

# Uma breve introdução à Teoria de Jogos com aplicações a Redes de Computadores

Edmundo de Souza e Silva, Daniel Ratton Figueiredo

## **Abstract**

*Resource sharing is one of the most important and challenging problems in distributed systems such as the Internet. Traditionally, this problem is dealt with by the use of well defined distributed protocols that are enforced and adopted by all users. In this scenario, such protocols are designed to offer good performance to all end-users and to guarantee that the resources are fairly shared. The design of such protocols are based on the development of mathematical models that are able to predict with reasonable accuracy the behavior of the system under investigation. Several techniques are commonly used in the modeling process, such as Queueing Theory and Markov chain based methods.*

*Lately, Internet end-users and applications are becoming more autonomous, and free from the rigid control enforced by the communication protocols. This behavior contrasts with the classical paradigm. Users or applications may devise their own strategy to acquire the system's shared resources. Since each user tries to optimize her own performance metrics independently of others, conflicts of interest arise. In this new scenario, the classical modeling techniques do not apply, since the users do not behave accordingly to any pre-defined rule. As a consequence, Game Theory emerges as a promising option to deal with such scenario and offers a set of rigorous mathematical tools to predict how the conflicts of interest are solved and the resulting system performance.*

*The objective of this text is to present the fundamentals of Game Theory and to illustrate how the theory can be applied to Computer Networks. Using simple examples, we study different scenarios that lead to conflicts of interest among system users. We show how Game Theory can be utilized to analyze each example and to predict how conflicts can be solved. We present important concepts and basic results of the theory as well as its limitations. Examples in the areas of routing, congestion control and wireless communication are discussed in order to highlight how conflicts among users emerge and how Game Theory can be applied to each problem.*

## **Resumo**

*Um dos problemas mais importantes de sistemas distribuídos, como a Internet, é o compartilhamento eficiente dos recursos disponíveis entre os diferentes usuários do sistema. Tradicionalmente, este problema é resolvido através de protocolos distribuídos bem definidos, projetados sob um consenso comum e adotado fielmente por todos os usuários. Neste cenário, tais protocolos podem*

*ser projetados para oferecer garantias de estabilidade, bom desempenho e equidade no compartilhamento de recursos. Esta análise é normalmente realizada através da construção de modelos matemáticos que possam avaliar e prever o comportamento do sistema. Várias teorias podem ser utilizadas para auxiliar o processo de modelagem, como a teoria de filas e cadeias de Markov.*

*Entretanto, usuários e aplicações de sistemas abertos, como a Internet, vêm se tornando cada vez mais autônomos, independentes de protocolos rígidos, o que quebra o paradigma clássico e faz surgir um novo cenário. Em particular, usuários ou aplicativos podem agir de forma estratégica ao se depararem com um recurso que precisa ser compartilhado, e ao tentarem obter as melhores condições para si próprio acabam gerando um conflito de interesses. Neste cenário, técnicas de modelagem clássica não se aplicam mais, pois usuários ou aplicativos deixam de seguir o modo operacional previsto pelo sistema. Dessa forma, a teoria dos jogos, que oferece um ferramental matemático rigoroso, surge como uma opção promissora para modelar tais interações e tentar prever como tais conflitos de interesse serão resolvidos.*

*O objetivo deste texto é introduzir os fundamentos da teoria dos jogos, e ilustrar como esta pode ser aplicada na área de redes de computadores. Através de exemplos simples, apresentamos diferentes cenários que levam a um conflito de interesses entre os usuários do sistema. Mostramos então como a teoria dos jogos pode ser aplicada para analisar tais situações, e prever como tais conflitos podem ser resolvidos. Desta forma, apresentamos alguns conceitos importantes e resultados básicos da teoria, bem como suas limitações. Exemplos específicos de problemas de roteamento, congestionamento e transmissão em redes sem fio são discutidos de forma a elucidar o surgimento do conflito de interesses entre usuários e a aplicabilidade da teoria dos jogos.*

## **2.1. Introdução**

### **O papel da modelagem**

A maioria das atividades que realizamos requer um planejamento cuidadoso e avaliação do impacto das possíveis opções disponíveis. *Modelos matemáticos* são estruturas abstratas (ex. equações matemáticas) freqüentemente usadas para ajudar a compreender o sistema em estudo, investigar opções, e prever a influência de fatores externos e outros parâmetros do sistema. Por exemplo, modelos econômicos podem ser empregados para nortear políticas governamentais, modelos atmosféricos nos permitem prever os impactos climáticos decorrentes da liberação de CO<sub>2</sub> na atmosfera, modelos biológicos podem servir para entender as relações entre causa e efeito de organismos vivos.

Os sistemas de computação são cada vez mais complexos, e a crescente dependência da sociedade dos computadores e sistemas de comunicação de dados torna essencial o desenvolvimento de equipamentos mais eficientes e mais confiáveis. Em conseqüência, o desenvolvimento destes sistemas deve ser precedido pela elaboração de modelos que nos ajudem a determinar como o sistema deve ser projetado para atender aos requisitos do projeto. Tomemos um exemplo de um sistema computacional genérico onde usuários submetem programas para ser executados. Tal sistema é constituído por uma série de

*recursos*, como a CPU, a memória, e os discos, que são *compartilhados* pelos programas em execução. Um modelo de *desempenho* é necessário para avaliar o impacto da *contenção* pelos recursos, isto é, entender como compartilhar os recursos equitativamente, de forma a otimizar o seu uso e diminuir o tempo de execução dos programas. Um requisito de projeto pode estabelecer uma eficiência mínima para o sistema durante um determinado intervalo de tempo. Este é o caso, por exemplo, de computadores de bordo de aeronaves. Estes sistemas são críticos e não podem falhar durante um voo (ou operar abaixo de um limite de eficiência), pois as conseqüências seriam trágicas. Como garantir, com algum grau de confiança, que o sistema a ser desenvolvido atenderá a esse requisito? Neste caso é necessário desenvolver um modelo de *confiabilidade* que possa simular as possíveis mudanças de estrutura do sistema, decorrentes de falhas de componentes, e o comportamento de técnicas de recuperação responsáveis pela substituição das partes inoperantes, assim como o uso de componentes redundantes. Além disso, modelos de desempenho fornecem a eficiência obtida para diferentes estruturas do sistema. Várias teorias estão disponíveis para auxiliar na construção de modelos de desempenho e/ou confiabilidade, sendo que a teoria de filas e cadeias de Markov são a base de muitos dos modelos de contenção de recursos e das mudanças estruturais dos sistemas.

A modelagem é usada, além da fase de projeto, durante a operação do sistema. Neste caso pode-se descobrir possíveis gargalos no uso de recursos decorrentes de variações da carga de trabalho, e avaliar novas técnicas para melhorar o desempenho.

A Internet é um excelente exemplo de um sistema extremamente complexo onde a modelagem matemática foi fundamental para guiar a adoção de novas técnicas de transmissão e avaliar protocolos que foram propostos e eventualmente adotados ao longo do tempo. Anteriormente ao seu nascimento, o mecanismo básico de transmissão de informação na Internet (*comutação por pacotes - packet switching*) teve a sua eficiência demonstrada pelo uso de modelos de filas. Depois de ser proposto e adotado em ampla escala, o protocolo para controle de transmissão de dados (TCP) foi bastante estudado e analisado utilizando modelos matemáticos. Tais estudos levaram a um melhor entendimento do seu desempenho e geraram sugestões de modificações que o tornasse mais eficiente. Muitas outras técnicas e protocolos existentes na Internet seguiram a mesma sistemática: modelos matemáticos foram a base para avaliar o seu desempenho. Modelos de protocolos de múltiplo acesso empregados em redes Ethernet e por rádio (e.g. *WiFi*), técnicas de escalonamento de pacotes em roteadores, técnicas de distribuição de tráfego, e algoritmos de planejamento de capacidade são apenas alguns exemplos dentre os inúmeros existentes.

### **Solução de conflitos pela imposição de regras rígidas**

O problema de contenção de recursos na Internet é resolvido pela adoção de protocolos distribuídos bem definidos impostos aos usuários. Para ilustrar esse conceito, consideremos o protocolo de transporte TCP, e um usuário, o Alberto, que acessa um sítio na Web para “baixar” um arquivo (transferir um arquivo do servidor Web para o computador local). Este usuário não consegue que o arquivo seja transmitido pelo servidor na taxa que ele bem entender. Fundamentalmente, a taxa de transmissão está limitada pela capacidade de transmissão dos canais de comunicação que ligam o servidor Web ao Alberto, e pelo controle exercido para evitar transbordamento dos *buffers* de dados entre os protocolos de transporte e aplicação (controle de fluxo). O protocolo TCP controla a taxa de transmissão de dados de forma a evitar que *congestionamentos* se formem nos canais de transmissão que ligam o servidor Web ao Alberto (controle de congestionamento). De maneira simplificada, este controle do TCP opera da seguinte forma: a taxa de transmissão é aumentada até que um congestionamento na rede seja detetado. Quando isto ocorre, a taxa de transmissão de dados é drasticamente reduzida de forma a tentar evitar as conseqüências prejudiciais de um congestionamento. Todos sabemos o que um congestionamento de tráfego de carros pode causar: imensos atrasos para ir de um ponto a outro. As conseqüências de um congestionamento de pacotes de dados na Internet não são muito diferentes das de congestionamento nas vias de uma cidade. Caso as regras rígidas do protocolo TCP não fossem seguidas por todos os usuários da rede, um congestionamento levaria a um enorme aumento do tempo total de transmissão dos dados e nosso usuário provavelmente não conseguiria baixar o seu arquivo.

Considere agora usuários interligados por uma rede sem fio, semelhante à que temos em nossas casas ou em nossos ambientes de trabalho. O canal de rádio que dá acesso a rede sem fio é compartilhado por todos os computadores ligados a esta rede. Quando o computador de um usuário transmite, fazendo uso do canal, outros usuários não podem transmitir, pois caso contrário haveria uma “mistura” dos dados enviados (*colisão*), e conseqüentemente toda a informação seria perdida. Ora, o problema em questão é como fazer com que esses usuários fisicamente distantes um do outro, cujo único meio para se comunicar é o próprio canal de rádio, consigam, de forma ordenada, fazer uso deste, minimizando as colisões decorrentes de acessos simultâneos. Protocolos de controle de acesso ao meio, como os da família 802.11, impõem regras rígidas de acesso, visando assim garantir um compartilhamento do canal de forma equilibrada e eficiente.

O mesmo ocorre com o roteamento de pacotes na Internet. Os algoritmos distribuídos de roteamento implementados nos roteadores forçam regras de escolha de caminhos pelos pacotes de informação, de forma a minorar congestionamentos, otimizar o tempo de transferência e o uso equilibrado dos canais de comunicação.

Para tornar mais claro o conceito descrito nos exemplos acima, retorne-

mos ao exemplo do nosso amigo Alberto, que deseja baixar seu arquivo de música de um sítio qualquer na Internet. Simplificaremos bastante o problema, mas mantendo sua essência. Considere que Alberto faz uso de um canal de transmissão compartilhado por outros usuários que também desejam baixar músicas do mesmo sítio. Suponha que, quando apenas Alberto acessa o sítio, os dados podem ser transmitidos a uma taxa próxima à capacidade do canal de transmissão compartilhado. Nosso amigo começa a transmissão sozinho. Rapidamente alcança a taxa máxima e fica bastante satisfeito. Entretanto, um segundo usuário, a Barbara, começa a baixar o seu arquivo. Evidentemente, ambos não podem obter a taxa máxima de transmissão, pois neste caso a capacidade do canal compartilhado seria ultrapassada. À medida que o TCP permite o aumento da taxa com que Barbara recebe os dados, um congestionamento fatalmente ocorre. O protocolo TCP, que regula as taxas de transmissão de Alberto e Barbara, reage sem demora e reduz drasticamente a taxa de ambos para limites tais que acabam com o congestionamento. Um modelo matemático simples pode nos fornecer a taxa com que cada usuário vai receber seus dados, e ainda mostrar que o canal será compartilhado de forma justa entre ambos. Alberto porém, não satisfeito com a taxa de recebimento dos seus dados, embora ciente que esta é a melhor e mais justa neste cenário, resolve estudar Redes de Computadores, e aprende que pode alterar a sua aplicação de forma a abrir duas conexões TCP para baixar uma mesma música. Ora, não há como diferenciar facilmente entre as duas conexões de um mesmo usuário e a terceira do outro. O espertinho Alberto acaba de contornar as regras de uso do TCP e irá terminar de baixar sua música praticamente na metade do tempo de Barbara.

### **Conflitos onde não há regras**

O controle rígido das aplicações tem garantido o bom funcionamento da Internet através dos anos. Entretanto, os usuários vêm se tornando cada vez mais autônomos e independentes dos protocolos da rede. Algumas aplicações de *streaming*, por exemplo, fazem uso do protocolo UDP, que não impõe nenhum controle na taxa de transmissão. Outras aplicações procuram determinar o próprio roteamento dos dados gerados, contornando o roteamento realizado pela rede. De uma forma geral, os usuários dessas aplicações procuram usar os recursos de rede de forma egoísta, buscando o melhor para si, sem se importar com o que ocorre com outras aplicações. Este cenário é bastante diferente do anterior, pois as regras rígidas da rede deixam de ser obedecidas.

Retornemos ao exemplo do nosso amigo Alberto que, com as modificações realizadas na sua aplicação, está conseguindo uma taxa de transmissão cujo valor é praticamente o dobro da de Barbara. Barbara, que não é boba, rapidamente aprende o truque e também modifica a sua aplicação para abrir mais conexões TCP. Ora, é claro o *conflito de interesse* que surge entre nossos dois usuários. O pior é que essa disputa entre Alberto e Barbara pode levar a uma situação em que ambos acabem obtendo uma taxa de transmis-

são efetiva inferior àquela que obteriam se a aplicação de cada um utilizasse normalmente uma única conexão TCP! (Um modelo simples de desempenho pode ser usado para comparar os cenários com uma ou várias conexões TCP abertas por usuário.) Será que Alberto e Barbara acabam percebendo que é mais vantajoso cooperar? O que acontece se ambos forem egocêntricos e apenas olharem para seus ganhos individuais? Modelos tradicionais de desempenho conseguem prever a taxa obtida, dada uma determinada situação (um usuário abrindo  $x$  conexões TCP e outro  $y$ , por exemplo), mas não são apropriados para lidar com o tipo de situação descrita. É neste contexto que a Teoria dos Jogos surge como promissor ferramental matemático para ajudar a entender como os conflitos de interesse que surgem em ambientes onde há recursos compartilhados podem ser resolvidos. A Teoria tem sido largamente aplicada em áreas como economia, filosofia, biologia e, mais recentemente, na computação.

É importante enfatizar as diferenças entre os resultados que modelos tradicionais de desempenho e confiabilidade podem fornecer e aqueles da Teoria de Jogos. Modelos de desempenho focam nos recursos (CPU, discos, canal de comunicação, etc.) compartilhados entre usuários (pessoas, programas, protocolos, pacotes de informação, etc.), e fornecem métricas tipo vazão (*throughput*), utilização de um recurso, tempo médio de espera, distribuição do número de usuários esperando para utilizar os recursos, entre outros. A carga de trabalho gerada por diferentes tipos de usuários a cada recurso do sistema, assim como o tempo de uso destes recursos, são variáveis aleatórias com distribuição conhecida ou estimada. Uma suposição básica destes modelos é que a maneira como os recursos são utilizados pelos usuários é bem estabelecida. Analogamente, os modelos de confiabilidade concentram-se nas mudanças estruturais de um sistema que ocorrem devido a falhas de seus componentes e as métricas a serem obtidas são usualmente: a probabilidade do sistema falhar em um dado intervalo - *confiabilidade*, e a razão entre o tempo em que o sistema está funcional e o tempo total de observação - *disponibilidade*. Neste caso o comportamento de falha dos componentes (ex. o tempo até uma falha, o tempo de reparo) é bem determinado por variáveis aleatórias, com distribuições estimadas, assim como as regras de substituição de componentes ou reparo dos mesmos. Existem ainda os modelos chamados de *performabilidade*, que visam calcular o desempenho de sistemas cujas estruturas são alteradas com o tempo.

Entretanto, como exemplificado acima, nem sempre o usuário está sujeito a regras de utilização dos recursos, e seu comportamento é determinado pelo ganho que ele pode obter ao usar uma determinada regra, de acordo com as opções disponíveis. O ganho, por exemplo, pode ser uma métrica de desempenho, como o tempo de recebimento de dados. As estratégias podem ser a forma pela qual um usuário tentará obter os recursos compartilhados. Em suma, o usuário pode agir de forma estratégica, deixando de ser um agente passivo do sistema. E se todos os usuários agem de forma estratégica, temos

um conflito de interesse entre eles. Uma questão importante é como solucionar este conflito. Entende-se por solução as estratégias a serem usadas por cada usuário e/ou a forma pela qual o sistema evolui com o tempo. No caso da disputa entre Alberto e Barbara discutida acima, será que ambos abririam mais e mais conexões TCP até o limite do sistema? Tal situação poderá levar ao colapso da rede, onde nenhum dos dois consegue baixar seus respectivos arquivos. Ou será que naturalmente eles acabariam por usar apenas uma conexão TCP?

Não é difícil imaginar inúmeros outros cenários de conflitos semelhantes aos de Alberto e Barbara. O leitor provavelmente já fez uso de alguma ferramenta de voz sobre IP, e teve a oportunidade de perceber que a qualidade da conversa pode variar muito. A qualidade da voz é inversamente proporcional à fração de pacotes descartados pela rede devido, em grande parte, à ocorrência de congestionamentos. É possível implementar técnicas de redundância para recuperar pacotes descartados, e com isso reduzir a queda de qualidade. O uso destas técnicas, entretanto, leva a um aumento da taxa de transmissão, o que pode piorar o congestionamento e a taxa de descarte de pacotes, diminuindo ainda mais a qualidade de voz. Por outro lado, técnicas de compressão de voz reduzem a taxa de transmissão, o que contribui para minorar situações de congestionamento. Mas quanto maior a taxa de compressão, pior a qualidade de voz. É evidente o compromisso existente, se couber ao usuário a escolha dessas alternativas. Como eles se comportarão?

Tais questões não podem ser respondidas utilizando a modelagem tradicional de desempenho. Para abordá-las é necessário utilizar uma teoria que represente matematicamente agentes estratégicos, como a Teoria dos Jogos.

A Internet é hoje um sistema altamente complexo, com múltiplos tipos de agentes independentes que muitas vezes possuem interesses próprios, tais como provedores de serviços, operadores de redes, usuários, aplicativos, etc. A Teoria dos Jogos tem se mostrado efetiva para modelar o conflito que surge dessas interações e para melhor compreender seus resultados. No restante deste texto iremos introduzir conceitos fundamentais da Teoria dos Jogos utilizando alguns exemplos simples, e então focar em situações de conflito que surgem na Internet. Discutiremos problemas de roteamento, congestionamento e transmissão em redes sem fio, de forma a elucidar a aplicabilidade da teoria.

## **2.2. Teoria dos Jogos**

A teoria dos jogos oferece um ferramental matemático para modelar e analisar situações onde dois ou mais agentes se confrontam. O objetivo da teoria é analisar tais situações de forma lógica e determinar como os agentes devem agir para resolver o confronto da melhor maneira possível. Uma situação de confronto é modelada através de um "jogo", cuja representação caracteriza de forma precisa tal situação. Um jogo possui minimamente as seguintes características:

- Dois ou mais jogadores.

- Um conjunto de estratégias para cada jogador. Uma estratégia equivale a uma opção de ação para o jogador.
- Uma associação entre as estratégias escolhidas por todos os jogadores e uma recompensa obtida por cada jogador.

Jogadores envolvidos em um jogo não são necessariamente pessoas, pois um jogo pode modelar situações de confronto entre países, animais, empresas, idéias, ou quaisquer outros agentes. As *estratégias* disponíveis para cada jogador correspondem a possíveis planos de ação, são as opções de ação que cada jogador possui. O conjunto de estratégias de um jogo pode variar de um jogador para o outro, e possuir minimamente duas estratégias ou até mesmo ser infinito. A estratégia escolhida por cada jogador determina o *resultado* do jogo. Uma *recompensa* numérica para cada jogador está associada a cada resultado do jogo. Esta recompensa modela o benefício ou a perda que cada jogador terá, caso o jogo termine com o respectivo resultado.

		Barbara	
		1	2
Alberto	1	3, 3	2, 5
	2	5, 2	1, 1

**Figura 2.1. Jogo de conexões TCP**

Para ilustrar um jogo, consideremos a situação de confronto entre dois usuários de uma rede, descrita no capítulo 2.1. Este jogo possui dois jogadores, Alberto e Barbara. Cada jogador possui duas estratégias, que são estabelecer uma ou duas conexões TCP para efetuar a transferência de um arquivo. A ilustração deste jogo está na figura 2.1, que representa a *matriz do jogo*.

Cada linha da matriz do jogo corresponde a uma estratégia disponível para o primeiro jogador (no caso, Alberto), e cada coluna corresponde a uma estratégia disponível para o segundo jogador (no caso, Barbara). Cada célula da matriz representa um possível resultado do jogo. O valor dentro da célula indica a recompensa que cada jogador receberá, caso o jogo termine com o respectivo resultado. O primeiro valor dentro da célula corresponde à recompensa do primeiro jogador, e o segundo valor a recompensa do segundo.

Observe que este jogo possui quatro resultados possíveis: os dois jogadores escolhem estabelecer uma conexão, os dois escolhem duas conexões, Alberto escolhe duas conexões e Barbara uma, e finalmente, Alberto escolhe uma conexão e Barbara duas. A recompensa em nosso exemplo é uma representação numérica da qualidade de serviço que cada jogador irá receber da



rede. Por exemplo, uma representação da vazão (throughput) que o jogador irá receber. Assim sendo, tanto Alberto quanto Barbara desejam obter a melhor qualidade de serviço possível.

Agora que você já conhece os detalhes do jogo, qual estratégia você escolheria se estivesse jogando como Alberto? Estabeleceria uma ou duas conexões para transferir o arquivo? Vamos analisar o jogo cuidadosamente para decidir a estratégia que Alberto deve escolher. Suponhamos que Alberto deseja obter a maior recompensa disponível no jogo (que é 5) e por isto decide estabelecer duas conexões. O problema é que Barbara pode pensar da mesma forma e também estabelecer duas conexões. Neste caso, o resultado do jogo será ruim para ambos, pois os dois receberão a menor recompensa disponível no jogo (que é 1). Sendo assim, Alberto decide estabelecer uma conexão, pois neste caso sua recompensa mínima será dois. Se Barbara pensar da mesma forma e também estabelecer uma conexão, o resultado do jogo dará a cada jogador uma recompensa com valor 3, que está entre a mínima e a máxima. Porém, se Alberto desconfiar que Barbara irá escolher uma conexão, então deve escolher duas, pois assim obterá a maior recompensa do jogo. Mas Barbara também pode pensar da mesma maneira, o que nos levaria novamente ao pior resultado do jogo! O que fazer então? O objetivo da teoria dos jogos é exatamente analisar tais situações e propor soluções para a resolução destes conflitos. Na seção 2.2.6 veremos uma análise mais detalhada deste exemplo e a prescrição da teoria para a solução.

Uma das premissas fundamentais da teoria é que cada jogador deseja obter para si próprio a maior recompensa possível. Porém, como acabamos de ver, o resultado do jogo não é determinado exclusivamente pela escolha de estratégia de apenas um jogador. Por outro lado, cada jogador pode influenciar o resultado do jogo ao escolher sua estratégia.

Uma outra consideração importante é a ordem em que os jogadores irão escolher suas estratégias. No exemplo acima, nossos jogadores escolheram suas estratégias de forma *simultânea*, pois tomaram esta decisão sem saber qual era a do outro jogador. Nos próximos exemplos, iremos adotar esta mesma forma de jogo. Entretanto, na seção 2.2.11 veremos jogos onde os jogadores irão escolher suas estratégias de forma *seqüencial*.

		Barbara	
		1	2
Alberto	1	3, 3	2, 4
	2	4, 2	5, 5

**Figura 2.2. Jogo de conexões TCP**

Vejamos o mesmo exemplo de confronto entre Alberto e Barbara, mas com recompensas diferentes associadas a cada resultado do jogo. Este novo exemplo está ilustrado na Figura 2.2. Qual estratégia você escolheria neste caso? Vamos analisar este novo jogo. O que Alberto deve fazer se Barbara escolher estabelecer uma conexão? Neste caso, Alberto deve estabelecer duas conexões, pois assim receberá uma recompensa maior (4 ao invés de 3). E o que fazer se Barbara escolher duas conexões? Neste caso, Alberto deve novamente escolher duas conexões, pois receberá uma recompensa maior (5 ao invés de 2). Observe que, em ambos os casos, Alberto deve escolher estabelecer duas conexões. Ou seja, independente do que Barbara escolher, será sempre melhor para Alberto escolher duas conexões. Você deve conferir que o mesmo argumento vale para Barbara. Ou seja, independente da escolha de Alberto, será sempre melhor para Barbara escolher duas conexões. Assim sendo, a solução deste jogo prescrita pela teoria é que ambos os jogadores escolham duas conexões.

Uma consideração importante são as recompensas atribuídas a cada jogador em cada possível resultado do jogo. Como vimos nos dois exemplos acima, tais recompensas são fundamentais para determinar como cada jogador deve proceder no jogo. Entretanto, a determinação do valor numérico das recompensas não faz parte da teoria dos jogos. A teoria simplesmente assume que tais valores refletem adequadamente a situação de confronto sendo modelada pelo jogo. Assim sendo, fica a critério da pessoa modelando o confronto determinar os valores específicos para as recompensas que melhor representam cada possível resultado do confronto em questão. Em alguns jogos esses valores são explícitos, como por exemplo no jogo de par-ou-ímpar. Porém, em muitos outros casos, essa tarefa é bem mais delicada e requer uma análise detalhada da situação de confronto. A *teoria da utilidade* estuda tanto propriedades de recompensas quanto a determinação numérica relativa.

### 2.2.1. Jogos em forma normal

Na seção anterior apresentamos alguns exemplos de confronto e vimos informalmente que jogos são definidos utilizando os conceitos de estratégias, resultados e recompensas. Nesta seção definiremos formalmente um jogo. Existe mais de uma maneira de se definir um jogo, mas a forma mais comum é chamada de *forma normal*, que apresentaremos agora.

**Definição 1** Um jogo  $\mathcal{J} = (N, S, \mathbf{u})$  em forma normal é composto de:

- Um conjunto  $N$  de jogadores.
- Um conjunto  $S_i$  de estratégias disponíveis para cada jogador  $i \in N$ . O conjunto de vetores  $S$  representa as estratégias disponíveis para todos os jogadores, com  $S = \times_{i \in N} S_i$  ( $S$  é o produto cartesiano dos conjuntos  $S_i$ ).
- Uma função de utilidade  $u_i : S \rightarrow \mathfrak{R}$  para cada jogador  $i \in N$ , que associa um valor numérico finito a cada resultado do jogo. O vetor  $\mathbf{u}$  representa

as funções de utilidade de todos os jogadores, com  $\mathbf{u} = \times_{i \in N} u_i$ .

O resultado  $\mathbf{s} \in S$  do jogo é um vetor formado pelas estratégias escolhidas por cada jogador, onde  $s_i \in S_i$  representa a estratégia escolhida pelo jogador  $i \in N$ .

O objetivo de cada jogador é maximizar sua recompensa, representada pela sua função de utilidade,  $u_i$ . Porém, observe que cada jogador decide apenas a sua estratégia  $s_i \in S_i$ , mas sua função de utilidade depende da estratégia escolhida por todos os outros jogadores. Assim sendo, ao maximizar sua função de utilidade, o jogador  $i \in N$  necessariamente precisa considerar como os outros jogadores irão jogar. Isto leva a um problema de maximização muito diferente dos problemas clássicos de otimização.

Para facilitar a apresentação, iremos adotar a seguinte notação vetorial: seja  $\mathbf{v}$  um vetor com dimensão  $n$ . Ou seja,  $\mathbf{v} = (v_1, v_2, \dots, v_n)$ . O vetor  $\mathbf{v}_{-i}$  representa o vetor  $\mathbf{v}$  sem a  $i$ -ésima posição. Ou seja,  $\mathbf{v}_{-i} = (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$  com  $i = 1, \dots, n$ . Além disso, a função de utilidade  $u_i$  do jogador  $i$  será representada por  $u_i(s_i, \mathbf{s}_{-i})$ , onde o primeiro parâmetro  $s_i \in S_i$  é a estratégia escolhida por  $i$ , e o segundo parâmetro  $\mathbf{s}_{-i} \in S_{-i}$  é um vetor com as estratégias escolhidas por todos os outros jogadores.

Por último, um jogo  $\mathcal{J}$  é considerado finito se o conjunto de jogadores,  $N$ , é finito. Caso contrário, o jogo é considerado infinito. Jogos infinitos podem ser mais suscetíveis ao tratamento analítico, e por isso são às vezes utilizados como aproximações de jogos com muito jogadores. Além disso, o conjunto de estratégias  $S_i$  disponível para cada jogador  $i \in N$  pode ser discreto ou contínuo. Jogos com conjuntos de estratégias contínuos também podem facilitar a análise.

Para exemplificar jogos em forma normal, enquadramos agora o segundo exemplo do jogo de conexões TCP, apresentado informalmente na seção anterior na sua forma normal. Temos então que  $N = \{A, B\}$ , onde  $A$  representa o jogador Alberto e  $B$  a jogadora Barbara. O conjunto de estratégias de ambos os jogadores é dado por  $S_A, S_B = \{1, 2\}$ , onde 1 e 2 representam, respectivamente, estabelecer uma ou duas conexões. A função de utilidade de Alberto é dada por:  $u_A(1, 1) = 3$ ,  $u_A(1, 2) = 2$ ,  $u_A(2, 1) = 4$ ,  $u_A(2, 2) = 5$ . Enquanto a função de utilidade de Barbara é dada por:  $u_B(1, 1) = 3$ ,  $u_B(1, 2) = 2$ ,  $u_B(2, 1) = 4$ ,  $u_B(2, 2) = 5$ .

Jogos com dois jogadores e conjunto de estratégias discreto são mais facilmente representados por uma matriz, conhecida como *matriz do jogo*. A matriz do jogo de conexões TCP está ilustrada na figura 2.2. Está claro que tal representação é mais conveniente que a definição formal acima, e assim sendo, utilizaremos a matriz do jogo em diversos outros exemplos.

### 2.2.2. Princípio da dominância

Considere o jogo de conexões TCP apresentado na seção anterior, representado pela matriz do jogo ilustrada na figura 2.2. Na análise deste jogo vimos

que Alberto deveria escolher a estratégia de estabelecer duas conexões, independente da escolha feita por Barbara. Vimos que escolher duas conexões é neste caso estritamente melhor do que escolher uma conexão. Dizemos então que a estratégia “duas conexões” *domina* a estratégia “uma conexão”. Mais formalmente, o conceito de estratégias dominadas é definido da seguinte forma: considere um jogo  $\mathcal{J} = (N, S, \mathbf{u})$ .

**Definição 2** Uma estratégia  $a \in S_i$  do jogador  $i \in N$  domina a estratégia  $b \in S_i$  do mesmo jogador, se  $u_i(a, s_{-i}) > u_i(b, s_{-i})$  para qualquer resultado  $s_{-i} \in S_{-i}$  do jogo.

Se o domínio não for estrito, ou seja,  $u_i(a, s_{-i}) = u_i(b, s_{-i})$  para algum resultado  $s_{-i} \in S_{-i}$  do jogo, então dizemos que a estratégia  $a$  *domina fracamente* a estratégia  $b$ .

O *princípio da dominância* estabelece que jogadores nunca devem escolher estratégias dominadas. O princípio é óbvio, uma vez que é possível obter uma melhor recompensa ao escolher a estratégia dominante. Assim sendo, estratégias dominadas podem ser eliminadas do jogo. A eliminação de estratégias dominadas deve ser feita por ambos os jogadores, o que pode levar a um processo iterativo de eliminação de estratégias dominadas. Em alguns casos, este processo termina deixando apenas uma única estratégia para cada jogador, que fica sendo então a solução do jogo.

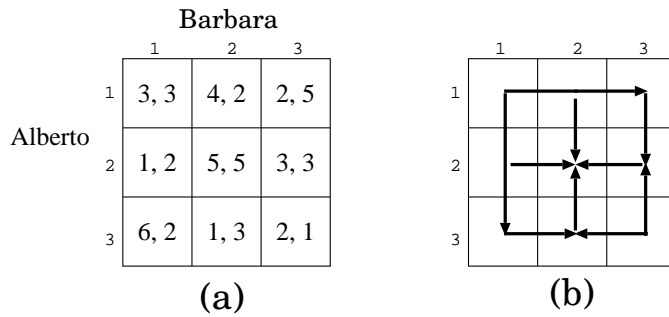
No exemplo do jogo de conexões TCP, a estratégia “uma conexão” para a jogadora Barbara também é dominada pela estratégia “duas conexões”. Assim sendo, tal estratégia deve ser eliminada. Repare que após a eliminação, ambos os jogadores possuem apenas uma estratégia: “duas conexões”, que neste caso é a solução do jogo.

Infelizmente, a eliminação de estratégias dominadas nem sempre deixa os jogadores com apenas uma estratégia. Aliás, muitos jogos nem sequer possuem estratégias dominadas, limitando a aplicabilidade deste tipo de análise, como por exemplo, o jogo de conexão TCP ilustrado na figura 2.1. Isto motiva a definição de outros conceitos de solução para os jogos.

### 2.2.3. Princípio de pontos de sela

Consideremos agora uma variação do jogo de conexões TCP, onde cada jogador possui três estratégias: estabelecer uma, duas, ou três conexões para transferir o arquivo. A matriz do jogo está ilustrada na figura 2.3. Nossos jogadores Alberto e Barbara precisam escolher suas respectivas estratégias. Como eles devem proceder?

Façamos novamente uma análise cuidadosa do jogo. Primeiramente, podemos verificar que nenhum dos dois jogadores deste jogo possui estratégias dominadas. Assim sendo, nenhuma estratégia pode ser eliminada. Resta-nos então fazer uma análise de hipóteses. Vamos começar imaginando que Alberto ache que Barbara irá escolher a estratégia 1 (ou seja, estabelecer uma conexão). O que Alberto deve fazer neste caso? Deve escolher a estratégia



**Figura 2.3. Jogo de conexões TCP e diagrama de movimento**

3, pois esta oferece a maior recompensa. Mas se Barbara achar que Alberto escolherá a estratégia 3, então ela deve escolher a estratégia 2, pois esta lhe dará uma maior recompensa. Porém, se Alberto achar que Barbara irá escolher a estratégia 2, ele também deve escolher a estratégia 2, pois dará a ele a maior recompensa. Finalmente, se Barbara achar que Alberto irá escolher a estratégia 2, então ela também deve escolher a estratégia 2, pois dará a ela a maior recompensa neste caso. Observe que há algo interessante sobre o resultado (2,2) do jogo. Em particular, nenhum dos dois jogadores prefere mudar de estratégia caso o outro esteja jogando a estratégia 2. Dizemos então que a estratégia (2,2) é um *ponto de sela* do jogo.

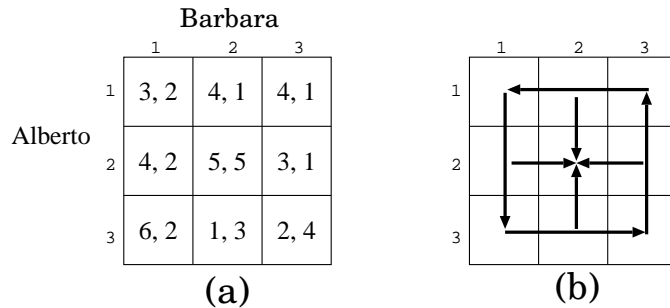
Informalmente, um ponto de sela é um resultado do jogo onde nenhum dos jogadores pode aumentar sua recompensa escolhendo isoladamente uma outra estratégia. Segue agora uma definição formal do ponto de sela de um jogo:

**Definição 3** O resultado  $s \in S$  de um jogo  $\mathcal{J} = (N, S, \mathbf{u})$  é um ponto de sela se e somente se para todo jogador  $i \in N$ ,  $u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i})$  para qualquer outra estratégia  $s'_i \in S_i$ .

O princípio de pontos de sela estabelece que jogadores devem sempre escolher estratégias que levem a um ponto de sela. Uma das razões por trás deste princípio está justamente na linha de raciocínio que nos levou a encontrar o ponto de sela no jogo acima. Sabendo que o outro jogador escolherá sua estratégia para maximizar sua recompensa leva os jogadores a convergirem no ponto de sela.

A análise acima que nos levou a encontrar o ponto de sela do jogo pode ser resumida com um diagrama de movimento. Este diagrama indica qual a direção de preferência dos jogadores quando a estratégia do outro jogador é fixada. O diagrama de movimento do jogo descrito acima está ilustrado na figura 2.3. As setas das colunas indicam a preferência de recompensa de Alberto, enquanto

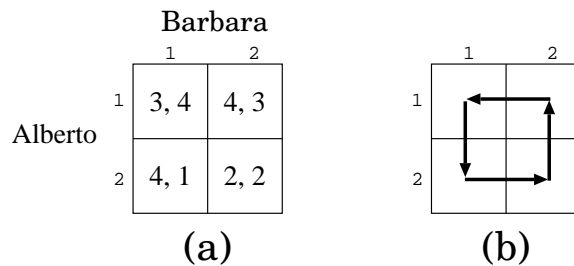
as setas nas linhas indicam a preferência de recompensa de Barbara. As setas apontam para a maior recompensa da linha ou coluna (neste caso cada linha ou coluna possui um único máximo). Podemos observar que o ponto (2,2) é um sumidouro (não há setas saindo deste ponto) se seguirmos as setas, o que caracteriza um ponto de sela neste diagrama. Uma outra observação é que todos os caminhos estabelecidos pelas setas do diagrama levam ao ponto de sela, independente de onde começamos a análise.



**Figura 2.4. Jogo de conexões TCP e diagrama de movimento**

Nem todos os jogos possuem um diagrama de movimento que sempre levam a um ponto de sela. Considere uma modificação da matriz do jogo acima ilustrada na figura 2.4. O respectivo diagrama de movimento deste jogo está ilustrado na figura 2.4. Observe que o resultado (2,2) continua sendo um ponto de sela do jogo. Porém, existe um ciclo no diagrama de movimento formado pelas setas externas.

#### 2.2.4. Estratégias mistas



**Figura 2.5. Jogo de conexões TCP e diagrama de movimento**

Nem todos os jogos possuem pontos de sela. Considere a variação do jogo

de conexões TCP representado pela matriz do jogo da figura 2.5. Neste exemplo cada jogador possui apenas duas estratégias. O diagrama de movimento deste jogo está ilustrado na figura 2.5, e podemos verificar que não existem estratégias dominadas nem pontos de sela neste jogo. Observe que se um jogador fixar uma estratégia, o outro pode explorar este fato para aumentar sua recompensa. Como proceder neste caso?

Uma das únicas opções para resolver este impasse é a utilização da randomização. Ou seja, os jogadores devem escolher suas estratégias de forma aleatória, e não fixá-las a priori. Por exemplo, os jogadores podem utilizar uma moeda para decidir a estratégia que irão escolher. Mais precisamente, os jogadores podem associar probabilidades fixas às estratégias disponíveis e então utilizar um dispositivo aleatório para decidir qual estratégia irão jogar no momento da escolha. Tal plano de ação é chamado de *estratégia mista*, pois envolve potencialmente mais de uma estratégia. Em contraste, *estratégias puras* envolvem apenas uma única estratégia, que não é escolhida de forma aleatória.

**Definição 4** Considere um jogo  $\mathcal{J} = (N, S, \mathbf{u})$ . Uma estratégia mista para o jogador  $i \in N$  é uma função  $p_i(\cdot)$  que associa a cada estratégia  $s_i \in S_i$  uma probabilidade de forma que  $\sum_{s_i \in S_i} p_i(s_i) = 1$ . O vetor  $p = \times_{i \in N} p_i$  é formado pelo produto cartesiano das funções  $p_i$ .

Observe que uma estratégia mista é simplesmente uma função de probabilidade de massa sobre o conjunto de estratégias do jogador.

Ao utilizarmos estratégias mistas, o resultado do jogo passa a ser aleatório, pois irá depender das escolhas efetivamente feitas pelos jogadores no momento do jogo. Entretanto, podemos facilmente associar uma probabilidade a cada possível resultado do jogo.

**Definição 5** Considere um jogo  $\mathcal{J} = (N, S, \mathbf{u})$  e o vetor  $p$  de estratégias mistas. A probabilidade que um determinado resultado do jogo  $s \in S$  ocorra é dado por  $\prod_{i \in N} p_i(s_i)$ .

Assim sendo, podemos definir o *valor esperado* da recompensa obtida em cada possível resultado do jogo. Este valor esperado é dado pela recompensa obtida em um determinado resultado do jogo ponderada pela probabilidade do resultado ocorrer. Eis sua definição formal:

**Definição 6** Considere um jogo  $\mathcal{J} = (N, S, \mathbf{u})$  e o vetor  $p$  de estratégias mistas. O valor esperado da recompensa obtida em determinado resultado do jogo  $s \in S$  para cada jogador  $i \in N$  é dado por  $u_i(s) \prod_{j \in N} p_j(s_j)$ .

Por último, dadas as estratégias mistas dos jogadores, podemos calcular o valor esperado da recompensa que será obtida por um jogador quando este escolher, de forma determinística, uma de suas estratégias.

**Definição 7** Considere um jogo  $\mathcal{J} = (N, S, \mathbf{u})$  e o vetor  $p$  de estratégias mistas. O valor esperado da recompensa obtida pelo jogador  $i \in N$  quando este escolhe a estratégia  $s_i \in S_i$  é dado por  $\sum_{(s_i, s_{-i}) \in S} u_i(s_i, s_{-i}) \prod_{j \in N, j \neq i, s_j \in S_{-i}} p_j(s_j)$ .

Esta definição de recompensa esperada será usada nos exemplos que seguem.

Dadas as estratégias mistas dos outros jogadores, como um jogador deve escolher suas estratégias? Considere o exemplo acima (figura 2.5) para ilustrar esta questão. Vamos supor que Barbara decidiu utilizar uma estratégia mista que atribui probabilidade  $1/2$  a cada uma de suas duas estratégias. Como Alberto deve jogar neste caso? Vejamos qual é o valor esperado da recompensa que ele obterá ao escolher cada uma de suas estratégias. Ao escolher a estratégia 1, a recompensa esperada será  $3.5$  ( $3 \times 1/2 + 4 \times 1/2$ ). Se ele escolher a estratégia 2, a recompensa esperada será  $3$  ( $4 \times 1/2 + 2 \times 1/2$ ). Assim sendo, Alberto deve jogar a estratégia 1, pois esta possui um valor esperado maior.

Isto nos leva ao *princípio do valor esperado de recompensas*, que estabelece que jogadores devem escolher estratégias que maximizam a recompensa esperada. Este princípio é apenas uma extensão do princípio de racionalidade que, como vimos, determina que jogadores estão sempre querendo maximizar seus benefícios.

Dada a discussão acima, como é que Barbara deve escolher sua estratégia mista? Observe que se Alberto jogar a estratégia 2, a recompensa esperada de Barbara será  $1.5$  ( $1 \times 1/2 + 2 \times 1/2$ ), que é menor do que ela obteria se escolhesse a estratégia 2 com probabilidade 1, que daria a ela no mínimo uma recompensa de valor 2. Podemos dizer que Barbara está em desvantagem ao escolher a estratégia mista  $(1/2, 1/2)$ . O ideal seria que Barbara escolhesse uma estratégia mista que não pudesse ser explorada por Alberto. Ou seja, uma estratégia mista na qual o valor esperado da recompensa recebida por Alberto é sempre o mesmo, independente da escolha de estratégia que ele fizer.

Esta estratégia mista pode ser determinada da seguinte forma. Vamos supor que Barbara irá escolher a estratégia 1 com probabilidade  $p$  e a estratégia 2 com probabilidade  $1 - p$ . Se Alberto escolher a estratégia 1, sua recompensa esperada será  $3 \times p + 4 \times (1 - p)$ , e ao escolher a estratégia 2, sua recompensa esperada será  $4 \times p + 2 \times (1 - p)$ . O objetivo de Barbara é determinar o valor de  $p$  tal que as duas recompensas esperadas tenham o mesmo valor. Ou seja, ela precisa determinar  $p$  tal que  $3 \times p + 4 \times (1 - p) = 4 \times p + 2 \times (1 - p)$ . A solução neste caso é  $p = 2/3$ . A estratégia mista que Barbara deve utilizar é então dada por  $(2/3, 1/3)$ . Observe que o valor esperado da recompensa recebida por Alberto é  $10/3$ , independente da estratégia que ele escolha.

Para determinar sua estratégia mista, Alberto deve proceder da mesma maneira. Pois se Barbara souber ou desconfiar de como ele irá jogar, poderá mudar sua estratégia para obter uma recompensa esperada mais alta. Por isto, Alberto precisa determinar sua estratégia mista de forma que ambas as opções de estratégia de Barbara levem à mesma recompensa esperada. Suponhamos que Alberto irá escolher a estratégia 1 com probabilidade  $q$ . Então, os valores



esperados das recompensas que Barbara obterá ao escolher as estratégias 1 ou 2, respectivamente, são  $4 \times q + 1 \times (1 - q)$  e  $3 \times q + 2 \times (1 - q)$ . Igualando as duas equações e resolvendo para  $q$  temos que  $q = 1/2$ . Assim sendo, Alberto deve utilizar a recompensa mista  $(1/2, 1/2)$ , que dará a Barbara uma recompensa esperada de  $5/2$ , independente da escolha de estratégia que ela fizer.

Observe que este jogo está em vantagem para Alberto, pois a sua recompensa esperada é  $10/3$ , enquanto a recompensa esperada de Barbara é  $5/2$ . Entretanto, ao utilizar estratégias mistas, Barbara conseguiu garantir um valor esperado maior do que o valor mínimo que conseguiria se escolhesse estratégias puras. Podemos verificar que o mesmo ocorre para Alberto.

### 2.2.5. Equilíbrio de Nash

O equilíbrio de Nash é um dos conceitos mais importantes da teoria dos jogos, pois define um conceito de solução para os jogos. Entretanto, o equilíbrio de Nash nada mais é do que um ponto de sela do jogo, onde nenhum jogador pode aumentar sua recompensa mudando unilateralmente de estratégia. Eis a definição formal:

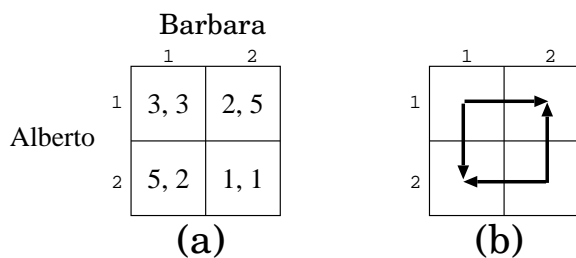
**Definição 8** Considere um jogo  $\mathcal{G} = (N, S, \mathbf{u})$  e um resultado  $\mathbf{s} \in S$  deste jogo. O resultado  $\mathbf{s}$  é um equilíbrio de Nash se para todo jogador  $i \in N$ ,  $u_i(s_i, \mathbf{s}_{-i}) \geq u_i(s'_i, \mathbf{s}_{-i})$  para qualquer outra estratégia  $s'_i \in S_i$ .

Confira que esta definição é equivalente à definição de ponto de sela. O ponto de sela recebeu o nome de equilíbrio de Nash em homenagem ao matemático John Nash, que em 1950 provou que todo jogo em forma normal possui ao menos um ponto de sela utilizando estratégias puras ou mistas. Assim sendo, de agora em diante iremos nos referir a pontos de sela como sendo equilíbrio de Nash do jogo.

O resultado obtido por Nash é importante, pois estabelece que todo jogo em forma normal possui ao menos um equilíbrio. Apesar do aspecto positivo, veremos na próxima seção que o equilíbrio de Nash possui diversos problemas como um conceito de solução do jogo.

### 2.2.6. Jogos com múltiplos equilíbrios

Infelizmente, jogos em geral podem possuir mais de um equilíbrio de Nash. Além disso, os resultados correspondentes aos múltiplos equilíbrios de Nash não são necessariamente equivalentes. É possível que um jogo com dois jogadores possua dois equilíbrios de Nash, mas que cada jogador prefira um dos equilíbrios. O que fazer neste caso? Este é um tema ainda sem alguma solução satisfatória mas, bastante discutido por pesquisadores da área. Uma das idéias para resolver este impasse é a eliminação de equilíbrios de Nash que possuam propriedades negativas, como por exemplo, que não sejam estáveis. Veremos mais detalhes sobre esta idéia na seção 2.3, ao discutirmos teoria dos jogos evolucionária.



**Figura 2.6. Jogo de conexões TCP e diagrama de movimento**

Voltamos agora finalmente ao primeiro jogo apresentado neste texto, o jogo de conexões TCP, ilustrado novamente na figura 2.6 juntamente com seu diagrama de movimento. Uma análise do jogo revela que o mesmo possui dois equilíbrios de Nash sob estratégias puras, dados pelos resultados Alberto escolhendo duas conexões e Barbara uma (2, 1), e Alberto escolhendo uma conexão e Barbara duas (1, 2). observe que Alberto prefere o primeiro equilíbrio de Nash, pois este o oferece uma recompensa maior. Ao mesmo tempo, Barbara prefere o segundo equilíbrio de Nash.

Há ainda um terceiro equilíbrio de Nash neste jogo, obtido quando utilizamos estratégias mistas tanto para Alberto quanto para Barbara (você deve calcular este equilíbrio). O valor esperado da recompensa neste caso é  $7/3$ , tanto para Alberto quanto para Barbara. Desta forma podemos dizer que este equilíbrio de Nash é mais justo que os outros dois, pois ambos os jogadores recebem a mesma recompensa. Assim sendo, podemos dizer que este equilíbrio é mais adequado como sendo a solução do jogo.

### 2.2.7. O dilema do prisioneiro

O Dilema do Prisioneiro é um dos jogos mais estudados e vastamente aplicado em diversas áreas do conhecimento, pois o jogo captura um problema fundamental entre o interesse individual e o interesse do grupo. A estória que ilustra o jogo é a seguinte: dois supostos criminosos parceiros, detidos por cometer um crime em conjunto, estão sendo interrogados simultaneamente em salas separadas. Se um deles apresentar provas que incriminem o colega e o outro não, o que implicou o parceiro terá sua pena reduzida para 2 anos, enquanto que o outro pegará a maior pena possível, 20 anos de cadeia. Caso nenhum acuse o outro, suas penas serão moderadas, de 5 anos, pois não há provas suficientes para condená-los completamente. Por outro lado, se os parceiros se acusarem mutuamente, as penas de ambos serão severas, de 10 anos. O jogo está ilustrado na figura 2.7, onde as estratégias disponíveis para cada jogador são “D” (delatar o parceiro) e “N” (não delatar o parceiro). Observe que a recompensa deste jogo indica o número de anos de cadeia, e

assim sendo, os jogadores preferem valores de recompensa menores.

		Prisioneiro 2	
		N	D
Prisioneiro 1	N	5, 5	20, 2
	D	2, 20	10, 10

**Figura 2.7. Jogo do dilema do prisioneiro**

O que os prisioneiros devem fazer? O que você faria no lugar de um deles? Uma simples análise do jogo revela que a estratégia “delatar” domina a estratégia “não delatar”. Ou seja, independente do que o outro prisioneiro fizer, é sempre melhor delatar o parceiro. Assim sendo, o resultado obtido quando os dois prisioneiros delatam-se mutuamente é o único equilíbrio de Nash do jogo. Mas observe que há algo curioso sobre este resultado: individualmente, é sempre melhor delatar. Porém, se ambos ficarem em silêncio, ambos sairão ganhando.

Isto ilustra o conflito entre o interesse individual e o interesse do grupo. A racionalização individual nem sempre leva ao melhor resultado para o grupo, ou até mesmo para o próprio indivíduo, conforme exemplificado no dilema do prisioneiro. Uma conclusão precipitada levaria a concluir que o melhor seria então não delatar o parceiro. Porém, como ter certeza de que o outro irá cooperar, pensando da mesma maneira e também não delatando? Este é um dos problemas fundamentais da teoria dos jogos: o estímulo à cooperação num ambiente onde a individualidade unilateral traz benefícios sobre a racionalização em grupo. Voltaremos ao dilema do prisioneiro na seção 2.2.18, quando mostraremos que sob determinadas condições, a individualidade pode levar a um melhor resultado para ambos os prisioneiros.

### 2.2.8. Ótimo de Pareto

No jogo do dilema do prisioneiro apresentado na seção anterior, vimos que o resultado obtido quando ambos jogadores escolhem a estratégia “não delatar” leva a uma recompensa melhor que a obtida com o equilíbrio de Nash (ver matriz do jogo na figura 2.7). Observe que ambos os jogadores preferem este resultado, pois oferece a ambos uma melhor recompensa. Este resultado do jogo é conhecido como *ponto ótimo de Pareto*. Informalmente, um resultado de um jogo é um ponto ótimo de Pareto se nenhum outro resultado oferece a todos os jogadores uma recompensa melhor. Segue a definição mais precisa:

**Definição 9** Considere um jogo  $\mathcal{J} = (N, S, \mathbf{u})$  e um resultado  $s \in S$  deste jogo.

O resultado  $s$  é um ponto ótimo de Pareto se não existir outro resultado  $s' \in S$  onde  $u_i(s') > u_i(s)$  para cada jogador  $i \in N$ .

Confira que o equilíbrio de Nash do jogo do dilema do prisioneiro não é um ponto ótimo de Pareto. Entretanto, todos os outros três resultados deste jogo são pontos ótimos de Pareto. Observe que o quantitativo “ótimo” não quer dizer que o resultado é superior a todos os outros resultados do jogo, e sim que não há um outro resultado que seja preferido por todos os jogadores.

O conceito de ótimo de Pareto está relacionado à racionalização em grupo, enquanto o conceito de equilíbrio de Nash está relacionado à racionalização individual.

### 2.2.9. Jogos com infinitas estratégias

Até agora vimos apenas jogos onde o número de estratégias disponível para cada jogador era finito, e em particular, bem limitado. Entretanto, nada impede que os jogadores possuam um número infinito de estratégias. Observe que esta generalização não afeta a definição do jogo, nem os conceitos já introduzidos, como o equilíbrio de Nash.

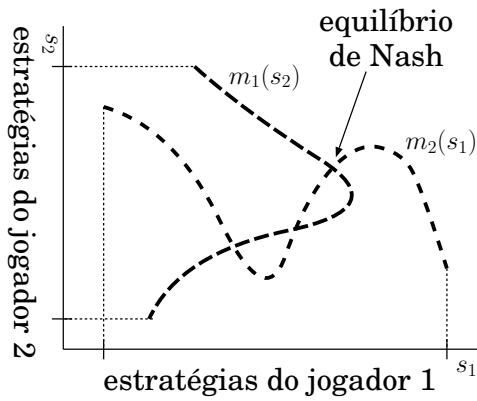
Jogos com um número infinito e contínuo de estratégias, em muitos casos, podem ser solucionados de forma analítica mais facilmente. Porém, um conceito fundamental para lidar com este tipo de jogo são *funções de melhor resposta*. Esta função fornece a melhor estratégia para um determinado jogador, dadas as estratégias escolhidas por todos os outros jogadores. Formalmente temos a seguinte definição:

**Definição 10** Considere o jogo  $\mathcal{J} = (N, S, \mathbf{u})$ . A função de melhor resposta  $m_i : S_{-i} \rightarrow S_i$  do jogador  $i \in N$  é definida por  $m_i(s_{-i}) = \arg \max_{s'_i \in S_i} u_i(s'_i, s_{-i})$ , para qualquer  $s_{-i} \in S_{-i}$ .

observe que a imagem da função de melhor resposta é o conjunto de estratégias do jogador em questão. Outra observação é que o conceito de melhor resposta não está limitado aos casos onde o conjunto de estratégias dos jogadores é infinito.

Uma propriedade importante é que o equilíbrio de Nash pode ser caracterizado pelas funções de melhor resposta do jogador. Observe que num equilíbrio de Nash, nenhum jogador pode aumentar sua recompensa mudando unilateralmente sua estratégia. Desta forma, a função de melhor resposta para cada jogador no equilíbrio de Nash fornece a própria estratégia que define o equilíbrio. Mais precisamente, o equilíbrio de Nash é um *ponto fixo* das funções de melhor resposta de todos os jogadores. Mais formalmente, temos a seguinte propriedade:

**Propriedade 1** Considere o jogo  $\mathcal{J} = (N, S, \mathbf{u})$  e as funções de melhor resposta  $m_i, i \in N$ . Um resultado do jogo  $s \in S$  é um equilíbrio de Nash se e somente se  $s_i = m_i(s_{-i})$  para todos os jogadores  $i \in N$ .



**Figura 2.8. Ilustração das funções de melhor resposta de um jogo com dois jogadores**

Esta equivalência entre o ponto fixo das funções de melhor resposta e o equilíbrio de Nash pode ser visualizada graficamente. Considere um jogo com dois jogadores com conjunto de estratégias infinito. A figura 2.8 ilustra as funções de melhor resposta dos dois jogadores. Os eixos  $x$  e  $y$  do gráfico representam, as estratégias dos jogadores 1 e 2, respectivamente. A curva  $m_1(s_2)$  é a função de melhor resposta do jogador 1. Observe que o domínio desta função é dado pelo conjunto de estratégias do jogador 2, enquanto a imagem é dada pelo conjunto de estratégias do jogador 1. O equivalente vale para a função resposta do jogador 2,  $m_2(s_1)$ . Os pontos onde as duas curvas se interceptam são pontos fixos mútuos, ou seja,  $m_1(s_2) = s_1$  e  $m_2(s_1) = s_2$ . Assim sendo, todas as interseções entre as duas curvas correspondem a equilíbrios de Nash do jogo.

### 2.2.10. Duopólio de Cournot

O duopólio de Cournot é um modelo econômico muito utilizado para modelar competição entre empresas que produzem o mesmo produto. Iremos aplicar este modelo à competição entre dois provedores de serviço de Internet de banda larga. Considere dois provedores que precisam decidir o quanto de banda passante eles devem comprar para se conectarem com a Internet. Chamaremos de  $b_i, i = 1, 2$  a quantidade de banda comprada pelos provedores. Obviamente, os provedores têm um custo ao comprar determinada quantidade de banda passante. Para simplificar, iremos assumir um modelo de custo simples, sem custo fixo, apenas proporcional à quantidade de banda comprada. Assim sendo, o custo para comprar uma quantidade de banda passante  $b$  será dado por  $c \times b$ , onde  $c$  é o custo por unidade de banda.

Precisamos agora considerar a demanda do mercado pelo serviço ofere-

cido pelos dois provedores. Para simplificar, iremos assumir que o mercado possui uma demanda fixa por banda passante, denotada por  $d$ . Ou seja, todos os possíveis clientes dos dois provedores consumirão  $d$  unidades de banda passante. Intuitivamente, o preço pago pelos usuários para acessar a Internet irá depender da quantidade de banda passante que os provedores comprarem e da demanda dos próprios usuários. Vamos adotar um modelo simples, onde o preço pago pelos usuários é dado por  $\max(0, d - b_1 - b_2)$ . Observe que o preço pago será zero, caso os provedores ofereçam mais banda passante do que a demanda existente no mercado. Outra observação é que os dois provedores oferecem o mesmo preço aos usuários, pois caso contrário os usuários mudarão para o provedor mais barato. Finalmente, a receita dos provedores é dado pelo que eles recebem menos o que gastam. Ou seja,

$$u_1(b_1, b_2) = b_1 \times \max(0, d - b_1 - b_2) - c \times b_1 \quad (1)$$

$$u_2(b_2, b_1) = b_2 \times \max(0, d - b_1 - b_2) - c \times b_2 \quad (2)$$

Observe que precisamos assumir que a demanda é maior que o custo marginal, ou seja, que  $d > c$ , pois caso contrário o custo seria muito alto e os provedores não conseguiriam obter lucro (sob o modelo de preço adotado) oferecendo tal serviço.

Suponhamos agora que os dois provedores precisam decidir simultaneamente o quanto de banda passante irão comprar. O que devem fazer? Podemos modelar a situação acima como um jogo de dois jogadores, onde as estratégias são a escolha da quantidade de banda  $b_i$  que cada um deve adquirir, e a recompensa é dada pela Equação 2. Qual é o equilíbrio de Nash deste jogo?

Iremos utilizar as funções de melhor resposta para obter o equilíbrio de Nash neste exemplo. Começaremos calculando as funções de melhor resposta para cada jogador. Por definição, temos que  $m_1(b_2) = \arg \max_{0 \leq b_1} u_1(b_1, b_2)$ , e substituindo:

$$m_1(b_2) = \arg \max_{0 \leq b_1} b_1 \times \max(0, d - b_1 - b_2) - c \times b_1 \quad (3)$$

$$m_2(b_1) = \arg \max_{0 \leq b_2} b_2 \times \max(0, d - b_1 - b_2) - c \times b_2 \quad (4)$$

Podemos calcular o valor exato das funções de melhor resposta simplesmente calculando o seu máximo do argumento. Observe que temos um simples problema de maximização e que para resolvê-lo, basta satisfazer as condições de primeira ordem. Ou seja, o máximo ocorre quando a derivada do argumento da função de melhor resposta é igual a zero. As derivadas dos argumentos  $a_1(b_1)$  e  $a_2(b_2)$  são, respectivamente:

$$\frac{da_1}{db_1} = d - c - 2b_1 - b_2 \quad (5)$$

$$\frac{da_2}{db_2} = d - c - 2b_2 - b_1 \quad (6)$$

Igualando a zero e resolvendo, temos:

$$b_1^* = \frac{d-c-b_2}{2} \quad (7)$$

$$b_2^* = \frac{d-c-b_1}{2} \quad (8)$$

Onde  $b_1^*$  e  $b_2^*$  correspondem ao máximo dos argumentos  $a_1(b_1)$  e  $a_2(b_2)$ , respectivamente. Ou seja, correspondem à função de melhor resposta. Temos então:

$$m_1(b_2) = \frac{d-c-b_2}{2} \quad (9)$$

$$m_2(b_1) = \frac{d-c-b_1}{2} \quad (10)$$

Sabemos que o equilíbrio de Nash ocorre no ponto fixo destas duas funções. Ou seja, quando  $m_1(b_2^o) = b_1^o$  e  $m_2(b_1^o) = b_2^o$ . Podemos calcular o ponto fixo  $b_1^o$  e  $b_2^o$  utilizando substituição simples e resolvendo a equação:  $b_1^o = m_1(m_2(b_1^o))$ . Aplicando a substituição, obtemos  $b_1^o = (d-c)/3$ . Resolvendo agora para  $b_2^o$ , obtemos  $b_2^o = (d-c)/3$ . O equilíbrio de Nash deste jogo é dado então pelas estratégias  $b_1 = (d-c)/3$  e  $b_2 = (d-c)/3$ . Neste ponto, o preço pago pelos usuários pelo acesso à Internet é de  $(d+2c)/3$ , e a quantidade total de banda comprada pelos provedores é dada por  $b_1^o + b_2^o = 2(d-c)/3$ .

É interessante comparar o resultado do jogo acima com um cenário onde apenas um provedor estivesse oferecendo acesso à Internet aos usuários. Qual seria o resultado neste caso? Qual preço os usuários pagariam pelo acesso? Para calcular o resultado neste cenário, basta fixar  $b_2 = 0$  no desenvolvimento acima, pois assim o segundo provedor não irá influenciar na decisão do outro provedor. Você deve conferir que, neste caso, o ponto ótimo para o provedor 1 é dado por  $b_1^o = (d-c)/2$ , o que leva ao preço de  $(d+c)/2$  a ser pago pelos usuários. Observe que este preço é *maior* que o preço a ser pago quando os dois provedores competem pelos usuários. Além disso, o total de banda comprado pelo provedor é menor neste caso. Ou seja, a competição entre os dois provedores é benéfica para os usuários, que pagam um preço mais baixo e recebem uma maior quantidade de banda agregada. O monopólio, neste caso, leva a um resultado inferior.

### 2.2.11. Jogos em forma extensiva

Todos os jogos descritos até agora consideram que os jogadores tomam suas decisões simultaneamente, sem antes saber qual será a decisão dos outros jogadores. Apesar deste modelo de jogo aplicar-se a diversos cenários, há muitas situações onde o jogo ocorre de forma seqüencial, com os jogadores alternando suas decisões. Tais jogos podem ser mais facilmente representados em *forma extensiva*.

Jogos extensivos são descritos utilizando o conceito de *árvore do jogo*. A árvore do jogo é formada por nós, que representam pontos onde um jogador

deve tomar uma decisão, e arestas, que representam as possíveis decisões que tal jogador pode tomar naquele ponto do jogo. Assim sendo, cada nó interno da árvore está associado ao jogador que deverá jogar naquele ponto do jogo. As folhas da árvore são chamadas de nós terminais, pois indicam o fim do jogo. Desta forma, cada nó terminal representa também um resultado do jogo. Assim como em jogos em forma normal, uma recompensa para cada jogador está associada a cada possível resultado do jogo.

A descrição formal de um jogo extensivo possui uma notação bastante complicada, e por isto evitaremos uma descrição mais rigorosa. Entretanto, adotaremos a seguinte definição:

**Definição 11** Um jogo  $\mathcal{J} = (N, \mathcal{T}, S, \mathbf{u})$  em forma extensiva é formado por:

- Um conjunto  $N$  de jogadores.
- Uma árvore  $\mathcal{T}$ , chamada de árvore do jogo.
- Um conjunto  $S_i$  de estratégias disponíveis para cada jogador  $i \in N$ . O conjunto de vetores  $S$  representa as estratégias disponíveis para todos os jogadores, com  $S = \times_{i \in N} S_i$ .
- Uma função de utilidade  $u_i : S \rightarrow \mathfrak{R}$  para cada jogador  $i \in N$ , que associa um valor numérico finito a cada resultado do jogo. O vetor  $\mathbf{u}$  representa as funções de utilidade de todos os jogadores, com  $\mathbf{u} = \times_{i \in N} u_i$ .

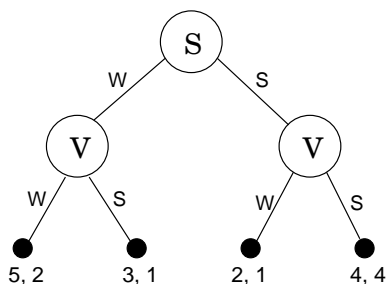
Jogos em forma extensiva também possuem equilíbrios de Nash. A definição 8 se aplica igualmente a este tipo de jogo, que como veremos em breve, também pode ser representado em forma normal.

Vejamos um exemplo de um jogo em forma extensiva. Suponhamos que os grupos de desenvolvimento do Skype e do Vivavoz<sup>1</sup>, dois aplicativos que possibilitam a transmissão de voz pela Internet, estão decidindo o padrão para codificação da voz a ser implementado pelo aplicativo. Os dois grupos possuem duas opções: adotar o padrão Windows Media Audio (WMA) ou o padrão aberto Speex. Vamos supor que os desenvolvedores do Vivavoz tomarão uma decisão apenas depois dos desenvolvedores do Skype terem decidido sua estratégia. O jogo pode então ser representado em forma extensiva, e está ilustrado na figura 2.9.

Observe que a árvore do jogo possui três nós internos e quatro nós terminais. Cada nó interno está representado com o jogador que deverá tomar uma decisão naquele ponto do jogo (i.e., letra “S” para Skype e letra “V” para VivaVoz). A cada ponto do jogo, cada jogador possui duas opções, adotar o protocolo  $W$  (Windows Media Audio, WMA) ou o protocolo  $S$  (Speex). Cada nó terminal da árvore representa um resultado do jogo, onde cada jogador recebe uma recompensa. A notação  $x, y$  denota que o primeiro jogador (Skype, neste caso) receberá a recompensa  $x$ , enquanto o segundo receberá a recompensa

<sup>1</sup> Vivavoz é um aplicativo para transmissão de voz pela Internet do conjunto de ferramentas *freemeeting* desenvolvido no LAND/UFRJ.





**Figura 2.9. Jogo do Skype contra Vivavoz**

y. Dado o jogo acima, qual decisão você tomaria se fosse desenvolvedor do Skype? E se fosse desenvolvedor do Vivavoz?

Antes de começar a analisar o jogo, precisamos definir uma estratégia no contexto de jogos em forma extensiva. Observe que a representação em forma extensiva requer que os jogadores tomem decisões em diversos pontos do jogo, mais especificamente, em todos os possíveis pontos do jogo onde o jogador poderá ter que jogar. Assim sendo, a escolha de estratégia de um jogador deve dizer exatamente o que este deve fazer em cada ponto do jogo, independente do que os outros jogadores escolherem. O conjunto de todas as estratégias disponíveis para um jogador corresponde então a todas possíveis combinações das opções disponíveis em cada ponto do jogo.

No exemplo acima, as estratégias disponíveis para o Skype são duas,  $\{W, S\}$ , pois ele possui apenas um ponto de tomada de decisão, no início do jogo. Entretanto, o jogador Vivavoz possui dois possíveis pontos de tomada de decisão: (i) que decisão tomar, caso o Skype escolha  $W$ ; (ii) que decisão tomar, caso o Skype escolha  $S$ . Assim sendo, a estratégia do Vivavoz é dada pelo par  $xy$ , onde  $x$  corresponde ao que fazer caso o Skype escolha  $W$  e  $y$  corresponde ao que fazer caso o Skype escolha  $S$ . Como tanto  $x$  quanto  $y$  podem assumir o valor  $W$  ou  $S$  (dado pela árvore do jogo), temos que o conjunto de estratégias para o Vivavoz é dado por  $\{WW, WS, SW, SS\}$ . Estas estratégias podem ser lidas da seguinte forma:

- $WW$  - escolher  $W$  independente do que o Skype adotar
- $WS$  - escolher exatamente o que o Skype adotar
- $SW$  - escolher exatamente o contrário que o Skype adotar
- $SS$  - escolher  $S$  independente do que o Skype adotar

Voltemos agora à análise do jogo. Podemos observar que o melhor para o Vivavoz seria que o Skype adotasse o padrão Speex, pois desta forma a recompensa para o Vivavoz, ao adotar também o padrão Speex, seria a mais alta possível. Porém, o melhor para o Skype é adotar o padrão WMA, caso o

Vivavoz também adote este padrão em seguida, pois receberia assim a maior recompensa possível. Observe que dado que o Skype adotou o padrão WMA, é melhor para o Vivavoz também adotar este padrão. Então, uma solução razoável para este jogo seria o Skype escolher a estratégia  $W$  e o Vivavoz escolher a estratégia  $WS$ . Como veremos em breve, este resultado é um equilíbrio de Nash, porém, este jogo possui ainda outros equilíbrios de Nash.

### 2.2.12. Equivalência entre as formas normais e extensivas

Todo jogo em forma extensiva pode ser representado em forma normal, e vice-versa. A forma extensiva apenas facilita a representação de jogos em seqüência, onde os jogadores alternam suas decisões. Para transformar um jogo de forma extensiva em forma normal, é necessário determinar o conjunto de jogadores, as estratégias disponíveis para cada jogador, os possíveis resultados do jogo, e finalmente, a recompensa para cada jogador em cada resultado do jogo. O conjunto de jogadores é idêntico em ambos os casos. O conjunto de estratégias de cada jogador é dado por todas as possíveis escolhas que o jogador pode tomar em cada ponto do jogo. Os possíveis resultados do jogo são dados pelos nós terminais. Finalmente, as recompensas são dadas pelos resultados que são obtidos quando cada jogador escolhe uma estratégia.

É importante notar que se o jogo possuir apenas dois jogadores, e o conjunto de opções para cada jogador em cada ponto do jogo for discreto, então o jogo equivalente em forma normal pode ser descrito por uma matriz.

		Vivavoz			
		WW	WS	SW	SS
Skype	w	5, 2	5, 2	3, 1	3, 1
	s	2, 1	4, 4	2, 1	4, 4

**Figura 2.10. Representação em forma normal do jogo do Skype contra Vivavoz**

Vejamos o exemplo do jogo entre o Skype e o Vivavoz. O jogo equivalente em forma normal está representado na figura 2.10. Observe que o primeiro jogador, Skype, está representado nas linhas da matriz, enquanto o segundo jogador, Vivavoz, está representado nas colunas. As estratégias do primeiro jogador (no caso, Skype) correspondem às linhas da matriz, enquanto as estratégias do segundo jogador (no caso, Vivavoz) correspondem às colunas da matriz. As recompensas para cada escolha de estratégia são obtidas através da árvore do jogo. Por exemplo, se o Skype escolher  $S$  e o Vivavoz escolher  $SW$ , então o resultado do jogo é dado pelo terceiro nó terminal da árvore, que possui recompensa 2, 1. Chegamos a este resultado pois ao escolher  $S$  o Skype posiciona o jogo no nó à direita (segundo nível da árvore). A estratégia

$SW$  do Vivavoz determina a escolha  $W$  quando o Skype escolher  $S$  (indicado pela segunda letra do par  $SW$ ). Isto então nos leva ao resultado indicado, com recompensa 2, 1. As recompensas dos outros resultados da matriz do jogo são obtidos da mesma forma.

### 2.2.13. Equilíbrios de Nash e ameaças inacreditáveis

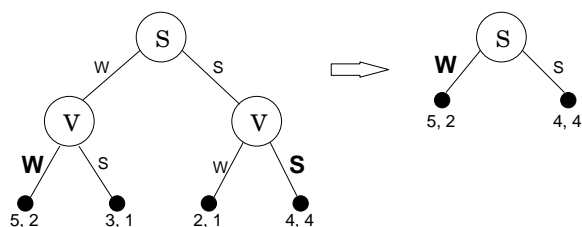
Podemos verificar que o jogo em forma normal entre o Skype e o Vivavoz possui três equilíbrios de Nash, dados pelas seguintes escolhas de estratégias:  $\{W, WS\}$ ,  $\{W, WW\}$  e  $\{S, SS\}$ . Na discussão do jogo em forma extensiva, apresentada acima, vimos que a escolha de estratégia  $\{W, WS\}$  parecia razoável como equilíbrio do jogo. Porque então o jogo possui estes outros equilíbrios? Isto ocorre por causa de *ameaças inacreditáveis* (ou ameaças sem créditos), conforme agora descrevemos.

Como o Skype deve agir caso o Vivavoz decida adotar o padrão Speex, independente do padrão adotado pelo Skype? Neste caso, seria melhor para o Skype adotar também o padrão Speex. Observe que este resultado é melhor para o Vivavoz, que obterá assim sua maior recompensa no jogo. Esta linha de raciocínio leva ao equilíbrio de Nash  $\{S, SS\}$ . Porém, não há motivos concretos para o Skype temer a ameaça do Vivavoz. Observe que se o Skype decidir adotar o padrão WMA, então é melhor para o Vivavoz também adotar o padrão WMA. Sendo o Vivavoz um jogador racional, que deseja maximizar seu ganho, a ameaça de decidir pelo Speex independente da escolha do Skype é inacreditável.

Concluimos assim que dentre os três equilíbrios de Nash existentes neste jogo, apenas um deles é “razoável”, o equilíbrio dado por  $(W, WS)$ . Os outros dois equilíbrios de Nash são ameaças inacreditáveis. É importante notar que, em geral, jogos em forma extensiva possuem muitos equilíbrios de Nash. Entretanto, a maioria destes equilíbrios são irrelevantes, conforme ilustrado em nosso exemplo simples. Uma técnica para descartar tais equilíbrios fúteis e identificar o equilíbrio de Nash mais adequado para o jogo é a *indução de trás-para-frente*, que apresentamos a seguir.

### 2.2.14. Indução de trás-para-frente

A indução de trás-para-frente é uma técnica que pode ser aplicada a jogos em forma extensiva para remover equilíbrios de Nash que não são razoáveis, como as ameaças inacreditáveis. Em alguns casos a técnica é capaz de preservar apenas um equilíbrio de Nash, eliminando todos os outros. A indução de trás-para-frente é um método iterativo que elimina um nível da árvore do jogo a cada iteração, começando pelos nós terminais, até que reste somente a raiz da árvore. Em cada iteração, os nós pais dos nós terminais se tornam nós terminais, e os nós terminais são eliminados da árvore do jogo. A recompensa associada a um nó pai é dada pela melhor opção disponível para o jogador naquele ponto, que corresponde a algum nó terminal. A opção escolhida pelo nó pai é anotada e o processo recomeça. Ao final, temos a estratégia que cada jogador deve seguir, que é dada pelas anotações feitas em cada nó da árvore.



**Figura 2.11. Exemplo de indução de trás-para-frente no jogo do Skype contra Vivavoz**

Vejam o exemplo do jogo entre o Skype e o Vivavoz. A figura 2.11 ilustra o algoritmo de indução trás-para-frente. Começando com a árvore original, eliminamos os nós terminais, tornando os pais destes nós novos nós terminais. A melhor escolha para o nó não-terminal à esquerda é  $W$ , o que leva a uma recompensa de 5,2. A melhor escolha para o nó não-terminal à direita é  $S$ , o que leva a uma recompensa de 4,4. A escolha de estratégia do jogador 2 neste ponto do jogo é então  $WS$ . O processo de indução inicia uma nova iteração, com a nova árvore de jogo ilustrada na figura. A melhor opção para o jogador 1 é a escolha  $W$ , que oferece a recompensa 5,2. Tendo chegado à raiz da árvore do jogo, o processo termina. Temos então, a estratégia  $\{W, WS\}$ , que como vimos, é um equilíbrio de Nash do jogo.

É importante ressaltar que a indução de trás-para-frente está fortemente relacionada com a eliminação de estratégias dominadas em jogos em forma normal, apresentada na seção 2.2.2. Em particular, a indução de trás-para-frente pode eliminar também estratégias fracamente dominadas.

### 2.2.15. Líderes e Seguidores

No jogo acima vimos que o Skype sai em vantagem em relação ao Vivavoz, pois ao escolher a estratégia  $W$  antes do Vivavoz fazer sua escolha, ele condiciona o Vivavoz a escolher também a estratégia  $W$ . Mas o que aconteceria se o Vivavoz jogasse primeiro? O leitor deve construir este jogo utilizando a árvore ilustrada na figura 2.9, e trocando a ordem de tomada de decisão. Qual equilíbrio de Nash sobrevive à indução de trás-para-frente neste novo jogo? Se você concluiu que o Vivavoz deve escolher a estratégia  $S$  e Skype escolher a estratégia  $WS$ , então você acertou. Observe que neste caso o resultado do jogo terá recompensa (4,4), o que é um resultado melhor para o Vivavoz.

O mais interessante neste caso é que a ameaça inacreditável do Vivavoz de jogar  $S$  independe do que o Skype fizer passar a ser crível. Ou seja, ao jogar primeiro e escolher a estratégia  $S$ , o Vivavoz se impõem no jogo e força o Skype a decidir também por  $S$ .

Este poder de jogar primeiro, neste caso, traz vantagens para ambos os jogadores. Observe que os dois preferem decidir primeiro, pois desta forma

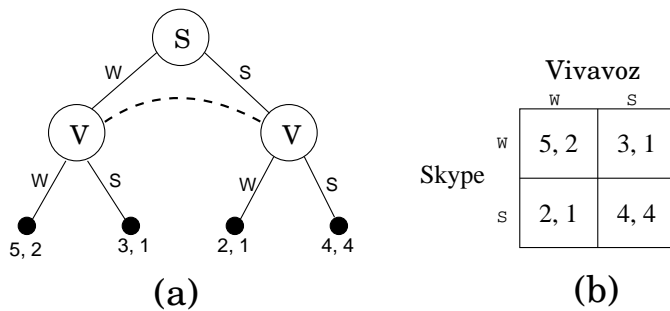
condicionam o outro jogador. Entretanto, nem sempre é vantajoso jogar primeiro. Há jogos onde ambos os jogadores preferem ser o segundo a decidir, pois assim poderão obter uma melhor recompensa.

### 2.2.16. Informação perfeita e imperfeita

Nos jogos extensivos que apresentamos até agora assumimos implicitamente que ao tomar uma decisão em algum determinado ponto do jogo, os jogadores conhecem perfeitamente todas as decisões tomadas anteriormente por todos os outros jogadores. Ou seja, as decisões tomadas por cada jogador a cada instante do jogo eram de conhecimento público no jogo. Este tipo de jogo extensivo é denominado *jogos com informação perfeita*, e pode ser definido da seguinte maneira:

**Definição 12** Um jogo extensivo é denominado jogo com informação perfeita se todos os jogadores possuem conhecimento preciso de todas as tomadas de decisão feitas por todos os outros jogadores em qualquer ponto do jogo.

Entretanto, nem todos os jogos extensivos são de informação perfeita. Por exemplo, é possível que em algum determinado ponto do jogo, dois jogadores tenham que tomar uma decisão simultânea. Neste caso, cada jogador desconhece a escolha feita pelo outro naquele momento do jogo. Um outro exemplo ocorre quando uma decisão tomada por um jogador não é revelada imediatamente ao outro jogador, que então terá de fazer sua escolha sem conhecimento da decisão tomada pelo outro. Este tipo de jogo extensivo é denominado *jogos com informação imperfeita*. Desta forma, todo jogo extensivo pode ser classificado como sendo de informação perfeita ou imperfeita.



**Figura 2.12. Jogo do Skype contra o Vivavoz com informação imperfeita**

Jogos extensivos com informação imperfeita podem ser representados graficamente unindo com um arco os nós da árvore do jogo onde o jogador terá de tomar sua decisão sem conhecimento da escolha feita pelo outro. Desta forma,

o jogador não sabe, no momento da decisão, em que ponto da árvore o jogo se encontra. Vejamos o exemplo do jogo do Skype contra o Vivavoz, onde o segundo jogador desconhece a decisão feita pelo primeiro. O jogo está ilustrado na figura 2.12.a. Repare que os nós da árvore do jogo onde o Vivavoz precisa tomar sua decisão estão unidos com um arco. Ou seja, o Vivavoz não sabe exatamente em que ponto da árvore ele se encontra no momento da tomada de decisão. Entretanto, ao tomar a decisão o jogo termina e então a decisão do Skype é revelada, levando ao resultado do jogo apropriado.

É importante ressaltar que a forma normal deste jogo extensivo também será alterada. Em particular, o Vivavoz não mais condiciona sua decisão na escolha feita pelo Skype. Isto porque o Vivavoz agora desconhece esta escolha. Assim sendo, o Vivavoz possui então apenas duas estratégias, ao invés de quatro. A representação em forma normal deste jogo está ilustrada na figura 2.12.b. Observe que o jogo possui dois equilíbrios de Nash em estratégias puras. Um destes equilíbrios possui a mesma recompensa que o equilíbrio de Nash obtido no jogo com informação perfeita.

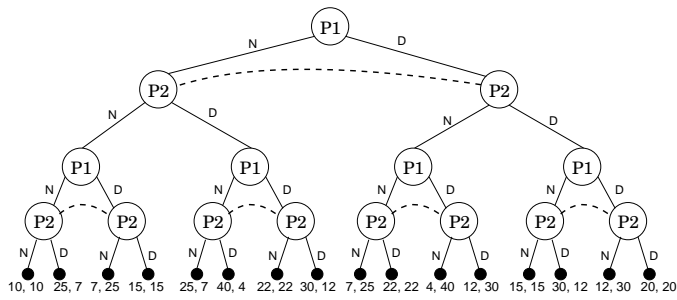
### 2.2.17. Jogos repetidos

Os jogos que apresentamos até agora são caracterizados por apresentarem um único confronto entre os jogadores. Ou seja, os jogadores jogam o jogo uma única vez e pronto! Os jogadores recebem suas recompensas e então seguem a vida sem nunca mais se confrontar novamente. Entretanto, há situações onde o mesmo confronto entre dois jogadores pode se repetir. Este tipo de situação é modelado utilizando o conceito de *jogos repetidos*.

É importante notar que um jogo jogado múltiplas vezes entre os mesmos jogadores pode possuir características bem diferentes do jogo jogado uma única vez. Por exemplo, os jogadores podem “aprender” a estratégia dos outros jogadores e tentar assim obter resultados que levem a recompensas maiores.

Um jogo repetido pode ser facilmente descrito em forma extensiva, onde o jogo base é repetido múltiplas vezes na árvore do jogo. Cada iteração do jogo base é chamado de *fase* do jogo. Um jogo onde os jogadores se confrontam  $k$  vezes com o mesmo jogo é um jogo repetido com  $k$  fases. A recompensa ao final do jogo é dada pelo soma das recompensas obtidas nas  $k$  fases do jogo.

Vejamos um exemplo do jogo do dilema do prisioneiro onde os dois suspeitos respondem agora por dois crimes conjuntos e serão interrogados simultaneamente duas vezes. Entretanto, o resultado da primeira interrogação é conhecido por ambos antes deles começarem a segunda interrogação. Esta situação pode ser modelada por um jogo repetido com duas fases. O jogo em forma extensiva está ilustrado na figura 2.13 e o jogo base em forma normal correspondente está ilustrado na figura 2.7. A recompensa ao final do jogo é a soma das recompensas obtidas em cada uma das duas fases do jogo (lembre-se que a recompensa aqui é medida em anos de cadeia). Observe que os dois prisioneiros terão de tomar uma decisão duas vezes, correspondentes ao primeiro jogo e ao segundo jogo base. Entretanto, a segunda decisão poderá



**Figura 2.13. Jogo do dilema do prisioneiro repetido duas vezes**

estar condicionada às decisões tomadas anteriormente, o que levará a um número muito maior de estratégias. Vamos definir o conjunto de estratégias para cada jogador. Na primeira decisão os jogadores não possuem nenhuma informação anterior sobre o jogo, e por isto possuem apenas duas alternativas: “D” (delatar) ou “N” (não delatar). Ao tomar a segunda decisão os jogadores têm conhecimento do resultado do primeiro jogo base, ao qual poderão condicionar suas decisões. Como o primeiro jogo base possui quatro resultados diferentes, a segunda decisão estará condicionada a cada um destes resultados, e pode ser representada por  $xyz$ , onde  $x, y, w$  e  $z$  são as decisões a ser tomadas quando o resultado do primeiro jogo base for “DD”, “DN”, “ND” e “NN”, respectivamente. A estratégia de cada jogador é então dada por  $\{v, xyz\}$ , onde  $v$  é a primeira decisão e  $xyz$  é a segunda decisão, condicionada no resultado do primeiro jogo base. Como em cada ponto de decisão o jogador possui duas alternativas (“D” ou “N”), o conjunto de estratégias possui  $2^5 = 32$  estratégias.

É importante ressaltar que em jogos extensivos, uma determinada estratégia de um jogador não está condicionada no desenrolar do jogo. A estratégia precisa fornecer uma escolha específica em cada ponto de decisão da árvore do jogo, independente se este ponto pode ou não ser alcançado dado a estratégia. Desta forma, uma dada estratégia contém decisões para partes da árvore do jogo que jamais serão alcançadas sobre a mesma estratégia. Por exemplo: considere a estratégia  $\{D, DDNN\}$  para o primeiro prisioneiro no jogo acima. Dado que a estratégia determina “D” para a primeira decisão, as escolhas “N” e “N” para a segunda decisão jamais serão alcançadas, uma vez que estas requerem que a primeira escolha seja “N”. Este é um problema fundamental de jogos extensivos, pois leva a um número muito grande de estratégias, sendo que muitas destas são equilíbrios de Nash de pouco significado, como as ameaças inacreditáveis que vimos na seção 2.2.13.

Voltando ao exemplo do jogo repetido, como você jogaria? Qual solução é recomendada pela teoria? Podemos fazer uma análise do jogo e constatar que há vários equilíbrios de Nash. Entretanto, podemos aplicar a técnica de

indução de trás-para-frente para obter um equilíbrio de Nash mais robusto. O equilíbrio de Nash neste caso é dado por sempre delatar, independente do que o outro fizer! O equilíbrio é dado pelas estratégias  $\{D, DDDD\}$  para ambos os jogadores, o que leva ao resultado 20,20, que não é um bom resultado para ambos os jogadores. Veremos porque isto ocorre.

Repare que independente do resultado dos jogos base anteriores, na última fase do jogo repetido, é melhor um jogador delatar o outro. Isto levará a uma pena acumulada menor para este jogador. Assim como é melhor delatar na última fase, é também melhor para um jogador delatar na penúltima fase, pois levará a uma recompensa acumulada menor. Seguindo este raciocínio acabamos chegando à primeira fase do jogo sempre delatando o outro jogador. Este resultado, aliás, é independente do número de fases. Assim sendo, mesmo em um cenário onde os prisioneiros se confrontam múltiplas vezes, o raciocínio individual leva a um resultado ruim para ambos. Porém, veremos em breve que isto nem sempre ocorre, em particular quando o futuro valer menos que o presente, mas for suficientemente importante.

### 2.2.18. Jogos repetidos infinitamente

Os jogos repetidos apresentados na seção anterior possuem um determinado número de fases, que corresponde ao número de confrontos entre os mesmos jogadores. Nesta seção, consideraremos jogos onde o número de confrontos é muito grande, em particular, infinito. O objetivo é modelar situações onde o confronto é constante e freqüente.

Um jogo repetido infinitamente pode ser representado em forma extensiva, onde a árvore do jogo é infinita. Entretanto, em cada fase do jogo repetido, os jogadores se confrontam no jogo base. A recompensa de cada jogador é dada pela recompensa acumulada em todas as fases anteriores. Diferentemente dos jogos repetidos finitos, a recompensa obtida em uma determinada fase será ponderada por um *fator de desconto*  $\delta$ , com  $0 < \delta < 1$ . O objetivo é modelar o fato de que uma recompensa obtida no futuro vale menos que uma recompensa igual obtida no presente. Em particular, uma recompensa de valor  $c$  obtida  $k$  fases no futuro, vale no presente  $c\delta^k$ . Assim sendo, se um jogador obtém uma recompensa de valor exatamente  $c$  em todas as fases do jogo, sua recompensa acumulada em qualquer instante do jogo é dada por:

$$\sum_{i=0}^{\infty} c\delta^i = \frac{c}{1-\delta} \quad (11)$$

Como vimos anteriormente, jogos extensivos possuem um conjunto muito grande de estratégias. Em jogos repetidos infinitamente, este conjunto é necessariamente infinito. Entretanto, podemos definir uma classe de estratégias bastante simples, conhecida como “estratégias de gatilho”.

**Definição 13** *Uma estratégia é denominada estratégia de gatilho se esta determina uma escolha fixa até que uma condição seja satisfeita e outra escolha fixa para o restante do jogo.*



Por exemplo, considere o jogo do dilema do prisioneiro repetido infinitamente. Em cada fase do jogo, os jogadores precisam decidir, simultaneamente, se delatam ou não o outro. A estratégia “não delatar enquanto o outro jogador não delatar e delatar para sempre após o outro delatar” é uma estratégia de gatilho. Outra estratégia de gatilho é “delatar até a vigésima fase e não delatar para sempre depois”. Estratégias de gatilho são bastante simples e por isto podem ser analisadas mais facilmente.

Veremos agora um resultado bastante importante que ilustra a idéia de que a repetição propicia a cooperação, mesmo quando interesses individuais dominam os interesses do grupo. Considere o jogo do dilema do prisioneiro repetido infinitamente. Temos o seguinte teorema:

**Teorema 1** *Se  $\delta$  for suficientemente grande, a estratégia de gatilho “não delatar enquanto o outro jogador não delatar e delatar para sempre após o outro delatar” escolhida por cada jogador é um equilíbrio de Nash onde o resultado do jogo é nunca delatar.*

A prova deste teorema é simples e vamos ilustrá-la com base no exemplo do jogo do dilema do prisioneiro. Para verificar que um determinado par de estratégias forma um equilíbrio de Nash, basta verificar que nenhuma outra estratégia é capaz de dar uma recompensa melhor a um dos jogadores. A recompensa acumulada obtida quando os dois utilizam a estratégia de gatilho acima é dada por  $\sum_{i=0}^{\infty} 5\delta^i = 5/(1-\delta)$ , uma vez que o resultado de todos os jogos base é nunca delatar. Vamos supor que um dos jogadores escolha outra estratégia, no caso o primeiro prisioneiro. Para que se obtenha um resultado diferente, esta outra estratégia necessariamente envolve a escolha “delatar” em pelo menos uma das fases. A partir desta fase, o segundo jogador irá delatar para o restante do jogo, independente do que o primeiro venha a jogar. Assim sendo, o melhor para o primeiro jogador é também delatar no restante do jogo. Sem perda de generalidade, considere a fase zero como sendo a fase onde o primeiro jogador delata pela primeira vez. A recompensa para o primeiro jogador nesta fase é 2, e em todas as fases seguintes será 10. Assim sendo, a recompensa acumulada do primeiro jogador, ponderada pelo fator de desconto é dada por  $2 + \sum_{i=1}^{\infty} 10\delta^i = 2 + 10\delta/(1-\delta)$ . Desta forma, a estratégia onde os dois nunca delatam é um equilíbrio de Nash se e somente se  $5/(1-\delta) < 2 + 10\delta/(1-\delta)$ , pois neste caso nenhuma outra estratégia é capaz de oferecer uma recompensa menor (pois estamos falando de anos de cadeia) ao primeiro prisioneiro. Resolvendo em função de  $\delta$ , temos que se  $\delta > 3/8$ , então a estratégia onde os dois nunca delatam é um equilíbrio de Nash. É fácil constatar que este resultado pode ser generalizado para quaisquer valores de recompensa para o dilema do prisioneiro.

Assim vimos, finalmente, que apesar dos interesses individuais, é possível que a repetição infinita de um mesmo jogo tenha como equilíbrio estratégias onde os jogadores cooperam mutuamente. Entretanto, isto só foi possível num cenário onde o jogo se repete infinitamente, onde o futuro é suficientemente

importante e sobre a ameaça de nunca mais cooperar. Mas infelizmente, esta estratégia de gatilho não é a única que leva a um equilíbrio de Nash. Aliás, existem infinitas outras estratégias que também formam equilíbrios de Nash deste jogo.

### **2.3. Teoria dos Jogos Evolucionária**

Nas seções anteriores estudamos os conceitos básicos da Teoria de Jogos clássica, incluindo o equilíbrio de Nash dos jogos. Entretanto, uma questão fundamental não foi discutida: de que forma um equilíbrio de Nash é atingido? Além disso, caso mais de um equilíbrio exista, qual será o escolhido? Note que a teoria clássica dos jogos assume que os jogadores, de alguma forma, sejam capazes de determinar o equilíbrio de Nash. Nada é dito entretanto, sobre como o processo pelo qual o jogo (que modela um sistema) atinge esse equilíbrio. A teoria dos jogos evolucionária surge como uma alternativa à teoria clássica para explicar o comportamento dos sistemas que, em geral, são dinâmicos, isto é, evoluem com o tempo.

Em um cenário dinâmico, os jogadores estão constantemente interagindo uns com os outros e adaptando as suas escolhas de acordo com o ganho que eles recebem. Um exemplo desta dinâmica, ou jogo evolucionário, é o descrito na seção 2.1, onde usuários de uma ferramenta VoIP disputam a capacidade de transmissão da rede.

Imagine que exista uma ferramenta de VoIP que permita ao usuário escolher a taxa de compressão de áudio e tipo de redundância para a recuperação de dados. Este é o caso do Vivavoz, do conjunto de ferramentas *freemeeting de vídeo-conferência*. Considere um cenário onde vários usuários do Viva-voz compartilhem os mesmos recursos de rede (ex. canais de comunicação). Suponha que, por *default*, o aplicativo VoIP inicie com a mais alta taxa de compressão disponível e sem a utilização de redundância nos dados. A taxa de transmissão inicial é portanto a mais baixa possível mas, por outro lado, a qualidade de voz será baixa, mesmo sem congestionamento. Como os usuários irão perceber a baixa qualidade, é esperado que eles procurem diminuir a taxa de compressão. O resultante aumento das taxas de transmissão pode causar congestionamento, e o ganho almejado de qualidade poderá ser nulo ou até negativo. Alguns usuários podem começar a utilizar redundância para contrabalançar as perdas de pacotes devido ao congestionamento. Estes terão sua qualidade melhorada, mas às custas da queda de qualidade dos usuários que ainda não utilizaram esse recurso. Esse cenário é perfeitamente possível nos dias de hoje, e não é nada óbvio qual será o comportamento dos usuários e a qualidade final da voz recebida por cada um. Caso os usuários atinjam taxas de transmissão elevadas pelo uso de menos compressão e mais redundância, o congestionamento resultante poderá implicar uma péssima qualidade de voz para todos. Por outro lado, se os usuários utilizarem uma taxa de compressão e redundância “adequadas”, é provável que a qualidade geral seja aceitável para todos. A pergunta é: qual a estratégia a ser usada pelos usuários? Haverá um

colapso na rede? Ou as estratégias adotadas convergirão para um ponto tal que a qualidade geral será aceitável? Será que alguns dos usuários receberão excelente qualidade enquanto que outros serão muito prejudicados? A Teoria dos Jogos evolucionária pode tentar equacionar essas perguntas.

A título de exemplo numérico, suponha que a ferramenta permita que o usuário escolha dois tipos de compressão de áudio, com taxas de 8 Kbps e 24 Kbps, ou ainda PCM (sem compressão), com taxa de 64 Kbps. Suponha ainda que o usuário possa adicionar redundância, ao transmitir em PCM, para recuperar possíveis perdas de pacotes. A redundância implementada é tal que dobra a taxa de transmissão. Em resumo, cada usuário dispõe de 4 escolhas: A, B, C, D, correspondendo a taxas de transmissão de 128 Kbps, 64 Kbps, 24 Kbps, 8 Kbps, respectivamente. No caso deste exemplo, taxas de transmissão maiores estão associadas a uma melhor qualidade do som, desde que não exista congestionamento no canal compartilhado. Suponha que o “jogo” comece com alguns usuários utilizando uma alta taxa de compressão, sem redundância (usuários da classe D) e outros estejam transmitindo na mais alta taxa possível (classe A, com redundância). Os usuários classe D ouvirão o som com qualidade ruim, e tenderão a diminuir a compressão, com conseqüente aumento das taxas de transmissão. Suponha ainda que o canal compartilhado não tem capacidade suficiente para suportar todos os usuários transmitindo na maior taxa possível (isto é, todos os usuários sendo de classe A). À medida que os usuários aumentam as suas taxas o canal congestionado, piorando a qualidade de voz para todos, devido ao descarte de pacotes na fila do canal congestionado. Cada classe de usuário será influenciada pelo congestionamento de acordo com a classe atual do conjunto. O sistema evoluirá com os usuários tentando adaptar suas escolhas de codificação e redundância para obter a melhor qualidade de áudio possível (a melhor recompensa individual). O problema é o de determinar o processo pelo qual os usuários escolherão suas estratégias ao longo do tempo (o grau de compressão e a redundância usada), de acordo com a qualidade do som recebida (a função utilidade). Existirá ou não um ponto de equilíbrio? Como os usuários adaptam as suas estratégias?

Alguns pontos devem ser enfatizados de forma a entender as diferenças entre a teoria clássica e a evolucionária. Na teoria evolucionária os participantes estão continuamente interagindo como no exemplo acima, e mudando as suas escolhas de acordo com a recompensa recebida. É parte do *framework* de modelagem determinar o processo de adaptação dos jogadores e de como estes alcançam um possível equilíbrio. Os jogadores selecionam suas estratégias de acordo com a história do jogo e possivelmente com a sua “classe”. Não é preciso supor que os jogadores joguem todos de forma racional, apenas para maximizar a sua estratégia. “Erros” são permitidos.

Um jogo evolucionário na sua forma mais simples consiste em jogadores com escolhas idênticas e funções de utilidade simétricas, como no caso do exemplo acima. A definição abaixo formaliza essas idéias.

**Definição 14** Um jogo evolucionário simétrico  $\mathcal{J} = (N, S, \mathbf{u}, \sigma(t)), \mathcal{D}$ , é composto de:

- Um conjunto  $N$  de jogadores, que são agrupados em classes para jogar, num total de  $M$ .
- Um conjunto de estratégias  $S$ , disponíveis para cada jogador, com  $k = |S|$ .
- Uma função utilidade  $u_i : S^M \rightarrow \mathfrak{R}$  para cada jogador  $i \in N$ .
- Um vetor de valores  $\sigma(t) = (\sigma_1(t), \dots, \sigma_k(t))$ , onde  $\sigma_i(t)$  é a fração da população que seleciona a estratégia  $s_i \in S$  no instante  $t$ .
- Um jogo base  $\mathcal{J}(t) = (M, S, \mathbf{u})$  (stage game no tempo  $t$ ).
- Um processo dinâmico de ajuste de estratégia que é uma função, talvez estocástica,  $\mathcal{D} : (\sigma(t), \mathcal{J}(t), t) \rightarrow \sigma(t + \Delta t)$  (o processo de seleção dinâmica).

Note que os jogadores são agrupados em  $M$  classes para o jogo base que é jogado de acordo com  $\sigma(t)$ . (Em todos os exemplos deste texto,  $M = k$ , isto é, o número de classes será igual ao das estratégias disponíveis.) Das recompensas obtidas (por exemplo a qualidade de voz no exemplo acima) os jogadores ajustam sua estratégia conforme a dinâmica  $\mathcal{D}$ . Não se espera que o equilíbrio de Nash seja adotado a longo prazo, exceto se existir uma explicação plausível da dinâmica do jogo. O processo de adaptação dos jogadores é o ponto-chave do problema.

Podemos usar várias ferramentas analíticas para ajudar-nos a compreender as interações de sistemas dinâmicos. Equações diferenciais e cadeias de Markov são duas destas ferramentas. Mostraremos seu uso nos exemplos que se seguem assim como a dinâmica através da qual um sistema do tipo acima evolui com o tempo.

Façamos simplificações na ferramenta Vivavoz (parte do sistema *freemee-ting*) para clarificar alguns dos conceitos acima. No sistema simplificado existem apenas dois usuários competindo pelo canal. Apenas as opções A, B e C estão disponíveis. A Tabela 2.3 indica as recompensas recebidas por usuários de cada classe, quando o sistema encontra-se em uma dos 6 estados possíveis, isto é, para os possíveis valores de  $\sigma$ . (Portanto, estado neste caso é o vetor cujo componente  $i$  indica o número de usuários aplicando a estratégia  $i$ . Note que o vetor de estado num instante  $t$  é  $N\sigma(t)$ , onde  $N = 2$  e constante no tempo.)

As recompensas não foram selecionadas ao acaso. Procurou-se escolher cada uma igual ao valor de MOS<sup>2</sup> semelhante ao que se obteria em cada uma das situações, para cada classe de usuários, para um dado cenário. Note que

<sup>2</sup> MOS (Mean Opinion Score) é uma medida de qualidade de voz que varia aproximadamente de 1 a 4,5, sendo que qualidades aceitáveis são acima de 3,5.

Estado $N\sigma$	Utilidade		
	classe A	classe B	classe C
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

Tabela 2.1. Recompensas obtidas do “jogo” do Vivavoz

quando dois usuários estão transmitindo a taxa máxima (estado  $(2,0,0)$ ), a qualidade é baixa  $(2,0)$  devido ao congestionamento. Quando um transmite a taxa alta e outro a taxa média, a qualidade de ambos aumenta relativamente ao estado  $(2,0,0)$ , pois o congestionamento é reduzido com a diminuição de taxa do segundo usuário. Por outro lado, o segundo terá qualidade inferior ao primeiro, pois está transmitindo com menor taxa (e portanto maior taxa de compressão de dados). As outras entradas da tabela foram escolhidas da mesma forma.

O jogo base em um instante  $t$ ,  $\mathcal{J}(t)$  é determinado supondo que pares formados pelos grupos de usuários repetidamente jogam este jogo de acordo com as recompensas acima. Note que as recompensas do jogo base dependem do estado no sistema em  $t$ ,  $N\sigma(t)$ . Suponha que  $N = 2$ . Então  $\mathcal{J}$  é:

	classe A	classe B	classe C
classe A	2,0; 2,0	3,0; 2,1	3,5; 2,3
classe B	2,1; 3,0	4,0; 4,0	3,8; 3,1
classe C	2,3; 3,5	3,1; 3,8	3,3; 3,3

Tabela 2.2. Jogo base do exemplo simplificado do Vivavoz

Considere que o processo dinâmico de escolha das estratégias  $\mathcal{J}$  seja tal que a taxa de transição (em indivíduos por unidade de tempo) de um usuário de uma classe para outra seja proporcional à diferença entre a recompensa deste usuário no estado presente em relação à obtida por este usuário na nova classe. Além disso, suponha que a probabilidade de dois usuários mudarem de classe ao mesmo tempo seja nula. A Figura 2.14 ilustra o diagrama de transição entre possíveis estados do sistema, assim como as taxas entre os estados deste jogo. Por exemplo, a taxa de transição de  $(2,0,0)$  para  $(1,1,0)$  é igual ao dobro da diferença entre as recompensas de um usuário na classe B em relação à sua na classe A, uma vez que existem dois usuários que poderiam mudar suas escolhas de A para B.

Pelo diagrama de transição da Figura 2.14, é fácil verificar que o sistema

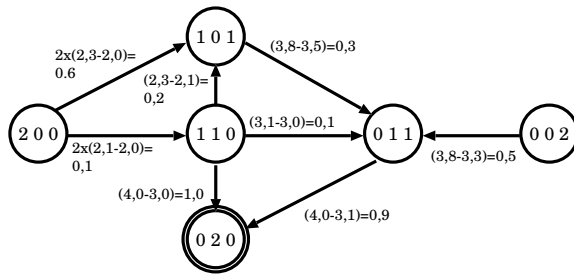


Figura 2.14. Exemplo da dinâmica de um sistema

evolui naturalmente para o estado  $(0, 2, 0)$ , isto é, com os dois usuários empregando a estratégia que usa as taxas médias de transição. O leitor atento poderá perceber da Figura 2.14 que, a partir de certos estados, mais de uma transição é possível, devido ao processo  $\mathcal{D}$  que foi definido. É possível associar uma probabilidade de transição proporcional às taxas de transição. Neste caso, o diagrama de transição de estados pode representar uma cadeia de Markov de tempo contínuo e espaço de estados discreto e finito. Retornaremos a esse ponto mais adiante

### 2.3.1. A dinâmica do replicador

Um conceito muito importante em sistemas dinâmicos é o do *replicador*. O replicador é uma entidade capaz de fazer cópias de si mesmo. Um sistema replicador é um conjunto de replicadores em um ambiente e uma certa estrutura de interação. O processo de evolução é tal que as estratégias de maior ganho são as que mais rapidamente se reproduzem. Além disso, existe a possibilidade de ocorrência de erros aleatórios (mutações) no processo de replicação, fazendo com que novas classes de replicadores possam surgir em uma população.

Suponha que exista uma população de indivíduos de  $M$  classes, onde indivíduos da classe  $i$  jogam uma determinada estratégia  $s_i$  e podem se reproduzir. Lembre-se que  $\sigma(t) = (\sigma_1(t), \dots, \sigma_M(t))$  é o vetor que fornece a fração de indivíduos em cada classe e portanto,  $\sum_i \sigma_i(t) = 1$ . Seja  $u_i(\sigma)$  a recompensa de indivíduos da classe  $i$  quando a população está dividida de acordo com  $\sigma$ . Suponha que, em um período  $\Delta t$ , um indivíduo de classe  $i$  aprenda com probabilidade  $\lambda \Delta t$  sobre a recompensa de um outro indivíduo da população (digamos um indivíduo de classe  $j$ ) selecionado aleatoriamente e migre para a classe  $j$  caso  $u_j(\sigma)$  seja maior que sua própria recompensa. É razoável supor que a probabilidade de migrar seja proporcional à diferença das recompensas, pois quanto maior a recompensa da outra classe, é de se esperar que maior seja o incentivo da mudança. Seja  $N_i(t)$  a variável igual ao número de indivíduos da classe  $i$  em  $t$  e  $N(t) = \sum_i^M N_i(t)$ .

Dado que  $i$  aprende sobre a recompensa de  $j$ <sup>3</sup>:  $N_i(t + \Delta t) = N_i(t) + \lambda \Delta t N_i(t) (u_i(\sigma(t)) - u_j(\sigma(t)))$ . Descondicionando de acordo com a probabilidade de escolher um indivíduo de classe  $j$ ,  $\sigma_j(t)$ , temos:

$$N_i(t + \Delta t) = N_i(t) + \lambda \Delta t N_i(t) \left( u_i(\sigma(t)) - \sum_{j=1}^M u_j(\sigma(t)) \sigma_j(t) \right) \quad (12)$$

Então, dividindo por  $N(t)$ :

$$\sigma_i'(t) = \lambda \sigma_i(t) (u_i(\sigma(t)) - \bar{u}(\sigma(t))), \quad (13)$$

onde  $\bar{u}(\sigma(t)) = \sum_{j=1}^M u_j(\sigma(t)) \sigma_j(t)$  é a recompensa média da população distribuída nas classes de acordo com  $\sigma(t)$  no instante  $t$ , e  $\sigma_i'(t)$  a derivada de  $\sigma_i(t)$ . (Como o parâmetro  $\lambda$  apenas altera a taxa de ajuste, e não a trajetória, faremos  $\lambda = 1$  no que se segue.) A equação (13) é chamada de *dinâmica do replicador*.

Note que a dinâmica descrita acima não inclui dinâmicas estocásticas, por exemplo, quando existe uma probabilidade de um indivíduo cometer erros e mudar para uma estratégia pior. Para modelar variações estocásticas precisamos usar processos estocásticos como cadeias de Markov.

Um jogo comumente usado para ilustrar a dinâmica do replicador é o denominado de *águia-pombo-burguês*. Neste jogo indivíduos lutam por território. Aqueles que adotam a estratégia *águia* são mais agressivos e lutam até conquistar o território ou ficar seriamente feridos. Os indivíduos *pombo*, por outro lado, não morrem, uma vez que fogem quando uma luta começa. Por outro lado, não conquistam território de uma *águia*, apenas de outro indivíduo com a mesma estratégia. Em um encontro de indivíduos com a mesma estratégia as chances de vitória são idênticas. Finalmente, os indivíduos *burgueses* se comportam como uma *águia* quando detém um território, mas como um *pombo* caso contrário. A probabilidade dos *burgueses* deterem território é 0,5. A tabela abaixo ilustra o jogo, onde  $v$  indica o ganho por conquistar território e  $c$  o custo de ficar ferido.

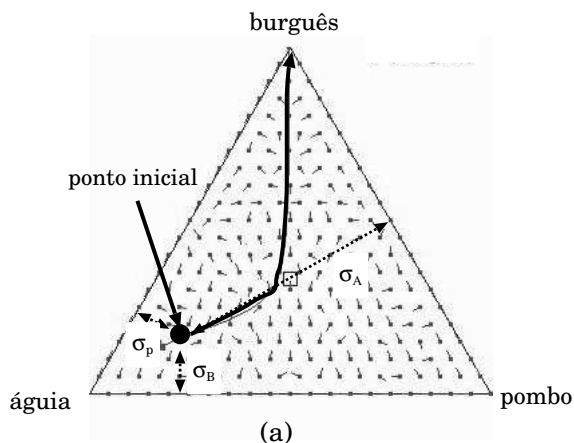
	Águia	Pombo	Burguês
Águia	$\frac{v-c}{2}, \frac{v-c}{2}$	$v, 0$	$\frac{3v}{4} - \frac{c}{4}, \frac{v-c}{4}$
Pombo	$0, v$	$\frac{v}{2}, \frac{v}{2}$	$\frac{v}{4}, \frac{3v}{4}$
Burguês	$\frac{v-c}{4}, \frac{3v}{4} - \frac{c}{4}$	$\frac{3v}{4}, \frac{v}{4}$	$\frac{v}{2}, \frac{v}{2}$

Considere a entrada *águia - águia* da tabela. Um indivíduo tem a metade das chances de ficar ferido, portanto o ganho será  $v/2 - c/2$ . Seja agora a entrada *burguês - águia*. Caso o burguês seja *águia*, o ganho é de  $(v - c)/2$ , idêntico ao de um confronto *águia - águia*. Caso contrário, o burguês se comporta

<sup>3</sup> Estamos abusando um pouco da notação neste ponto, para simplificar a exposição.  $N_i(t)$  é uma variável aleatória neste caso, mas que assume o valor da expressão com probabilidade 1.

como pombo e o ganho é zero. Dessa forma, como as chances do burguês utilizar a estratégia águia são de 50%, o ganho de um indivíduo burguês é de  $0,5(v - c)/2 + 0$  e o da águia  $0,5v + 0,5(v - c)/2 = (3v - c)/4$ . O restante da tabela é explicado da mesma forma.

Suponha que o benefício de uma vitória ( $v$ ) seja igual a 50, e um ferimento grave tenha custo  $c = 100$ . A Figura 2.15 ilustra o resultado da dinâmica do jogo pela simulação da equação 13 quando a população é iniciada por 75% de águia, 11% de pombos e 14% de burgueses<sup>4</sup>. Nesta figura, a fração de



**Figura 2.15. Dinâmica do replicador para o jogo da águia-pombo-burguês**

usuários usando cada estratégia é indicada pelas coordenadas baricêntricas (dadas pela distância do ponto a cada uma das três faces do triângulo). Note que mesmo com uma alta população de águias, o sistema converge para toda população agindo como burguês.

Façamos agora um paralelo entre a dinâmica do replicador e o jogo da ferramenta Vivavoz. Quais as principais diferenças e semelhanças? Seria possível aplicar a equação (13) para aquele problema? A primeira dificuldade no caso do jogo da ferramenta VoIP seria o de determinar o MOS para uma população de  $N$  (grande) usuários, cuja fração em cada classe seja dada por  $\sigma(t)$  em um instante  $t$ . Supondo que não seja difícil determinar o MOS em função de  $\sigma(t)$  para um dado  $N$  (o que não é verdade), então a equação 13 poderia, em princípio, ser aplicada, uma vez que seria fácil calcular o MOS médio no estado  $\sigma(t)$  e saberíamos o MOS da classe  $i$  dado  $\sigma(t)$ . Entretanto, o modelo que resulta na equação (13) tem uma limitação básica. O fluxo de usuários de

<sup>4</sup> A simulação foi feita com o applet em [www.land.ufrj.br/~sadoc/impa](http://www.land.ufrj.br/~sadoc/impa)



uma classe  $i$  para uma outra no instante  $t$  se dá apenas quando a recompensa (MOS, no caso) da classe  $i$  for inferior à recompensa média em  $t$ . Da mesma forma, a classe  $i$  ganha usuários quando a recompensa desta for superior à média atual. A taxa de ganho ou perda é proporcional à fração de usuários naquele instante na classe  $i$ . Este comportamento deve ser contrastado com aquele do modelo indicado no diagrama da Figura 2.14. Neste último, o fluxo de usuários, de  $i$  para qualquer classe, é proporcional ao número de usuários na classe  $i$ , como na dinâmica do replicador. Mas o fluxo para a classe  $i$  a partir de  $j$  contrasta com a dinâmica do replicador, pois é proporcional ao número de usuários da classe  $j$ . Além disso, o fluxo de  $i$  para  $j$  é proporcional ao ganho de recompensa que a classe  $i$  teria pela mudança, e não a média da recompensa em  $t$ .

### 2.3.2. Estratégias evolucionariamente estáveis

O conceito de equilíbrio de Nash pode ser refinado na teoria evolucionária, para determinar o comportamento da população a longo prazo, dando origem à estratégia evolucionariamente estável (*evolutionary stable strategy* - ESS). Informalmente, uma população adotando estratégias segundo uma distribuição  $\sigma$  é ESS se ela não é vulnerável a invasões por indivíduos (perturbações) que conseqüentemente irão alterar a distribuição da população pelas estratégias.

Considere um jogo em sua forma normal com estratégias  $S = \{s_1, \dots, s_M\}$  e recompensas  $u_i(\sigma), u_j(\sigma)$  quando dois jogadores adotam a estratégia  $s_i$  e  $s_j$ , respectivamente, no estado  $\sigma$ . Lembre-se que um jogo base simétrico possui funções utilidades simétricas como no caso do jogo da Tabela 2.2, e podemos representá-lo por uma matriz, a matriz do jogo. Seja  $\Phi = (\phi_{i,j})$  esta matriz que representa então as recompensas do jogador linha, uma vez que as do outro jogador são obtidas simplesmente pela transposta de  $\Phi$ . O jogo se dará entre membros escolhidos de uma população, sendo que a fração de seus indivíduos adotando uma estratégia  $s_i$  é dada pela  $i$ -ésima entrada do vetor  $\sigma$ . Seja  $\rho_{i,\sigma}$  o valor esperado da recompensa ganha pelo jogador da classe  $i$  (i.e. que adota a estratégia  $s_i$ ) quando joga contra uma população, cujo estado é dado por  $\sigma$ . Condicionando na estratégia do jogador da população escolhido para o jogo com o de classe  $s_i$  temos:

$$\rho_{i,\sigma} = \sum_{j=1}^M \phi_{i,j} \sigma_j = \mathbf{e}_i \Phi \sigma^T.$$

( $\sigma^T$  é a transposta de  $\sigma$  e  $\mathbf{e}_i$  é o vetor, cuja  $i$ -ésima componente é 1 e as demais 0.) O valor esperado da recompensa de um indivíduo de uma população distribuída em classes de acordo com  $\sigma$  e escolhido ao acaso é:

$$\rho_{\sigma,\sigma} = \sum_{i=1}^M \rho_{i,\sigma} \sigma_i = \sum_{i,j=1}^M \phi_{i,j} \sigma_i \sigma_j = \sigma \Phi \sigma^T,$$

uma vez que a probabilidade de se escolher um membro de classe  $i$  desta população é  $\sigma_i$ .

Suponha que uma fração  $\varepsilon$  da população passe a jogar a estratégia  $s_k$ . Então a nova população será:

$$\sigma' = (1 - \varepsilon)\sigma + \varepsilon e_k.$$

A recompensa média de um indivíduo classe  $i$  da população distribuída de acordo com  $\sigma'$  é:

$$\rho_{i,\sigma'} = e_i \Phi \sigma'^T = e_i \Phi [(1 - \varepsilon)\sigma^T + \varepsilon e_k^T] = (1 - \varepsilon)\rho_{i,\sigma} + \varepsilon e_i \Phi e_k^T.$$

Logo a recompensa média de um indivíduo “mutante” é  $\rho_{k,\sigma'}$ .

Encolhendo-se aleatoriamente um indivíduo da população distribuída de acordo com  $\sigma$  para jogar contra uma população distribuída com  $\sigma'$  obtemos, após descondicionar nesta escolha ao acaso,

$$\rho_{\sigma,\sigma'} = \sigma \Phi \sigma'^T.$$

Portanto  $\rho_{\sigma,\sigma'}$  é a recompensa média de um jogador da população que não mudou de estratégia.

Uma população pode ser invadida por indivíduos de classe  $k$  se:

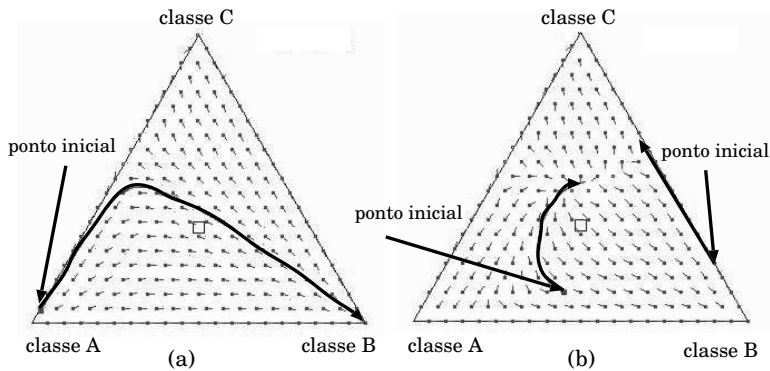
$$\rho_{k,\sigma'} \geq \rho_{\sigma,\sigma'},$$

isto é, a recompensa média de um invasor  $k$  (um mutante) jogando contra a nova população é maior ou igual que a recompensa média da população passível de ser invadida contra a nova população.

**Definição 15** *ESS. Um estado  $\sigma$  é ESS se uma das duas condições é verdadeira: (1) ou a média da recompensa da população não-mutante é maior do que aquela da população mutante, quando ambas jogam contra a população não mutante,  $\rho_{\sigma,\sigma} > \rho_{\sigma',\sigma}$ ; (2) caso essas recompensas sejam idênticas, então a média da recompensa da população não-mutante é maior do que aquela da população mutante, quando ambas jogam contra membros da população mutante,  $\rho_{\sigma,\sigma} > \rho_{\sigma',\sigma'}$ .*

Um estado  $\sigma$  é ESS se não existem mudanças na fração da população adotando as estratégias ao longo do tempo. Uma ESS é “pura” se uma estratégia é predominante sobre as outras. No caso da Figura 2.15 a estratégia pombo é claramente ESS. O leitor deve verificar que uma ESS é um equilíbrio de Nash.

Como outra ilustração de ESS usemos o jogo base dado pela Tabela 2.2, obtida do exemplo da ferramenta Vivavoz do conjunto *freemeeting*. Embora no caso do exemplo VoIP discutido anteriormente a matriz do jogo base mude com o estado do sistema, vamos agora assumir que a matriz é a mesma em todo instante de tempo e usar as equações da dinâmica do replicador para modelar o sistema (esta simplificação é válida quando o número de jogadores é muito grande). A Figura 2.16(a) mostra o resultado da simulação da equação 13 a partir do ponto inicial (0,95;0,005;0,045). Neste caso existe uma estratégia



**Figura 2.16. Dinâmica do replicador para o jogo da Tabela 2.2**

ESS. Note que, neste exemplo, a maioria dos usuários começa usando a maior taxa de transmissão. Com o tempo os usuários passam a usar a estratégia de menor taxa, e finalmente todos acabam convergindo para o ponto de taxa média. (Enfatizamos que este jogo não é aquele ilustrado no exemplo original do Vivavoz para um grande número de usuários. Apenas adotamos a tabela de recompensas para o caso de dois usuários para compor um novo jogo dado pela dinâmica do replicador.)

Alteremos agora o jogo de forma a que a recompensa obtida quando os usuários se utilizam da taxa média seja 2,4 ao invés de 4,0. A Figura 2.16(b) mostra a trajetória da simulação para dois pontos iniciais:  $(0,5;0,4;0,1)$  e  $(0,0;0,8;0,2)$ . Neste caso existem dois pontos de atração. Qualquer trajetória que inicie com uma população de usuários na classe A convergirá para  $(0,24;0,275;0,484)$ . Caso contrário convergirá para  $(0,0;0,417;0,583)$ . O leitor deve verificar se existe um ponto ESS segundo a definição apresentada.

É importante determinar as relações entre equilíbrio de Nash, estratégia ESS e pontos de equilíbrio. Como existem várias propostas de dinâmicas evolucionárias essas relações ainda são objeto de estudo. O teorema abaixo apresenta algumas dessas relações:

**Teorema 2** *Caso o jogo evolucionário satisfaça a dinâmica do replicador, temos que:*

- Se  $\sigma^*$  é um equilíbrio de Nash do jogo evolucionário, então  $\sigma^*$  é um ponto fixo<sup>5</sup> na dinâmica do replicador
- Se  $\sigma^*$  é um equilíbrio evolucionário, então é um equilíbrio de Nash

<sup>5</sup> Lembre-se que  $x_0$  é um ponto fixo de uma função  $f(x)$  se  $x_0 = f(x_0)$  e a seqüência  $x, f(x), f(f(x)), \dots$  converge para  $x_0$ .

O leitor deve procurar mostrar as relações indicadas pelo Teorema 2, e ainda outras relações possíveis, por exemplo, se um equilíbrio de Nash é um equilíbrio evolucionário da dinâmica do replicador. Este último é um ponto muito importante da teoria evolucionária.

### **2.3.3. Sistemas com dinâmicas estocásticas**

A dinâmica do replicador é determinística. Podemos também considerar dinâmicas estocásticas onde o comportamento dos jogadores é modelado por um processo estocástico. Cadeias de Markov são uma ferramenta poderosa para caracterizar a dinâmica de sistemas que não podem ser capturados por equações diferenciais. Já introduzimos o leitor neste conceito.

Na verdade, o exemplo ilustrado pelo diagrama da Figura 2.14 é uma cadeia de Markov, se assumirmos que os eventos que provocam as transições entre estados são tais que o intervalo entre eles tem duração dada por uma distribuição exponencial. Note que existe um ponto de equilíbrio neste sistema, e sua evolução é dada pelas probabilidades dos jogadores adotarem as estratégias com o tempo, isto é, pelas probabilidades de estado do sistema. No caso deste exemplo, temos uma cadeia de Markov de tempo contínuo e espaço de estado discreto. Identicamente podemos modelar um sistema dinâmico por uma cadeia de Markov de tempo discreto. As estratégias são modeladas pela matriz taxa de transição ou matriz de probabilidade de transição entre estados, para o caso contínuo ou discreto, respectivamente.

### **2.4. Aplicações a Redes de Computadores**

Nesta seção iremos apresentar alguns problemas específicos que surgem na área de Redes de Computadores, quando mais de um usuário ou agente disputa por recursos escassos. Veremos como a teoria dos jogos vem sendo aplicada a estes problemas e discutiremos alguns resultados recentes.

Começaremos discutindo problemas de roteamento egoísta, pois este é um tópico bastante discutido na literatura, tanto na área da computação, como também em outras áreas onde a abstração de redes está presente, como a rede rodoviária.

#### **2.4.1. Problemas de roteamento**

No problema de roteamento egoísta em uma rede de computadores, usuários ou agentes decidem o caminho que seus pacotes irão tomar pela rede. Em geral, a rede possui uma topologia fixa, onde os enlaces possuem características bem determinadas. Cada enlace da rede está associado a uma função que determina o retardo dos pacotes que atravessam o enlace. Tal função depende de diversas propriedades do enlace, tais como a capacidade de transmissão ( $C$ ), a taxa de chegada de pacotes ( $\lambda$ ), e o retardo de propagação ( $P$ ). Os usuários desejam enviar tráfego de um determinado ponto da rede a outro. Entretanto, ao invés da rede determinar a rota que os pacotes deverão tomar, o roteamento será determinado pelos próprios usuários. Obviamente, cada usuário deseja minimizar o tempo que seus pacotes levarão para atra-

vessar a rede, sem se preocupar com o bem-estar comum. Temos assim, o *roteamento egoísta*.

Veremos dois importantes exemplos que motivam problemas que surgem com o roteamento egoísta. O primeiro exemplo foi dado por Pigout e o segundo é o famoso paradoxo de Braess. Nos exemplos que se seguem, a função de retardo dos enlaces da rede será dada por  $r(x) = x/C^2 + P$ , onde  $x$  é a taxa de chegada de pacotes no enlace (em pacotes por segundo),  $C$  é a capacidade de transmissão do enlace (em pacotes por segundo), e  $P$  é o retardo de propagação do enlace (em segundos).

#### 2.4.1.1. Exemplo de Pigout

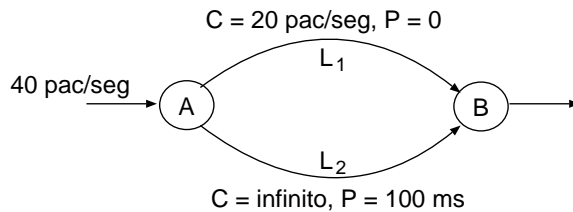


Figura 2.17. Exemplo de Pigout

Considere a rede ilustrada na Figura 2.17, onde o tráfego inserido no ponto A possui dois caminhos alternativos para chegar ao ponto B. O caminho  $L_1$  atravessa um único enlace, cuja função de retardo é dada por  $r(x) = x/400$ . O caminho  $L_2$  também atravessa um único enlace, cujo retardo é dado por  $r_2(x) = 0.1$ . Repare que o retardo do caminho  $L_2$  não depende da taxa de chegada de pacotes ao enlace, o que representa o fato do enlace possuir uma capacidade de transmissão muito alta (i.e., infinita). A taxa agregada de chegada de pacotes no ponto A é de 40 pacotes por segundo (pps). Qual é o retardo do roteamento egoísta?

Consideremos que a taxa de chegada de pacotes no ponto A pertence a 40 usuários, onde cada usuário possui uma taxa de 1 pps (equivalentemente, poderíamos considerar 80 usuários, onde cada um possui uma taxa de 0.5 pps). Por onde os usuários devem rotear seus pacotes? Vamos denotar por  $x_1$  e  $x_2$  a taxa de pacotes atravessando os enlaces  $L_1$  e  $L_2$ , respectivamente. Temos que  $x_1 + x_2 = 40$ . Repare que em equilíbrio, os dois enlaces devem possuir o mesmo retardo. Caso contrário, o usuário que estiver enviando seu tráfego no enlace com maior retardo pode alterar a rota e diminuir o retardo. Qual é o retardo do sistema em equilíbrio?

Como os dois retardos necessariamente são iguais em equilíbrio, temos que  $r_1(x_1) = r_2(x_2)$ . Assim sendo,  $x_1/400 = 0.1$ , ou seja,  $x_1 = 40$ . Como  $x_1 + x_2 =$

40, temos ainda que  $x_2 = 0$ . O retardo de todos os pacotes é então de 100 ms. Repare que em equilíbrio apenas o enlace  $L_1$  é utilizado pelos usuários. Isto ocorre, pois caso um ou mais usuários escolham o enlace  $L_2$ , o retardo que seus pacotes receberão será de 100 ms, enquanto o enlace  $L_1$  oferecerá um retardo menor. Logo, os usuários que escolheram o enlace  $L_2$  preferem o enlace  $L_1$ .

Uma observação interessante é que o retardo obtido no roteamento egoísta não é o menor retardo possível. Se o roteamento for determinado pela rede de forma centralizada, é possível então que o retardo seja menor. Por exemplo, se apenas metade dos usuários utilizar o enlace  $L_2$  e a outra metade for forçada a utilizar o enlace  $L_1$ , então estes últimos terão um retardo de 100 ms, enquanto os primeiros terão um retardo de 50 ms. Observe que esta solução diminui o retardo para metade dos usuários sem aumentar o retardo para nenhum outro.

#### 2.4.1.2. Paradoxo de Braess

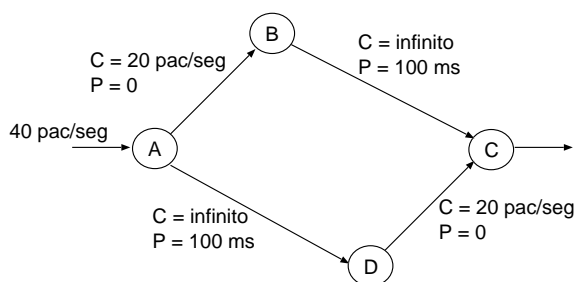
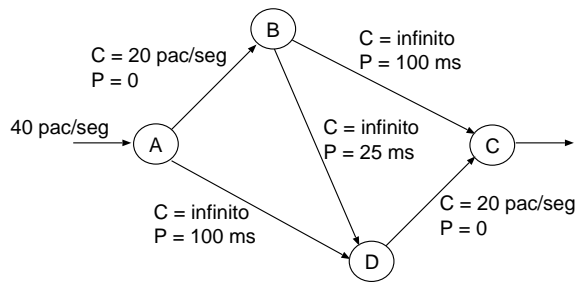


Figura 2.18. Paradoxo de Braess: topologia original

Considere a rede ilustrada na Figura 2.18. Assim como no exemplo anterior, usuários desejam transportar seus pacotes do ponto A ao ponto C da rede. Há dois caminhos disponíveis:  $L_1 = A \rightarrow B \rightarrow C$  e  $L_2 = A \rightarrow D \rightarrow C$ . Cada caminho atravessa dois enlaces com diferentes propriedades. Entretanto, repare que o enlace  $A \rightarrow B$  é equivalente ao enlace  $D \rightarrow C$ , enquanto o enlace  $B \rightarrow C$  equivale ao enlace  $A \rightarrow D$ . O retardo de um caminho é dado pela soma dos retardos dos enlaces que formam o caminho. Assim sendo, a função de retardo dos caminhos  $L_1$  e  $L_2$  é dada por  $r_1(x_1) = r_{AB}(x_1) + r_{BC}(x_1) = x_1/400 + 0.1$  e  $r_2(x_2) = x_2/400 + 0.1$ , respectivamente, onde  $x_1$  e  $x_2$  são as taxas de pacotes utilizando os caminhos  $L_1$  e  $L_2$ , respectivamente. Observe que  $x_1 + x_2 = 40$ . Qual é o retardo do roteamento egoísta?

Como vimos antes, em equilíbrio o retardo nos dois caminhos necessariamente é igual. Assim, temos que  $r_1(x_1) = r_2(x_2)$ . Assim temos que  $x_1/400 + 0.1 = x_2/400 + 0.1$  e  $x_1 + x_2 = 40$ . Substituindo, obtemos a solução

$x_1 = 20$  e  $x_2 = 20$ , o que leva a um retardo de 150 ms. Repare que neste caso, esta é a melhor solução possível, mesmo que o roteamento seja determinado de forma centralizada pela rede.



**Figura 2.19. Paradoxo de Braess: topologia com enlace adicionado**

Considere agora a adição de um novo enlace  $BD$  na topologia acima, dando origem a uma nova topologia ilustrada na Figura 2.19. Este novo enlace possui capacidade de transmissão muito alta (i.e., infinita), e um tempo de propagação mais baixo que os outros enlaces da rede. Intuitivamente, é de se esperar que o retardo dos pacotes na rede irá diminuir. O que você acha? Vamos calcular o retardo do roteamento egoísta.

Temos agora um novo caminho disponível para os usuários, que é dado por  $L_3 = A \rightarrow B \rightarrow D \rightarrow C$ . Denotaremos por  $x_3$  a quantidade de tráfego utilizando o caminho  $L_3$ . Repare que a taxa de chegada ao enlace  $AB$  é dada agora por  $x_1 + x_3$ , enquanto a taxa de chegada ao enlace  $DC$  é dada por  $x_2 + x_3$ . Assim sendo, as funções de retardo são dadas por  $r_1(x_1, x_3) = (x_1 + x_3)/400 + 0.1$ ,  $r_2(x_2, x_3) = (x_2 + x_3)/400 + 0.1$  e  $r_3(x_1, x_2, x_3) = (x_1 + x_3)/400 + 0.025 + (x_2 + x_3)/400$ . Em equilíbrio, necessariamente os três retardos são idênticos. Assim sendo, temos que  $r_1(x_1, x_3) = r_2(x_2, x_3) = r_3(x_1, x_2, x_3)$ . Além disso, devido à conservação de fluxo de pacotes, temos que  $x_1 + x_2 + x_3 = 40$ . Resolvendo este sistema de equações lineares, utilizando substituição simples, obtemos  $x_1 = 10$ ,  $x_2 = 10$  e  $x_3 = 20$ . Esta solução resulta em um retardo de 175 ms, que é *maior* que o retardo obtido *sem* a adição do novo enlace. O que deu errado?

Repare que um roteamento centralizado imposto pela rede pode reduzir o retardo do roteamento egoísta. Por exemplo, se o roteamento centralizado simplesmente não utilizar o novo enlace, *todos* os usuários obterão um retardo menor (e não apenas alguns, como no exemplo de Pigout).

A moral deste exemplo, e também do exemplo de Pigout, é que a interação entre usuários egocêntricos pode levar a um resultado pior para todos, quando comparado com uma solução centralizada. Porém, isto não é novidade para nós, pois como vimos em muitos exemplos, o conflito de interesses em muitos casos leva a um resultado pior para os usuários. Sob esta perspectiva, o pa-

radoxo de Braess pode ser visto como mais uma manifestação do dilema do prisioneiro.

### 2.4.1.3. O preço da anarquia

Vimos nos exemplos anteriores que o retardo do roteamento egoísta pode ser maior que o retardo do roteamento centralizado, imposto pela rede. Uma questão interessante é quantificar esta diferença. O quão pior é o roteamento egoísta?

Para responder a esta pergunta, é necessário primeiramente definir uma métrica para avaliar as diferenças entre os dois paradigmas (egoísta e centralizado). Considere a razão entre o retardo obtido com o roteamento egoísta em equilíbrio e o menor retardo obtido com roteamento centralizado. Esta razão, que é necessariamente maior ou igual a 1, é conhecida como o *preço da anarquia*.

Em sua tese de doutorado, Tim Roughgarden estudou o preço da anarquia em problemas de roteamento egoísta. Ele demonstrou uma série de resultados interessantes, entre eles, que o preço da anarquia, quando as funções de retardo dos enlaces são lineares, é de no máximo  $4/3$ . Surpreendentemente, este resultado não depende da topologia da rede, nem da matriz de tráfego atravessando a rede, e a ineficiência máxima pode ser atingida com o simples exemplo de Pigout.

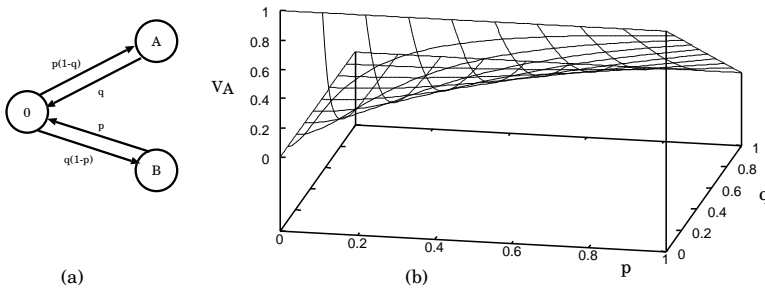
### 2.4.2. Problemas em redes sem fio

Conforme delineado na seção 2.1, um problema de redes sem fio é o de achar uma forma eficiente e equitativa de se fazer uso do canal de rádio compartilhado entre várias estações. Caso não haja nenhuma regra rígida de controle, o objetivo dos usuários será o de obter para si a máxima capacidade de transmissão do meio compartilhado. Restringiremos o texto à apresentação de problemas simples mas que ilustrem os principais conceitos.

A maioria dos leitores já deve ter estudado o protocolo Aloha de transmissão. É um protocolo muito simples, totalmente descentralizado, de acesso ao meio de comunicação. Considere a versão sincronizada (*slotted*) deste protocolo, onde o tempo é dividido em intervalos de mesmo tamanho (ranhuras no tempo ou *slots*) de comprimento suficiente para transmitir um pacote de dados. Uma estação pronta para transmitir espera a próxima ranhura e envia um pacote com probabilidade 1. Caso duas ou mais estações transmitam simultaneamente, uma colisão ocorre e ambas transicionam a um estado pós-colisão, onde passam a transmitir um pacote com probabilidade  $p$  a cada slot. Em geral  $p < 1$  para evitar excessivas colisões. Por outro lado, um valor pequeno que  $p$  diminuirá a vazão das estações. É fácil construir modelos de desempenho para calcular a vazão obtida para valores conhecidos de  $p$ .

Suponha agora que as estações que compartilham o meio possam escolher os valores de  $p$ , visando única e exclusivamente otimizar o seu ganho. O





**Figura 2.20. (a) Cadeia de Markov do sistema com duas estações; (b) vazão da estação  $E_A$  em função de  $p$  e  $q$**

conflito de interesse é evidente. Estudemos este conflito, supondo que temos apenas duas estações,  $E_A$  e  $E_B$ . Para obter a vazão do sistema, construiremos um modelo de desempenho utilizando uma cadeia de Markov. Neste exemplo a cadeia possui três estados, conforme mostrado na Figura 2.20. O estado rotulado como 0 indica que houve colisão e as duas estações tentarão transmitir a cada ranhura de tempo com probabilidades  $p$  e  $q$ , respectivamente. Assim que a estação  $E_A$  transmite com sucesso, o modelo transiciona para o estado  $A$ . Neste estado  $E_A$  transmite a cada intervalo com probabilidade 1, até que  $E_B$  transmita também e haja então uma colisão. O comportamento do sistema no estado  $B$  é simétrico àquele no estado  $A$ , para a estação  $E_B$ .

Deste modelo, podemos calcular facilmente a vazão  $v_A$  e  $v_B$  de cada estação em função de  $p$  e  $q$ , uma vez que  $v_i = \pi_i$  ( $i = A, B$ ), onde  $\pi_i$  é a fração de visitas ao estado  $i$  do modelo:

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

É interessante observar que, se as duas estações adotarem o mesmo protocolo, i.e.  $p = q$ , então a vazão máxima será dada por  $v_A = v_B = 1/3$ , o que ocorre quando as probabilidades de transmissão tendem a zero. Neste caso, o sistema é bastante injusto, pois uma estação que ganha o meio de comunicação permanece um tempo muito grande transmitindo até haver uma colisão.

Para evitar esse problema, podemos colocar um limite para a vazão de cada estação de forma a dar chance de a outra transmitir. Analogamente, poderíamos limitar o total de transmissões (com ou sem sucesso) de cada estação por ranhura,  $C_A = v_A + p\pi_B$  ( $C_B = v_B + q\pi_A$ ), por exemplo, de forma a limitar a energia média gasta para transmitir. Usaremos o primeiro limite neste exemplo, portanto  $v_A, v_B \leq C$ .

Estudaremos agora o que acontece quando as estações podem unilateralmente modificar seus parâmetros  $p$  e  $q$ , competindo pela capacidade do canal.

Esta competição pode ser modelada por um jogo de Stackelberg.

**Definição 16** *Um jogo de Stackelberg é um jogo extensivo de dois rivais e com informação perfeita onde o líder escolhe uma estratégia do seu conjunto  $S$  e o seguidor, informado sobre a escolha do líder, escolhe em seguida a sua, do seu conjunto de opções.*

Suponha que limitemos a vazão de ambos os nós a 0,2, e o líder seja a estação  $E_A$ . O líder escolhe um valor de  $p$ , digamos, para efeito deste exemplo,  $p = 0.1$ . Em seguida,  $E_B$  escolhe um valor de  $q$  que maximiza  $v_B$  dada a restrição. No caso,  $q = 0.0666$ . De posse do resultado de  $E_B$ ,  $E_A$  escolhe a sua melhor estratégia, dada a restrição, que no caso é  $p = 0.0446$ , e assim por diante. Neste caso, não é difícil verificar que as estações alcançam a vazão máxima permitida a cada uma. Aumentemos esse valor limitante para 0,5. O leitor poderá verificar que  $E_A$  receberá aproximadamente 45% do canal, enquanto que a outra nada obterá.  $E_B$  torna-se menos agressiva tentando maximizar a sua vazão. Claramente neste caso é preferível ser o líder. Note ainda que, mesmo quando a estação  $E_A$  domina o canal, ela não conseguirá ocupar 100% da capacidade. Este é o preço pago pela disputa entre as duas estações. O leitor deve experimentar com outros limites de vazão, e estudar a evolução do sistema.

Note também que, como a estação  $E_A$  sabe que  $E_B$  tentará maximizar o seu ganho,  $E_A$  pode escolher de antemão o melhor valor de  $p$ . Portanto,  $E_A$  pode escolher  $p$  prevendo o que  $E_B$  irá fazer.

Agora que você tem uma idéia melhor do jogo, iremos formalizar um pouco o problema descrito acima. Nosso objetivo é resolver o jogo de Stackelberg. Em primeiro lugar, note que as estações estão resolvendo o seguinte problema de otimização:

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

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

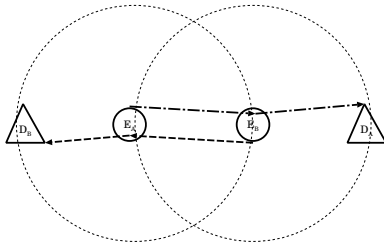
**Problema do líder  $E_A$ :** o líder sabe que o seguidor tentará maximizar a sua escolha, e portanto ele deve de antemão escolher o valor do seu parâmetro da forma:

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

Para resolver o jogo inicialmente resolvemos o problema do seguidor para todo valor de  $p$ , i.e. uma das estratégias do líder. O líder então decide a sua estratégia ótima de acordo com as possíveis respostas dadas pelo seguidor. Esse procedimento é chamado de indução traz-para-frente, conforme estudamos anteriormente.

Outros exemplos semelhantes existem. É sabido que ao aumentar a potência de transmissão aumenta-se a probabilidade de recuperação do pacote caso haja uma colisão. Como existe um custo de transmissão, o leitor pode imaginar facilmente um jogo em que os jogadores podem escolher níveis de potência além de transmitir ou não um pacote, sendo que o sucesso é função do nível de transmissão das estações envolvidas na colisão.

O exemplo acima é típico de sistemas de múltiplo acesso. Vários outros exemplos relativos a redes sem fio existem. No que se segue, mencionamos o exemplo denominado *Dilema do Retransmissor*.



**Figura 2.21. Dilema do Retransmissor**

Suponha que existam duas estações  $E_A$  e  $E_B$ , tentando transmitir para seus respectivos destinos  $D_A$  e  $D_B$ , conforme indicado na Figura 2.21. Na figura, os círculos pontilhados indicam o alcance de cada estação, e conseqüentemente,  $E_A$  não atinge o seu destino diretamente. Para enviar um pacote de dados para  $D_A$ , este necessariamente precisa ser retransmitido por  $E_B$ . O mesmo ocorre com  $E_B$ . Uma estação, ao transmitir um pacote, recebe uma recompensa  $r$  caso este chegue ao seu destino. Por outro lado, o custo de transmissão é  $c$  que representa a energia despendida pela estação para realizar a transmissão. Obviamente, cada estação deseja maximizar a sua vazão enquanto minimiza o seu custo. Caso uma estação descarte um pacote da outra, visando economizar energia, esta última pode fazer o mesmo e portanto, se esse comportamento persistir, nenhuma das duas conseguirá transmitir. O conflito é claro. Com base nos seus conhecimentos, você pode formular um ou mais jogos que modelem tal disputa.

O leitor atento já deve ter observado que o dilema do transmissor é semelhante ao do dilema de um sistema P2P, onde existe um custo envolvido em deixar arquivos para ser compartilhados na rede. Por outro lado, sistemas P2P em geral incluem mecanismos de incentivo para os chamados “bocas-livres” (*free-riders*).

### 2.4.3. Problemas de congestionamento

Evitar a ocorrência de congestionamentos em uma rede de comunicação de dados é um dos problemas mais importantes da área, tem sido extensiva-

mente estudado há décadas, e continua sendo um tópico atual de pesquisa. Um congestionamento pode ocorrer quando as demandas por recursos, por exemplo um canal de comunicação, são maiores que a capacidade deste. Tipicamente, o controle de congestionamento é realizado impondo-se ao usuário regras rígidas que limitam a taxa de transmissão quando um congestionamento é detectado. Entretanto, como abordado na seção 2.1, existem cenários onde usuários não estão sujeitos a tais protocolos. O problema foi introduzido naquela seção através de um exemplo que mostra uma competição entre usuários de uma ferramenta VoIP que usa o protocolo UDP (portanto livre das regras rígidas do TCP), e permite variar a taxa de compressão e tipo de redundância para os dados da voz recebida. A seção 2.3 emprega o mesmo exemplo para ilustrar a dinâmica de um jogo na disputa pelo canal de comunicação, de forma a obter a melhor qualidade de voz possível. Vários artigos na literatura abordam o problema de controle de congestionamento baseado em uma abordagem de Teoria de Jogos. Nesta seção apresentaremos algumas idéias básicas através de exemplos simples.

Como primeiro exemplo, suponha que um conjunto de usuários esteja disputando um canal de comunicação com capacidade  $C$ . Suponha ainda que cada usuário pode ajustar livremente a sua taxa de transmissão, de forma a otimizar uma métrica de desempenho. Naturalmente cada usuário irá tentar obter a maior vazão possível com o menor tempo de resposta. Obviamente, gostaríamos de tentar entender o que acontece neste cenário. Existirá um equilíbrio de Nash? Qual a dinâmica do sistema?

Para ilustrar o problema, considere a disputa entre dois usuários apenas. Adotemos como função utilidade a razão entre a vazão e o tempo médio de espera do usuário. Esta função é conhecida como potência (*power*) da fila do canal. Supondo a disciplina FCFS (*first come first served*) ou PS (*processor sharing*), chegadas Poisson, e que o tempo médio de serviço dos pacotes é idêntico para cada usuário (e dado por uma distribuição exponencial com média  $s$ ), o tempo médio de espera do usuário  $i$  é  $s/(1-\rho)$  onde  $\rho = (\lambda_1 + \lambda_2)s$ , e  $\lambda_i$  é a taxa de chegada do usuário  $i$ . Portanto:

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

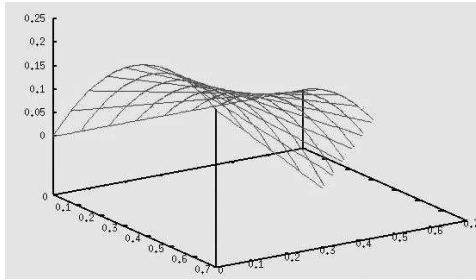
com a restrição  $(\lambda_1 + \lambda_2)s < 1$ .

Suponha que o conjunto  $S_i$  de estratégias do usuário  $i$  seja composto de apenas duas taxas: 0,1 e 0,499. A Figura 2.22 ilustra o jogo em sua forma normal para  $s = 1$ . Caso ambos os jogadores adotem a estratégia agressiva, de alta taxa, a recompensa de ambos será muito baixa, pois o sistema estará bastante congestionado, com utilização próxima a 1. Por outro lado, um jogador unilateralmente pode reduzir sua taxa e obter maior ganho. Se assim o fizer beneficiará o outro também. O leitor já deve ter percebido que tratamos de jogos semelhantes a esse em seções anteriores (por exemplo na seção 2.2.6). Quantos equilíbrios de Nash existem então?

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

**Figura 2.22. Jogo de usuários disputado um canal**

Consideremos a seguir que os jogadores podem adotar qualquer valor para a sua vazão, obviamente com a restrição de a utilização global ser inferior a 1. A Figura 2.23 mostra a função utilidade para um dos usuários. Como modelar a evolução do jogo?



**Figura 2.23. Função utilidade**

Suponha, por exemplo, que um dos jogadores (jogador 1) escolha a sua taxa de transmissão, digamos  $\lambda_1 = 0,1$ . O segundo jogador tentará escolher a sua estratégia para otimizar a sua recompensa, e assim por diante. A função de utilidade é simples e portanto, é fácil verificar que o sistema converge para o ponto  $\lambda_1 = \lambda_2 = 1/3$ . O leitor pode também observar que o mesmo resultado é obtido se o jogador 1 escolher uma alta taxa de transmissão, digamos  $\lambda_1 = 0,9$ .

A evolução pode ser modelada como um jogo de Stackelberg, como já vimos na seção 2.4.2. Para resolver o jogo de Stackelberg lembre-se que, em primeiro lugar, o segundo jogador irá otimizar a sua recompensa, seja qual for a escolha do primeiro. Derivando-se a função utilidade do segundo em relação a  $\lambda_2$  obtemos:  $\lambda_2^*(\lambda_1) = (1 - \lambda_1)/2$ , onde  $\lambda_2^*$  é o valor ótimo da taxa do segundo jogador em função da escolhida pelo primeiro. Ora, como o primeiro jogador sabe da escolha do segundo, ele otimiza a sua função utilidade:  $\lambda_1^*(\lambda_1) = (1 - \lambda_2)/2$ . Então,  $\lambda_1^*(\lambda_1^*) = 1/3$ . Portanto, neste exemplo, existe apenas um ponto de equilíbrio.

É interessante observar que, no ponto de equilíbrio, cada jogador re-

cebe uma recompensa igual a  $1/9$ . Portanto, a recompensa total do sistema ( $2/9 = 0,22$ ) é inferior à recompensa obtida se ambos os jogadores utilizassem taxas iguais a  $0,25$ , pois, neste caso, a recompensa total seria de  $0,25$ . Essa diferença de recompensas é o preço de não cooperar, como já enfatizamos anteriormente.

Retornemos agora ao exemplo do conjunto de ferramentas *freemeeting*, onde usuários de voz podem escolher a taxa de recebimento dos dados e tipo de redundância. O problema básico de congestionamento já foi descrito anteriormente na seção 2.3. No que se segue, nosso objetivo é o de descrever um modelo que capture a dinâmica do sistema. No modelo, um número finito de usuários compartilha um canal que representa o gargalo do sistema. Cada usuário pode escolher dinamicamente a taxa de compressão da voz e o tipo de redundância, procurando maximizar a qualidade da voz recebida, independentemente dos outros. Suponhamos que os usuários possuem a mesma função utilidade. Esta é obtida de um modelo de desempenho que modela o canal de comunicação, quando os usuários transmitem dados com taxas determinadas pelo método de compressão e redundância escolhida. O modelo de desempenho possibilita-nos calcular o descarte de pacotes ocorrido pelo congestionamento e espaço limitado de *buffer* de canal. Este, por sua vez, fornece os dados de entrada para que se calcule a qualidade da voz recebida por cada usuário, por exemplo pela utilização do E-model da ITU.

Os usuários jogam o seguinte jogo simétrico:  $\mathcal{J} = (N, S, \mathbf{u})$ , onde o conjunto de estratégias de cada usuário é dado pelo algoritmo de compressão e tipo de redundância usados. Podemos modelar o processo de mudança de estratégias dos usuários ao longo do tempo por uma cadeia de Markov (CM) de tempo contínuo  $\mathcal{X} = \{X(t) : t \geq 0\}$ . Desta forma, um estado  $e_i$  da CM é uma  $n$ -upla, cuja  $j$ -ésima entrada é igual ao número  $n_j^{(i)}$  de usuários utilizando a estratégia  $s_j$ , no estado  $e_i$ , dentre as  $M$  possíveis.

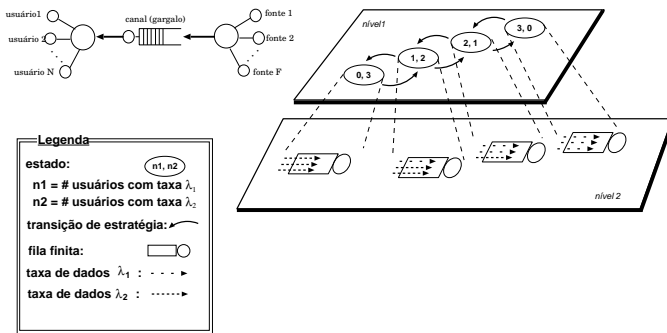
A transição entre estados da CM determina o processo dinâmico de mudanças de estratégia. Para determinar estas transições, devemos levar em consideração que cada usuário deve mudar sua estratégia procurando melhorar a qualidade da voz por ele percebida. Além disso, os usuários não reagem instantaneamente a mudanças do ambiente e podem alterar suas estratégias a cada instante de forma assíncrona. Como as recompensas são simétricas, utilizemos a seguinte taxa de transição de  $e_i$  para  $e_j$ :

$$\begin{cases} n_l^{(i)} \times [u_m(e_j) - u_l(e_i)] & \text{se } u_m(e_j) > u_l(e_i) \\ n_l^{(i)} \times \varepsilon & \text{caso contrário} \end{cases} \quad (14)$$

onde  $u_l(e_i)$  é a recompensa obtida pelos jogadores que utilizam a estratégia  $s_l$  quando o estado do sistema é  $e_i$ . A escolha dessas taxas de transição atende às especificações acima. Quanto maior a recompensa obtida pela troca de estratégia de um usuário, mais rapidamente este alterará sua estratégia. O modelo permite que erros possam ocorrer, isto é, que um usuário possa avaliar

incorretamente a qualidade da voz e escolher uma estratégia que proporcione uma qualidade inferior. (Estes erros devem ser pouco prováveis, e ocorrem com taxa proporcional a  $\varepsilon$ , um parâmetro do modelo.)

A Figura 2.24 representa o modelo do sistema. O leitor já deve ter per-



**Figura 2.24. O modelo do sistema**

cebido os dois níveis de modelagem: a de desempenho para o cálculo das recompensas, e a CM para representar as mudanças de estratégia com o tempo. Utilizando uma ferramenta de modelagem, por exemplo o Tangram-II<sup>6</sup>, o leitor poderá simular o sistema e verificar a sua dinâmica.

Como o modelo em questão é ergódico, existe uma única solução que fornece as probabilidades em estado estacionário. A pergunta é: qual a relação desta solução com equilíbrios de Nash? Quando  $\varepsilon \rightarrow 0$ , é possível relacionar os estados do conjunto que recebe probabilidade positiva com equilíbrios de Nash. Embora a prova deste resultado esteja além do propósito deste texto, um exemplo servirá para clarear os conceitos.

Considere um cenário onde 16 usuários podem escolher entre codecs G.723.1 e G.729, que correspondem a taxas de transmissão de 6.4 Kbps e 8.0 Kbps, respectivamente. Seja a taxa de transmissão do canal gargalo igual a 160 Kbps, e o tamanho do *buffer* 10K. O modelo neste caso tem um único equilíbrio de Nash, correspondendo ao estado onde cinco usuários escolhem a taxa mais baixa e o restante a mais alta. (Este é o estado, cuja probabilidade em estado estacionário é igual a 1.0 quando  $\varepsilon \rightarrow 0$ .) Este estado, entretanto, não é o equilíbrio de Pareto, que fornece uma qualidade maior para todos os usuários. O equilíbrio de Pareto é o estado onde todos os usuários transmitem a taxa mais baixa.

Com este tipo de modelo, pode-se estudar o comportamento do sistema ao variarmos a capacidade de armazenamento do *buffer*, por exemplo, e verificar as estratégias escolhidas pelos usuários de acordo com estas alterações. O

<sup>6</sup> Tangram-II está disponível em [www.land.ufrj.br](http://www.land.ufrj.br) - ferramentas.

leitor interessado deve brincar com o modelo e verificar outras propriedades interessantes.

## 2.5. Referências bibliográficas

Teoria dos jogos é um assunto bastante farto na literatura, com muitos livros publicados nos diferentes níveis de aprendizado e com enfoques diferentes<sup>7</sup>. O assunto também é farto na literatura acadêmica, onde artigos sobre os mais variados aspectos da teoria e sua aplicação continuam sendo publicados<sup>8</sup>. Desta forma, uma revisão bibliográfica mais abrangente torna-se uma tarefa difícil e extensa. Entretanto, oferecemos aqui algumas referências importantes de aplicações da teoria dos jogos a redes de computadores, bem como alguns livros-texto usados na preparação deste tutorial.

### Livros-texto e referências gerais à Teoria dos Jogos

- **Game Theory and Strategy**, por P. Straffin, 1993.  
Um livro introdutório com muitos exemplos simples e explicações bem intuitivas.
- **Game Theory Evolving**, por H. Gintis, 2000.  
Um livro introdutório centrado na solução de problemas. Mais de 100 jogos são apresentados e muitos são resolvidos e discutidos. Também cobre a área de teoria dos jogos evolucionária e contém discussões interessantes sobre o significado de equilíbrios em jogos.
- **A Course in Game Theory**, por M. Osborne e A. Rubinstein, 1994.  
Um livro mais avançado, no nível de pós-graduação. Bastante técnico com definições e teoremas rigorosos.
- **Game Theory.net**, <http://www.gametheory.net>.  
Sítio na Web para educadores, alunos e pesquisadores que têm interesse na área de teoria dos jogos. Contém muito material educacional.

### Referências sobre roteamento egoísta

- “Competitive routing in multiuser communication networks”, A.Orda, R.Rom, N.Shimkin, IEEE/ACM Transactions on Networking, vol. 1, no. 5, pp. 510.521, 1993.
- “How Bad is Selfish Routing?”, T.Roughgarden e E.Tardos, Journal of the ACM, 49(2):236–259, 2002.
- “Selfish Routing with Atomic Players”, T. Roughgarden, ACM-SIAM Symposium on Discrete Algorithms (SODA), 2005.
- “On Selfish Routing in Internet-Like Environments”, L.Qiu, Y.R.Yang, Y.Zhang, S.Shenker, ACM SIGCOMM 2003.

<sup>7</sup> Uma busca no sítio amazon.com por “Game Theory” retorna mais de 1.000 livros!

<sup>8</sup> R. Aumann e T. Schelling receberam o Prêmio Nobel em economia de 2005 por suas contribuições no entendimento de cooperação e conflito através da teoria dos jogos.



- **Selfish Routing and the Price of Anarchy**, por T.Roughgarden, 2005.  
Um livro inteiro sobre roteamento egoísta, bastante formal com muitos resultados e provas, e também muitas referências.

#### Referências sobre controle de congestionamento

- “Making greed work in networks: a game-theoretic analysis of switch service disciplines”, S.Shenker, IEEE/ACM Transactions on Networking, Vol.3, no.6, 1995.
- “An Evolutionary Game-Theoretic Approach to Congestion Control”, D.Menasché, D.Figueiredo, E. de Souza e Silva, Performance Evaluation, 62(1-4), 2005.
- “Efficiency loss in a network resource allocation game”, R.Johari e J.N. Tsitsiklis, Mathematics of Operations Research 29 (3): 407-435 (2004).
- “Selfish behavior and stability of the internet: a game-theoretic analysis of TCP”, A.Akella, S.Seshan, R.Karp, S.Shenker, C.Papadimitriou, ACM SIGCOMM 2002.
- “TCP Connection Game: A Study on the Selfish Behavior of TCP Users”, H.Zhang, D.Towsley, W.Gong, IEEE ICNP 2005.
- **To Queue or not to Queue: Equilibrium Behavior in Queueing Systems**, por M.Haviv e R.Hassin. Kluwer's International Series, 2003.  
Um livro abrangente sobre *jogos em filas*, com muitas referências para a literatura nesta área.

#### Referências sobre redes sem fio

- “Game theory and the design of self-configuring, adaptive wireless networks”, A.B. MacKenzie e S.B. Wicker, IEEE Communications Magazine, 39 (11), 2001.
- “Stability of Multipacket Slotted Aloha with Selfish Users and Perfect Information”, A.B. MacKenzie e S.B. Wicker, IEEE INFOCOM 2003.
- “Modeling and Analysis of Generalized Slotted-Aloha MAC Protocols in Cooperative, Competitive and Adversarial Environments”, R.T.B.Ma, V.Misra e D.Rubenstein, IEEE Intern. Conf. on Distributed Computing Systems (ICDCS'06).
- “An Evolutionary Game Perspective to ALOHA with Power Control”, E.Altman, N.Bonneau, M.Debbah, G.Caire, 19th International Teletraffic Congress, 2005.
- “On Selfish Behavior in CSMA/CA Networks”, M.Cagalj, S.Ganerwal, I.Aad and J.-P.Hubaux, IEEE INFOCOM 2005.
- “CDMA Uplink Power Control as a Noncooperative Game”, T.Alpcan, T.Basar, R.Srikant, E.Altman, Wireless Networks 8:659-670 (2002), Kluwer Academic Press.

- **Security and Cooperation in Wireless Networks:** Thwarting malicious and selfish behavior in the age of ubiquitous computing, por Levente Buttyán e J.-P. Hubaux. Cambridge University Press, 2007. Parte do livro é dedicada ao estudo do comportamento egoísta em redes sem fio, onde alguns problemas específicos são apresentados e discutidos.

### **Agradecimentos**

Este trabalho foi parcialmente financiado pelo CNPq e FAPERJ. Agradecemos a Daniel S. Menascé, cuja tese de mestrado, co-orientada pelos autores, deu origem a um artigo em controle de congestionamento baseado em Teoria dos Jogos. Uma das aplicações aqui tratadas baseia-se naquele trabalho. O grande interesse de Daniel no assunto, aliado aos estudos de doutorado de um dos autores, foi muito importante para criar a base de onde este texto foi inspirado. Agradecemos também aos alunos de graduação do DCC/UFRJ Fabrício M. Ferreira, Erico D. V. Prezzi, Renato R. Fischer, Vinicius N. Borges, e ao revisor anônimo do JAI'2007 pela leitura e comentários ao texto. Nossos agradecimentos também vão para o EPFL/Suíça e a UMass/USA.