Controlo Distribuído de Agentes Não-Cooperantes

Exploradores e Traidores

Pedro Mariano
plsm@det.ua.pt
http://labmag.di.fc.ul.pt/~plsm

Departamento de Electrónica,
Telecomunicações e Informática
Universidade de Aveiro

Primeira Parte

Teoria de Jogos

Conceitos gerais

- Jogador
- Estratégia
- Ganho
- Utilidade
- Jogo a n-jogadores
- Agente procura maximizar a sua utilidade
- Qual a estratégia que deve seguir?

Exemplo – 1

Dilema do Prisioneiro

- Jogo a dois jogadores
- Duas acções: C e T

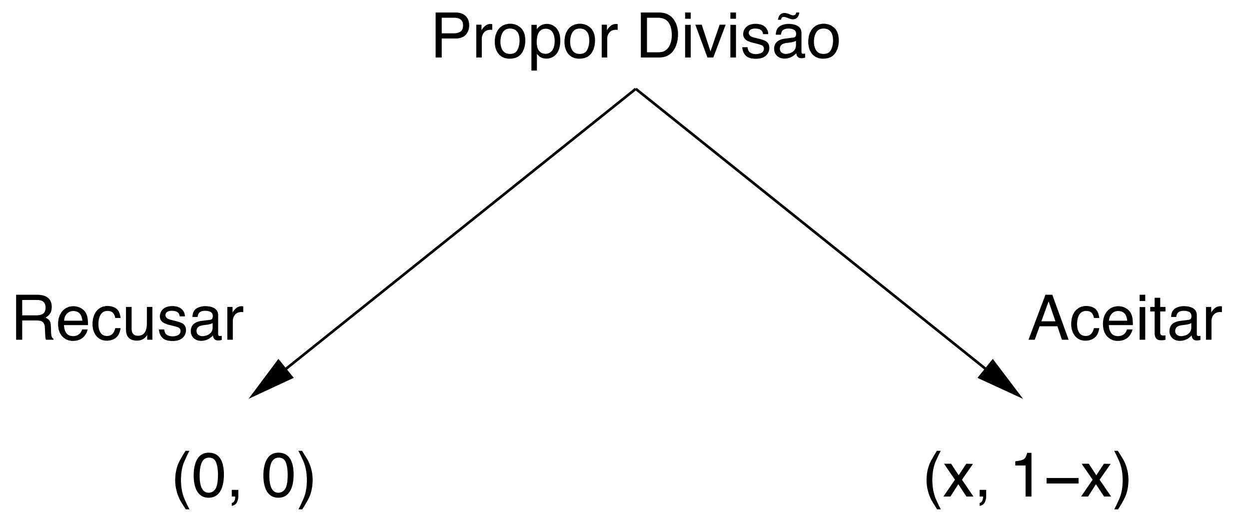
	C	T
C	3	0
T	5	1

ganho do agente que escolhe linhas

Exemplo – 2

Ultimato

- Jogo a dois jogadores
- O Ditador propõe uma divisão x de uma quantia em dinheiro



Exemplo – 3

Give-Take

- Jogo a dois jogadores
- Agentes têm estado
  - Agente com o recurso pode dá-lo ou não fazer nada
  - Agente sem o recurso pode tirá-lo ou não fazer nada

com recurso		
	N	T
N	p_r	$-c_{st}$
G	b_g	$b_g - c_{st}$

sem recurso		
	N	T
	0	$p_r - c_{pt}$
	p_r	$p_r - c_{pt}$

Qual a Estratégia a Utilizar? − 1

- Algoritmo *min-max*
  - Minimizar as perdas
  - No jogo Dilema do Prisioneiro devo escolher a acção T
  - No jogo Ultimato
    - O Ditador deve escolher a divisão mais favorável a ele
    - O outro jogador deve aceitar qualquer proposta

Qual a Estratégia a Utilizar? − 2

- Equilíbrio de Nash

$$\mathbf{s}^\star = (s_1^\star, s_2^\star, \ldots, s_n^\star)$$

- Conjunto de estratégias tal que nenhum agente tem um incentivo para mudar a sua estratégia

$$u_i(\mathbf{s}^\star) \geq u_i(s_1^\star, \ldots, s_i, \ldots, s_n^\star) \quad \forall s_i \in \mathbb{S}_i$$

Qual a Estratégia a Utilizar? – 3

- Óptima de Pareto
- Conjunto de estratégias que maximiza os ganhos dos agentes

$$\max_{\mathbf{s}} \sum_i u_i(\mathbf{s})$$

Estes conceitos apresentados nos 3 últimos acetatos dizem respeito a situações em que o jogo é jogado uma única vez. Quer o jogo seja iterado ou não. Mas podemos considerar uma situação em que temos não um conjunto de jogadores para jogar o jogo, mas sim uma população de agentes sujeita a um regime evolutivo.

Jogos Evolucionários

- População de estratégias que varia ao longo do tempo
- A proporção de uma estratégia depende de
  - Do seu ganho: $u_i(\mathbf{s})$
  - Do ganho médio de todas as estratégias: $\overline{u}$
- Progressão de cada estratégia em relação ao tempo:

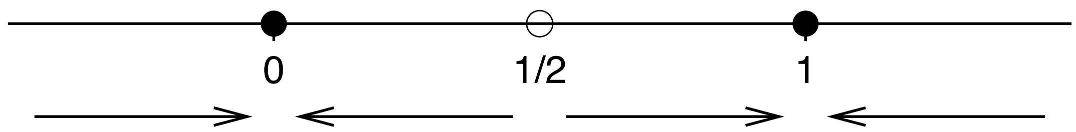
$$\frac{\delta x_i}{\delta t} = x_i(u_i - \overline{u})$$

  Equação de Replicação

- Pressupõe que todas as estratégias interagem umas com as outras

Os Jogos Evolucionários estendem o conceito de jogos a uma população de estratégias sujeita a um regime evolucionário. As proporções das estratégias variam conforme o seu desempenho no jogo. Uma equação diferencial modela a variação das proporções das estratégias ao longo do tempo. Os pontos fixos desta equação representam as Estratégias Evolucionarmente Estáveis.

Exemplo Pontos Fixos



- $x = 0$ e $x = 1$ são pontos fixos estáveis
- $x = 1/2$ ponto fixo instável

Um ponto fixo estável da equação de replicação representa um estado estável da população: para perturbações pequenas a população retorna ao ponto fixo.

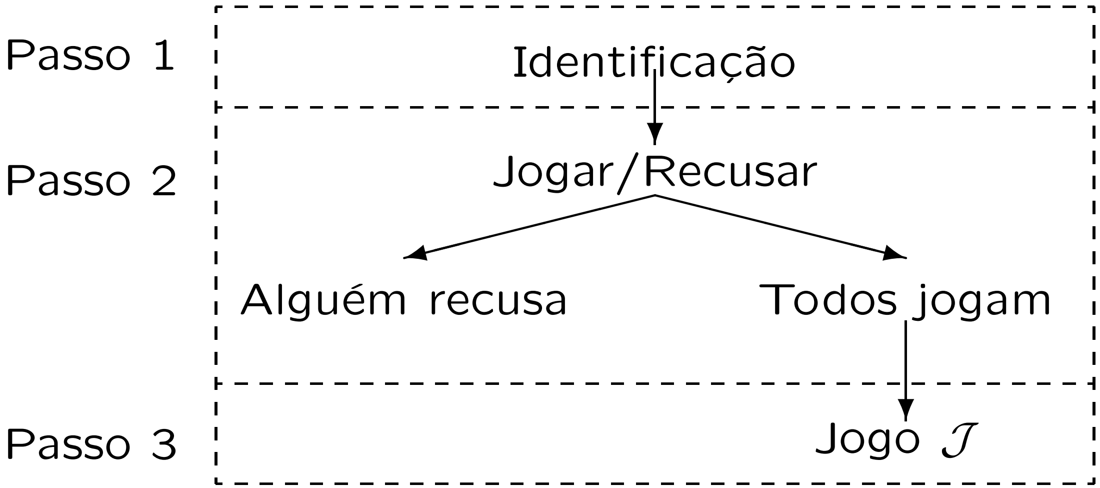
No exemplo mostra-se uma representação dos pontos fixos de uma equação com x=0 e x=1 pontos fixos estáveis e x=1/2 ponto fixo instável.

Com isto termino uma breve apresentação da Teoria de Jogos.

Segunda Parte

Jogo de Recusa

- Meta-Jogo aplicável a qualquer jogo
- Primeira fase onde os jogadores decidem se jogam ou não
- Recusa em jogar implica ganho u_{RC} para todos os jogadores



O jogo de recusa foi desenvolvido com vista a alcançar o objectivo deste trabalho. É um meta-jogo aplicável a qualquer jogo J. Numa primeira fase os agentes identificam-se e depois anunciam se querem jogar J ou não. Caso algum recuse, então todos os agentes obtêm um ganho dito uRC que é menor que qualquer ganho no jogo J.

Análise Teórica – Ganhos

- Dois grupos de agentes

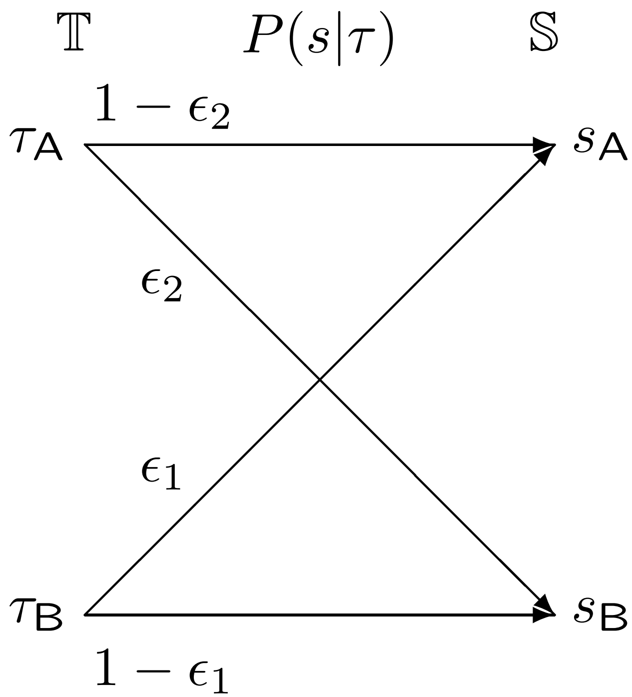
  A– grupo dos agentes não-cooperantes

  B– grupo dos agentes cooperantes

jogadores	utilidade
todos A	$u^{\mathcal{R}}_{\mathsf{AA}} = u^{\mathcal{J}}_{\mathsf{AA}}$
misto	$u^{\mathcal{R}}_{\mathsf{AB}} = (1-\epsilon_1)u_{RC} + \epsilon_1\, u^{\mathcal{J}}_{\mathsf{AB}}$
	$u^{\mathcal{R}}_{\mathsf{BA}} = (1-\epsilon_1)u_{RC} + \epsilon_1\, u^{\mathcal{J}}_{\mathsf{BA}}$
todos B	$u^{\mathcal{R}}_{\mathsf{BB}} = \epsilon_2\, u_{RC} + (1-\epsilon_2)u^{\mathcal{J}}_{\mathsf{BB}}$

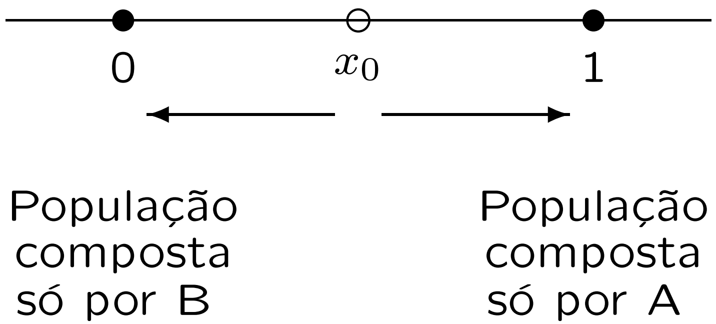
Relativamente à análise teórica considerei que existem dois grupos de agentes: os A são os agentes não-cooperantes enquanto os B são compostos pelos cooperantes. A análise visa verificar quais as condições em que os agentes B são capazes de serem os dominantes na população. Interessa saber a proporção de enganos que os B podem fazer: quando jogam com os A e nenhum deles recusa; quando jogam entre si e pelo menos em deles recusa jogar. Daí nesta tabela o valor do ganho dos B depender dos parâmetros ϵ_1 e ϵ_2.

Análise Teórica – Identificação

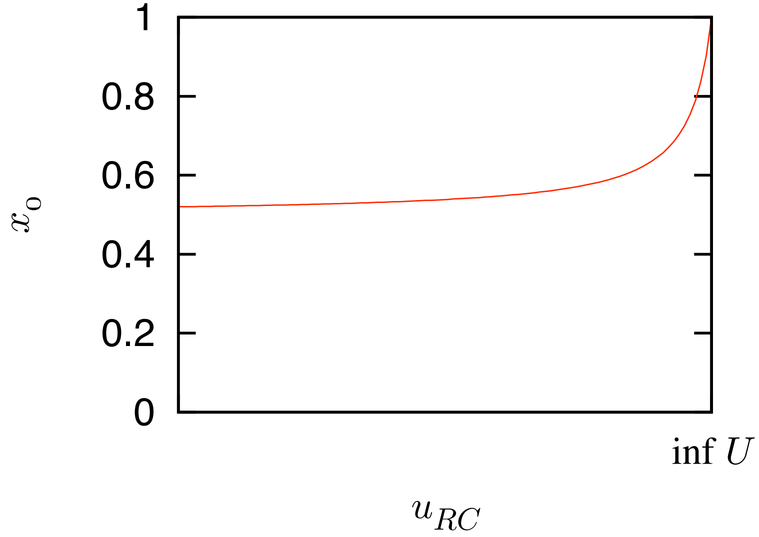


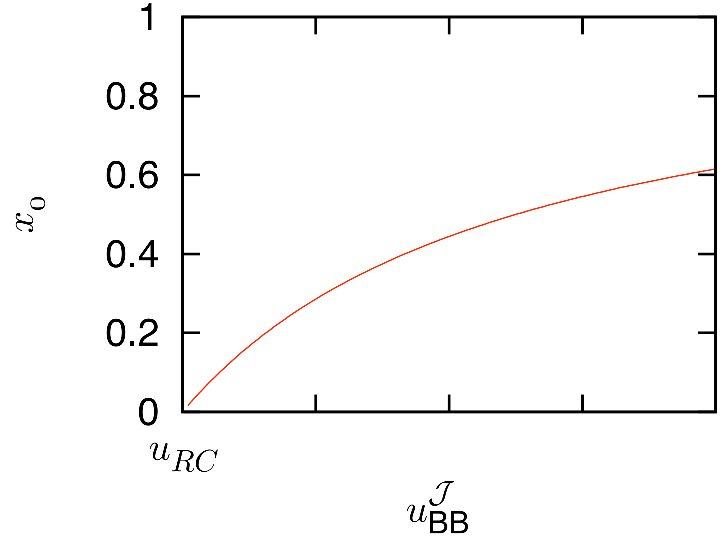
Os parâmetros ϵ_1 e ϵ_2 podem ser relacionados com a identificação da estratégia de um agente dado o seu tipo.

Análise Teórica – Ponto Fixo



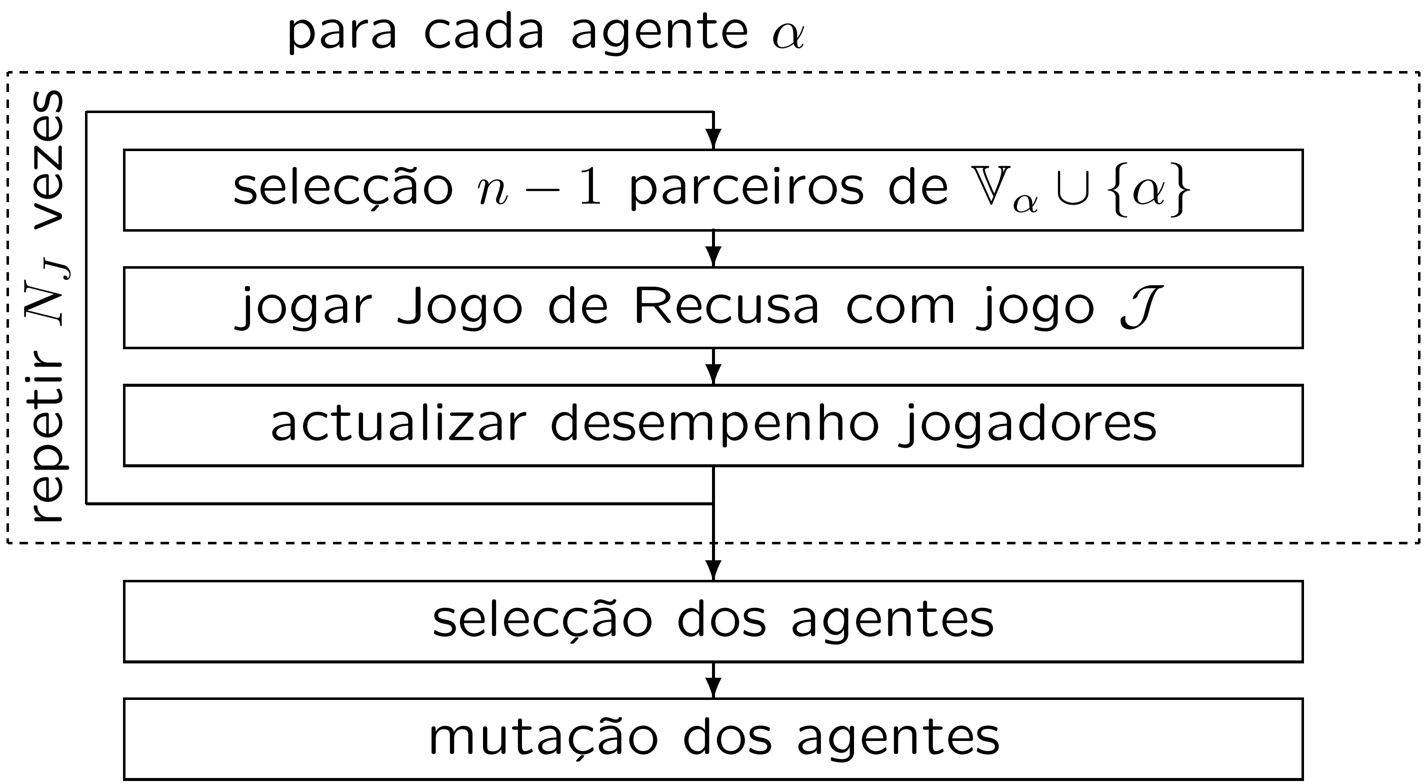
$$x_0 \equiv \frac{1}{1 + \dfrac{u^{\mathcal{R}}_{\mathsf{AA}} - u_{RC}}{u^{\mathcal{J}}_{\mathsf{BB}} - u_{RC}}}$$





Neste figura podemos ver que se a população é composta só por A ou só por B então mantém-se neste estado. O ponto x0 é aquele que divide os casos em que a população tende para um dos estados que mencionei. Podemos ver que o valor de x0 depende, assumindo que há distinção entre os A e os B, de uRC e do ganho uJBB (este é o ganho que os B obtêm quando jogam o jogo J entre si). O ganho de uRAA pode ser qualquer. Neste gráfico podemos ver que se diminuirmos uRC então quer dizer que eu preciso de ter 50 por cento + 1 agentes B para que a população passe a ser composta só por B. Devo escolher este valor se a assunção anterior não se verificar.

Análise Experimental

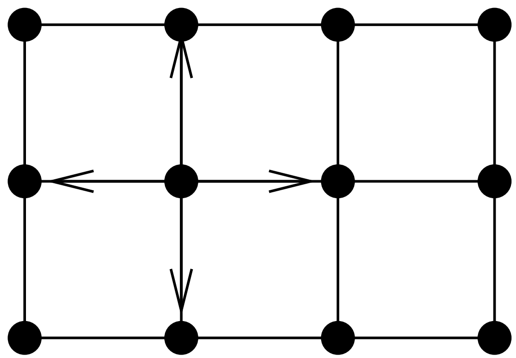


Relativamente à análise experimental neste acetato podemos ver o esquema utilizado: para cada agente são seleccionados n-1 jogadores, joga-se o jogo de recusa com o jogo J e por fim são aplicados os operadores evolucionários.

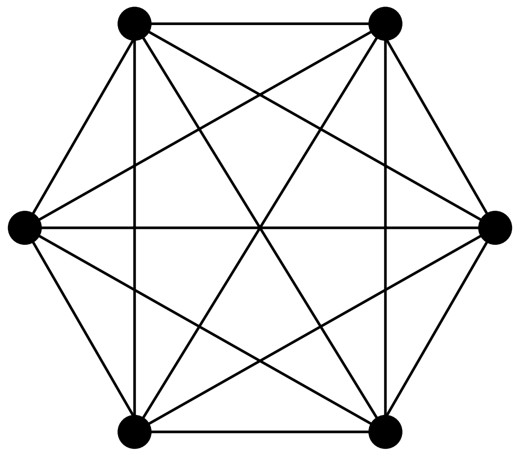
Com esta abordagem pretendemos analisar a dinâmica do jogo ao longo do tempo: saber se o jogo tem estratégias evolucionarmente estáveis; quais as condições em que as EES são capazes de resistir (uma taxa de mutação demasiado alta pode eliminar uma EES da população).

População

- Estrutura da população influencia:
  - Com quem um agente interage
  - Argumentos dos operadores evolucionários
- Estruturas utilizadas



lattice



bag

Composição do Agente

- Estratégia s
- Tipo $\tau \in \mathbb{T} = \{1, 2, \ldots\}$
- Relação entre estratégia e tipo

$$\tau \leftarrow (\varphi(s) + U(0, x)) \bmod |\mathbb{T}|$$

- Vector de utilidades indexado por tipo

$$\mathbf{v}(\tau) = (v(1), v(2), \ldots)$$

- Vector de número de jogos indexado por tipo

$$\mathbf{j}(\tau) = (j(1), j(2), \ldots)$$

- Parâmetros globais
  - Jogos exploratórios p
  - Limite de jogar/recusar σ

Aqui neste acetato podemos ver qual a definição de agente que foi utilizada neste trabalho. Tendo em conta a nossa assunção de que os agentes interagem no contexto de um jogo, um dos parâmetros do agente é a estratégia s que utilizam. Adicionalmente temos o tipo que é utilizado na fase de identificação do Jogo de Recusa. O tipo é função da estratégia adicionado de um ruído uniforme. Deste modo ao variarmos o parâmetro x conseguimos ter agentes onde é possível distinguir perfeitamente as estratégias dado um tipo. Claro que se as estratégia têm parâmetros reais ou são estocásticas, há sempre alguma indefinição quanto à estratégia associada a determinado tipo.

O vector de utilidades e de número de jogos indexado por tipo são utilizados na fase de jogar/recusar do Jogo de Recusa.

Agente e Informação sobre Parceiros

- Informação de suporte à decisão de jogar ou recusar
- Média dos ganhos obtidos em função do tipo do agente

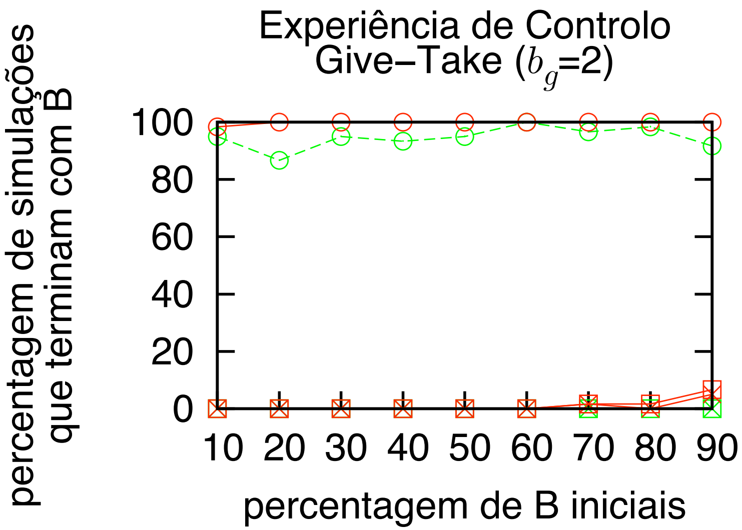
$$v(\tau) \leftarrow \frac{v(\tau)\, j(\tau) + u}{j(\tau) + 1}$$

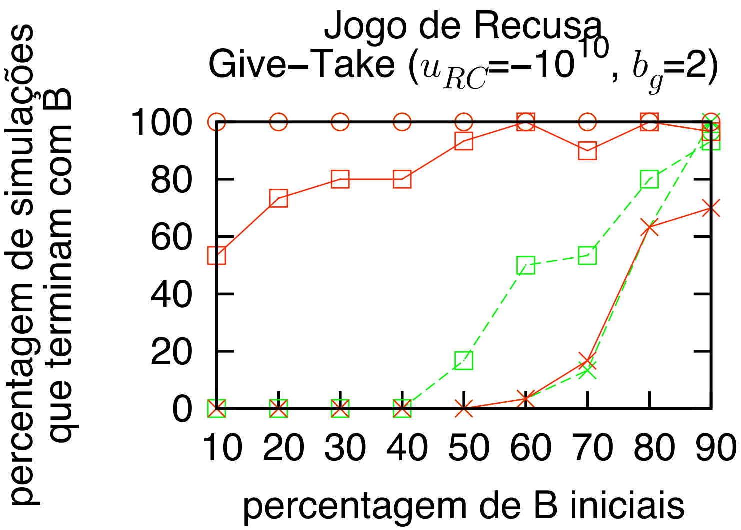
- Joga ou recusa se:

$$\begin{cases} j(\tau) \leq p & \text{joga} \\ j(\tau) > p & \begin{cases} v(\tau) < \sigma & \text{recusa} \\ v(\tau) \geq \sigma & \text{joga} \end{cases} \end{cases}$$

Aqui mostramos o algoritmo utilizado pelos agentes para decidirem se jogam ou se recusam. Existe um período inicial, representado pelo parâmetro p, que representa o número inicial de jogos que um agente joga sempre com qualquer tipo. Após esse número ser atingido, o agente recusa se a informação que juntou for inferior a um limite, representado pelo parâmetro σ.

Resultados – 1





- Estrutura da população lattice e bag
    - Influencia: interacção, operadores evolucionários
- Agentes A
    - Parâmetros gerados aleatoriamente e podem recusar ou jogar sempre
    - Parâmetros pré-definidos

Analisei diversos parâmetros quer relativos ao Jogo de Recusa quer a outros aspectos. Neste acetato apresento à esquerda o resultado de uma experiência de controlo onde não é utilizado o jogo de recusa, e à direita é utilizado o jogo de recusa. Estes gráficos mostram a percentagem de simulações que terminam só com agentes B em função de: PRIMEIRO, percentagem inicial de B's; SEGUNDO, estrutura da população (que determina com quem um agente interage e o domínio dos operadores evolucionários); e TERCEIRO, parâmetros dos agentes A (parâmetros calculados aleatoriamente, ou pré-definidos). Podemos ver que com o jogo de recusa há um incremento do número de simulações que terminam só com agentes B.

Resultados – 2

Percentagem de condições em que há correlação entre um parâmetro e a presença de B na última geração

	u_{RC}	p	estrat. id.
Give-Take $b_g = 0$	18,5	15,7	2,8
Give-Take $b_g = 2$	31,5	13,0	0,0
IPD	15,0	13,0	0,9
D. Lenhador	18,5	18,8	0,9

- Valor de u_{RC}: $\inf U$, -10^{10}
- p, número de jogos iniciais que um agente nunca recusa
- Relação entre Estratégia e Identidade
- Jogos: Give-Take, Dilema do Prisioneiro Iterado, Dilema do Lenhador

Neste outro acetato apresento um quadro com a percentagem de simulações onde o teste da hipótese de existência de correlação entre cada um dos três parâmetros uRC, p (número de jogos iniciais que um agente não recusa) e Estratégia-Tipo (relação entre a estratégia e o tipo), e a presença ou não de B na última geração. Podemos ver que é o parâmetro uRC aquele que maior influencia tem na presença ou não de B na última geração.

Trabalho Futuro

- Outras estruturas da população
- Ligações que podem surgir ou desaparecer
- Número de agentes variáveis

Tanto a análise teórica como a prática considerou que os agentes interagem uns com os outros. Só na análise prática é que foram consideradas outras topologias de interacção. O trabalho

Conclusões

- Jogo de Recusa aumenta as hipóteses de uma população passar a ser composta por um único grupo
- O grupo com maior ganho entre si não precisa de ter uma proporção grande na população
- O grupo dos cooperantes tem vantagem com o Jogo de Recusa
  - Requer a menor proporção inicial

Concluindo esta apresentação, mostrei que o jogo de recusa aumenta as hipóteses de uma população ser composta unicamente por agentes cooperantes, sendo que estes não precisam de uma proporção grande para poder alcançar este objectivo. Os cooperantes são beneficiados com o Jogo de Recusa.

Pergunta − 1

Considere o seguinte jogo:

	D	C
G	0,1	4,0
P	0,1	2,2

Em cada entrada o primeiro valor representa o ganho do jogador de linhas

Suponha que joga 2 jogos com alguma das pessoas que está sentada perto de si. Qual o seu ganho acumulado nos dois jogos?

Pergunta – 2

Desenhe o grafo em que cada nó representa um aluno e cada ligação entre dois nós indica se esses dois alunos estudam ou trabalham em conjunto. Desenvolva um jogo que modele uma situação esclarecimento de dúvidas entre dois alunos: por exemplo um aluno tem uma dúvida e pode pedir ajuda a outro aluno que pode responder correcta ou incorrectamente. Ao invés de procurar ajuda, o aluno pode procurar em livros o que consome tempo. Desenvolva pelo menos dois modelos.