

RACIONALIDADE ESTRATÉGICA E INSTITUIÇÕES*

Celia de Andrade Lessa

Apresentação

As escolhas e decisões dos indivíduos em sociedade são freqüentemente decisões "coletivas", no sentido de que os resultados de suas ações isoladas são reconhecidamente dependentes de decisões e ações de outros indivíduos com os quais eles interagem. Este tipo de interação pode ser teoricamente representado como *interação estratégica* e modelado como um *jogo*. Adicionalmente, o fato de que os resultados sejam qualificadamente coletivos, ao passo que as decisões são estritamente individuais, requer, neste caso, a elaboração de uma noção de *racionalidade estratégica* atribuível aos indivíduos em tais circunstâncias, na medida em que suas possíveis ações são contingentes às ações possíveis dos outros com os quais interagem.

Neste artigo pretendo elaborar conceitualmente, e não tecnicamente, a idéia de *interação estratégica*, a partir de sua formulação por autores como von Neumann e Morgenstern (1944) e Anderson e Moore (1962), com ênfase no aspecto provocativo e inovador contido na visão

de interdependência complexa destes autores. Além disso, o artigo discute aspectos relacionados à visão de *racionalidade estratégica*, que pretende resolver a interação estratégica em termos da racionalidade dos indivíduos e de uma noção de equilíbrio. Em particular, focalizam-se aqui as condições de conhecimento normalmente atribuídas a agentes estratégicos, já que estas condições desempenham um papel crucial na possibilidade de resolução da interação *social* (como interação estratégica) em termos da racionalidade dos *indivíduos*. Neste contexto, torna-se relevante saber, por exemplo, em que medida são os indivíduos capazes de conhecer antecipadamente as ações uns dos outros.

Duas conclusões seguem-se: em primeiro lugar, problemas relacionados às condições de cognição revelam o lado problemático da resolução da interação estratégica em termos da racionalidade dos indivíduos, sugerindo que a definição de uma racionalidade estratégica é ainda precária; em segundo lugar, este artigo considera que a visão de interdependência complexa presente nas formulações

originais de von Neumann e Morgenstern (1944), bem como nas de Anderson e Moore (1962), requer a suposição de uma propriedade social ou interacional da racionalidade a partir da qual pode ganhar sentido a idéia de racionalidade estratégica dos indivíduos. Esta suposição, por sua vez, sugere a necessidade de uma reflexão de natureza filosófica sobre as condições de possibilidade de conhecimento dos indivíduos em sociedade, um tema de exploração rarefeita na literatura convencional de teoria dos jogos.

O artigo está dividido em cinco partes. A primeira seção introduz o problema da interação estratégica segundo as perspectivas seminais e levemente diferenciadas de von Neumann-Morgenstern e Anderson-Moore; a teoria dos jogos aparece nessas perspectivas como uma tentativa de formalização de uma visão original e persuasiva da interação social, crítica à visão comportamentalista prevalecente nas teorias sociais e econômicas nas décadas de 40 e 50. Ao mesmo tempo, pode-se perceber nessas mesmas perspectivas aspectos paradoxais e tensões que virão a ser marcantes nos desenvolvimentos posteriores de conceitos de "solução" na teoria dos jogos, em particular o potencial "conflito" entre os objetivos "relevância" (fidelidade em relação à complexidade do objeto) e "determinação" (desenvolvimento de conceitos de solução), ou ainda entre condições de racionalidade "fortes" e condições de informação "irrealistas", característico do debate entre as visões moderna e "clássica" da teoria dos jogos.¹

Na segunda seção procede-se a um exame das condições de conhecimento necessárias à definição de uma racionalidade estratégica, com especial ênfase na condição de "conhecimento comum" (*common knowledge*). Esta última é elaborada tanto em termos de conhecimento comum *stricto sensu* no que se refere às informações necessárias para se compreender e decidir em uma determinada situação estratégica ou jogo (as regras do jogo), quanto em termos do conhecimento comum da racionalidade uns dos outros que se atribui aos atores em uma situação estratégica.

Dificuldades com a condição de conhecimento comum em termos de informação são desviadas com o

desenvolvimento, na terceira seção do artigo, da condição de conhecimento comum em termos apenas da racionalidade dos agentes, a partir do exame do argumento de Robert Aumann e da noção de racionalidade bayesiana. A abordagem deste autor substitui supostos irrealistas de informação por condições fortes de racionalidade que presumivelmente melhor acomodam a incerteza prevalecente em interações de natureza estratégica. A solução de Aumann, entretanto, a meu ver repete a equação da interação social em termos comportamentalistas, pois considera a incerteza como variável exógena, não oferecendo, desta forma, uma boa representação da interação estratégica (e da racionalidade estratégica) tal qual visualizada por von Neumann-Morgenstern e Anderson-Moore.

Finalmente, na quarta seção são brevemente descritas algumas alternativas à visão de Aumann, as quais examinam possíveis suplementações à idéia de uma racionalidade estratégica compreensiva. Nos comentários finais, sugiro que a noção de racionalidade estratégica precisa ser complementada não apenas mediante a adição de contexto aos problemas de escolha, mas também, e principalmente, pelo exame da questão do conhecimento de um ponto de vista filosoficamente mais bem fundado.

Interação estratégica

1ª aproximação: decisões "coletivas"

Segundo Hardin (1988), a teoria dos jogos é a melhor representação que possuímos de uma interação estratégica que se caracteriza pelo fato de que toda vez que eu tento alcançar algum objetivo em sociedade, constato que o que *nós* fazemos é mais importante do que aquilo que *eu* faço. Uma das implicações que se segue deste fato, diz Hardin, é a de que não podemos saber se uma escolha que fizemos foi certa ou errada, moralmente ou em termos dos objetivos que tínhamos em mente, até conhecermos o resultado final que irá depender das escolhas de outros com quem interagimos. Em particular, uma escolha que

tenhamos feito tem apenas o poder de restringir o conjunto de resultados possíveis.

A teoria dos jogos, contudo, em suas formulações originais, aparece já comprometida com a idéia de prover uma representação formal e uma solução racional única para o problema da escolha estratégica na qual um certo número de pessoas encontra-se coetaneamente envolvido. Assim sendo, passamos a relatar a visão de von Neumann e Morgenstern do problema da interação estratégica bem como do caminho para as soluções únicas, pois usualmente atribui-se a eles o estabelecimento das fundações de uma *teoria* dos jogos no âmbito dos problemas de decisão interativos.²

2ª aproximação: o problema de von Neumann-Morgenstern

Sempre que o resultado de uma ação de um indivíduo depende das ações dos outros com quem este interage na mesma medida em que depende de suas próprias ações (*e os indivíduos estão cientes desta dependência*), suas atividades demonstram um componente estratégico e, logo, se adequam à definição de von Neumann-Morgenstern (1944) de um jogo. (Bicchieri, 1993, p. 8; ênfase minha)

Segundo Mirowski (1992), von Neumann esteve envolvido, nos anos 20, em dois projetos de pesquisa diferentes: um deles, o conhecido programa de pesquisa do matemático alemão David Hilbert (1826-1943), que visava ao desenvolvimento de uma axiomática compreensiva para a matemática, e o segundo, o projeto de pesquisa da física quântica. O primeiro projeto sofreu de forma indelével com a prova de Gödel em 1930, a qual demonstrou que um sistema lógico não pode ser simultaneamente completo e consistente. Von Neumann acusou o golpe abandonando o primeiro projeto em sua forma original.³ O segundo projeto, entretanto, estava destinado ao sucesso e produziu efeitos sobre o trabalho matemático de von Neumann, especialmente no desenvolvimento dos jogos.

Na verdade, o desenvolvimento por von Neumann de uma matemática para lidar com jogos testemunhou uma certa persistência de sua ligação com o programa de Hilbert, menos óbvia mas nem por isso menos profunda, se não com o projeto de uma axiomática ao menos com o espírito mais geral daquele programa, cujo objetivo

genérico era estender a linguagem formal matemática ao maior número possível de campos do conhecimento humano, na tentativa de delinear uma teoria geral da racionalidade. Segundo Leonard (1995, p. 733), citando Weyl, a adoção generalizada do formalismo matemático representaria, basicamente, a conversão da questão *daverdade* científica na questão da *consistência*.

Por outro lado, o próprio programa de pesquisa da física quântica requeria o desenvolvimento de novas ferramentas matemáticas. Claramente, a nova física, e as relações originais entre observador-observado que propunha, requeria uma nova matemática: a física infra-atômica sugeria a introdução de relações estocásticas nos níveis mais básicos dos postulados físicos, em particular no próprio sistema observador-observado, que não poderia mais ser considerado um sistema fechado, já que a simples observação afetava o que era observado. Desenvolvimentos na teoria dos conjuntos e nos métodos combinatórios seguiram-se às tentativas de representação formal do novo fenômeno físico.

Indo mais além, von Neumann acreditava que tanto o mundo natural quanto o mundo social compartilhavam a característica de promover uma relação complexa entre o observador e o observado.⁴ Esta perspectiva o afastava das visões comportamentalistas acerca do mundo social, as quais, não obstante professarem um monismo metodológico, adotavam o princípio inverso de um sistema observador-objeto "fechado" como sendo uma boa representação dos dois mundos, onde o objeto de investigação, quer natural, quer social, poderia ser analisado de forma isolada, não sendo afetado pela observação.

A virada metodológica da parte de von Neumann ensejou a elaboração de uma nova ferramenta de medição necessária para aferir o "novo" fenômeno, que havia sofrido uma mudança substancial: sistemas cientificamente definidos acomodavam um elemento subjetivo em seu interior e isto, evidentemente, requeria uma nova percepção daquilo que se ignora; vale dizer, o que se ignora não pode mais ser considerado externo e alheio aos nossos

próprios esforços de compreender, como supunham os comportamentalistas.

Nesse contexto, pode-se então entender a importância que a noção de probabilidade, como ferramenta matemática aparentemente adequada ao trabalho de aferição de sistemas abertos, passa a assumir no trabalho de von Neumann. Ela parece derivar da visão de ignorância implícita na definição de um sistema aberto. Dessa forma, a noção de probabilidade, o desenvolvimento de medidas de valor "esperado" e a noção de estratégias mistas se fazem acompanhar pela idéia de que a ignorância não é algo externo, não é ignorância *sobre* aquilo que é observado mas em relação à influência da própria observação sobre aquilo que se observa. É precisamente esta influência ou os seus efeitos que o observador quer medir.

Pode-se, no entanto, argumentar que não é de todo improvável o surgimento de um "problema de intersubjetividade", ou seja, um conflito entre percepções distintas de um mesmo evento. Esta hipótese, entretanto, é afastada por von Neumann tanto em sua física quanto em sua sociologia, isto é, em termos tanto de observações de fenômenos físicos quanto de distintas percepções de uma mesma situação estratégica no mundo social. Von Neumann sugere que, no primeiro caso, o primeiro observador deve ser utilizado como ferramenta de medição para o segundo; quanto à situação estratégica, está convencido da existência de uma racionalidade única. Neste sentido, o instrumental formal desenvolvido por von Neumann serve de suporte para o seu projeto original (ainda que modificado) de construir uma teoria inequívoca da racionalidade para situações cujo modelo é um jogo, onde toda ação está condicionada em alguma medida pela expectativa das reações que ela pode engendrar.

O interesse de von Neumann por problemas econômicos já era acentuado quando de seu encontro com Morgenstern em Princeton, no ano de 1938, o que pode ser estabelecido por uma cuidadosa revisão do debate em teoria econômica nos anos 30. Contudo, a colaboração com Morgenstern e, em particular, a visão de economia professada por este proveram campo fértil para a cultura dos jogos, nutrindo o

aparato formal de von Neumann com novas e interessantes intuições.

Como se sabe, Morgenstern pertencia à escola austríaca de economia, cuja principal característica consiste na ênfase no problema da limitada capacidade de previsão dos agentes econômicos numa sociedade extensa, em virtude da interdependência generalizada, e, correlativamente, em indicar a oportunidade de uma noção de equilíbrio econômico adequada a este cenário.⁵

Como von Neumann, Morgenstern também procurava uma analogia mais adequada entre o mundo social e o mundo natural do que a oferecida pela mecânica clássica.⁶ As imagens mecânicas do equilíbrio geral deveriam ser substituídas por outras que não assumissem previsão perfeita, pois a mensagem mais interessante do programa de pesquisa walrasiano, segundo Morgenstern, era precisamente a idéia de um mundo econômico interdependente, e esta estaria a requerer uma representação ou uma noção de equilíbrio mais adequada.

Conceitos usuais de mínimos e máximos deveriam, segundo Morgenstern, ser substituídos por conceitos alternativos que pudessem capturar a interdependência das decisões dos agentes, como o "minimax". Ademais, a idéia de "consistência de planos" entre os agentes deveria ser reformulada de modo a representar os agentes econômicos não como tomadores de decisão isolados, restritos apenas por crenças e máximas incondicionais, como nas escolhas paramétricas, e sim como tendo seu comportamento restrito por máximas condicionais, ou seja, condicionado por suas crenças quanto ao comportamento dos outros. O interesse de Morgenstern na formalização destes problemas ensejou sua colaboração com von Neumann.

A equação da interação estratégica, nos termos propostos por von Neumann e Morgenstern, consistia na conciliação de dois elementos aparentemente contraditórios, quais sejam, a incerteza e a imprevisibilidade que, segundo eles, eram características da interação humana e a técnica matemática que permitiria a representação formal desta interação. Tratava-se, para von Neumann, de elaborar uma

teoria da racionalidade (estratégica) que pudesse ser publicamente conhecida e que este conhecimento "público"² não provocasse nenhum desvio por parte dos agentes com relação às prescrições da teoria portanto, uma teoria imune a problemas de recursividade e de regressão ao infinito. Morgenstern, de modo similar, estava interessado em incrementar a relevância da teoria a partir do desenvolvimento de conceitos de equilíbrio mais aptos a refletir a natureza complexa e, portanto, incerta das atividades econômicas. Esperava-se que os novos conceitos de solução fossem capazes de "descontar" os efeitos da incerteza sobre as regras dos jogos e sobre o comportamento dos jogadores.

3ª aproximação: o problema de Anderson-Moore

[Um jogo tem a capacidade de] alterar-se repentinamente *tendo em vista obstruir nossos esforços para alcançar uma solução* [...] (Anderson e Moore, 1962, p. 413; ênfase minha)³

Anderson e Moore (1962) também recorrem à imagem de um *jogo* para expressar o que eles consideram ser a representação mais adequada dos fenômenos sociais, os quais envolvem a interação entre atores sociais, em oposição à representação oferecida pelas abordagens comportamentalistas. Para melhor desenvolver o seu ponto de vista, eles contrastam o método comportamentalista com o da teoria dos jogos, usando uma analogia com o contraste entre um jogo e um enigma (*puzzle*).

Assim sendo, um *puzzle* caracteriza uma situação de incerteza exógena, em que há algo que se ignora e cujo conhecimento implica a solução do problema, e o processo pelo qual este conhecimento é adquirido, bem como a própria solução, não afetam os termos do problema. Em termos contemporâneos, o *puzzle* é uma representação adequada de um problema típico da teoria da decisão, um ramo da teoria da escolha racional sob incerteza.

Um jogo, contrariamente, representaria uma situação de incerteza endógena, na qual as próprias tentativas de se alcançar a solução afetariam os termos do problema que se quer solucionar. Segundo Anderson e Moore, este é tipicamente o caso das interações entre indivíduos no

mundo social. Harsanyi (1982) afirmara que o esforço da teoria dos jogos deveria ser na direção de transformar um jogo em um problema de teoria da decisão, um *puzzle* com solução determinada. Contudo, Anderson e Moore insistem em que há diferenças por demais profundas entre as duas situações de sorte a estabelecer limites às ambições de qualquer teoria em fornecer soluções determinadas. A presença de intencionalidade é talvez o aspecto mais interessante apontado por ambos a emprestar complexidade problemática à interação. Assim, a possibilidade de conversão de um jogo em um *puzzle* torna-se no mínimo desinteressante.

A presença de intencionalidade complexifica o horizonte de previsão necessário à racionalidade dos indivíduos, uma vez que torna a aleatoriedade uma hipótese menos atraente para dar conta do tipo de incerteza a ser engendrado em interações entre atores sociais, dificultando, portanto, a "estratégia" teórica de "descontar" a incerteza. Diferentes intenções poderiam ser invocadas para dar conta das possíveis reações dos indivíduos às possíveis ações de outros com quem interagem em situações estratégicas, assim como as mais variadas motivações poderiam ser atribuídas a estes últimos em suas possíveis escolhas estratégicas. O fato de os atores estarem conscientes de que são objeto das conjecturas de outros leva-os a adotar comportamentos que podem ter múltiplos significados.² Como saber? E, talvez mais interessante: como sabê-lo sem afetar os próprios parâmetros do problema de decisão?

Nesse caso, a teoria dos jogos vale mais como uma heurística do que como uma teoria de conceitos matemáticos de solução para situações estratégicas. Indo mais além, a teoria dos jogos pertenceria ao conjunto de teorias que circulam no mundo social, afetando indelevelmente este mundo ao mesmo tempo que refletindo conceitos dele, conceitos sociais cujas raízes dificilmente poderiam ser encontradas por meio de um exame dedutivo das exigências da racionalidade ou da psicologia dos indivíduos.

Em resumo, a persistência de intencionalidade nas interações entre os indivíduos poderia turvar seus horizontes de previsão de forma a dificultar o tratamento da incerteza como aleatoriedade, por um lado; por outro, qualquer teoria sobre a interação social, mesmo a teoria dos jogos, cuja consciência mais atenta para a complexidade dessa interação é sua própria *raison d'être*, é habitante deste mundo, e nessa condição não poderá ser completa como teoria deste mundo. Esta última observação tem implicações negativas sobre a condição de publicidade da teoria, na forma de conhecimento comum (*common knowledge*), um conceito que apresentarei na próxima seção.

Em conclusão, aparentemente temos aqui dois programas de pesquisa divergentes: um que se aprofunda na busca de soluções determinadas para os jogos e outro que reivindica o menos ambicioso objetivo de descrever situações estratégicas sem solucioná-las, remetendo a um universo teórico que ultrapassa o ambiente puramente dedutivo e determinístico. Na realidade, reivindico neste artigo que o programa formal (por convenção) só funciona se acoplado ao programa informal (também por convenção); além disso, é possível considerar algumas possibilidades de elevação do *status* teórico do programa informal, o que é apontado na quarta seção do artigo com a incorporação de uma breve consideração de ordem filosófica.

Racionalidade estratégica (i): o suposto do conhecimento comum

Na busca de soluções determinadas para situações estratégicas, considera-se usualmente¹⁰ que o suposto do conhecimento comum é a condição de conhecimento relevante. Nesse caso, a possibilidade de existência de um equilíbrio em um jogo estaria também a depender do cumprimento da cláusula de conhecimento comum.¹¹ Esta afirmação precisa ser qualificada, o que faço a seguir.

Em primeiro lugar, deve-se ter claro a definição de conhecimento comum (CC, daqui para diante).

Informalmente, um evento é de conhecimento comum quando ele é conhecido pelas pessoas, todas as pessoas sabem disso, e todas sabem que todas sabem disso, e assim por diante; resumidamente, um evento e é CC de ordem n toda vez que (todo mundo sabe)ⁿ e for verdade.¹²

Em segundo lugar, há vários sentidos atribuíveis à condição de conhecimento comum. O sentido original desta condição foi elaborado pelo filósofo David Lewis (1969) no contexto de seu trabalho sobre convenções, onde ele pretendia estabelecer um argumento lógico que justificasse a possibilidade de coordenação social *à la* Hume, isto é, sem recorrer ao artifício de um contrato social ou acordo explícito entre as partes envolvidas. Lewis conclui ser racional para os indivíduos a adesão às convenções uma vez que de convenções se tratar a coordenação social, isto é, se as pessoas preferirem coordenar suas ações a não coordenar, se os pontos focais através dos quais elas passam a coordenar suas ações tiverem origem arbitrária, não resultando de um acordo explícito, e, finalmente, mas não menos importante, se estes pontos focais bem como aquelas preferências forem de conhecimento comum. Pessoas que preferem coordenar suas ações a não coordenar, que são comumente conhecidas por esta preferência, que ademais conhecem e se conformam às convenções que engendrarão a coordenação, sendo comumente conhecidas por estas crenças e disposições, promoverão coordenação social sem a necessidade de um contrato explícito e de sanções.

Temos duas questões importantes aqui. A primeira diz respeito à origem arbitrária das convenções, que pode tanto produzir "bons" quanto "maus" equilíbrios; isto é, as pessoas podem se conformar tanto a convenções que resultarão em cooperação efetiva quanto àquelas que resultarão em prejuízo generalizado, e neste caso a condição de CC não será suficiente para engendrar cooperação. A segunda questão é: em que sentido podemos dizer que crenças e disposições das pessoas são de conhecimento comum, no âmbito de nossa definição?

Com relação à última questão, Dupuy (1989) sugere dois sentidos possíveis. O primeiro interpreta a condição de CC

como uma espécie de "senso comum", refletindo certas idéias convencionais que as pessoas compartilham de forma mais ou menos igual. Esta interpretação, entretanto, não é a usual na teoria dos jogos. O segundo sentido sugerido por Dupuy parece corresponder melhor ao uso que se faz da condição de CC na teoria, e trata-se da idéia de que os jogadores conhecem a teoria dos jogos e tomam suas decisões de acordo com o que a teoria prescreveria para o tipo de jogo em que eles estão envolvidos. Este último sentido é, ainda, elaborado de duas formas diferentes, correspondendo à versão clássica ou moderna da teoria: no primeiro caso as decisões dos atores são tidas como independentes uma das outras e os atores são presumivelmente possuidores de informação completa; no segundo caso as decisões são de tal forma interdependentes que configuram uma situação de incerteza generalizada.

Jogos não-cooperativos e o conceito de equilíbrio de Nash

O aspecto atraente do modelo de convenção de Lewis resulta da sugestão de que a coordenação das ações dos indivíduos adviria não de um acordo explícito entre eles feito, o que implicaria problemas de visibilidade e monitoramento das ações, mas sim do desejo dos indivíduos de conformar-se ao comportamento do outro; seguir a convenção seria então racional, dada a preferência pela conformidade por parte dos indivíduos. Porém, conformar-se ao que os outros estão fazendo pode redundar em desastre coletivo, vide o comportamento de mercados especulativos quando há uma súbita reversão do estado das expectativas (ver Keynes, 1973; Dupuy, 1989; Orléan, 1989).

O modelo de jogos não-cooperativos baseia-se em estrutura de preferências distinta. Em princípio, um equilíbrio em tais jogos refletiria o tipo de coordenação que os indivíduos seriam capazes de engendrar sem recorrer a sanções externas, de um lado, e sem abdicar de sua racionalidade, de outro. Na realidade, a estrutura de preferências implícita nos jogos não-cooperativos representa uma situação na qual as pessoas estão prontas a cooperar com as outras apenas na medida em que a cooperação não conflitar com sua racionalidade individual

(maximizadora de utilidade); o equilíbrio representaria a compatibilização dos planos e das decisões dos indivíduos, a qual resultaria, por sua vez, de esforços consistentes de maximização de utilidade por parte dos indivíduos tomados isoladamente.

Os jogos não-cooperativos enfatizam, pois, o aspecto de compatibilização dos planos e decisões de agentes racionais, e portanto da possibilidade de um *equilíbrio* emergir, ao invés da questão da possibilidade da *cooperação*. Esta constatação já nos coloca um problema interessante que é o da *justificação*¹³ dos conceitos de solução em teoria dos jogos, isto é, um conceito de solução deveria ser expressivo da racionalidade da escolha individual, do equilíbrio e da eficiência das ações. Seria interessante examinar, então, o conceito de solução clássico dos jogos não-cooperativos, o equilíbrio de Nash (1951), tendo em vista a pedra de toque da *justificação*.

Como se sabe, os jogos não-cooperativos de soma não-zero encontraram uma solução com o desenvolvimento de um conceito de equilíbrio por Nash em 1951. Desde então, segundo Kreps (1987), o equilíbrio de Nash tem sido considerado condição necessária para que um "acordo" entre pessoas racionais seja auto-imposto, isto é, dispense a intromissão de coerção externa. Nessa categoria classificaria-se, por exemplo, a própria coordenação social resultante da interação entre indivíduos racionais "não-cooperativos" (isto é, para os quais a estratégia cooperativa é não-dominante): a coordenação social entre tais indivíduos só seria possível se passasse pelo teste de Nash. Apesar de necessário, o teste de Nash não seria suficiente para assegurar a cooperação, pois uma situação de não-cooperação generalizada por parte dos indivíduos também passaria no teste. E ainda, segundo Kreps, em alguns jogos não-cooperativos muitos equilíbrios de Nash poderiam surgir e, na falta de um critério de seleção claramente definido, as decisões dos agentes baseadas nestes equilíbrios poderiam redundar em desequilíbrio e irracionalidade coletiva.

Um exame mais minucioso do conceito de Nash indica, entretanto, que, em acréscimo às suas lacunas quanto a

atender aos critérios de justificação de um conceito de solução em teoria dos jogos, nomeadamente, escassas perspectivas de eficiência e equilíbrio único e determinado das soluções, o desenvolvimento deste conceito não liberou a teoria dos jogos do peso de supostos substantivos. Ou seja, a adoção, por razões de parcimônia, de uma estrutura de preferências por parte dos indivíduos que não concede à cooperação um papel proeminente como era o caso nos jogos de coordenação ou convenção analisados por Lewis não implicou a dispensa de supostos substantivos, dessa vez confinados à hipótese implícita quanto à mútua compatibilização das crenças de indivíduos racionais. Se não, vejamos.

Segundo Kreps, um equilíbrio de Nash em estratégias puras seria "um perfil de estratégias s tal que nenhum jogador isolado, pela alteração de sua parte em s , pode obter uma maior utilidade *se os outros mantêm suas estratégias*." (Kreps, 1987, p. 584; ênfase minha).

O equilíbrio de Nash caracteriza-se como o perfil de estratégias que representam a melhor resposta ao que os agentes consideram como *certo* que os outros que com eles interagem irão fazer; em particular eles consideram que agentes racionais não jogarão estratégias dominadas. Esta crença encontra respaldo na idéia de racionalidade dos agentes, que indica o comportamento maximizador de utilidade como o comportamento racional, bem como na idéia de conhecimento comum desta condição, vale dizer, os agentes numa interação estratégica sabem que são racionais, sabem que isto é conhecido, e assim por diante, o número de rodadas requerido correspondendo ao número de eliminações de estratégias dominadas possível no jogo. Está excluída, por exemplo, a possibilidade de erro por parte dos outros jogadores.

Implicitamente, portanto, as crenças racionais que sustentam a escolha da melhor resposta por parte de um jogador revelam simultaneamente uma compatibilidade mútua das crenças dos jogadores envolvidos na situação estratégica, como apontado por Heap e Varoufakis (1995). Numa situação estratégica, portanto, o comportamento racional por parte de um indivíduo caracteriza-se pelas

decisões que ele toma baseado em conjecturas ou crenças racionais no sentido forte, isto é, crenças que são internamente consistentes e ao mesmo também intersubjetivamente consistentes, pois são compatíveis com as crenças simetricamente cultivadas pelos demais agentes em interação.

Este tipo de exigência quanto às crenças associado ao conceito de Nash tem despertado reações na teoria dos jogos. Uma das mais famosas é o conceito de solução de "racionalizabilidade" (*rationalizability*), desenvolvido simultaneamente por Pearce (1984) e Bernheim (1984), segundo o qual é racional a decisão de um agente que puder ser "racionalizável", no sentido de as crenças que a sustentem serem *internamente* consistentes. Outro conceito de solução é o dos equilíbrios correlacionados (*correlated equilibria*), desenvolvido por Aumann (1987b), onde se substitui a hipótese de crenças mutuamente compatíveis pela assim chamada doutrina de Harsanyi, das probabilidades anteriores comuns, o que explico mais adiante. Seja como for, a noção de Nash aparece como forte demais no que se refere às crenças dos agentes, revelando-se portanto pouco "robusta" (segundo o jargão) em sua incapacidade de oferecer uma regra de decisão para o caso de erros ou desvios em relação às estratégias de equilíbrio.

Em síntese, a solução de Nash para os jogos não-cooperativos pode ser considerada, por um lado, fraca demais por sua incapacidade de fornecer equilíbrios únicos ou determinados¹⁴ e, em alguns casos, equilíbrios convincentes,¹⁵ além de oferecer esperança diminuta para a emergência de cooperação,¹⁶ e por outro, forte demais na medida em que implica a suposição de que existem crenças racionais que serão inequivocamente perseguidas por agentes racionais e cuja *rationale* é única. A solução parece ser, pois, duplamente má heurística, pois a hipótese heróica que a sustenta a de que crenças internamente consistentes são também mutuamente consistentes é incapaz de prover resultados únicos e determinados.

Tipos e erros

A possibilidade de existência de informação incompleta levou Harsanyi (1967/1968) a desenvolver um modelo de transformação de um jogo com informação incompleta em um jogo de informação completa (na realidade, de informação *imperfeita*), modelo que ficou conhecido na literatura como a "transformação de Harsanyi". Este modelo deu origem ao conceito de equilíbrio de Nash bayesiano. De acordo com este modelo, os jogadores lidarão com a incerteza informacional através da atribuição de probabilidades aos "tipos" que podem ser assumidos pelos seus contendores. A noção de racionalidade bayesiana implica que, na busca de maximização de utilidade, os agentes são capazes de atribuir probabilidades aos diferentes estados de natureza onde suas decisões potencialmente se darão, e, em particular, probabilidades epistêmicas (ver Good, 1990), números de probabilidades que refletirão suas crenças pessoais quanto à ocorrência daqueles estados mesmo na ausência de evidências estatísticas. Segundo Heap e Varoufakis (1995, p. 29),

Tudo o que é necessário é que o jogador mantenha expectativas prévias comuns em relação ao seu oponente [...] quanto à probabilidade de este se revelar um tipo ou outro de jogador, e o jogo torna-se assim um jogo de informação completa.

A adoção de probabilidades subjetivas parece permitir a definição de um comportamento racional mesmo sob informação incompleta. Nesta circunstância, o comportamento racional é aquele consistente com as restrições impostas pela adoção de probabilidades subjetivas.¹⁷

Com relação à possibilidade de que os agentes errem em suas escolhas estratégicas, Harsanyi (1982) oferece uma nova *rationale* para o uso de estratégias mistas, que as considera como o instrumento que reflete a sensibilidade dos jogadores quanto à possibilidade de que seus pares cometam pequenos erros ou desvios em relação ao perfil de estratégias de equilíbrio.

A questão é: como antecipar comportamentos que não são esperados da parte de indivíduos racionais? Qual a política

racional a adotar diante de possíveis "irrationalidades" cometidas por estes indivíduos?

De acordo com Aumann (1987b), a incerteza é caracteristicamente generalizada na interação estratégica, sendo de tal modo pervasiva que se torna quase inócuo acentuar o caráter independente das escolhas pessoais, como o faz a teoria dos jogos clássica. Dessa maneira, a informação que cada pessoa possui não é tão importante quanto as estimativas que ela faz quanto ao que supõe que os outros sabem sobre as informações de que dispõe, pois é com base nestas conjecturas que os seus oponentes tomarão decisões, as quais por sua vez afetarão as decisões dela. A epígrafe ao artigo de Aumann é esclarecedora a este respeito:

Se algum poder nos desse o dom
De nos ver como os outros nos vêem!

A ênfase nas informações possuídas pelos agentes, e em particular na condição de informação completa, bem como no conhecimento comum dessas informações, característica da teoria dos jogos na sua fase "clássica", estará sendo agora superada por uma abordagem que admite a incerteza estratégica como característica inerente aos jogos, e na qual os conceitos de equilíbrio de Nash bem como equilíbrio de Nash bayesiano serão considerados como casos particulares.

Racionalidade estratégica (ii): conhecimento comum da racionalidade bayesiana, segundo Robert Aumann

Em sua discussão dos equilíbrios correlacionados, Aumann alarga o espaço de incerteza de modo a abranger todos os aspectos que dizem respeito à tomada de decisão por parte do agente em um jogo, inclusive a informação de que dispõe, excetuados o suposto de racionalidade e de conhecimento comum desta. Na realidade, os supostos informacionais são substituídos pelo suposto forte de racionalidade bayesiana.¹⁸

Os jogadores são racionais na medida em que são maximizadores de utilidade em ambiente de incerteza e, correspondentemente, se conformam às condições de consistência impostas a tais escolhas, tais como propostas por Leonard Savage (1954). Basicamente, isto significa que os jogadores são capazes de atribuir probabilidades a todos os eventos possíveis que compõem seu problema de decisão, probabilidades subjetivas ou pessoais, no sentido de Savage.

O conjunto de crenças dos jogadores quanto aos parâmetros do jogo conforma um conjunto de "estados do mundo", a cada um dos quais eles atribuem uma probabilidade, vale dizer, a incerteza refere-se a que estado do mundo irá vigorar. Cada estado do mundo, por sua vez, reflete um equilíbrio correlacionado (ou não-independente), isto é, uma situação onde os parâmetros estão fixados e o perfil de estratégias de equilíbrio é, portanto, dedutível do comportamento maximizador de utilidade dos agentes, bem como do conhecimento comum desta condição.

A idéia de uma escolha pessoal *independente*, observa Aumann, já não faz tanto sentido quando o problema de decisão estratégica é modelado dessa forma, pois a racionalidade da escolha pessoal de um agente é contingente à escolha pessoal do outro jogador com quem interage. Em particular, a escolha pessoal de um jogador não depende apenas de uma conjectura sua quanto à ação possível do outro, mas também, e talvez mais importantemente, de sua estimativa sobre a conjectura do outro quanto à sua própria escolha.

Probabilidades, regra de Bayes e probabilidades anteriores comuns

É interessante examinar, a esta altura, as noções de probabilidade, decisão racional num contexto de incerteza, e de incerteza propriamente dita, tais quais vislumbradas por Aumann.

Como dissemos, a noção de racionalidade bayesiana implica que os agentes racionais sejam capazes de atribuir probabilidades subjetivas a todos os aspectos sujeitos a

incerteza do problema de decisão que enfrentam. Na realidade, é também suposto que probabilidades subjetivas atribuídas por diferentes agentes aos mesmos eventos convergem objetivamente graças à intercessão de algum mecanismo, no caso a regra de Bayes. Aumann procura encontrar respaldo para essa visão nos desenvolvimentos da teoria da probabilidade por Savage. Entretanto, como veremos, Savage reconhece ao menos três diferentes interpretações para o significado do conceito matemático de probabilidade.

Segundo Savage, o conceito matemático de probabilidade é crucial para a definição das condições de consistência de qualquer ação em face da incerteza. Contudo, diz ele, ainda que o conceito axiomático de probabilidade seja bastante claro, existe muita discordância com relação à sua interpretação. Nesse particular, Savage identifica três classes de interpretações da noção matemática de probabilidade: (a) a objetivista (ou freqüentista), que entende a probabilidade como exprimindo repetições observadas de um evento; (b) a personalista (ou subjetivista, visão esposada por Savage), que interpreta a probabilidade como medida do grau de confiança de um indivíduo na verdade de uma proposição particular, permitindo variação deste grau entre indivíduos razoáveis confrontados com a mesma evidência factual; e (c) a necessária (ou lógica), para a qual a probabilidade mede a extensão em que um conjunto de proposições confirma a verdade de outro, por simples necessidade lógica, desconsiderando opiniões.

Com esse mapa de interpretações, podemos inferir que a visão bayesiana implica uma conjunção das leituras subjetivista e necessária de probabilidade, na medida em que ela aceita que as probabilidades são geradas subjetivamente mas requer que elas converjam para uma medida objetiva, que não é nem empírica nem resulta de um acordo intersubjetivo, mas sim de uma necessidade lógica, apriorística.

Savage tece críticas importantes à visão necessária, tanto em termos de sua resistência mínima à evidência factual quanto em termos de sua plausibilidade:

É atraente supor que, se dois indivíduos na mesma situação, tendo os mesmos gostos e dispondo das mesmas informações, agem razoavelmente, eles agirão da mesma maneira. Tal acordo, crença na qual repousa a visão necessária (em oposição à personalística) de probabilidade, certamente vale a pena ser procurado. Pessoalmente, acredito que isto não corresponda nem mesmo grosseiramente à realidade [...] Eu insisto em que, até que o contrário seja demonstrado, devemos estar preparados para julgar o raciocínio inadequado para o alcance do acordo completo. [...] Pode ser, e de fato eu assim o creio, que haja um elemento na decisão, para além do gosto, sobre o qual, assim como o gosto, não possa haver disputa. (Savage, 1954, p. 7)

Na verdade, Savage também observa que, em que pese a persistência de um elemento irredutivelmente subjetivo na estimativa pessoal de probabilidade, crenças pessoais que se mostram erradas quando confrontadas com testes de consistência devem ser corrigidas. De qualquer forma, esta afirmação de Savage não equivale a subscrever à visão objetivista, e muito menos à visão necessária de probabilidade.

Em qualquer caso, o problema de decisão ao qual Savage se refere, e para a solução do qual supõe-se que indivíduos racionais sejam capazes de proceder a cálculos probabilísticos, é o de escolha sob incerteza quanto a eventos *externos* à escolha. Pode-se justificadamente perguntar se há sentido em estimar probabilidades necessárias quando há interação entre agentes e ambiente, como ocorre na interação social; nesta, as estimativas de probabilidade refletem as expectativas ou crenças dos agentes concernentes ao comportamento dos outros, os quais, por sua vez, também tecem expectativas com relação às ações dos primeiros. Ou seja, a incerteza que suscita o cálculo probabilístico diz respeito a eventos endógenos. É lícito, então, questionar se existe, neste contexto, algum mecanismo pelo qual crenças são geradas e revistas de modo a se tornarem crenças racionais e, portanto, convergirem.

Para os "bayesianistas"¹⁹ existem dois mecanismos que garantem tanto a consistência interna das crenças, tornando-as racionais portanto, quanto a consistência intersubjetiva entre elas: a regra de Bayes e o suposto de probabilidades anteriores comuns (*common priors assumption*).

No contexto da teoria da decisão, onde não há incerteza endógena, a regra de Bayes é interpretada como um mecanismo de revisão de estimativas incorretas, uma vez que os agentes vão sendo providos com novas e melhores informações; no contexto da interação estratégica, com incerteza endógena portanto, onde as próprias estimativas afetam o que é estimado, a possibilidade de convergência das crenças entre agentes racionais estratégicos seria garantida pela assim chamada "doutrina de Harsanyi". Segundo esta doutrina, indivíduos que possuam as mesmas informações devem tecer as mesmas crenças, conclusão que se segue à premissa de que apenas diferenças em informação podem justificar diferentes crenças. Esta doutrina é também conhecida como o suposto das probabilidades anteriores comuns (PAC).

Normalmente, não bayesianistas (mas bayesianos alguns deles, como Binmore) reagem à estratégia de interpretar a regra de Bayes como um mecanismo de aprendizado (Binmore, 1993) ou como a descrição sintética de um processo mental racional em termos de um processo inferencial (Heap e Varoufakis, 1995). Segundo Binmore, a regra de Bayes dificilmente pode ser vista como um processo de aprendizado genuíno, porque quando os agentes passam a rever suas crenças da maneira indicada pela regra, o processo de aprendizado real, no sentido de compreensão da experiência vivenciada, já terá terminado. O processo de aprendizado seria entendido, então, como um processo passivo de absorção de dados não ambíguos, e com isso perde-se a perspectiva, se não adversária, ao menos mais problemática de perceber como a própria percepção dos fatos pode modificar estes fatos e produzir efeitos que nos interessem representar teoricamente.

Heap e Varoufakis (1995), por sua vez, parecem advertir para o fato de a simplificação redundar em simploriedade, ao rejeitarem a redução que os bayesianistas fazem do processo mental racional em termos exclusivamente de um processo inferencial, por fornecer uma imagem parcial e mecânica de racionalidade. Pode-se inferir da reação destes autores que esta imagem de racionalidade exclui os processos de *juízo* pelos quais regulamos a aplicação das regras de inferência lógica: "[Se a razão não passa de um

conjunto de regras de inferência] é possível que algum conjunto finito de regras contenha regras para sua própria aplicação a todas as circunstâncias possíveis?" (Heap e Varoufakis, 1995, p. 59).

Com relação ao suposto de probabilidades anteriores comuns, Aumann (1976) afirmara que com probabilidades anteriores iguais as pessoas não podem "concordar em discordar" mesmo que elas tenham partido de contextos de informação assimétrica; conhecendo as probabilidades posteriores umas das outras quanto ao mesmo evento, ao longo do tempo elas tenderiam a rever seus preconceitos originais e aproximar-se das "verdadeiras" probabilidades. Mais genericamente, o único motivo que levaria pessoas racionais a divergir em suas estimativas de probabilidade seria o fato de terem acesso a diferentes informações; na ausência desta diferença elas tenderiam a exibir PAC.

De acordo com Morris (1995), o suposto das PAC não é tão inofensivo quanto parece à primeira vista. Em primeiro lugar, a co-extensividade entre crenças e informações poderia ser questionada, isto sem nem mesmo recorrer à questão da interpretação da informação, e meramente fazendo referência ao trabalho de Savage, que desenvolveu sua noção de probabilidade pessoal por sobre a saudável premissa de que pessoas diferentes expostas às mesmas evidências são capazes de entreter crenças diferentes sem que isso implique falta de racionalidade da parte delas. É claro que Savage acrescenta, como já dissemos, que este desacordo não poderá durar em face dos testes de consistência que as propriedades numéricas da probabilidade naturalmente impõem. Entretanto, diz Morris, estamos aqui confrontados com estimativas de probabilidades referentes a eventos *endógenos*, vale dizer, eventos cuja probabilidade de ocorrer é afetada pelas ações dos indivíduos cujo comportamento se quer modelar. Neste caso, o sentido de PAC torna-se ainda mais obscuro, pois trata-se de estimativas de probabilidades do comportamento dos outros (o qual, por sua vez, é afetado pelas crenças quemantemos quanto a ele), ao invés de algum evento externo, como a meteorologia.²⁰

A introdução do suposto das probabilidades anteriores comuns por Aumann parece visar garantir que mesmo em ambiente de incerteza estratégica (aquela engendrada pelo próprio raciocínio estratégico dos jogadores) e ignorância generalizada quanto aos principais parâmetros do jogo pode-se construir uma noção de equilíbrio. Recorde-se que, em comparação com o conceito de Nash clássico, a noção de equilíbrio correlacionado defendida por Aumann é menos forte em seu suposto quanto às crenças de agentes racionais, pois não requer, como aquele, conhecimento certo quanto à estratégia escolhida pelo outro jogador, ao não atrelar as condições de consistência interna das crenças de um jogador às condições de consistência recíproca das crenças dos jogadores que interagem. Ao mesmo tempo, a noção de equilíbrio de Aumann parece endereçar-se ao tipo de incerteza que Anderson e Moore tinham em mente quando sugeriram que a mera presença de intencionalidade na escolha dos agentes torna a coordenação entre eles teoricamente intratável.

Se pudermos aceitar o suposto das PAC e ao mesmo tempo interpretar a regra de Bayes da maneira indicada por Aumann, teremos conseguido desembaraçar a noção de racionalidade de seu incômodo corolário da intencionalidade, e assim ingressado na fase moderna da teoria dos jogos, deixando para trás (ou emprestando novas conotações a) o seu período "iluminista"; em particular, teremos encontrado o procedimento racional pelo qual as crenças de indivíduos racionais se tornam alinhadas.

Desde logo, a pergunta que se impõe é: qual será então a nova "agência" do modelo de Aumann?

Um mundo aberto e o observador externo

Num mundo de incerteza generalizada, a suposição de escolhas independentes ou ação intencional e consciente é de pouca valia. Como na epígrafe anteriormente mencionada, os indivíduos que interagem apresentam grande interesse em conhecer sua imagem, tal qual percebida pelos outros, e, nesse sentido, um indivíduo pode se pensar como no mínimo dois, quer dizer, ele pode hipoteticamente destacar-se de si mesmo e considerar "de

fora" quantas pessoas diferentes se pode supor que ele seja. Para cada uma delas, vale dizer, dados certos parâmetros e um certo nível de informação que se atribui a cada uma dessas "pessoas", há uma estratégia de equilíbrio que ele pode "escolher" dada a escolha dos outros; ou seja, dentro de cada um destes "estados do mundo", o indivíduo escolhe sua estratégia ótima, aquela que maximiza sua utilidade esperada, e tudo o que ele precisa supor sobre os outros é que eles são racionais bayesianos e que isto é de conhecimento comum.

Voltando à idéia da bipartição do indivíduo, o indivíduo "de fora" considera quantas "pessoas" diferentes ele pode ser (correspondentes às distintas conjecturas que os outros jogadores fazem acerca das informações que ele possui) e atribui uma distribuição de probabilidades a estes n estados do mundo; em cada um deles, um equilíbrio correlacionado ou não-independente emerge como um corolário do suposto de maximização da utilidade esperada e do conhecimento comum desta condição.

Ora, dentro de cada estado do mundo, os jogadores "escolhem" no sentido da teoria dos jogos clássica; isto é, dada a escolha dos outros, cada um elege sua estratégia ótima. Contudo, que estado do mundo irá prevalecer e portanto que pessoa o jogador se revelará é matéria ignorada pelos jogadores, e esta situação se expressa na distribuição de probabilidades atribuída pelo indivíduo "de fora" àqueles estados.

A idéia de um observador externo é explicada por Aumann como um certo ponto de vista que qualquer um é capaz de ocupar e que fornece uma visão da totalidade das possíveis contingências que cercarão sua decisão. Do ponto de vista teórico, o observador externo é justificado por Aumann como um artifício para se lidar com a ignorância radical permitida pelo modelo; ele permitiria a especificação de um modelo totalmente inclusivo (*all-inclusive*).²¹

Neste contexto, então, a incerteza traduz-se em uma distribuição de probabilidades subjetivas atribuída por um observador externo ao conjunto de contingências possíveis que compõem os ambientes de decisão dos indivíduos. Em

contraste com a noção clássica de Nash, e mesmo com o equilíbrio de Nash bayesiano, o modelo de Aumann representa uma genuína expansão do espaço de incerteza, pois acomoda um amplo espectro de crenças quanto às crenças dos outros jogadores; não obstante, a cláusula de crenças iniciais iguais sob incerteza representa uma importante (e questionável) condição a garantir a consistência intersubjetiva entre elas.

Em síntese, a interpretação peculiar de incerteza e de probabilidade oferecida por Aumann tem o efeito de resolver um jogo ao transformá-lo em um problema de decisão sob incerteza (exógena), ou um *puzzle*, na terminologia sugestiva de Moore. Há um efeito adicional ao qual se deve fazer referência que é o fato de o conceito de solução de Aumann não se enquadrar no repertório tradicional de justificativas em teoria dos jogos, pois não há indicação teórica de como serão selecionados equilíbrios únicos dentre os múltiplos equilíbrios correlacionados.

Uma série de questões emerge como consequência da abordagem de Aumann. Uma primeira seria: como justificar a posição do observador externo nos termos da visão inicial de Aumann de incerteza estratégica?

Dois problemas apresentar-se-iam de imediato, o primeiro de natureza lógica e o outro epistemológico. Do ponto de vista epistemológico, se não há nenhum suposto metafísico excluindo a pessoa "de fora" do próprio sistema de crenças ao qual ela tenta dar expressão e atribuir probabilidades, não há como desconsiderar as conjecturas dessa pessoa como meras conjecturas; quer dizer, as crenças do "meta"-jogador também deveriam constituir um dos estados do mundo possíveis, e assim também com relação às crenças do meta-meta-jogador e assim sucessivamente. Entretanto, se o conjunto de estados do mundo fosse finito, tal situação não seria logicamente admissível: como fazer caber dentro desta lista a própria lista e a lista da lista, e assim sucessivamente?

Não obstante, o conjunto de estados do mundo poderia concebivelmente ser aberto ou incompleto, e esta parece ser a sugestão de Aumann (1987) quando insinua que, em

sua abordagem, perde todo o sentido a distinção entre estados "internos" ou "externos". Neste caso, entretanto, estaríamos fora do campo onde o uso de probabilidades seria justificável, ao menos segundo a abordagem canônica de Savage (1954), para quem o uso de probabilidades seria legítimo apenas em um universo fechado (onde se pode "olhar antes de saltar"), em oposição a um universo aberto (onde se pode "cruzar as pontes apenas quando as alcançamos") (ver também Binmore, 1993).

Adicionalmente, caberia ainda justificar a possibilidade prática de tal capacidade por parte dos jogadores, os quais, ainda que radicalmente ignorantes, são dotados de ilimitada inteligência cognitiva e computacional. Se o modelo estiver no caminho certo, o (meta)-jogador dotado de hiper-racionalidade terá assumido o lugar do cientista behaviorista que imaginava o fenômeno social como materialmente alheio a si, e a interação estratégica terá sido resolvida em termos de um experimento behaviorista.

Em um mundo menos-que-perfeito, a imagem sugerida por Aumann na epígrafe ao seu artigo de 1987 poderia acrescentar mais nuances e meios tons do que plena luz: o fato de que temos de levar em consideração não tanto o que sabemos que sabemos a nosso respeito quanto aquilo que supomos ser conhecido dos outros não nos permite tomarmos os outros como espelhos onde podemos ver claramente; a imagem que os outros nos devolvem de nós mesmos deve sem dúvida ser afetada pela imagem que fazemos deles, e ela já será obsoleta enquanto tentamos fixá-la. Em termos práticos, poderemos estar sendo inundados por um excesso de informação que será inútil como guia para a decisão racional; em termos epistemológicos, deveremos ser concebidos como agentes hiper-racionais (cujas crenças convergirão) capazes de antecipar todas as contingências possíveis e de definir uma política incondicional (uma distribuição de probabilidades) para lidar com elas. No primeiro caso, o modelo não é restritivo o suficiente; no segundo, é restritivo em demasia.

Uma suspeita permanece, e parece constituir boa razão para investigação posterior, qual seja: a de que estimar a ignorância é apenas possível quando nossos esforços para

medi-la são suficientemente independentes dela. A hipótese de interdependência entre conhecimento e ignorância será desenvolvida, de modo preliminar, ao fim da próxima seção.

Instituições: indeterminação, natureza e filosofia social

A adoção de supostos de racionalidade fortes, como na visão bayesiana, tem encontrado certa resistência entre os teóricos dos jogos. Uma dessas reações, aqui representada pelas posições do teórico alemão Fritz Scharpf, tem a intenção de prover uma nova justificativa, em bases filosoficamente mais plausíveis, para a teoria dos jogos clássica, vale dizer, para a manutenção do suposto de informação completa. Uma outra, a qual também se faz referência, é a da teoria dos jogos evolucionária, cujo trabalho seminal é o livro de 1982 de John Maynard Smith, que faz objeção à aparente uniformização dos agentes econômicos promovida pela adoção do suposto de racionalidade plena. Ambas as reações adotam a idéia de racionalidade limitada por parte dos agentes individuais e sugerem que algum mecanismo externo (à racionalidade) suplementará o comportamento individual de forma a propiciar a emergência de equilíbrio e, possivelmente, também de cooperação.

Segundo Fritz Scharpf, o suposto de racionalidade forte, isto é, de ilimitada capacidade computacional por parte de seres humanos, é muito mais implausível do que a condição de informação completa, algo que pode ser de alguma forma socialmente construído. Ele adota, então, o suposto de racionalidade limitada, e considera que a interação social é capaz de engendrar mecanismos que permitam melhorar o nível de informação dos indivíduos, ampliando assim o grau de previsibilidade na sociedade.

Estes mecanismos de interação podem ser individual ou socialmente construídos de modo a aumentar o grau de previsibilidade mútua, e são as instituições, regras, convenções e hábitos que podem permitir a redução dos

custos de informação bem como a ampliação de opções estratégicas exequíveis. Estes mecanismos, diz ele, podem substituir a condição de conhecimento comum pela de "senso comum"; eles tornam a coordenação possível ainda que não necessariamente eficiente.

Além dos mecanismos interindividuais que podem ser engendrados pela interação entre os indivíduos e que têm como efeito elevar o grau de previsibilidade recíproca na vida social, Scharpf faz referência ao fato de que as escolhas dos indivíduos são condicionadas por um elemento subjetivo, qual seja, as percepções dos atores e suas preferências entre os resultados possíveis, as quais não podem ser explicadas em termos do comportamento auto-interessado sem qualificação adicional. Em outras palavras, as motivações dos atores em suas escolhas são também elas condicionadas, externa e internamente, por critérios de relevância e correção, que são socialmente construídos, como, por exemplo, normas de justiça universalística, moralidade e equidade, ou ainda em termos de papéis definidos socialmente e identidades coletivas às quais freqüentemente os atores se referem quando fazem sua avaliação auto-interessada dos resultados (Scharpf, 1990, p. 485).

Scharpf não desconhece as inúmeras tentativas, no campo da teoria dos jogos, de deduzir as instituições endogenamente, como resultado da escolha racional dos indivíduos. Sua reação a essas tentativas é a de que para se obter acordos auto-impostos, como quer a solução endógena, supostos muito irrealistas deveriam ser incorporados, tais como: baixo custo na escolha dos parceiros com quem interagir, baixo custo de monitoramento da *performance* passada de potenciais parceiros, custos de informação zero na descoberta de percepções e preferências idiossincráticas dos parceiros.²²

Levar em consideração aqueles mecanismos, Scharpf reconhece, eleva a complexidade da teoria, mas esta parece ser escolha melhor do que construir uma teoria dedutiva baseada em supostos tão implausíveis.

O ponto de vista da teoria dos jogos evolucionária (Binmore, 1993 e 1994; Kirman, 1992 e 1995) é sob certos aspectos bastante diferente daquele professado por Scharpf, ainda que o ponto de partida comum seja a descrença em relação à acurácia do suposto de *full-blown rationality* por partes dos indivíduos.

Na realidade, a visão de interação social sustentada pelos teóricos evolucionários, ainda que estratégica, difere bastante da abordagem tradicional da teoria dos jogos não-cooperativos (tanto em sua versão clássica quanto moderna), representando uma reação à visão do sistema econômico pela lente do indivíduo representativo.²³

Contra a imagem do indivíduo representativo, bem como contra a uniformização do agente econômico empreendida pela teoria dos jogos não-cooperativos, Kirman (1995) visualiza o sistema econômico como interação entre indivíduos heterogêneos, com racionalidade limitada, cujo comportamento adaptativo, na base da tentativa e do erro, freqüentemente imitativo, será finalmente sancionado pela seleção das estratégias que se evidenciarem *ex-post* melhor adaptadas ao ambiente. Portanto, o fenômeno que deve ser teoricamente deslindado é o de como indivíduos heterogêneos, com racionalidade limitada, e que interagem ao nível local, são capazes de promover uma interação global dando origem à coordenação no sistema como um todo. Os agentes econômicos, limitados por sua localização no sistema bem como por sua racionalidade menos-que-perfeita, tentam se adaptar ao ambiente mediante comportamentos imitativos ou processos de tentativa e erro; neste processo emergirão estratégias e o acaso selecionará dentre elas as que mostrarem vantagens adaptativas.

Em síntese, as duas alternativas dentro do campo da teoria dos jogos aqui sucintamente apresentadas possuem a característica comum de supor que a racionalidade individual instrumental é tipicamente insuficiente para engendrar tanto equilíbrio quanto cooperação entre os indivíduos, sendo necessário, em conseqüência, modificar o suposto de racionalidade e complementá-lo com algum mecanismo externo. A motivação de Scharpf parece ser a

do irrealismo dos supostos, ao passo que a de Kirman, a visão distinta do objeto que a teoria quer explicar; em outras palavras, Scharpf parece advogar uma *teoria* mais complexa (menos determinada e mais densa descritivamente), enquanto Kirman parece sugerir uma teoria do *objeto* complexo.

Enquanto Scharpf reconhece a importância da construção de mecanismos sociais para o incremento da previsibilidade no mundo social, renuncia no mesmo ato à possibilidade de teorização quanto a esses mecanismos, indicando os limites da abordagem dedutiva no que diz respeito a esta questão. A perspectiva da teoria dos jogos evolucionária remete, por outro lado, a uma explicação "evolucionária" da emergência desses mecanismos, abdicando, contudo, de uma investigação mais aprofundada das chances de ocorrência dessas "emergências", indicando uma possível "naturalidade" nelas, reunindo as formas sociais e naturais de vida num mesmo não filosofável campo.

Ainda que se possa reconhecer a extrema dificuldade em se deduzir as instituições endogenamente, vale dizer, da escolha racional de indivíduos, esta situação não nos coloca necessariamente diante da fatalidade de localizá-las na natureza, qual mutações genéticas de origem e conseqüências *ex-ante* desconhecidas. Podemos pensar as instituições, como de resto as regras, hábitos e tradições na vida social, como formas sociais de conhecimento cuja origem pode ser localizada precisamente na ignorância dos indivíduos. Podemos compreender as instituições, pois, dentro de uma determinada visão de mundo social que se admira não das possíveis falhas e erros dos indivíduos, mas sim do conhecimento que eles venham a possuir: "A situação aqui parece ser a seguinte: antes que possamos explicar por que as pessoas cometem erros, devemos primeiramente explicar por que alguma vez elas acertaram." (Hayek, 1949, p. 34).

À pergunta "o que podemos conhecer?" Hayek responde: muito pouco, em conseqüência da complexidade peculiar do mundo social. Na realidade, sempre em virtude dessa complexidade, quanto mais conhecemos menos conhecemos, pois a nossa busca bem-sucedida por porções

adicionais de conhecimento tem como conseqüência a amplificação, por efeito de composição, da quantidade de conhecimento acumulado no todo que a sociedade é. Segundo Hayek, nosso conhecimento está limitado em sentido não trivial pela nossa ignorância radical, de raiz portanto, que se refere ao fato de que ocupamos no mundo social posições que são inerradicavelmente parciais.

No contexto de ignorância radical, os atores sociais são essencialmente formuladores de hipóteses ou "teorias" uns sobre os outros; portanto, também em sua ignorância esses atores são produtores de conhecimento a ser adquirido por outros que com eles interagem. No entanto, estas conjecturas recíprocas poderiam ser incomensuráveis (não há por que assumir uma convergência entre elas sem pôr a perder a própria visão de mundo social como um mundo de conhecimento e de mudança constante) não fosse pela existência de regras abstratas que os indivíduos adquirem, muitas das vezes de modo não perceptível e freqüentemente inarticulável, e que os habilitam a comunicar-se e a fazer sentido de suas ações recíprocas.

Para Hayek, trata-se, na verdade, de um sistema de regras de vários níveis, das mais abstratas às menos, das mais inarticuláveis às mais conscientes, englobando a linguagem, os costumes, as tradições e hábitos, as instituições de modo geral. Quanto mais abstratas e não-conscientes são as regras, mais sujeitas estarão a interpretações particulares, idiossincráticas dos indivíduos, e, portanto, plausivelmente contraditórias. No entanto, são a estas que recorremos em intercursos com estranhos na vida ordinária, quando desejamos compreender o que eles estão fazendo ou adivinhar-lhes os desígnios. Nessas ocasiões, atribuímos sentido ao que os outros estão fazendo em analogia com o modo como julgamos deva-se agir em tais circunstâncias, quanto às regras que a elas nos parecem aplicáveis. Dessa forma, fazemos uso das regras e as modificamos constantemente em reação às circunstâncias que nos põem em intercuro com os outros. Adquirimos e produzimos conhecimento no mesmo ato em que tentamos fazer sentido daquilo que os outros estão fazendo ou de suas intenções, antes de reagirmos.

Da mesma forma, para Hayek, nossas ações são de algum modo interações *ex-ante*, uma vez que nossas finalidades e os meios de as alcançar são de algum modo constituídos ou melhor, contidos e restritos pelo sistema de regras que, como vimos, não é de modo algum um padrão de referência estático. O que é o interesse próprio de um indivíduo (assim como os meios de o satisfazer) sem esse sistema de referências a conferir-lhe sentido, sistema que evolui no tempo e é essencialmente intersubjetivo? O que se ganha em termos de compreensão da vida social quando se focaliza sobre a escolha racional dos indivíduos tendo em vista a visão de mundo social que aqui se apresenta? Ao invés de postular a racionalidade da escolha, esta perspectiva discute precisamente as condições de possibilidade de tal suposto, concentrando-se no exame das formas de conhecimento socialmente produzidas, raramente articuláveis ao nível dos indivíduos, sem as quais a própria racionalidade destes torna-se impensável.

Comentários finais

De acordo com o percurso feito neste artigo, o programa dedutivo da teoria dos jogos, qual seja, fornecer uma representação axiomática da interação estratégica em termos de uma racionalidade estratégica apenas, é incompleto. Em particular, as tentativas aqui relatadas de resolver o problema da incerteza estratégica o fato de que os agentes devem estimar as ações e reações uns dos outros às suas escolhas recíprocas antes de escolher têm implicado a adoção de supostos substantivos de difícil justificação e de baixa densidade filosófica, como a doutrina de Harsanyi quanto à origem das crenças, ou o uso da regra de Bayes como um mecanismo de aprendizado.

Em minha leitura, as sugestões de Scharpf e Kirman caminham na direção de aninhar o programa dedutivo em uma reflexão teórica mais ampla que possa fornecer elementos para se pensar as condições de atendimento dos supostos requeridos pelo programa dedutivo. Assim, por exemplo, Scharpf sugere que o suposto de informação completa pode ganhar plausibilidade se a informação for

pensada como algo produzido socialmente, de forma direta, via instituições e regras, ou indireta, através da construção da percepção e da própria identidade dos agentes. Kirman, por outro lado, sugere que uma racionalidade incompleta dos agentes econômicos não pode explicar a interação sistêmica entre eles, sendo necessário admitir-se a produção de efeitos emergentes das interações locais entre agentes limitadamente racionais cujo comportamento vai sendo selecionado ("racionalizado") em função de seu sucesso adaptativo em termos globais.

Evidentemente, as alternativas aqui examinadas apresentam suas próprias fraquezas. Pode-se, por exemplo, legitimamente questionar a posição de Scharpf quanto ao papel limitado da teoria e considerar-se que a linha de demarcação que ele propõe entre o que é teorizável e o que não é seja demasiado próxima do ponto de partida. Alguns desejariam, por exemplo, conceber uma teoria dos mecanismos que aumentam a previsibilidade na sociedade. Com relação à proposta de Kirman, pode parecer inaceitável conceder à seleção, um mecanismo cego que abdica em demasia do controle da razão sobre os fenômenos sociais, a importância central que ela assume em sua abordagem.

Entretanto, considero que há algo de bastante atraente nestas propostas, que é a convicção comum de que a racionalidade humana possui também um componente social (e histórico), algo que, se por um lado pode tornar nossos esforços de teorização mais problemáticos, por outro os torna certamente mais interessantes e persuasivos. Considere, por exemplo, que ao invés de tomar como dadas as condições de conhecimento dos agentes, refletíssemos sobre elas, investigando, em particular, *como* os agentes adquirem conhecimento em sociedade, quais são os processos pelos quais eles tornam suas ações mais bem informadas, partindo da premissa inversa da teoria convencional de que a condição "bíblica" dos indivíduos é a ignorância.

Conviria qualificar esta ignorância para diferenciá-la daquela tratada por Aumann, a qual se referia à falta de informação suficiente quanto às informações e estimativas

possuídas por outros agentes. A ignorância a que me refiro aqui diz respeito à posição ontológica do indivíduo, posição da qual ele não pode escapar em virtude de sua condição na sociedade, de ser apenas parte e de não ser dotado de perspectiva panóptica.²⁴ Na falta dessa capacidade panóptica e sinóptica, os indivíduos em sociedade recorrem, nem sempre de modo consciente e articulado, a formas de conhecimento social, condensadas nas regras, instituições e tradições. O que, contudo, torna a solução parte do problema é o fato de que assim agindo os indivíduos acabam produzindo mais conhecimento a ser controlado, pois no ato de consumir aquelas formas de conhecimento social eles adicionam suas interpretações a elas, de modo que, ao adquirir conhecimento os indivíduos simultaneamente produzem mais conhecimento a ser adquirido. Segundo esta visão, a distinção entre informação e crença, no sentido de Aumann, perde nitidez, mas, em compensação, ganha-se uma explicação para isso.

Seja como for, esta é apenas uma alternativa possível e, para os fins deste artigo, não constitui a mensagem central. Talvez a mensagem mais interessante seja o convite à reflexão sobre os supostos da teoria dos jogos em bases filosoficamente mais bem fundadas, mediante a cooperação entre uma discussão mais explícita desses supostos e o programa dedutivo da teoria, mantendo-se fiel à visão de interação complexa proposta por von Neumann-Morgenstern e Anderson-Moore. Esta cooperação poderia estabelecer bases mais sólidas para a compreensão da possibilidade da escolha racional dos indivíduos. É claro que sempre se pode apostar em novos desenvolvimentos matemáticos que possam tornar a visão de von Neumann determinada (ver Tan e Werlang, 1992). A aposta deste artigo é no sentido do *making sense* da teoria dos jogos e, portanto, da crença em um diálogo interdisciplinar com a Filosofia,²⁵ tanto em termos da busca de justificativas para nossos conceitos e argumentos quanto em termos de confrontarmos a teoria com uma visão do mundo social.

NOTAS

¹ O divisor de águas entre as visões clássica e moderna da teoria dos jogos localizar-se-ia na incorporação, por esta última, de probabilidades subjetivas. Ver Aumann (1987b).

² Relatos pormenorizados podem ser encontrados em Mirowski (1992) e Leonard (1995), de onde extraí os principais elementos da colaboração entre von Neumann e Morgenstern. Antecedentes da teoria dos jogos remontam a Cournot (1838) e Edgeworth (1881). Ver Aumann (1987a) para uma história resumida dos desenvolvimentos mais importantes na teoria desde os anos 1910 até 1986; ver também, a este respeito, Mirowski (1992).

³ Na ocasião, von Neumann declarou o seguinte: "The very concept of 'absolute' mathematical rigor is not immutable. The variability of the concept of rigor shows that something else besides mathematical abstraction must enter into the makeup of mathematics." (*apud* Mirowski, 1992, p. 122).

⁴ Von Neumann (*apud* Mirowski, 1992) expressou e generalizou esta intuição no que diz respeito à natureza dos fenômenos sociais da seguinte forma: "any event could be looked at as a game of strategy provided one looked at the effect it had on the participants".

⁵ Morgenstern expressou essas convicções da seguinte forma: "always there is exhibited an endless chain of reciprocally conjectural reactions and counter-reactions. This chain can never be broken by an act of knowledge but always through an arbitrary act a resolution [...] Unlimited foresight and economic equilibrium are thus irreconcilable with one another." (*apud* Mirowski, 1992, p. 129, nota de rodapé).

⁶ Segundo Morgenstern, "social sciences have the peculiarity of being able to affect their object of study. The prediction of the astronomer can have no effect on the subsequent movement of the stars, but that of the economists *can* change economic events." (*apud* Leonard, 1995, p. 741).

⁷ Trata-se, na verdade, da condição mais estrita de "conhecimento comum", como explico mais adiante.

⁸ No original: "[a game has the capacity to] change suddenly on us *with a view to thwarting our efforts at a solution* [...]".

² A propósito, Anderson e Moore referem-se aos resultados de questionários aplicados entre estudantes cujas respostas foram intencionalmente "deturpadas" em reação à própria pesquisa. As motivações que levaram os estudantes a "deturparem" suas respostas foram as mais variadas, incluindo desde propósitos definidos até a motivação anárquica de "perturbar" a pesquisa.

¹⁰ Exceções serão examinadas mais adiante no texto, sendo a mais saliente a da teoria dos jogos evolucionária.

¹¹ A ressalva deve ser feita em relação a jogos como o "dilema dos prisioneiros", em que a eliminação de estratégias dominadas é suficiente para indicar o vetor de estratégias de equilíbrio.

¹² Para uma definição técnica, a referência seminal é Aumann (1976).

¹³ Este termo me foi sugerido por Spyros Vassilakis, em comunicação pessoal.

¹⁴ A este respeito, vale a referência à literatura de seleção de equilíbrios. Ver Kreps (1990, p. 418). Kreps refere-se às dificuldades próprias desse programa de pesquisa, que tem se mostrado bastante inconclusivo em sua tentativa de responder à questão: "how can we reason about what will happen conditional on things that are not meant to happen in the first place?"

¹⁵ Ver, por exemplo, o caso do *Chain Store Paradox*, reportado por Selten (1978).

¹⁶ Aqui vale a referência ao programa de pesquisa de jogos repetidos, onde a cooperação emerge como uma estratégia racional. Entretanto, em face da persistência de equilíbrios não-cooperativos, segundo a avaliação de Binmore (1993, p. 357), o resultado deste programa tem sido o de mostrar

que "cooperation is not necessarily irrational when the Prisoner's Dilemma is repeated an indefinite number of times".

¹⁷ As condições para escolhas consistentes sob incerteza são expostas em Savage (1954).

¹⁸ Nas palavras de Aumann (1987b, p. 2): "indeed, in our treatment, the players do *not* in general know how others are playing. We assume only that it is common knowledge that all players are Bayesian utility maximizers, that they are rational in the sense that each conforms to the Savage theory."

¹⁹ Binmore (1993 e 1994) cunhou o termo *Bayesianismists* para designar aquelas pessoas que usam inadequadamente a regra de Bayes como um mecanismo de aprendizado. Para ele, a regra de Bayes é apenas um artifício para a computação de probabilidades posteriores; em particular, ela calcula a probabilidade condicional $p(E/F)$ como a razão entre a probabilidade conjunta $p(E \text{ e } F)$ e a probabilidade anterior $p(F)$.

²⁰ O problema formal, segundo Morris, "is that assuming a common, logical, prior about endogenous events makes the logical relation (if one existed) between information and priors self-referential." (Morris, 1995, p. 236). Se a relação lógica entre informação e estimativa inicial é pouco clara, a diferença material, por assim dizer, entre elas parece sê-lo menos ainda. Tome-se o seguinte exemplo sugerido por Morris (1995, p. 240): "If you are an 'expert' on U.S. politics, I may well want to alter my beliefs, on learning your beliefs, about who will win the presidential election, even if you have already told me all relevant 'information' in the usual sense. Your 'expert prior' is an information signal for me, so we must interpret it as information."

²¹ Nas palavras de Aumann: "Player *i* must take the ignorance of the other players into account when deciding on his own course of action, and he cannot do this if he does not explicitly include in the model signals other than the one he *knows* he got. The 'outside observer' [...] is thus a surrogate for the ignorance of the system as a whole the

lack of common knowledge [of information in our sense] of the signals received by each player." (Aumann, 1987b, p. 8).

²² Este último suposto, por exemplo, diz ele, confia em demasia "on the communicative capacities of real life actors who would need to recollect, specify, transmit, and correctly interpret all decision premises that are potentially relevant in characteristically ambiguous interaction situations." (Scharpf, 1990, pp. 482-483).

²³ Na visão de Kirman (1992, p. 132): "In reality, individuals operate in very small subsets of the economy and interact with those with whom they have dealings. It may well be that out of this local but interacting activity emerges some sort of self organization which provides regularity at the macroeconomic level."

²⁴ Neste último caso, uma "falha" de conhecimento por parte dos indivíduos não pode ser remediada com a administração de doses maiores de informação, pois serão informações possuídas por uma parte e que apenas acrescentarão maior complexidade ao todo. Sobre a relação partes-todo em sistemas complexos ver Lewin (1993).

²⁵ Na linha de um diálogo com a teoria da ação comunicativa de Habermas, ver Johnson (1993).

BIBLIOGRAFIA

ANDERSON, A.R. e MOORE, O.K. (1962), "Some puzzling aspects of social interaction". *The Review of Metaphysics*, XV, 3, 59.

AUMANN, R. (1976), "Agreeing to disagree". *The Annals of Statistics*, 4, 6: 1.236-1.239.

_____. (1987a), "Game theory". *The new palgrave a dictionary of economics*, Londres, The Macmillan Press Limited.

_____. (1987b), "Correlated equilibrium as an expression of Bayesian rationality". *Econometrica*, 55, 1: 1-18.

BERNHEIM, D. (1984), "Rationalizable strategic behaviour". *Econometrica*, 52: 1.007-1.028.

BICCHIERI, C. (1993), *Rationality and coordination*. Cambridge, Cambridge University Press.

BINMORE, K. (1993), "De-Bayesing game theory", in K. Binmore, A.P. Kirman e P. Tany (eds.), *Frontiers of game theory*, MIT Press.

_____. (1994), *Playing fair: game theory and the social contract*. Cambridge, Mass., The MIT Press.

COURNOT, A. (1838), *Recherches sur les principes mathématiques de la théorie des richesses*. Paris, Hachette.

DUPUY, J.-P. (1989), "Common knowledge, common sense". *Theory and Decision*, 27: 37-62.

EDGEWORTH, F. (1881), *An essay on the application of Mathematics to the moral sciences*. Nova York, August M. Kelley.

GOOD, I.J. (1990), "Subjective probability", in J. Eatwell, M. Milgate e P. Newman (eds.), *The new palgrave utility and probability*, Londres, The Macmillan Press Ltd.

HARDIN, R. (1988), *Morality within the limits of reason*. Chicago, The University of Chicago Press.

HARSANYI, J.C. (1967/1968), "Games with incomplete information played by 'Bayesian' players. Parts I, II and III". *Management Science*, 14: 159-182, 320-334 e 486-502.

_____. (1982), *Papers in game theory*. Dordrecht, D. Reidel Publishing Company [1a. ed. 1973].

HAYEK, F.A. (1949), *Individualism and economic order*. Londres, Routledge & Kegan Paul Ltd.

- HEAP, S.P.H. e VAROUFAKIS, Y. (1995), *Game theory a critical introduction*. Londres/Nova York, Routledge.
- JOHNSON, James. (1993), "Is talk really cheap? Prompting conversation between critical theory and rational choice". *American Political Science Review*, 87, 1, March.
- KEYNES, J.M. (1973), "The general theory of employment, interest, and money", in *The collected writings of John Maynard Keynes*, vol. 7, Londres, Macmillan/Cambridge University Press.
- KIRMAN, A.P. (1992), "Whom or what does the representative individual represent?". *Journal of Economic Perspectives*, 6, 2, Spring: 117-136.
- _____. (1995), "Economies with interacting agents". *Discussion Paper*, Rheinische Friedrich-Wilhelms-Universität Bonn, n. A-500.
- KREPS, David M. (1987), "Nash equilibrium". *The new palgrave a dictionary of economics*, vol. 3, Londres, The Macmillan Press Limited.
- _____. (1990), *A course in microeconomic theory*. Nova York, Harvester Wheatsheaf.
- LEONARD, R.J. (1995), "From parlor games to social science: von Neumann, Morgenstern, and the creation of game theory 1928-1944". *Journal of Economic Literature*, XXXIII, June: 730-761.
- LEWIN, R. (1993), *Complexity life at the edge of chaos*. Londres, Phoenix.
- LEWIS, D. (1969), *Convention: a philosophical study*. Cambridge, Mass., Harvard University Press.
- MAYNARD SMITH, J. (1982), *Evolution and the theory of games*. Cambridge, Cambridge University Press.
- MIROWSKI, P. (1992), "What were von Neumann and Morgenstern trying to accomplish?", in E.R.Weintraub (ed.), *Toward a history of game theory*, Durham/Londres, Duke University Press.
- MORRIS, S. (1995), "The common prior assumption in economic theory". *Economics and Philosophy*, 11: 227-253.
- NASH, J. (1951), "Non-cooperative games". *Annals of Mathematics*, 54: 286-295.
- ORLÉAN, A. (1989), "Mimetic contagion and speculative bubbles". *Theory and Decision*, 27: 63-92.
- PEARCE, D. (1984), "Rationalizable strategic behavior and the problem of perfection". *Econometrica*, 52: 1.029-1.050.
- SAVAGE, L. (1954), *The foundations of Statistics*. Nova York, Wiley.
- SCHARPF, F.W. (1990), "Games real actors could play the problem of mutual predictability". *Rationality and Society*, 2, 4: 471-494.
- SELTEN, R. (1978), "The chain store paradox". *Theory and Decision*, 9: 127-159.
- TAN, T.C.C. e WERLANG, S.R.C. (1992), "On Aumann's notion of common knowledge: an alternative approach". *Revista Brasileira de Economia*, 46, 2: 151-166.
- VON NEUMANN, J. e MORGENSTERN, O. (1944), *The theory of games and economic behavior*. Princeton, Princeton University Press.

* Gostaria de agradecer os comentários dos professores Alan Kirman e Spyros Vassilakis, do European University Institute, e do professor Jaques Kerstenetzky, da Universidade Federal Fluminense, a uma versão anterior deste artigo. Erros, imprecisões e opiniões remanescentes são, naturalmente, de minha inteira responsabilidade.