

TEORIA DOS JOGOS E APRENDIZADO DE MÁQUINA

Estudos Iniciais

André Filipe de Moraes Batista

Disciplina de Aprendizagem de Máquina – UFABC

2010

TEORIA DOS JOGOS

- Ramo da matemática aplicada
 - estuda situações estratégicas onde agentes escolhem diferentes ações na tentativa de melhorar seu retorno
- Ou seja, a teoria dos visualiza qualquer ambiente multiagentes como um jogo
 - em que qualquer agente dado precisará considerar as ações de outros agentes
 - e o modo como essas afetam seu próprio bem estar
 - ambientes cooperativos e competitivos
 - Similar à teoria da decisão
 - Estuda decisões que são tomadas em um ambiente onde vários jogadores interagem
 - Estuda as escolhas de comportamentos ótimos quando o custo e benefício de cada opção não são fixos, mas isso depende, sobretudo, da escolha dos outros indivíduos



INTELIGÊNCIA ARTIFICIAL

○ IA

- Área de pesquisa que investiga formas de habilitar o computador a realizar tarefas nas quais, até o momento, o ser humano tem um melhor desempenho
 - Muitas das técnicas utilizadas são baseadas no raciocínio humano, imitando ou simulando certos aspectos do pensamento e comportamentos inteligentes, realizando ações para atingir seus objetivos
- Resolução de problemas
 - Um dos processos fundamentais para a maioria das aplicações de IA
 - 2 tipos:
 - Procedimento determinístico → algoritmo
 - Procedimentos não determinísticos → busca de uma solução
 - Busca: gerar e testar



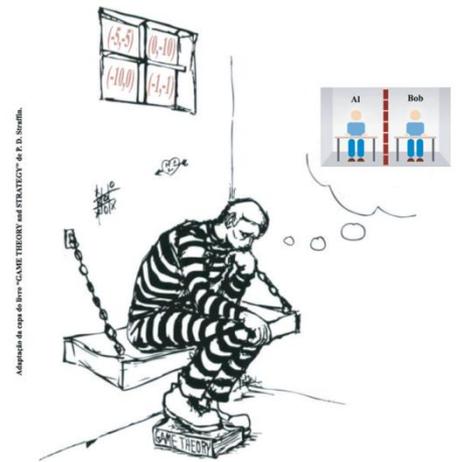
APRENDIZADO DE MÁQUINA

- IA
 - Máquinas só podem ser consideradas inteligentes quando forem capazes de aprender coisas novas e se adaptarem a novas situações, em vez de simplesmente fazer o que foi mandado
- Importantes características das entidades inteligentes
 - Adaptação a novos ambientes e resolução de novos problemas
- Jogos e AM
 - Xadrez – pioneiros na utilização de técnicas de AM
 - Técnica: avaliação estatística
 - Primeira abordagem
 - Alan Turing
 - Valores para peças, de acordo com utilidade para o jogo
 - Soma-se as brancas (B) e as pretas (P)
 - Utilizada-se coeficiente B/P para tomar decisão



DILEMA DO PRISIONEIRO

- Famoso problema da teoria dos jogos
- Jogo não cooperativo
- 2 criminosos são pegos e a polícia tem evidências para mantê-los presos por 1 ano. Porém não para condená-los, os presos são colocados em celas separadas, para que haja um acordo prévio
- As decisões são simultâneas e um não sabe nada sobre o outro



		Prisioneiro 2	
		Confessa	Não confessa
Prisioneiro 1	Confessa	7 / 7	0 / 10
	Não confessa	10 / 0	1 / 1



DILEMA DO PRISIONEIRO

- **O raciocínio do prisioneiro 1 será o seguinte:**
 - – não sei o que o prisioneiro 2 fará;
 - – se ele não confessar será melhor para mim confessar porque ficarei livre em vez de pegar um ano;
 - – se ele confessar também será melhor para mim ter confessado pois pegarei sete anos em vez de dez anos;
 - – sem pensar no que o prisioneiro vai fazer é melhor para mim confessar.
-
- **O outro prisioneiro pensará da mesma maneira.**
-
- Em ambos os casos, a **estratégia dominante** (o que é melhor para um jogador independentemente do que o(s) outro(s) façam) **os levaria a confessarem**. Temos um **equilíbrio de nash**

Equilíbrio de Nash representa uma situação em que nenhum jogador pode melhorar a sua situação dada a estratégia seguida pelo jogador adversário.

VÍDEO

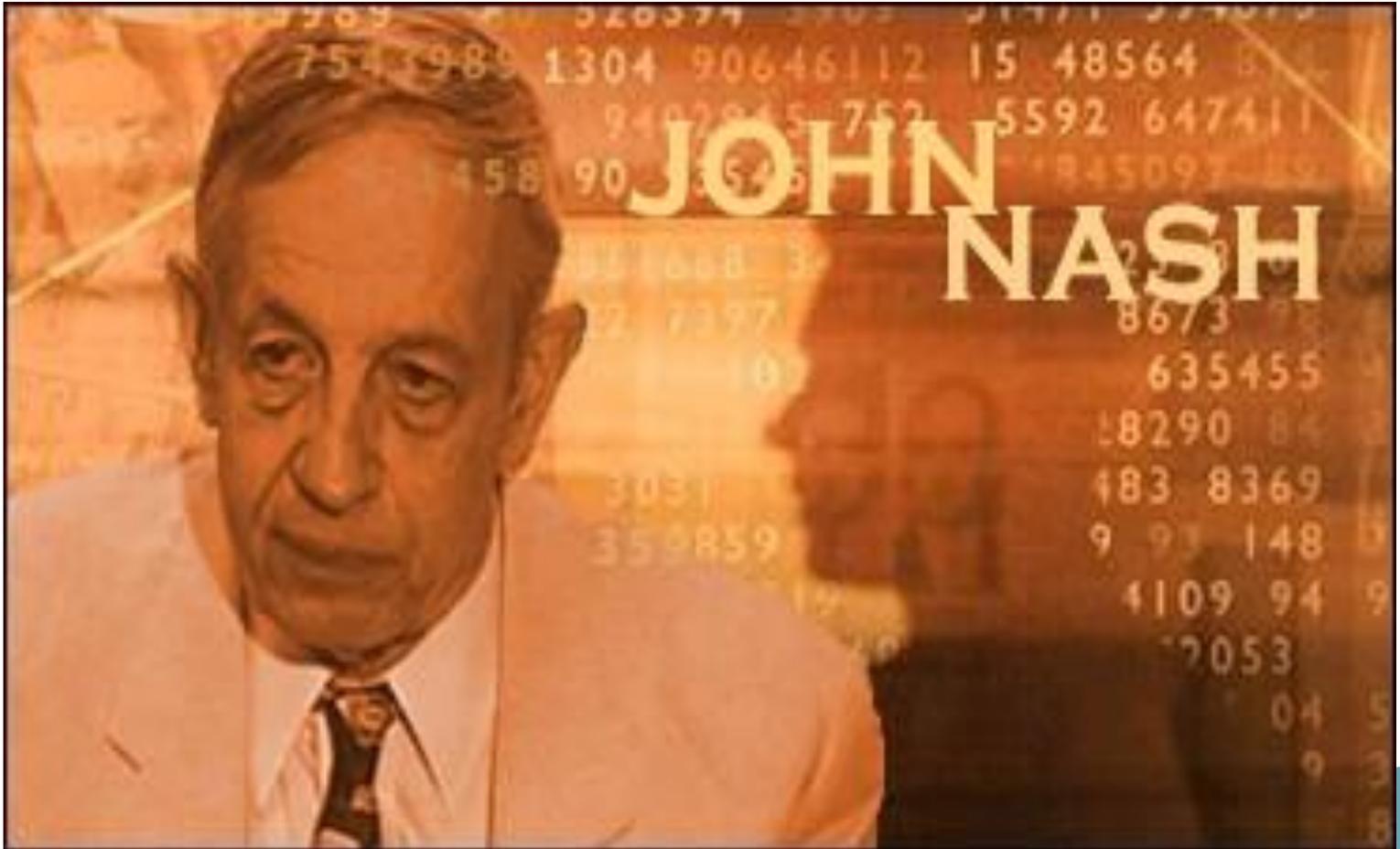
O DILEMA DOS PRISIONEIRO



18626004_mp4_h264_aac.mp4



JOHN FORBES NASH JR



JOHN FORBES NASH JR

- **John Forbes Nash Jr.** (13 de junho de 1928, Bluefield, Virgínia Ocidental) é um matemático que trabalhou na **Teoria dos Jogos** e na Geometria Diferencial.
- Recebeu em 1994 o Prêmio Nobel de Economia. Formado pela Universidade de Princeton, em 1950, com a tese ***Non-Cooperative Games*** (**Jogos Não-Cooperativos**, publicada em 1951).
- Nesta tese, Nash **provou a existência de ao menos um ponto de equilíbrio** em jogos de estratégia para múltiplos jogadores.



JOHN FORBES NASH JR

- Escreveu mais três artigos que consolidaram o chamado "**programa de Nash**" para solução de **jogos estratégicos**:
- ***The Bargaining Problem*** (O Problema da Barganha, 1949);
- ***Equilibrium Points in N-Person Games*** (Pontos de Equilíbrio em Jogos de N-Pessoas, 1950) e
- ***Two-Person Cooperative Games*** (Jogos Cooperativos de Duas Pessoas, 1953).



AM E JOGOS

- Como aprender em jogos?
 - Xadrez
 - Para movimentos mais complexos, é mais difícil decidir quais movimentos contribuíram para a vitória e quais para a derrota
 - Supõe-se que a máquina execute um movimento e tenha sido ruim, mas depois, por causa de um erro do oponente, a máquina acaba por ganhar a partida
 - O sistema comete um erro de creditar vitória ao movimento ruim.
 - O problema de decidir qual séria de ações é na verdade responsável por um determinado resultado é chamado de **problema de atribuição de crédito**
 - Presente em muitos outros problemas de AM



AM E JOGOS

- Ampla escala de fenômenos

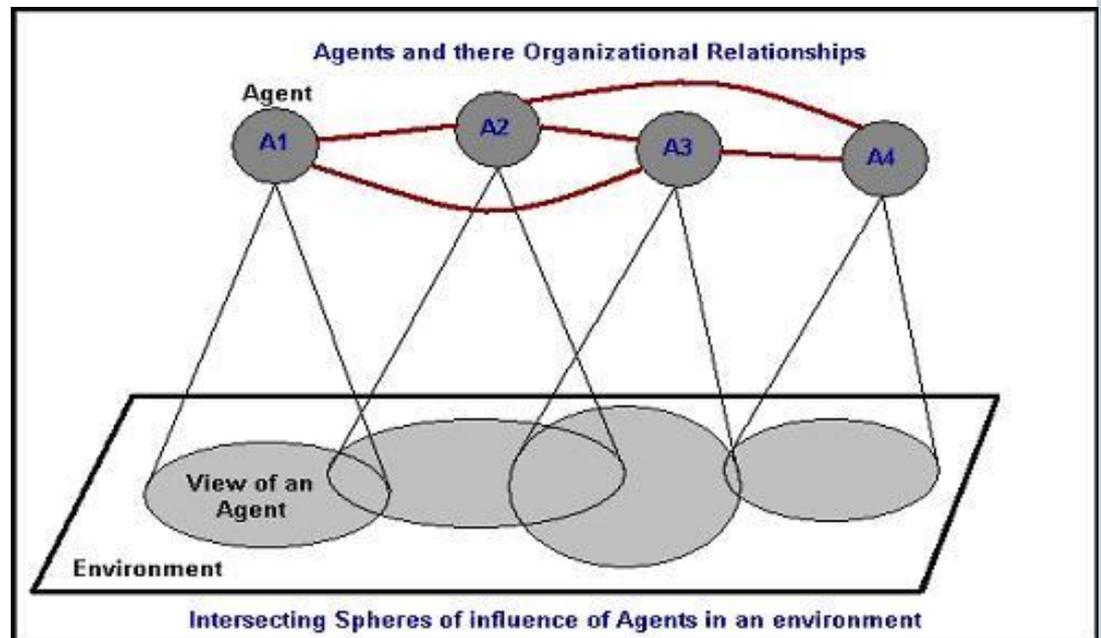


- Refinamento de habilidades
 - Pessoas melhoram a execução de suas tarefas na prática
- Aquisição de Conhecimentos
 - Por si só inclui muitas atividades
 - Conhecimento é normalmente adquirido através de experiência

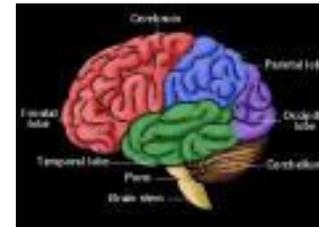
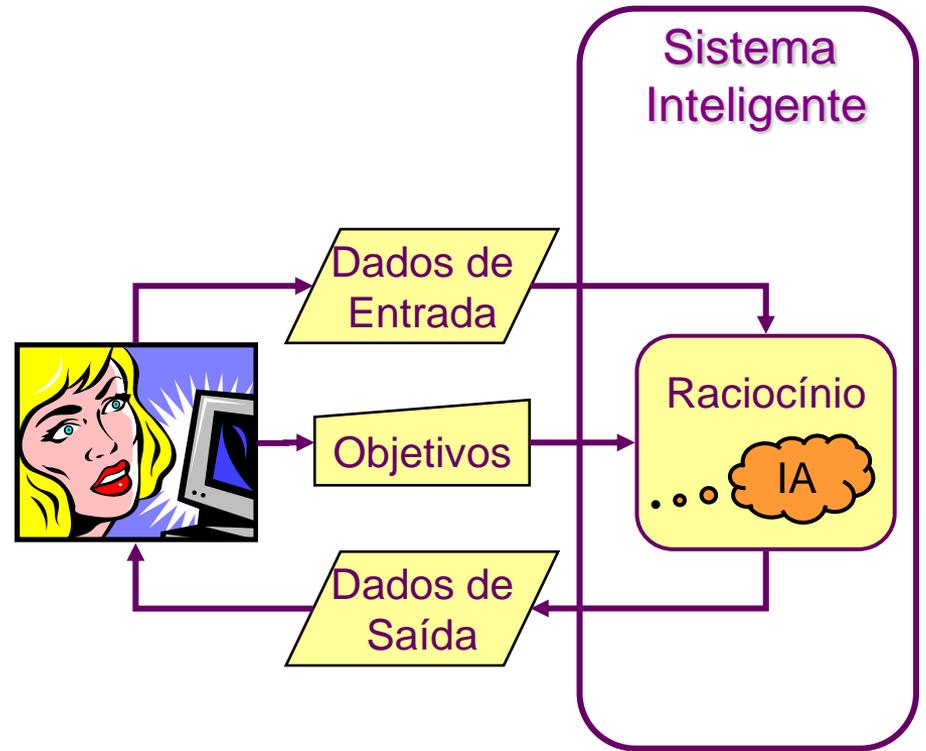
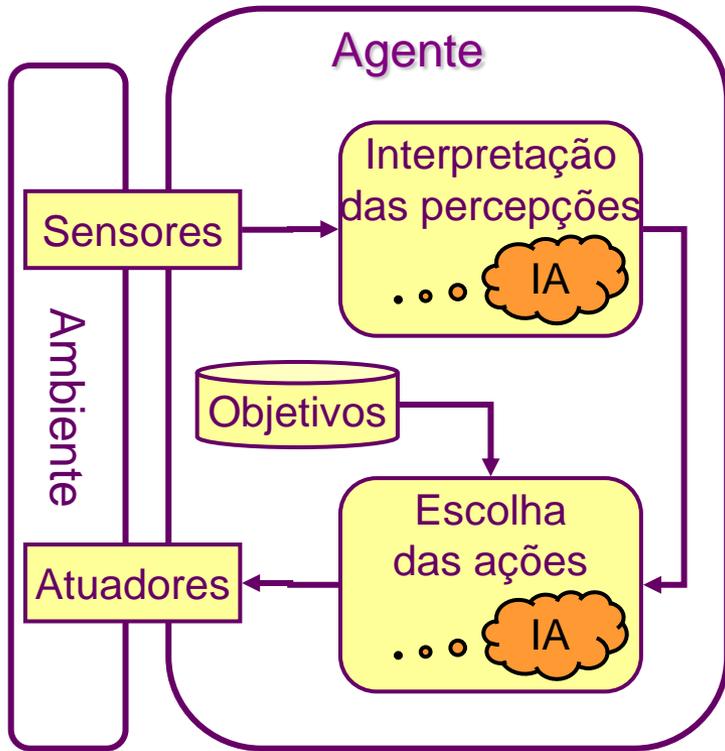


APRENDIZAGEM DE MÁQUINA COLETIVA

- Abordagem Multiagentes
- Sistemas Multiagentes
 - Uma rede flexível de resolvedores de problemas, que **interagem** para resolver problemas que estão além de suas capacidades individuais ou de seus conhecimentos sobre o problema
 - Agentes tentam resolver seus problemas de aprendizado, enquanto outros agentes também o fazem. Algo semelhante com **jogos**?



APRENDIZAGEM DE MÁQUINA COLETIVA



O QUE UM SISTEMA DEVE POSSUIR PARA APRENDER COLETIVAMENTE?

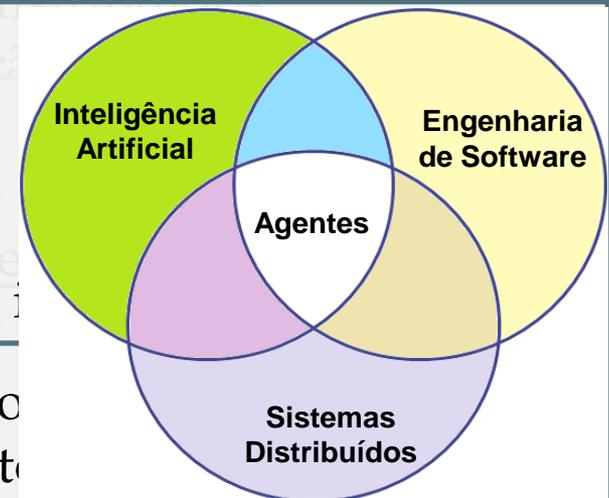
- Autonomia (IA)
 - raciocínio, comportamento guiado por objetivos
 - reatividade



- Adaptabilidade & aprendizagem (IA)
 - Requer máquina de inferência e base de conhecimento
 - Capacidade de adaptação a situações novas, para as quais não foi fornecido todo o conhecimento necessário com antecedência
- Duas implementações
- IA + técnicas avançadas de sistemas distribuídos:

Requer:

- Interface com rede
 - Protocolos de segurança
 - Suporte a código móvel
 - Essencial em agentes de exploração da rede
-
- Requer interface com sistema operacional
 - Essencial em filtragem, monitoramento





NASA's agents are designed to achieve the goals and intentions of the designers, not merely to respond to predefined events, so that they can react to unimagined events and still ensure that the spacecraft does not waste fuel while keeping to its mission.

IBM is building its autonomous agents with the intelligence to respond to changing conditions.



FACT

Agent Oriented Software has



DaimlerChrysler implemented an agent-based system on one factory floor to allow individual workpieces to be directed dynamically around the production area. The intention was to implement **flexible manufacturing** to meet rapidly changing operations targets, and the result was a **20% increase in productivity** on average. The resulting prototype system was in day-to-day operation for five years up to the end of the life-cycle of the targeted product.

FACT
26



VIDEO

PESQUISAS ATUAIS EM SMAS



YouTube- _Reactive Teaming for Game Characters (Final Version)?.mp4



APRENDIZADO EM AMBIENTES COMPETITIVOS: BUSCA COMPETITIVA

- Em um ambiente multiagentes
 - Qualquer agente dado precisará considerar as ações de outros agentes e o modo como essas afetam seu próprio bem estar
 - Imprevisibilidade de outros agentes pode introduzir muitas contingências no processo de resolução de problemas do agente
 - **Ambiente Competitivo**
 - Metas dos agentes estão em conflito
 - Problema de busca concorrente, a.k.a jogos!
 - Forma de aprendizado e decisão rápida: Minimax



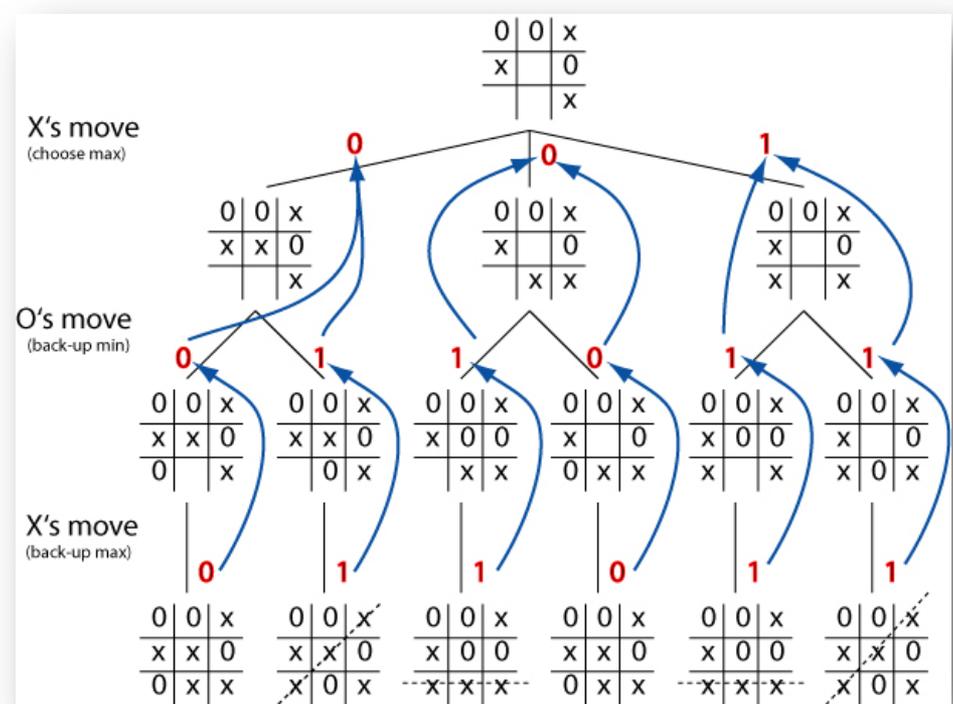
MINIMAX

- Procedimento de busca em profundidade
- Idéia:
 - Iniciar na posição corrente
 - Usar gerador de movimentos plausíveis
 - Conjunto de possíveis posições sucessoras
 - Aplicar função de avaliação estática a essas posições
 - Atribui-se um valor, que representa a qualidade de cada uma dessas posições
 - Retorna-se os valores para o estado inicial para escolher a jogada que resultará no estado mais promissor
- Ambientes Determinísticos vs. Não Determinísticos
 - Uso juntamente com aprendizagem por reforço



MINIMAX

- Lance:
 - Duas jogadas: MAX e MIN
- MAX
 - Tenta maximizar o ganho em suas jogadas
- MIN
 - Tenta minimizá-los
- Soma-Zero (se um ganha, o outro perde)
 - A jogada de um adversário tenta ser minimizada através de jogadas subsequentes que neutralizem seus ganhos
- A busca é feita para determinar a estratégia ótima para MAX
 - MAX deve encontrar uma estratégia de contingência que especifique o movimento de MAX no estado inicial, e depois os movimentos de MAX nos estados resultantes de cada resposta possível de MIN, e depois os movimentos de MAX nos estados resultantes de cada resposta possível de MIN a esses movimentos e assim por diante
 - A **estratégia ótima** leva ao melhor resultado considerando que o adversário sempre faz a melhor jogada possível

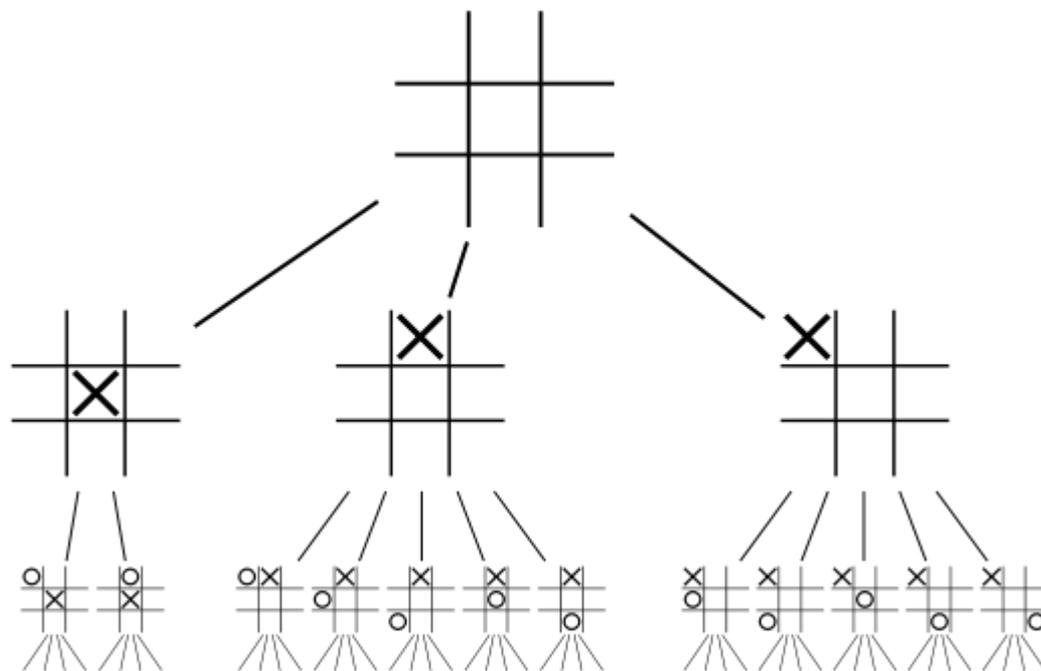


MINIMAX

- Caminha em profundidade na árvore
- Se a profundidade é m e em cada estado existem b jogas possíveis, a ordem de complexidade é $O(b^m)$
 - Algoritmo não é prático para jogos reais
 - Serve para análise matemática e como base para algoritmos mais eficientes
- Pode ser estendido para jogos com múltiplos jogadores
- É possível otimizá-lo
 - Poda alfa-beta, cortando ramos desnecessários para a tomada de decisão



EXECUÇÃO DO MINIMAX - JAVA



MINIMAX E APRENDIZADO POR REFORÇO

- Aprendizado por reforço
 - Baseado em um mecanismo de punição e recompensa
 - Tentativas e os erros são disciplinados por um supervisor, que fornece ao aprendiz um sinal de retorno
- Minimax + Aprendizado por Reforço
 - Alguns trabalhos já utilizam esse algoritmo com outras técnicas de AM
 - Para evitar explosão de consumo de memória, utilizado como mecanismo de treinamento (principalmente em jogos como xadrez)
 - Após treinamento, jogadores utilizam aprendizado por reforço



COMO APRENDER COM JOGOS?

- *Abordagens Ingênuas*

- Idéia básica: agentes realizam adaptação em seu mecanismo de decisão, ignorando o fato de que os demais estão também realizando estas adaptações.
- **1. Jogador fictício:** O Agente observa a frequência média de tempo de ações tomadas por outros agentes.

$$prob(action\ k) = \frac{\#times\ k\ observed}{total\ \# observations}$$

O agente então escolhe a melhor ação a ser tomada para vencer o oponente.

- Possível variante: poderação exponencial – carater recente



COMO APRENDER COM JOGOS?

○ 2. Teoria dos Jogos Evolucionária:

○ Modelo de Replicação Dinâmica:

- Uma população de agentes homogêneos
- A porcentagem de agentes jogando com uma determinada estratégia crescerá na proporção de quão boa é a performance da estratégia na população
- Os agentes da população são aleatoriamente pareados para disputarem um jogo simétrico: um jogo no qual ambos agentes possuem o mesmo conjunto de estratégias possíveis e recebem as mesmas recompensas para as ações



COMO APRENDER COM JOGOS?

REPLICATOR DYNAMICS (CONTINUAÇÃO)

- Seja $\phi^t(s)$ o número de agentes usando a estratégia s no tempo t .
Definimos:

$$\theta^t(s) = \frac{\phi^t(s)}{\sum_{s' \in S} \phi^t(s')}$$

Como a **fração de agentes jogando com s no tempo t** .

- A utilidade esperada para um agente jogando com s no tempo t é:

$$u^t(s) \equiv \sum_{s' \in S} \theta^t(s') u(s, s')$$

- Onde $u(s, s')$ é a utilidade recebida por um agente jogando com s contra um agente jogando com s' .

- Evolução da População

$$\phi^{t+1}(s) = \phi^t(s)(1 + u^t(s))$$



CONSIDERAÇÕES

○ Teoria dos Jogos

- Área de pesquisa que ganhou forte interesse, principalmente após as definições de **probabilidade**
- Quando do aprendizado em sociedade, a mesma pode ser utilizada como uma forma de adaptação de escolhas e mecanismos de decisão
- Área de pesquisa sendo descoberta por *engenheiros de software (aplicações distribuídas)*
- Mecanismo auxiliar dos processos de *Machine Learning*



REFERÊNCIAS

In Eduardo Alonso, editor, *Adaptive Agents: LNAI 2636*. Springer Verlag, 2003.

Learning in Multiagent Systems: An Introduction from a Game-Theoretic Perspective

José M. Vidal

University of South Carolina, Computer Science and Engineering,
Columbia, SC 29208
vidal@sc.edu

Abstract. We introduce the topic of learning in multiagent systems. We first provide a quick introduction to the field of game theory, focusing on the equilibrium concepts of iterated dominance, and Nash equilibrium. We show some of the most relevant findings in the theory of learning in games, including theorems on fictitious play, replicator dynamics, and evolutionary stable strategies. The CLRI theory and n -level learning agents are introduced as attempts to apply some of these findings to the problem of engineering multiagent systems with learning agents. Finally, we summarize some of the remaining challenges in the field of learning in multiagent systems.

1 Introduction

The engineering of multiagent systems composed of learning agents brings together techniques from machine learning, game theory, utility theory, and complex systems. A designer must choose carefully which machine-learning algorithm to use since otherwise the system's behavior will be unpredictable and often undesirable. Fortunately, we can use the tools from these areas in an effort to predict the expected system behaviors. In this article we introduce these techniques and explain how they are used in the engineering of learning multiagent systems.

The goal of machine learning research is the development of algorithms that increase the ability of an agent to match a set of inputs to their corresponding outputs [7]. That is, we assume the existence of a large, sometimes infinite, set of examples E . Each example $e \in E$ is a pair $e = \{a, b\}$ where $a \in A$ represents the input the agent receives and $b \in B$ is the output the agent should produce when receiving this input. The agent must find a function f which maps $A \rightarrow B$ for as many examples of A as possible. In a controlled test the set E is usually first divided into a *training set* which is used for training the agent, and a *testing set* which is used for testing the performance of the agent. In some scenarios it is impossible to first train the agent and then test it. In these cases the training and testing examples are interleaved. The agent's performance is assessed on an ongoing manner.

When a learning agent is placed in a multiagent scenario these fundamental assumptions of machine learning are violated. The agent is no longer learning

Teoria dos Jogos: As origens e os fundamentos da Teoria dos Jogos

Alexsandra Neri de Almeida
UNIMESP - Centro Universitário Metropolitano de São Paulo
Novembro/2006

Sinopse: A teoria dos jogos é a aplicação da lógica matemática no processo de tomada de decisões nos jogos, utilizada, na economia, na política, na guerra e caracterizadas, como nos jogos, por conflitos de interesse determinando a melhor estratégia para cada jogador.

Introdução

A teoria dos jogos tem a finalidade de prever os movimentos dos outros jogadores, sejam eles concorrentes ou aliados, através dessa teoria os jogadores se posicionam da melhor forma para obter o resultado desejado.

O objetivo da teoria dos jogos é entender a lógica na hora da decisão e ajudar a responder se é possível haver colaboração entre os jogadores, em quais circunstâncias o mais racional é não colaborar e quais estratégias devem ser adotadas para garantir a colaboração entre os jogadores.

A teoria dos jogos, por meio da matemática, equaciona os conflitos, onde o foco são as estratégias utilizadas pelos jogadores.

O objetivo desse trabalho é estudar a origem da Teoria dos Jogos e apresentar brevemente seus fundamentos e aplicações.

As Origens da teoria dos Jogos

Devido à falta de interesse científico, até 1920, não haviam análises técnicas adequadas para estudar estratégias de jogos.

Os jogos de tabuleiros, dados, cartas, ou em geral, os jogos de salão, divertem a humanidade desde a formação das primeiras civilizações, por colocarem as pessoas em situações nas quais vencer ou perder dependem das escolhas feitas no início das partidas, sendo assim, o jogo se tornou uma ferramenta para o desenvolvimento das pessoas, mas só despertou interesse após muito tempo, com o surgimento da teoria da probabilidade.

Os estudos sobre a teoria da probabilidade tiveram início com o filósofo, matemático e físico francês Blaise Pascal, juntamente com o matemático francês Fermat, através desses estudos desenvolveram a teoria da probabilidade em jogos de azar utilizando regras matemáticas.

Em seguida Antoine Augustin Cournot (1801-1877), matemático francês, com estudo da análise do ponto de equilíbrio nas estratégias de jogos, formalizou um conceito específico de equilíbrio, ou seja, aplicados em casos particulares, que mais tarde foi generalizado por John Forbes Nash Jr.

Mas o marco inicial da teoria dos jogos foi quando John Von Neumann (1903-1957), matemático húngaro-americano, provou o teorema minimax, segundo este teorema há sempre uma solução racional para um conflito bem definido entre dois indivíduos cujos interesses são completamente opostos, teorema deixado aberto pelo

