Teoria de Jogos Evolucionária

Edmundo de Souza e Silva - Daniel Ratton Figueiredo

Universidade Federal do Rio de Janeiro
Programa de Engenharia de Sistemas e Computação - COPPE
Departamento de Ciência da Computação do IM

- Assume jogadores são racionais
- Maximizam suas recompensas e sabem que os outros jogadores fazem o mesmo
- Nada é dito sobre como os jogadores atingem o equilíbrio de Nash é atingido
- Caso mais de um equilíbrio exista, como ele é atingido?

- **Tenta explicar o comportamento dos sistemas que, em geral, evoluem com o tempo**
- **Jogo repetido infinitas vezes**
- **Jogadores possuem uma dinâmica de adaptação de estratégia**
  - jogadores podem mudar de estratégia ao longo do jogo, de acordo com o ganho que eles recebem
  - Lembra do jogo do acesso a canal sem-fio de ontem?
- **Adaptação visa melhorar o desempenho do jogador**
- **Jogadores não necessariamente são racionais**
- **Tenta estudar a convergência do processo adaptativo**

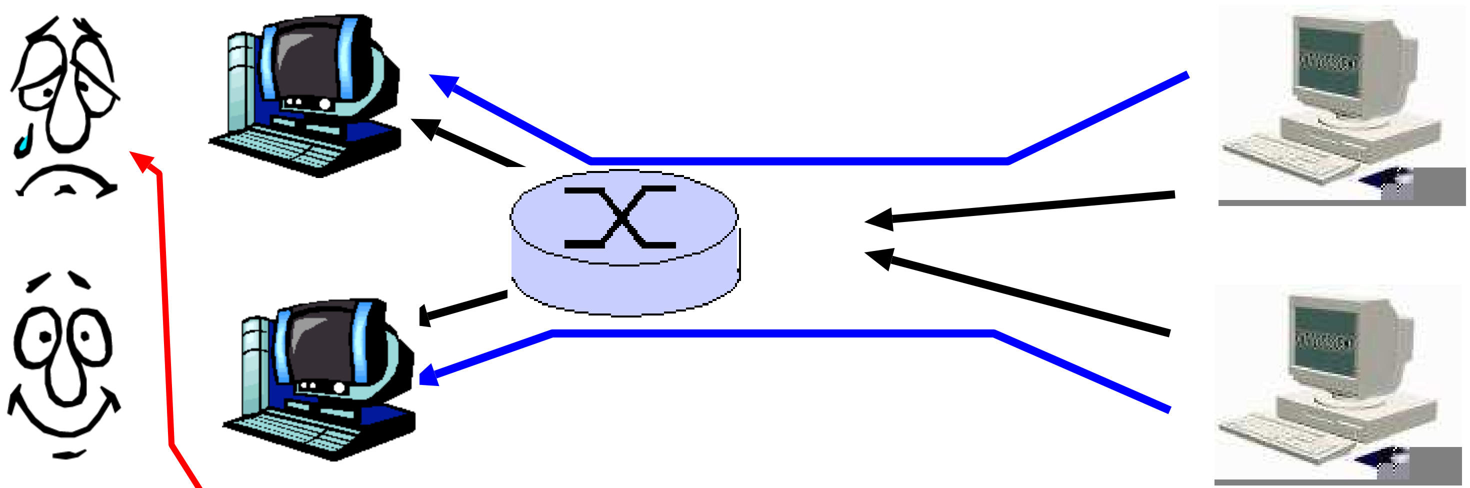
Exemplo

- Dois irmãos compartilham seu canal de acesso a Internet em casa.
- Ambos decidem escutar música via internet, ao mesmo tempo
- Ambos estão usando a versão do VivaVoz que permite escolher a taxa de recebemento do áudio
  - Podem escolher 3 taxas diferentes de recepção: 24Kbps, 64Kbps, 128Kbps
- Quanto maior a taxa, melhor a qualidade do som
- Mas... o canal compartilhado não tem capacidade suficiente para suportar as 2 conexões a taxa mais elevada

- **Ambos estão usando a versão do VivaVoz que permite escolher a taxa de recebemento do áudio**
  - Podem escolher 3 taxas diferentes de recepção: 24Kbps, 64Kbps, 128Kbps

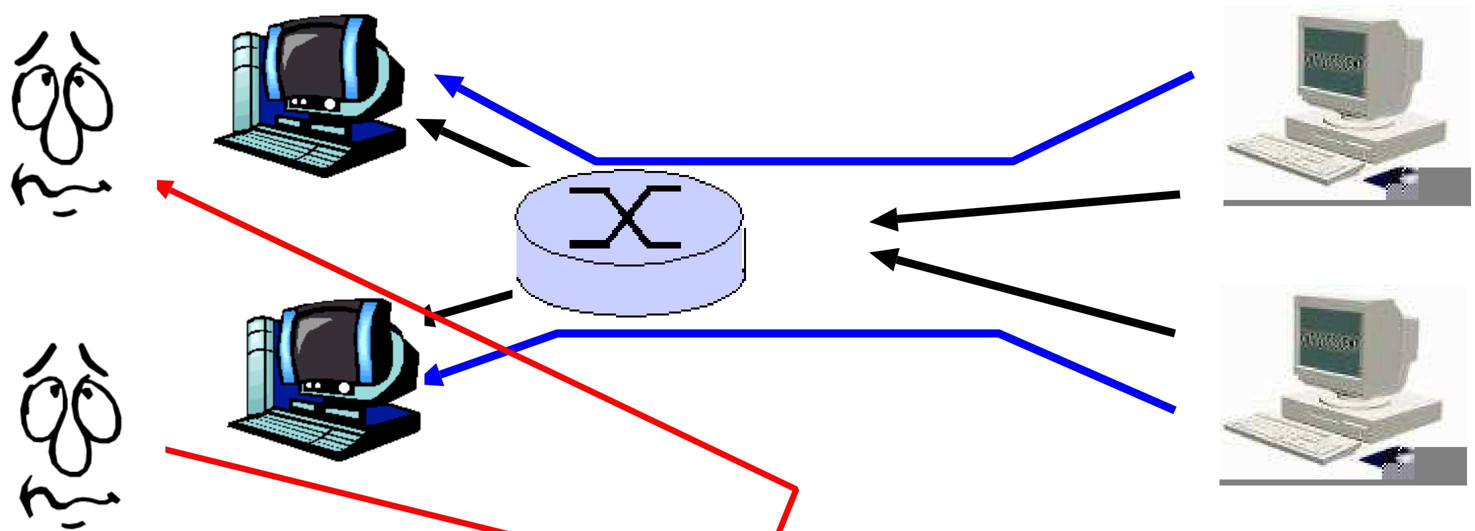
estado	Qualidade (MOS)		
$N\sigma$	128 Kbps	64 Kbps	24 Kbps
2 0 0	2,0	—	—
1 1 0	3,0	2,1	—
0 2 0	—	4,0	—
1 0 1	3,5	—	2,3
0 1 1	—	3,8	3,1
0 0 2	—	—	3,3

Exemplo

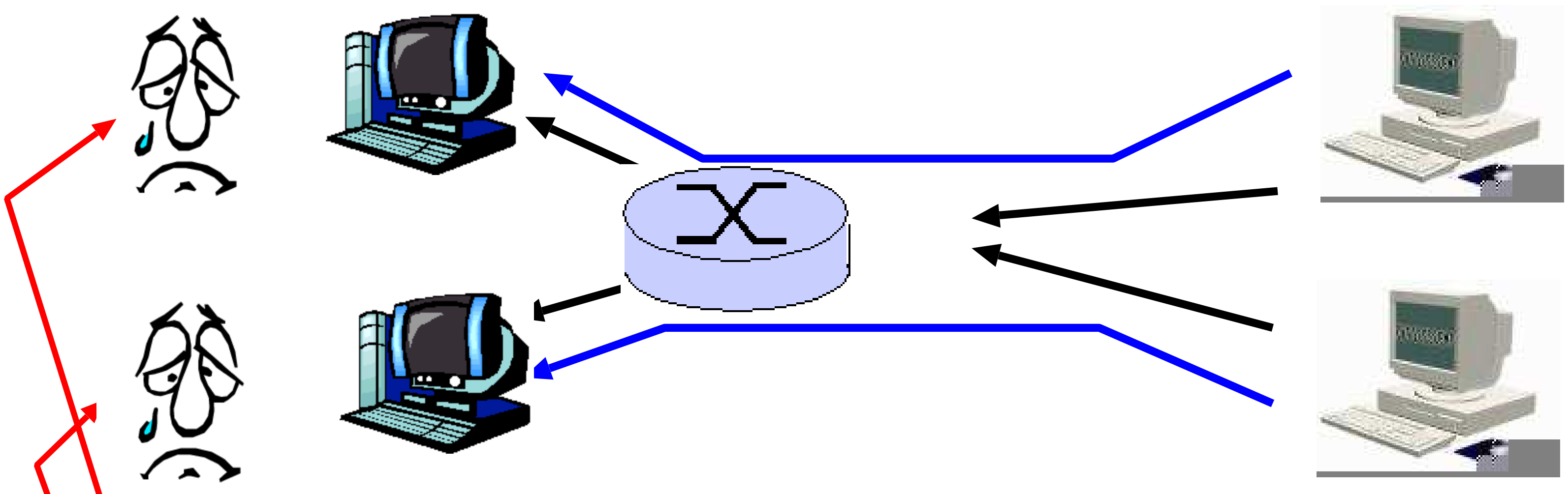


estado $N\sigma$	Qualidade (MOS) 128	64 Kbps	24 Kbps
2 0 0	2,0	—	—
1 1 0	3,0	2,1	—
0 2 0	—	4,0	—
1 0 1	3,5	—	2,3
0 1 1	—	3,8	3,1
0 0 2	—	—	3,3

Exemplo



estado $N\sigma$	Qualidade (MOS)		
	128	64 Kbps	24 Kbps
2 0 0	2,0	—	—
1 1 0	3,0	2,1	—
0 2 0	—	4,0	—
1 0 1	3,5	—	2,3
0 1 1	—	3,8	3,1
0 0 2	—	—	3,3



estado	Qualidade (MOS)		
$N\sigma$	128	64 Kbps	24 Kbps
2 0 0	2,0	—	—
1 1 0	3,0	2,1	—
0 2 0	—	4,0	—
1 0 1	3,5	—	2,3
0 1 1	—	3,8	3,1
0 0 2	—	—	3,3

estado $N\sigma$	Qualidade (MOS) 128	64 Kbps	24 Kbps
2 0 0	2,0	—	—
1 1 0	3,0	2,1	—
0 2 0	—	4,0	—
1 0 1	3,5	—	2,3
0 1 1	—	3,8	3,1
0 0 2	—	—	3,3

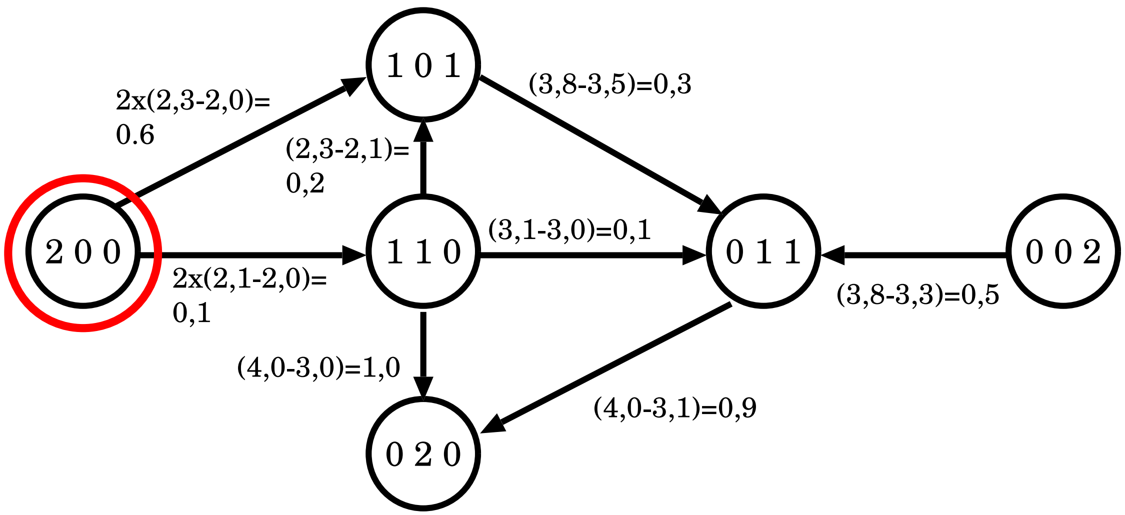
	128 Kbps	64 Kbps	24 Kbps
128 Kbps	2,0; 2,0	3,0; 2,1	3,5; 2,3
64 Kbps	2,1; 3,0	4,0; 4,0	3,8; 3,1
24 Kbps	2,3; 3,5	3,1; 3,8	3,3; 3,3

- **Como representar o processo dinâmico?**
- **Melhor resposta por ser difícil (como saber?)**
- **Idéia**
  - Mudar para estratégia que oferece algum ganho
  - mudança proporcional ao ganho

Diagrama de Transição

- **Processo dinâmico:**
  - taxa de transição (indivíduous/tempo) proporcional: differença de ganho

	128 Kbps	64 Kbps	24 Kbps
128 Kbps	2,0; 2,0	3,0; 2,1	3,5; 2,3
64 Kbps	2,1; 3,0	4,0; 4,0	3,8; 3,1
24 Kbps	2,3; 3,5	3,1; 3,8	3,3; 3,3

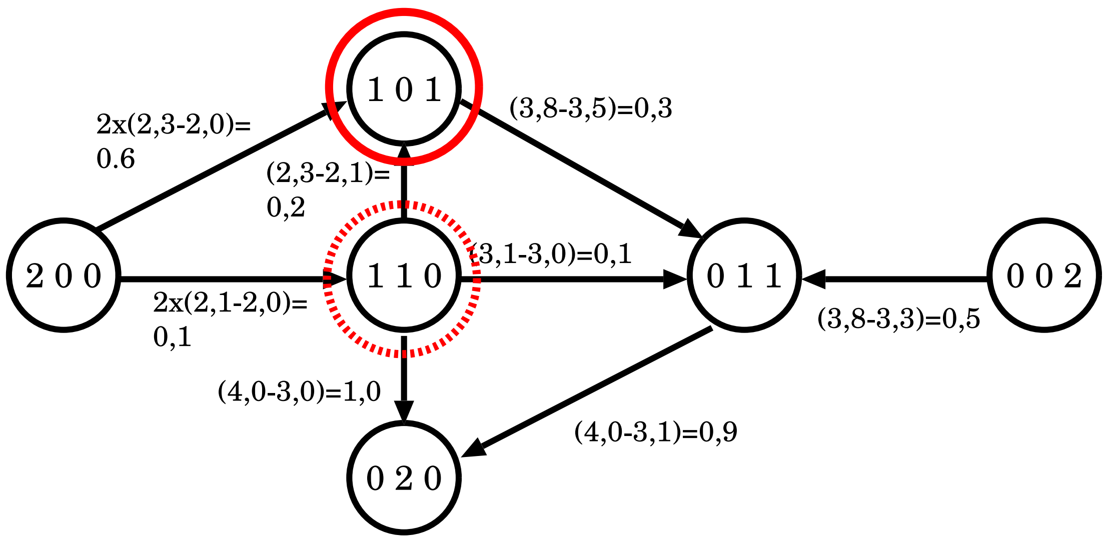


- **Construir diagrama de transição**

Processo dinâmico:

- taxa de transição (indivíduous/tempo) proporcional: differença de ganho

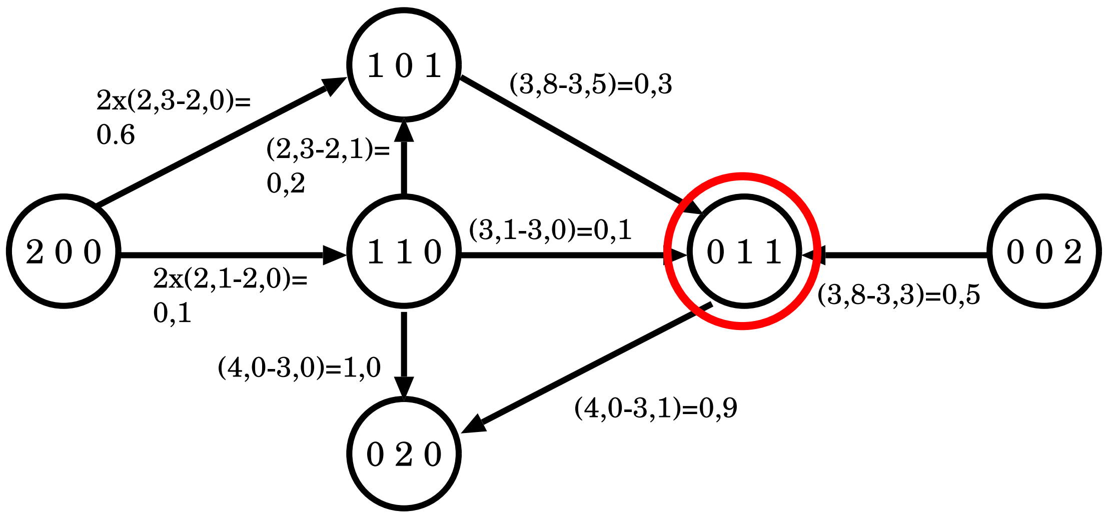
	128 Kbps	64 Kbps	24 Kbps
128 Kbps	2,0; 2,0	3,0; 2,1	3,5; 2,3
64 Kbps	2,1; 3,0	4,0; 4,0	3,8; 3,1
24 Kbps	2,3; 3,5	3,1; 3,8	3,3; 3,3



Processo dinâmico:

- taxa de transição (indivíduous/tempo) proporcional: differença de ganho

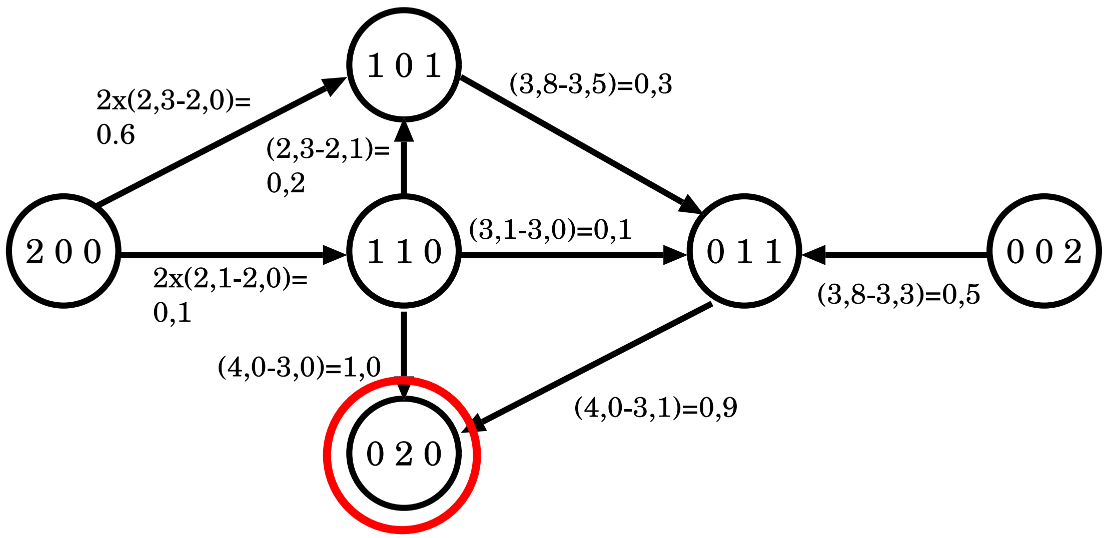
	128 Kbps	64 Kbps	24 Kbps
128 Kbps	2,0; 2,0	3,0; 2,1	3,5; 2,3
64 Kbps	2,1; 3,0	4,0; 4,0	3,8; 3,1
24 Kbps	2,3; 3,5	3,1; 3,8	3,3; 3,3



Processo dinâmico:

- taxa de transição (indivíduous/tempo) proporcional: differença de ganho

	128 Kbps	64 Kbps	24 Kbps
128 Kbps	2,0; 2,0	3,0; 2,1	3,5; 2,3
64 Kbps	2,1; 3,0	4,0; 4,0	3,8; 3,1
24 Kbps	2,3; 3,5	3,1; 3,8	3,3; 3,3



- **Assumir milhares de jogadores (infinito)**

- **Estado: fração de jogadores que adotam cada uma das estratégias no instante t:** $<\sigma_1, \sigma_2, \ldots, \sigma_M>$

- **Tempo contínuo (jogo está sendo jogado continuamente)**
  - Qual é a recompensa de um jogador que adote a estratégia s ?

- **recompensa de um indivíduo que adota a estratégia i:** $u_i(\sigma)$

- **Supor: em Δt indivíduo um indivíduo aprende sobre a recompensa de outro indivíduo com probabilidade $\lambda\Delta t$**

Escolher um jogador para jogar aleatoriamente:

$$N_i(t+\Delta t) = N_i(t) + \lambda\Delta t N_i(t)\left(u_i(\sigma(t)) - \sum_{j=1}^{M} u_j(\sigma(t))\sigma_j(t).\right)$$

Aumento (ou diminuição) da populção proporcioanal a diferença de recompensa

Dividindo por N(t)

$$\frac{N_i(t+\Delta t)}{N(t)} = \frac{N_i(t)}{N(t)} + \frac{\lambda \Delta t N_i(t)}{N(t)} \left(u_i(\sigma(t)) - \sum_{j=1}^{M} u_j(\sigma(t))\sigma_j(t). \right)$$

$$\sigma_i'(t) = \lambda \sigma_i(t) \left(u_i(\sigma(t)) - \overline{u}(\sigma(t)) \right),$$

DINÂMICA DO REPLICADOR

$$\sigma_i'(t) = \lambda\sigma_i(t)\left(u_i(\sigma(t)) - \overline{u}(\sigma(t))\right),$$

- **Estratégias com recompensas menor ou maior que média**
  - diminuem ou aumentam fração da população
- **Indivíduos adotam estratégias que possuem recompensas mais altas**
- **Sistema dinâmico é determinístico (equações diferenciais)**
- **Estratégia não possui variação se recompensa é igual a média**
  - Estudos na área médica
- **Equilíbrio**: $\sigma_i' = 0$ **para todo i**

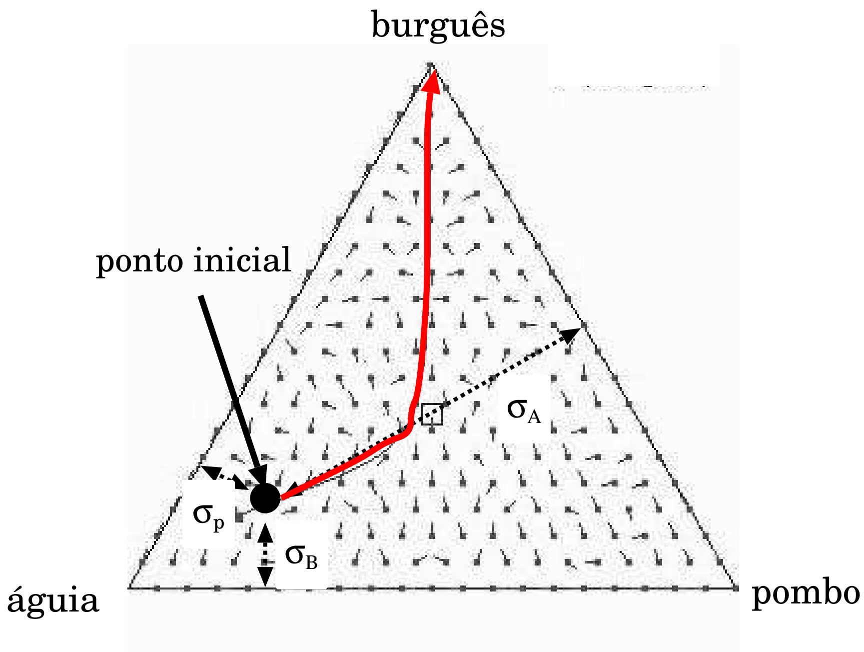
Jogo da águia-pombo-burguês

	Águia	Pombo	Burguês
Águia	$\frac{v-c}{2}, \frac{v-c}{2}$	$v, 0$	$\frac{3v}{4} - \frac{c}{4}, \frac{v-c}{4}$
Pombo	$0, v$	$\frac{v}{2}, \frac{v}{2}$	$\frac{v}{4}, \frac{3v}{4}$
Burguês	$\frac{v-c}{4}, \frac{3v}{4} - \frac{c}{4}$	$\frac{3v}{4}, \frac{v}{4}$	$\frac{v}{2}, \frac{v}{2}$

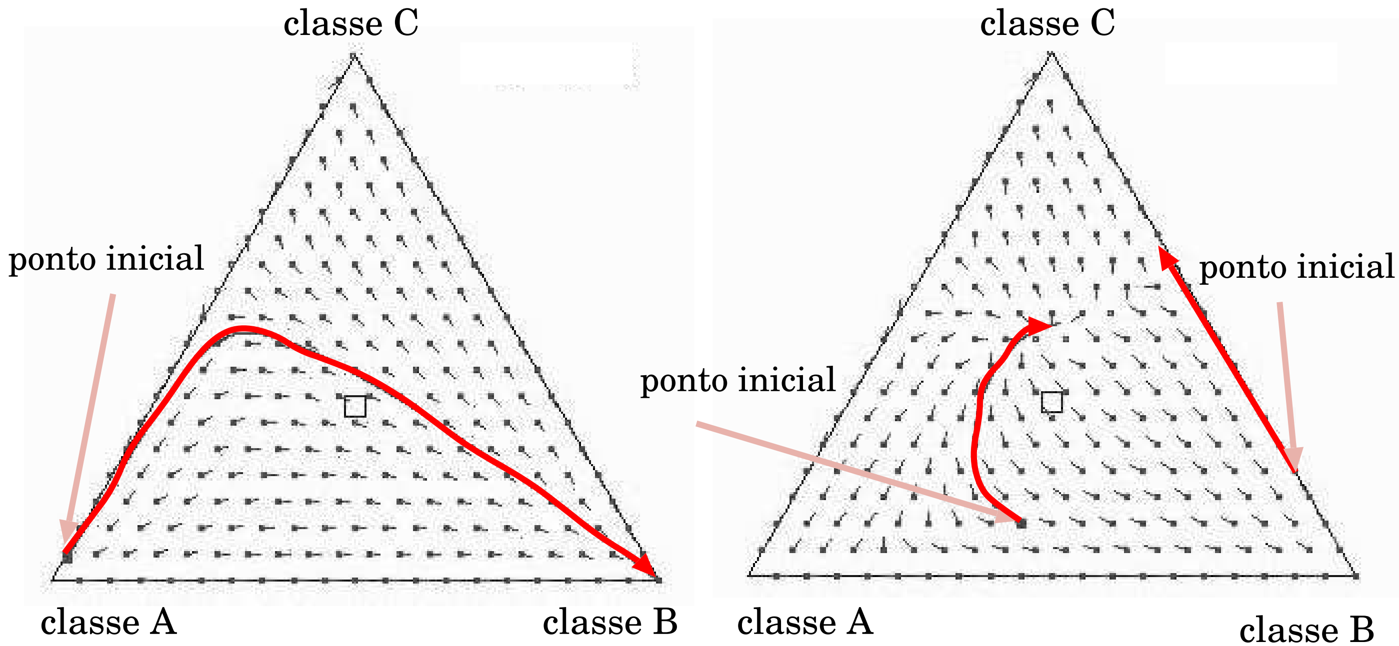
Qual é o equilíbrio?

Exemplo

	Águia	Pombo	Burguês
Águia	$\frac{v-c}{2}, \frac{v-c}{2}$	$v, 0$	$\frac{3v}{4} - \frac{c}{4}, \frac{v-c}{4}$
Pombo	$0, v$	$\frac{v}{2}, \frac{v}{2}$	$\frac{v}{4}, \frac{3v}{4}$
Burguês	$\frac{v-c}{4}, \frac{3v}{4} - \frac{c}{4}$	$\frac{3v}{4}, \frac{v}{4}$	$\frac{v}{2}, \frac{v}{2}$



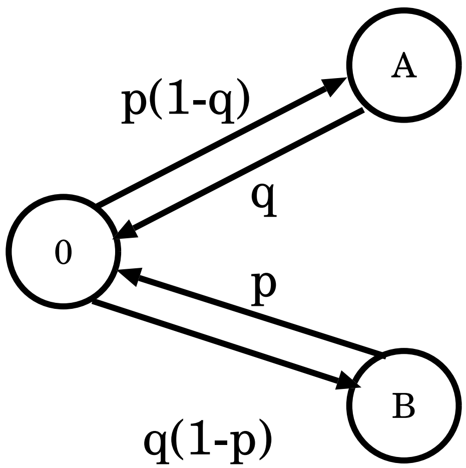
- **Estratégia evolucionariamente estável:**
  Informalmente, uma população adotando estratégias segundo uma distribuição σ é ESS se ela não é vulnerável a invasões por indivíduos (perturbações) que conseqüentemente irão alterar a distribuição da população pelas estratégia.



- **Se σ^* é um equil evolucionário então é um equil de Nash**

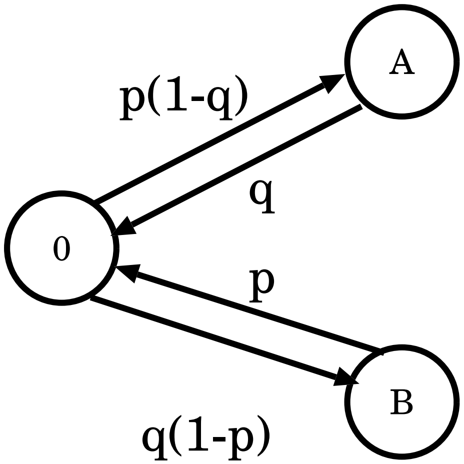
Exemplo: Rede sem Fio

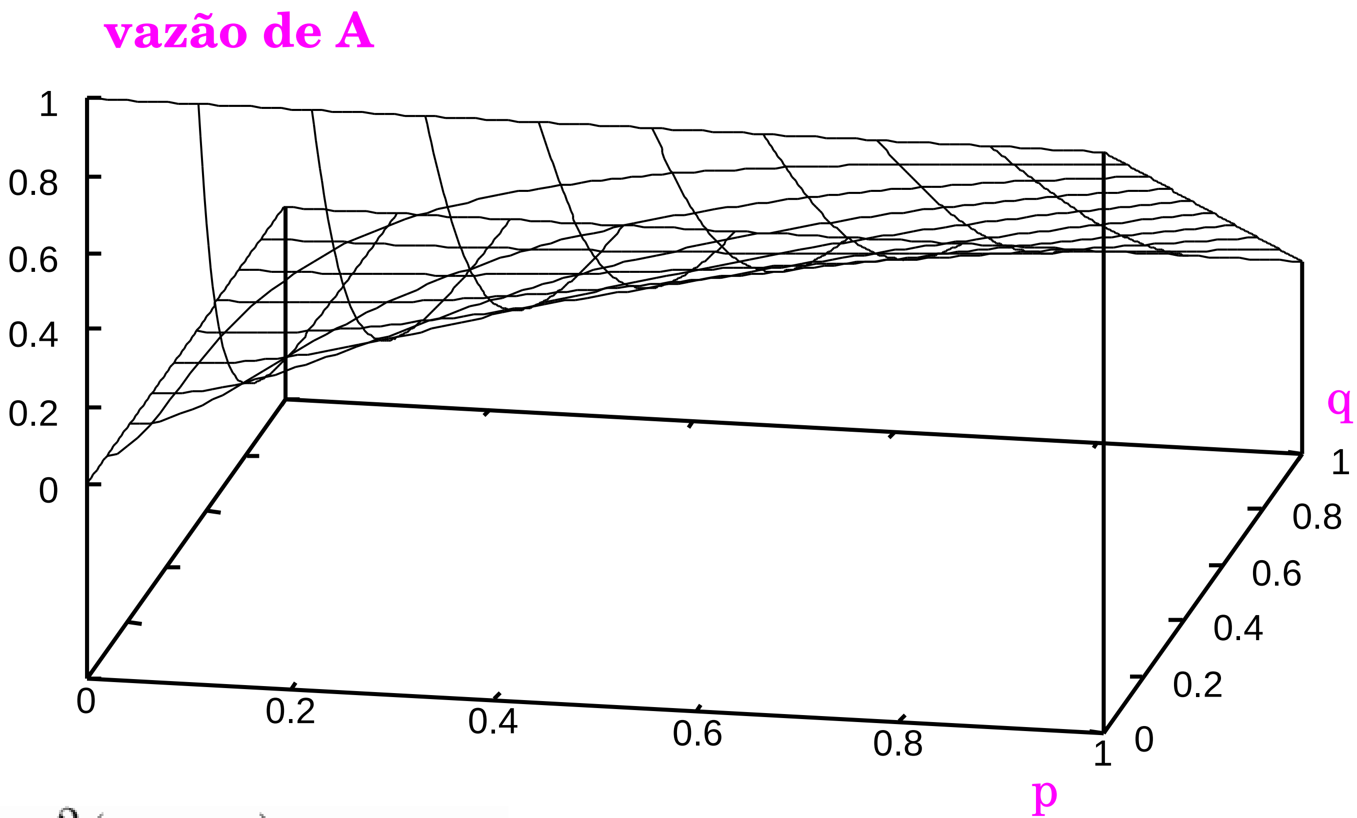
- **Já vimos esse problema... Vamos simplificar...**
- **Slotted Aloha**



- Colisão
- Estação A fica transmitindo até que B transmita
- Estação B fica transmitindo até que A transmita

Vazão: pode-se facilmente calcular (se você usar o Tangram-II!!!!)





$$v_A(p,q) = \frac{p^2(1-q)}{p^2+q^2+pq-p^2q-pq^2}$$

- **Se p=q --> vazão máxima = 1/3 (para que valor???)**
- **O que acontece???**

sistema MUITO injusto...

Por que???

Exemplo: Rede sem Fio

- Se p=q --> vazão máxima = 1/3 (para que valor --> p=q-->0
- sistema MUITO injusto...

A estação que ganha o canal permanece com ele...

⇩

VAMOS COLOCAR UM LIMITE NA VAZÃO DE CADA ESTAÇÃO

$$v_A, v_B \leq C$$

- **Como modelar a competição???**

JOGO DE STACKELBERG

Um jogo de Stackelberg é um jogo extensivo de dois rivais e com informação perfeita onde o líder escolhe uma estratégia do seu conjunto e o seguidor, informado sobre a escolha do líder, escolhe em seguida a sua, do seu conjunto de opções.

Exemplo: limitar vazão a 0,2, líder: E_A

- E_A escolhe p=0,1 -> E_B escolhe valor que maximiza a sua vazão, dada a restrição -> q=0,0666; E_A escolhe ...
  Ambas alcançam a vazão máxima permitida

- **E se a vazão máxima for 0,5? --> neste caso é preferível ser o líder... mas mesmo assim não ocupa 100% da banda**
  PREÇO DA ANARQUIA!!!

Exemplo: Rede sem Fio

- **Como E_A sabe que E_B irá maximizar a sua vazão depois da sua escolha...**

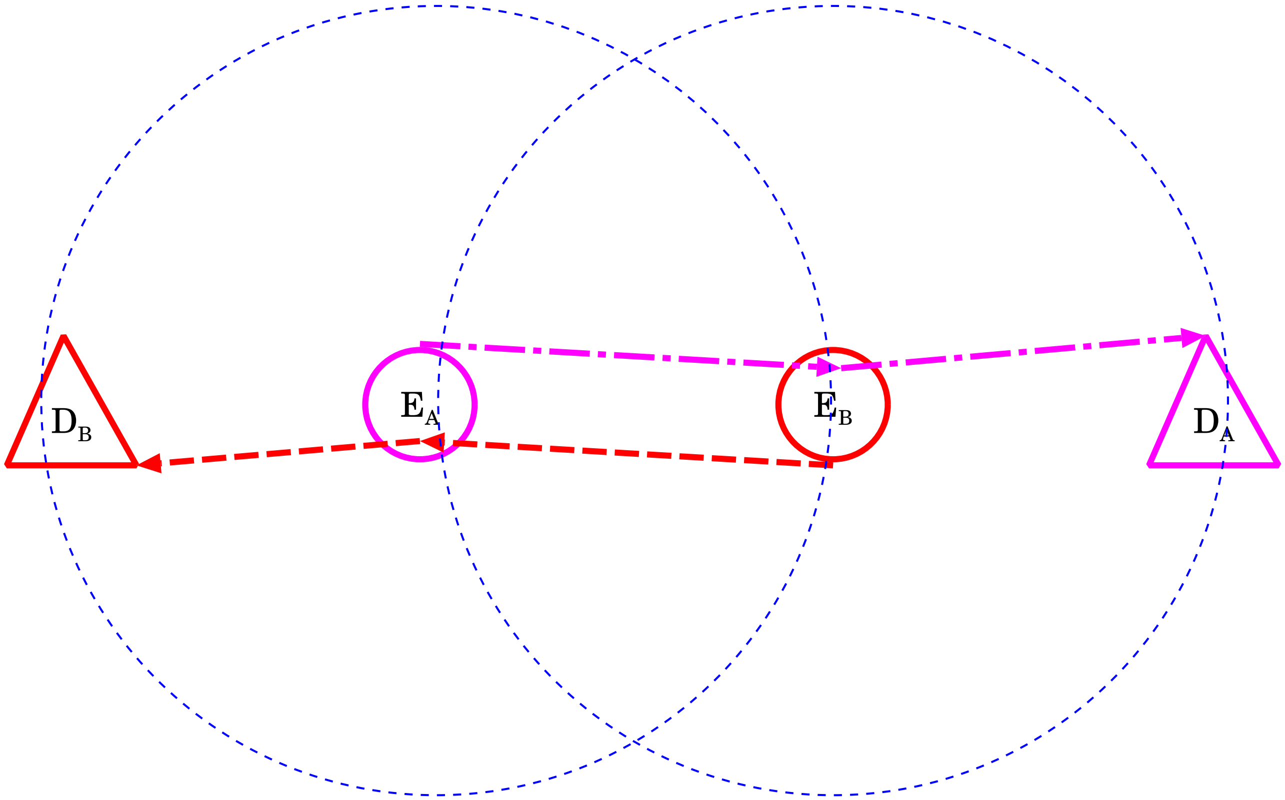
- **Problema do seguidir E_A: o seguidor conhece a estratégia do líder, isto é, o parâmetro p escolhido por ele.**
  Então apenas tenta otimizar a sua vazão dado p:

$$\text{Escolha } q^*(p) \text{ tal que } q^*(p) = \arg\max \{v_B(p,q)\}$$
$$\text{Sujeito a } v_B(p, q^*(p)) \leq C$$

- **Problema do líder E_A: o líder sabe que o seguidor tentará maximizar a sua escolha, e portanto ele deve de antemão escolher o valor do seu parâmetro da forma:**

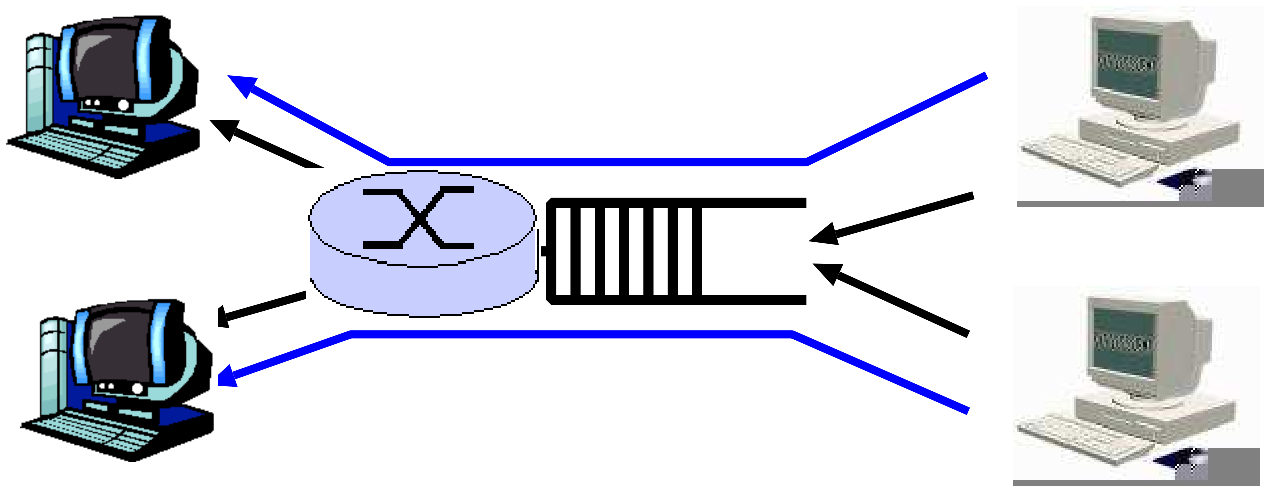
$$\text{Escolha } p^* \text{ tal que } p^* = \arg\max \{v_A(p, q^*(p))\}$$
$$\text{Sujeito a } v_A(p^*, q^*(p^*)) \leq C$$

Transmitir gasta energia...



Usuário ajusta taxa de transmissão -> otimizar desempoenho

- MÉTRICA: POWER (vazão/delay)

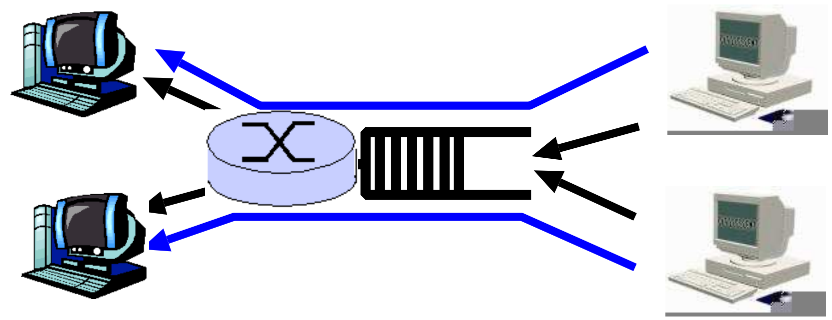


$$\mathbf{u}_i(\lambda_1, \lambda_2) = \frac{\lambda_i}{s}(1 - (\lambda_1 + \lambda_2)s) \qquad \text{para} \quad i = 1, 2$$

- Suponha estratégias: taxas 0,1 e 0,499

restrição:

$(\lambda_1 + \lambda_2)s <$



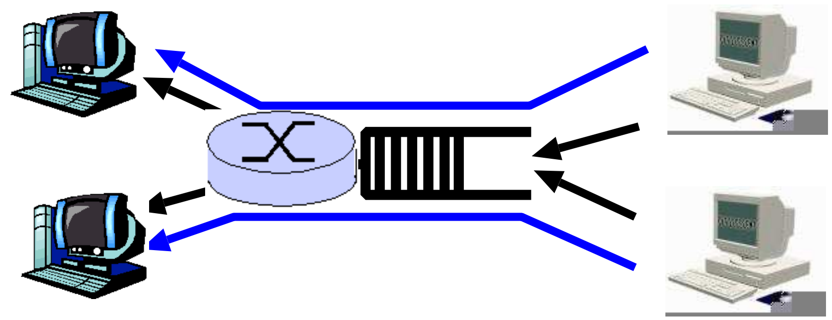
		usuário 1	
		0,1	0,499
usuário 2	0,1	0,08; 0,08	0,04; 0,2
	0,499	0,2; 0,04	10^{-4};10^{-4}

- Como modelar???

- **Jogador 1 escolhe estratégia, então jog 2, ...**

restrição:

$(\lambda_1 + \lambda_2)s \leq$

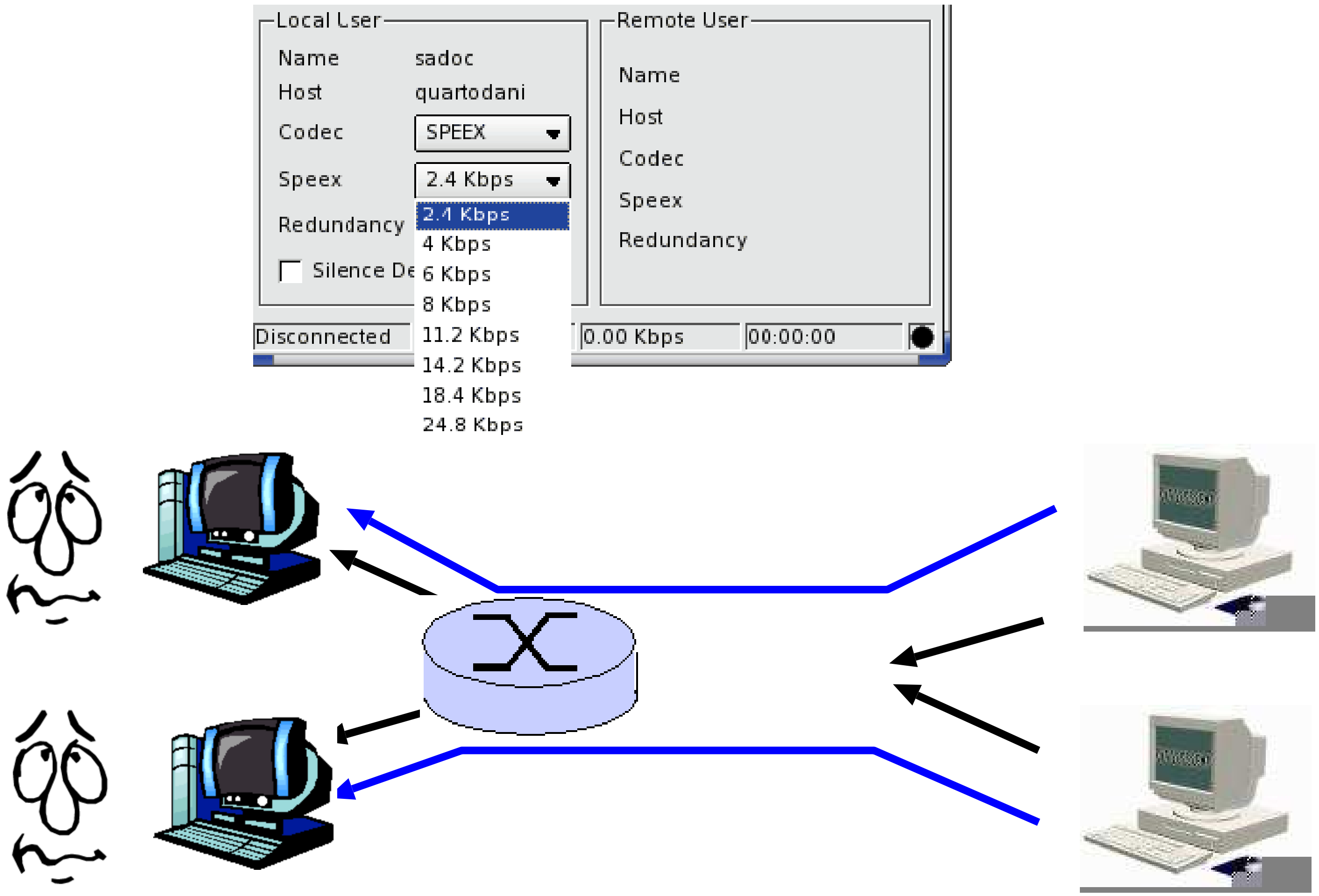


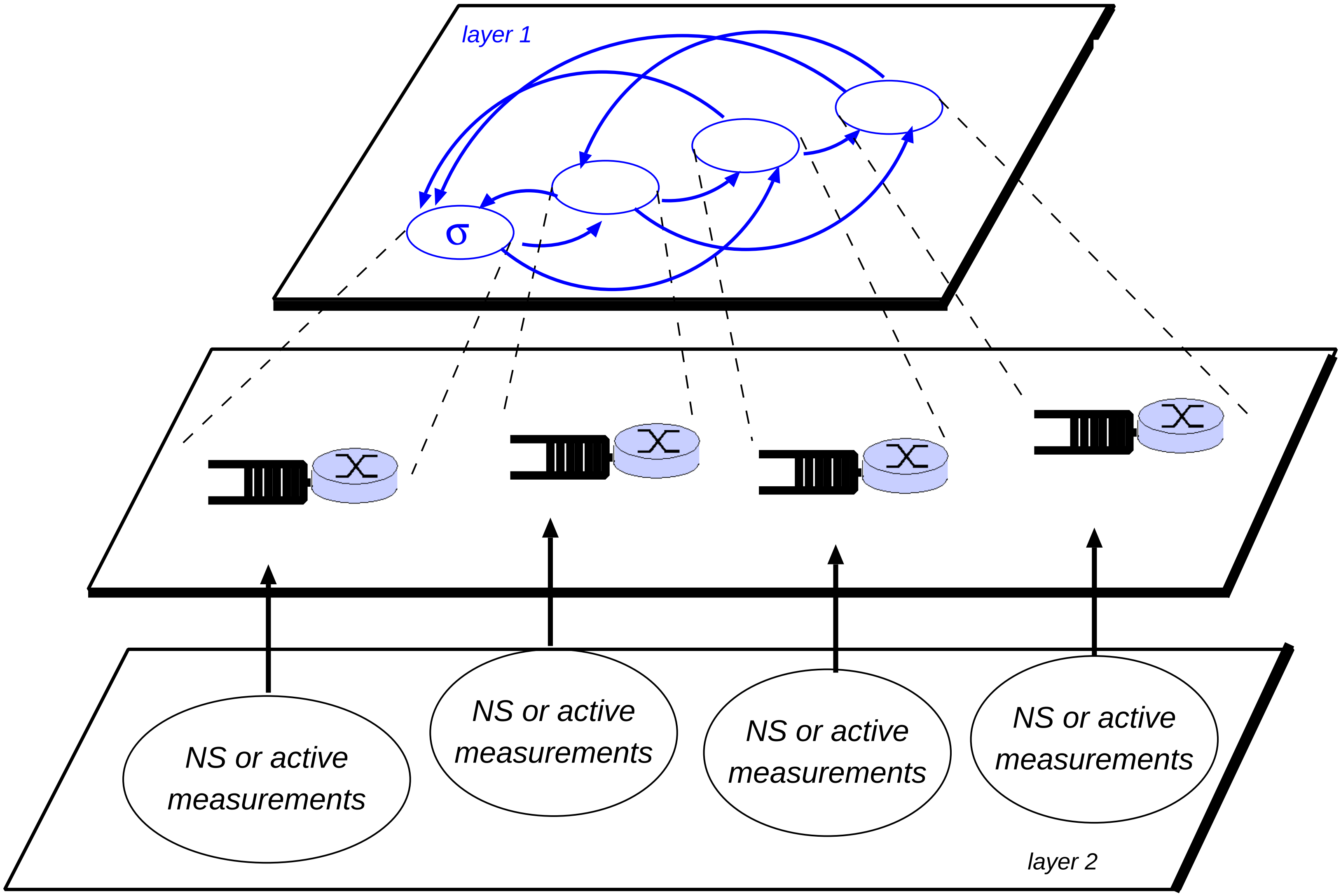
usuário 1

usuário 2		0,1	0,499
	0,1	0,08; 0,08	0,04; 0,2
	0,499	0,2; 0,04	$10^{-4};10^{-4}$

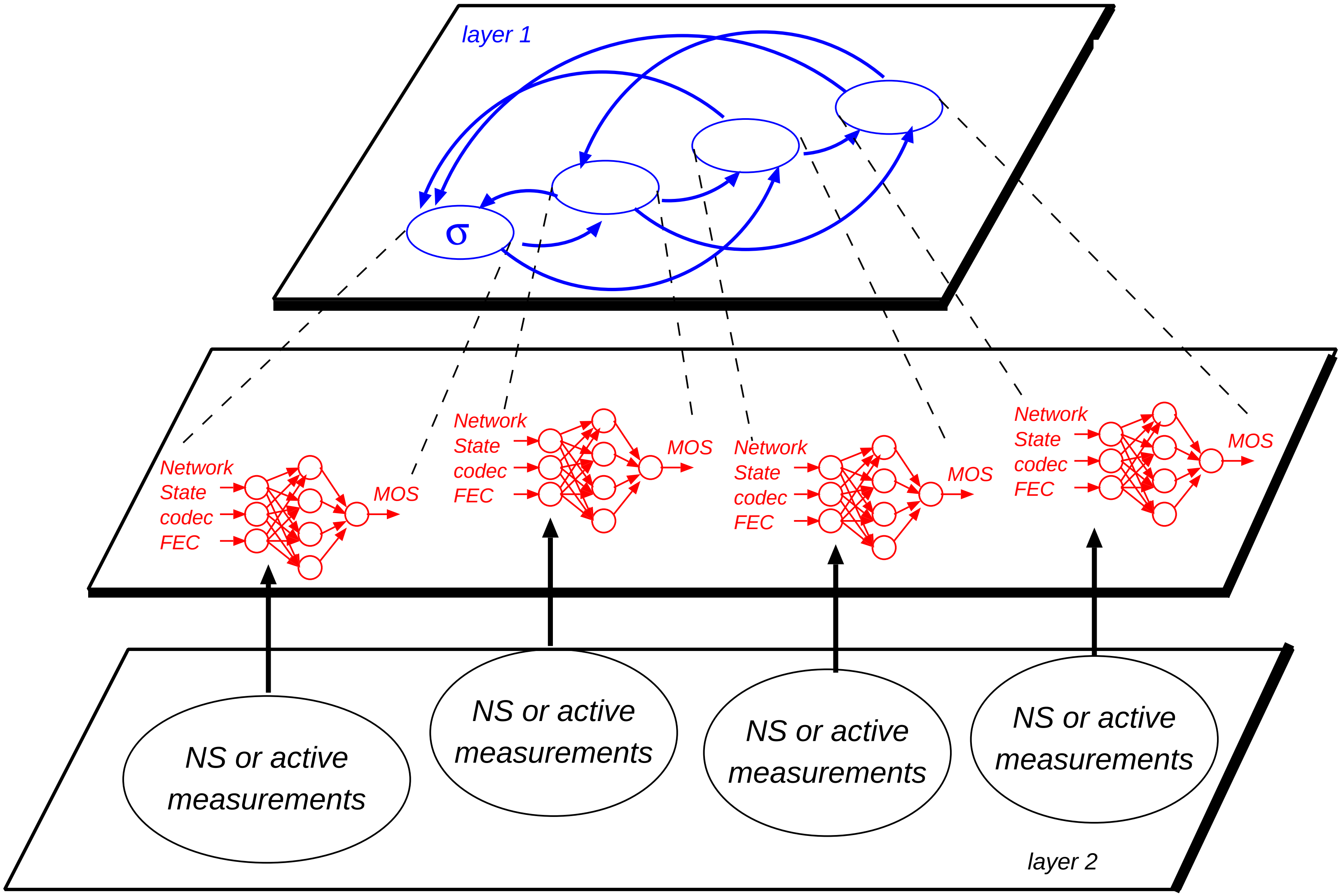
- **Convergência:** $\lambda_1 = \lambda_2 = 1/3$
recompença total: 1/9*2 = 0,22 ---
mas cooperação --- recompença 0,25 (taxas = 0,25)

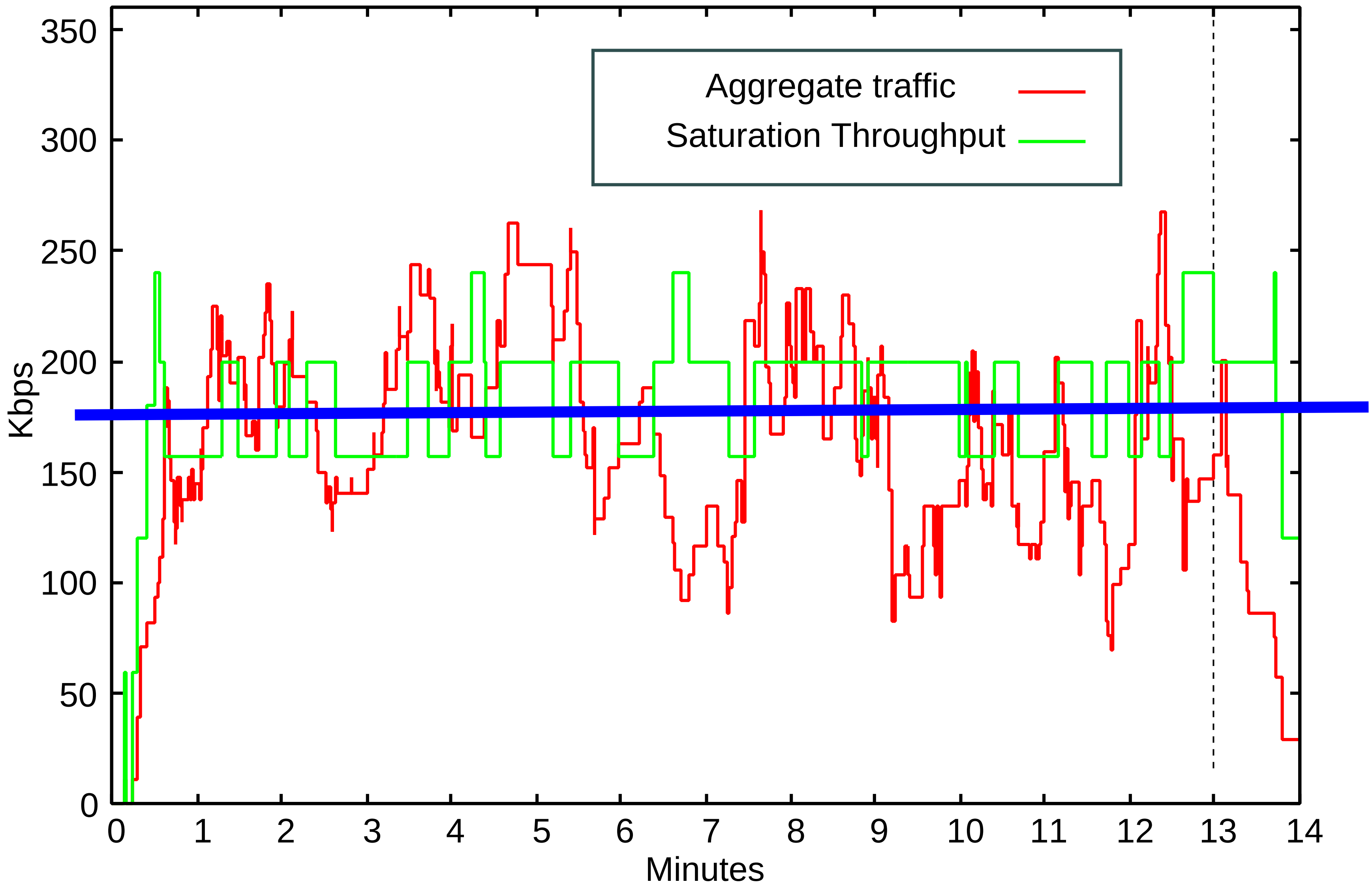
Problemas do VivaVoz




layer 1
σ
NS or active measurements
NS or active measurements
NS or active measurements
NS or active measurements
layer 2

Teoria dos Jogos Evolucionária




Aggregate traffic
Saturation Throughput
Kbps
0
50
100
150
200
250
300
350
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Minutes