrêa da Silva [2] (Supervisor)
1: fabiohirano@gmail.com
2: Departamento de Ciência da Computação, fcs@ime.usp.br

Introdução

Na Teoria dos Jogos, jogos com Informação Imperfeita são aqueles em que o jogador não tem pleno conhecimento sobre os resultados de suas jogadas, uma vez que não conhece as estratégias do adversário. Usualmente, encontrar uma estratégia ótima para tais jogos é um problema NP-Difícil e, portanto, uma abordagem heurística se faz necessária para o desenvolvimento de suas soluções e algoritmos.

Um exemplo de jogo com informação imperfeita que permite tal abordagem é o jogo de cartas *Bridge*. As tentativas de automatização do jogo de Bridge, em geral, envolvem diminuir o espaço de jogadas possíveis, dividindo o problema em sub-problemas para cada um dos naipes.

O presente trabalho parte da tentativa de formalizar o jogo de Bridge (e outros jogos com informação imperfeita) com um modelo que se assemelhe à maneira humana de jogar. O "Modelo da Melhor Defesa"é resultado dessa modelagem. Mostraremos que o conceito de *Ponto de Equilíbrio* para um jogo com informação imperfeita de 2 jogadores é bem definido para esse modelo e descreveremos o algoritmo de *Minimização Exaustiva de Estratégias* que encontra tais pontos. O trabalho é baseado em [FB98], com as definições de Teoria dos Jogos formalizadas em [NM44].

Conceitos de Teoria dos Jogos

Um *Jogo na Forma Extensiva* é um modelo para jogos em que a ordem das jogadas é importante. Um jogo para 2 jogadores nessa forma possui:

- Uma árvore finita T que descreve todo o jogo. Cada um de seus nós n é um movimento que pode ser aleatório ou definido por um dos dois jogadores. O primeiro nó, p_0 é o primeiro movimento do jogo.
- Uma partição $\mathcal{P}(n)$ dos nós de T em 3 conjuntos. Para cada nó n, $\mathcal{P}(n)$ define se o próximo movimento a partir de n é do jogador 1, do jogador 2 ou é um movimento aleatório.
- Para cada nó n tal que $\mathcal{P}(n)$ indique um movimento aleatório, uma distribuição de probabilidade $\pi(n, i)$ entre todos os seus i filhos.
- Um refinamento $\mathcal{I}_k(n)$ em k sub-conjuntos de cada conjunto definido por $\mathcal{P}(n)$. O jogador não possui informações suficientes para distinguir entre os nós de cada $\mathcal{I}_k(n)$.
- Para cada jogador i, uma função utilidade K_i que define, para cada folha de T, o valor que aquela sequência de movimentos representa para o jogador i.

A árvore T pode ser muito grande. Para um jogo da velha, por exemplo, essa modelagem gera uma árvore com $9! = 362880$ nós. Existe, porém, uma modelagem mais simples e equivalente à Forma Extensiva, chamada *Forma Normal*. Nela definimos, para cada jogador, um espaço de estratégias, de modo que sabemos qual decisão o jogador tomará após cada jogada do adversário. Formalmente, definimos essa lista de estratégias, em um jogo para 2 jogadores, como duas funções reais $K_1(x_1, x_2)$ e $K_2(x_1, x_2)$ que representa o ganho que cada jogador tem quando o jogador i realiza o movimento x_i.

Pontos de Equilíbrio de um jogo na forma normal são os pares de estratégia $(\widehat{x_1}, \widehat{x_2})$ tais que

$$K_1(\widehat{x_1}, \widehat{x_2}) \geq K_1(x_1, \widehat{x_2}), \forall x_1 \in X_1$$
$$K_2(\widehat{x_1}, \widehat{x_2}) \geq K_2(\widehat{x_1}, x_2), \forall x_2 \in X_2$$

Um *Jogo de Soma Zero* é um jogo tal que

$$K_1(x_1, x_2) + K_2(x_1, x_2) = 0, \forall x_1 \in X_1, \forall x_2 \in X_2$$

Teorema do Min-Max

Seja Γ um jogo de 2 jogadores e soma zero. Nesse jogo, cada jogador define seu movimento sem conhecer o movimento do adversário. Seja Γ_1 uma variação de Γ em que o jogador 1 escolhe sua movimento antes do jogador 2. Como o jogador 1 está em desvantagem em relação ao 2 nessa variante, chamaremos Γ_1 de *Jogo Minorante de* Γ. Uma segunda variação Γ_2 é variação dual de Γ_1 e, portanto, é chamada de *Jogo Majorante de* Γ.

Tome o jogo Γ_1. Para cada jogada $\widehat{x_1} \in X_1$, o jogador 2 escolherá x_2 tal que minimize a função $K_1(\widehat{x_1}, x_2)$, ou seja, que minimize o ganho do jogador 1. Dessa maneira, o jogador 1 já sabe qual será a jogada escolhida pelo jogador 2 e, portanto, o ganho do jogador 1 em Γ_1 será

$$v_1 = \max_{x_1} \min_{x_2} K_1(x_1, x_2)$$

Analogamente, se estivermos trabalhando com a versão majorante de Γ, teremos um ganho, para o jogador 1, de

$$v_2 = \min_{x_2} \max_{x_1} K_1(x_1, x_2)$$

Em [NM44] é mostrado que o valor v, que é o ganho esperado para o jogador 1 no jogo Γ tem como limitantes os valores de v_1 e v_2. Essa afirmação, que vale para jogos em geral, é refinada no caso de jogos de soma zero.

Teorema 1 *Dado um jogo de soma zero* Γ *com 2 jogadores, temos que*

$$\max_{x_1} \min_{x_2} K_1(x_1, x_2) = \min_{x_2} \max_{x_1} K_1(x_1, x_2)$$

onde $K_1(x_1, x_2)$ *é a função utilidade para o jogador 1.*

Bridge como um Jogo com Informação Imperfeita

O Bridge é um jogo de cartas para 4 jogadores, divididos em 2 duplas. Utilizam-se 52 cartas, com 13 de cada um dos 4 naipes. A mecânica do jogo baseia-se no conceito de vasas, que consiste num conjunto de 4 cartas, jogadas cada um por um dos jogadores. O primeiro jogador escolhe uma das cartas em sua mão e os demais devem escolher uma carta para jogar e fazê-lo no sentido horário a partir do primeiro. Caso um dos jogadores tenha em mãos alguma carta do mesmo naipe da carta jogada pelo primeiro jogador ele deve, obrigatoriamente, "responder"o naipe aberto. Caso contrário, a escolha de cartas é livre. O vencedor da vasa é o jogador que jogar a maior carta do mesmo naipe que o primeiro jogador, exceto nos casos em que um "trunfo"foi escolhido antes do início da partida. Nesse caso, a maior carta do naipe do trunfo vence a vasa e, caso não hajam trunfos em jogo, respeita-se a regra anterior.

Em termos de Teoria dos Jogos, podemos entender o Bridge como um jogo de soma zero para 2 jogadores, pois podemos considerar cada uma das duplas como um jogador apenas. Escolheremos trabalhar com $K_1(x_1, x_2)$ e, portanto, o jogador 1 será considerado MAX, pois está tentando maximizar o valor da função utilidade. Já o jogador 2 será chamado de MIN. Como os jogadores jogam alternadamente, podemos entender o jogo em árvores de decisão de 3 níveis MAX-MIN-MAX. O jogador MAX fará seu movimento tentando prever o que MIN fará. Por sua vez, MIN escolherá o movimento baseado no que ele acredita que MAX fará. Portanto a decisão de MAX é baseada no que MAX acredita que MIN acredita que MAX fará (decisão em 3 níveis).

Modelo de Melhor Defesa

Basendo-se em conhecimento de jogadores de Bridge, é possível conhecer a melhor jogada a partir de determinadas situações. Além disso, podemos calcular a probabilidade de um movimento ser o majorante, dependendo da distribuição das cartas restantes de determinados naipes. Com base nesse conhecimento, é possível formalizar algumas características do Bridge modelado como um jogo de Informação Imperfeita em que o jogador MAX inicia o jogo:

1. MIN conhece a jogada de MAX e pode prever a próxima jogada
2. MIN escolhe sua estratégia baseado na estratégia de MAX
3. A estratégia de MAX é pura, isto é, ele, de fato, escolhe um movimento (e não associa uma probabilidade a todos os movimentos)

Qualquer jogo transformado em um jogo que tenha as 3 características acima é um jogo na *Forma de Melhor Defesa*

Minimização Exaustiva de Estratégia

O algoritmo abaixo devolve a melhor estratégia para o jogador MAX no modelo de melhor defesa, na forma extensiva, de um jogo de soma zero para 2 jogadores.

Algoritmo 1 • *Crie um conjunto* S *de estratégias do jogador 1, formado por* q*-uplas em que a* i*-ésima entrada representa o caminho escolhido em todos os nós* n *para quem* $\mathcal{I}_1(n) = i$

- *Para cada* $s_j \in S$*, calcule* $E_j = esm(t, s_j)$
- *Devolve as estratégias para quem* s_j *é máximo*

No algoritmo, $esm(t, S)$ é definido pela tabela abaixo:

Condição	Resultado
t é uma folha	$K_1(t)$
$\mathcal{P}(node(t)) = 2$	$\min_{t_i \in sub(t)} esm(t_i, s)$
$\mathcal{P}(node(t)) = 1$	$esm(t_i, s)$, onde i é o $\mathcal{I}_1(node(t))$-ésimo elemento de s
$\mathcal{P}(node(t)) = 0$	$\sum_{t_i \in sub(t)} \pi(node(t_i)) esm(t_i, s)$

Conclusão

Apesar do algoritmo Minimização Exaustiva de Estratégia utilizar-se da Forma Extensiva de um Jogo, foi necessário o estudo de outras modelagens de Jogos com Informação Imperfeita. Para a modelagem na Forma de Melhor Defesa, fez-se necessário o Teorema do Min-Max a fim de garantir que a estratégia do jogador MAX seja pura e, para demonstra-lo, precisamos do jogo na Forma Normal. Por fim, a equivalência da Forma Normal e da Forma Extensiva garante que o algoritmo funciona para a Forma Extensiva.

Referências

[FB98] Ian Frank e David Basin. Search in games with incomplete information: a case study using bridge card play. *Artificial Intelligence*, 100(1-2):87–123, 1998.

[NM44] John Von Neumann e Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.