

A Volta de Ulisses: Anotações sobre a Lógica de Planos e Compromissos

ANA MARIA BIANCHI
ROBERTA MURAMATSU*

Ulysses's Return: Notes on the Logic of Plans and Commitments.

Abstract: This paper criticizes the conventional theory of choice for being grounded on a minimal set of rationality axioms. We claim that this theory does not take due account of the fact that agents are driven by motives other than the pursuit of material self-interest. Our departure point is logic of commitments and planned action, which helps us to identify some puzzles in the conventional theory of choice. As a way out, we discuss the Kantian perspective and the notions of metapreference and metaranking. We then build a model of choice which points to the possibility of a systematic treatment of higher order preferences and incommensurable objectives.

Key-words: Rationality, Commitments, Metapreferences.

JEL Classification: B41, D01.

* Respectivamente Professora do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA/USP). E-mail: amafbian@usp.br; e Doutoranda da Universidade Erasmus de Rotterdam. E-mail: rmuramatsu@uol.com.br. Este artigo originou-se de reflexões feitas pelas autoras no período em que Roberta Muramatsu estava redigindo sua dissertação de mestrado, sob orientação de Ana Maria Bianchi. Foi apresentado no Encontro Nacional da Anpec em 2000 e no encontro da History of Economics Society, realizado em 2001 na Universidade de Wake Forest. Agradecemos os comentários dos professores Mauro Boianovsky, da Universidade de Brasília, e Maurice Lagueux, da Universidade de Montréal, que atuaram como debatedores. Devemos um agradecimento especial ao Prof. Juan Moldau, que poucos meses antes de seu falecimento discutiu extensivamente conosco seu modelo de objetivos múltiplos irredutíveis, tendo feito importantes sugestões para o aperfeiçoamento da quarta sessão deste artigo. Cabem as advertências costumeiras. Submetido setembro 2003; Aceito março 2004.

“Radical simplifications are allowable in science so long as they don’t go against the essence of a given problem.”
(Oskar Morgenstern)

1. INTRODUÇÃO

Muitos economistas contemporâneos estão se dedicando a investigar se a abordagem econômica alcançou progresso explicativo, no sentido de extensão de escopo e articulação causal entre variáveis (Mäki 1994, 2001). Essa linha de investigação é motivada pelo reconhecimento de que, a despeito da relevância da teoria da escolha tradicional, pressupostos psicológicos minimalistas limitam o valor explicativo que a mesma tem sido capaz de exibir ao longo de sua história. De caráter essencialmente normativo e tributária de Savage (1954), essa teoria abdica de um exame aprofundado das forças que impulsionam o comportamento.¹ Em alguma medida, isso é assim porque a teoria da escolha racional funda-se em um conjunto de axiomas que supostamente garantem escolhas internamente consistentes. Ancora-se assim na postura metodológica explícita na passagem abaixo:

Is it not patently unrealistic to suppose that individuals consult a wiggly utility curve... that they can compute the expected utility... this objection is not strictly relevant. The validity of this assertion does not depend on whether the individuals know the precise odds... but solely on whether it yields sufficiently accurate predictions about the class of decisions with which the hypothesis deal... Whatever the psychological mechanism whereby individuals make choices, these choices appear to display some consistency... described by our utility hypothesis (Friedman e Savage 1948: 297-8)

De acordo com os autores, a escolha da pessoa revela sua preferência, e qualquer elemento que possa desencadear ou manter certos cursos de ação é redutível ao denominador comum da utilidade. Torna-se assim possível prescindir de uma descrição acurada dos mecanismos e processos geradores do comportamento. O importante seria a consistência de escolha que se revela na hipótese de utilidade, manifesta nas alternativas efetivamente selecionadas pelos agentes, independentemente dos mecanismos psicológicos que as desencadeiam.

Ora, essa visão de racionalidade tende a reduzi-la a um cálculo hedônico. Hollis e Sugden (1993: 4) enxergam nesse tipo de abordagem a influência do utilitarismo de Bentham, segundo o qual o princípio de maximização de utilidade

¹ Segundo Hollis e Sugden (1993: 28), os axiomas de Savage afastam a possibilidade de explicações em que “a fonte ou caráter das satisfações afetam a atitude do agente em relação às conseqüências, ou em que princípios são introduzidos nas descrições de fatos”.

contém todas as condições necessárias para a racionalidade, que equivale (e, portanto, se reduz) à maximização da utilidade esperada.²

Argumentaremos neste artigo que o conteúdo psicológico minimalista característico da visão ortodoxa de ação racional limita o valor explicativo dessa teoria. Nosso ponto de partida é o reconhecimento de uma categoria de comportamentos que os economistas têm dificuldade de explicar à luz da teoria da utilidade esperada, que são as escolhas rotuladas de “compromissos” e “planos”. Elas constituem um verdadeiro quebra-cabeça para uma teoria que se funda na noção de que a escolha efetiva revela o esquema de preferências do agente. Por quê? Porque compromissos e planos podem requerer escolhas contrapreferenciais, que trazem para primeiro plano a questão de encontrar uma maneira de acomodar as lacunas potenciais entre o que é preferido pelo agente e o que ele efetivamente escolhe. Deslindar esse enigma é essencial para entender os motivos que levam pessoas racionais a manter promessas e agir em conformidade com planos mesmo quando contrariam suas preferências auto-interessadas, bem como a acomodar instâncias em que se verificam confiança mútua e coordenação social.

O artigo está dividido em quatro partes. Após esta introdução, dedicamos a sessão 2 à reconstituição da perspectiva de David Gauthier (1996) sobre a racionalidade de planos e compromissos, que ilumina algumas implicações problemáticas da concepção ortodoxa. Gauthier defende a necessidade de introduzir no arcabouço da teoria da escolha razões para agir que não se restringem às preferências auto-interessadas do agente, na busca de um refinamento de sua base motivacional. Contudo, argumentamos que sua abordagem é insuficiente para apreender a complexidade das decisões tomadas pelo agente, bem como o papel dos planos e compromissos como mecanismos de coordenação social. A sessão 3 é dedicada ao conceito teórico de metapreferências, concebido como tentativa de acomodar a *rationale* do compromisso moral. Foi essa a perspectiva adotada por economistas como Sen (1982) e Hirschman ([1985]1992) para explicar por que os agentes decidem contrariar suas preferências auto-interessadas. Na sessão 4 discutimos um modelo de escolha com objetivos múltiplos irreduzíveis, construído a partir de Moldau (1988, 1993), que representa um esforço na direção de um tratamento sistemático da estrutura hierárquica de motivações introduzida na sessão anterior. Os limites dessa formulação serão também criticamente examinados. A última sessão do artigo alinhava alguns comentários finais.

2. PLANOS E COMPROMISSOS

A incipiente psicologia da escolha subjacente à teoria da escolha racional não nos permite entender por que as pessoas se comportam do jeito como costumam.

² Como veremos mais adiante, esta não é a visão de Hume, aqui adotada. Hume considera a razão como um mero instrumento para atingir os objetivos individuais, determinados por disposições internas ou estados psicológicos (desejos e crenças).

mam fazê-lo. Isso motiva o pesquisador a questionar a tarefa de descrever a realidade econômica por meio da teoria da utilidade esperada, que se tem mostrado impotente para explicar fenômenos de cumprimento de promessas, confiança mútua e coordenação social. (Frank 1988 e 1992; Thaler 1992, entre outros)

Nesse contexto, o ensaio de Gauthier (1996) sobre planos e compromissos nos ajuda a identificar dificuldades na teoria da escolha. Gauthier toma como ponto de partida de uma estratégia revisionista a introdução no corpo da análise de motivos para agir que não se restringem àqueles baseados em preferências. Há situações em que o agente atua em conformidade com um plano, mesmo que, ao longo das etapas de implementação do mesmo, os resultados esperados do desvio tornem-se mais atraentes para ele do que a persistência nas escolhas definidas anteriormente. O autor justifica seu pleito por uma perspectiva menos parcimoniosa a partir da possibilidade de aumentar o poder explicativo da teoria, fugindo a uma camisa de força conceitual que pode distorcer o entendimento:

To decide whether there is conceptual room for choice to override preference ... is really to decide whether the best explanatory and justificatory accounts of our choice behavior require such room whether we understate an agent's behavior better if we interpret her as choosing on the basis of plan-based as well as directly preference-based considerations. To assume from the outset that she can have only one kind of reason for choice is not to acknowledge an independently given conceptual requirement on our understanding of choice behavior but to impose a conceptual strait-jacket that may prove to distort our understanding. (Gauthier 1996: 238-9)

2.1. Sobre a racionalidade de planos e compromissos

Tendo tomado como premissa central o princípio da maximização da utilidade esperada, como a abordagem tradicional acomoda o fenômeno do compromisso e da ação planejada? Para ilustrar, tomemos um exemplo muito conhecido, que é a história de Ulisses e das sereias. Com a ajuda dessa metáfora, Gauthier corretamente enfatiza os pontos de identidade entre planos e compromissos (ou, mais precisamente, pré-compromissos),³ no sentido de que ambos envolvem escolhas predeterminadas. A metáfora de Ulisses nos ajudará a argumentar que a teoria ortodoxa não pode de fato explicar as forças motrizes do comportamento humano. Ela é emblemática de um tipo de situação em que ocorre conflito entre o interesse imediato do agente e algum tipo de interesse de longo prazo, mais genuíno, durável ou legítimo, como argumenta Schelling (1992: 169).

Que razões motivam o herói grego em sua volta da guerra de Tróia? Chegar

³ Elster (2000: 4), entre outros, reserva a denominação de “pré-compromisso” para instâncias de auto-constrangimento tais como a simbolizada pelo mito de Ulisses e as sereias, no sentido de restrições que o agente impõe a si mesmo em troca de um benefício esperado futuro. Jogar fora a chave do automóvel na entrada da festa (para evitar voltar bêbado para casa) é outro exemplo de um mecanismo de pré-compromisso, assim como qualquer ato de evitar tentação. Eles ilustram estratégias de auto-constrangimento que contrabalançam o comportamento induzido por fraqueza de vontade (acrasia).

são e salvo à Grécia, para reencontrar sua Penélope e trazer glória a seu povo, ou desfrutar do canto das sereias? Ulisses partiu para sua empreitada com o plano de voltar diretamente à Grécia, independentemente das tentações que pudessem motivá-lo a rever seu plano. Por outro lado, ele estava bem consciente de que suas preferências podiam mudar assim que ouvisse as sereias. Nessa etapa de sua trajetória ele seria impulsionado pelo desejo de alcançá-las, o que implicaria uma revisão de seu plano inicial. Em termos econômicos, podemos dizer que Ulisses tomou a decisão de adotar um plano contando com um retorno esperado pelo menos tão bom quanto um curso de ação alternativo. Analogamente, Ulisses implementaria seu plano se o retorno esperado associado a sua execução fosse pelo menos tão alto quanto o retorno esperado da ação de revê-lo, sucumbindo à tentação das sereias. Para dar conta dessas possibilidades, Gauthier sugere a consideração de duas versões do tomador de decisões: o *Ulisses esclarecido* (sofisticado) e o *Ulisses míope*.

Em sua primeira versão, Ulisses só adota um plano se ele for viável, ou seja, se o retorno esperado no final de sua viagem não induzi-lo em nenhum momento de seu horizonte de escolha a rever esse plano. Isso pressupõe que o agente age como se estivesse equipado com um grande volume de habilidade cognitiva. Em poucas palavras, Ulisses só adota o plano se ele puder ser mantido ao longo do tempo, descartando outros cursos de ação que possam conflitar com o mesmo. Assume-se que Ulisses tem antevisão perfeita e pode atribuir valores de utilidade para cada curso de ação que percebe. Curiosamente, porém, esse Ulisses sofisticado do ponto de vista cognitivo deve pagar um alto preço por sua vontade de atingir escolhas intertemporalmente consistentes: para assegurar sua fidelidade ao plano, independentemente das alternativas com as quais venha a defrontar-se, Ulisses precisa ser amarrado ao mastro do navio. Esse é um mecanismo de pré-compromisso.⁴

Para dar sentido ao dilema acima, Gauthier (1996: 223) distingue dois conjuntos de razões: (a) as que dispõem o agente a adotar um plano; e (b) aquelas que o dispõem a levar adiante seu plano, ou seja, a agir em conformidade com o mesmo. Nesse caso, o agente esclarecido antecipa o conflito potencial entre esses dois tipos de razão. Isso leva Ulisses a preferir ser amarrado — na realidade, como argumentam Ainslie e Haslam (1992a e b), ele prefere não ter a chance de escolher no futuro.

Não muitos tomadores de decisão de carne e osso agirão como o Ulisses esclarecido. Como explicar situações em que os agentes não conseguem levar adiante seu plano? Uma alternativa é introduzir o Ulisses míope, na terminologia de Gauthier. Tomadores de decisão míopes, embora capazes de antecipar o futuro e formular projetos de longo prazo, são incapazes de concretizá-los: no momento seguinte àquele em que planejam um determinado curso de ação, a utilidade es-

⁴ Como relata Elster (2000: 80), mecanismos de pré-compromisso podem ter efeitos sinalizadores indesejáveis e, assim, colocam um problema especial. No exemplo emblemático, o fato de Ulisses pedir para ser amarrado ao mastro pode evidenciar aos outros sua fraqueza de vontade.

perada de levar adiante seu plano torna-se igual ou inferior à utilidade esperada de desviar-se dele. Embasbacado pela sedução das sereias, Ulisses tem sua visão de longo prazo obscurecida.

Em nossa opinião, as duas versões de Ulisses criadas por Gauthier são insuficientes para retratar a complexidade do processo de tomada de decisão. Parece-nos mais razoável conceber uma gama de tipos de tomadores de decisão, cujas habilidades cognitivas não se coadunam com as versões extremas do Ulisses míope e do esclarecido. Um desses tipos alternativos seria o agente *resoluto*, personagem ao qual Gauthier não parece dar a devida importância. O autor argumenta que a escolha do Ulisses resoluto é moldada por sua expectativa de que a conformidade ao plano será uma razão suficiente para determinar sua volta à Grécia, mesmo que suas preferências posteriores mudem. (Gauthier 1996: 228) Ao contrário do Ulisses sofisticado, que se enxerga no ponto futuro a mercê de suas preferências (e por isso mesmo pede para ser amarrado ao mastro), o agente resoluto é motivado por razões para agir que estão além de suas preferências. O plano em si mesmo pode constituir uma razão para agir, ditada por imperativos categóricos, como veremos na próxima sessão. Assim, o estudo da lógica de planos e compromissos ilumina o fato de que as preferências não são as únicas razões para agir.

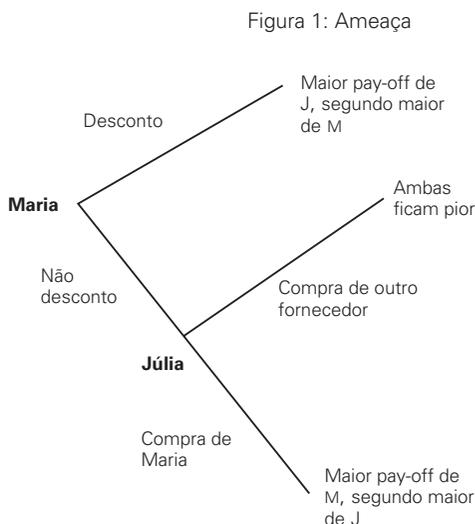
Isso posto, passamos a comentar a ocorrência de situações em que a ação planejada não envolve mudanças de preferências, à luz de um exemplo também inspirado em Gauthier.

Maria fornece todas as bugangas consumidas por Júlia. Esta última está consciente de que sua aquisição é parte importante da receita de Maria. Assim, Júlia tem um incentivo para pedir a Maria um desconto de 10%, ameaçando mudar de fornecedor caso seu pedido não seja atendido. À primeira vista, Júlia está certa em fazer a ameaça, porque há grande chance de Maria aceitá-la. Suponhamos, porém, que Júlia sabe que fazer negócio com outro fornecedor acarretará um acréscimo de 5% no preço de sua compra. Quando essa informação adicional é acrescentada, fica difícil entender porque Júlia faz uma ameaça que não tem interesse em levar adiante. A idéia é que ela espera que Maria aceite conceder-lhe o desconto. Mas, o que acontecerá se Maria recusar?

Este segundo exemplo pode ser definido como uma ameaça, que tem uma estrutura semelhante ao chamado “jogo de promessa” desenvolvido por Sugden (1991). Consideremos por um instante que Júlia é um tomador de decisão míope. Nesse caso, ela será movida por suas preferências correntes e não considerará o fato de que agir em conformidade com seu plano inicial, caso Maria venha a rejeitar sua pedido de desconto, lhe será prejudicial. Se, por outro lado, Júlia for uma tomadora de decisão sofisticada, inteligente o bastante para antecipar que seu plano é inviável (fazer negócio com outro fornecedor não paga a pena), sua ameaça é um mero blefe ou “conversa barata”. (Gauthier 1996: 233) Suas razões para adotar o plano, induzindo Maria a responder de tal forma a maximizar seu retorno esperado, não coincidem com suas razões para levá-lo adiante. Júlia apenas implementaria a ameaça planejada se estivesse ligada a um compro-

misso, ou seja, se algum imperativo interno a convencesse a cumprir a ameaça, a despeito do prejuízo que viria a sofrer.

O diagrama na Figura 1 resume a interação entre Maria (M) e Júlia (J).



Fonte: Adaptado a partir de Gauthier 1996: 234

A teoria da escolha tradicional é incapaz de explicar por que Júlia implementa um plano que contraria suas preferências, ou por que Maria eventualmente concede o desconto, mesmo sabendo que Júlia estaria pior comprando de outros. Isso é também um enigma para a teoria dos jogos tradicional, que assume que os indivíduos são maximizadores da utilidade esperada e exibem a chamada “racionalidade de conhecimento comum” (em inglês, *common knowledge rationality*, CKR), ou seja, assumem que toda a informação relevante é dada e transparente para todos os jogadores. Não é melhor para Maria manter o preço inalterado, uma vez que sabe que a ameaça de Júlia não tem credibilidade? Sua resposta seria diferente caso ela antecipasse que Júlia poderá concretizar a ameaça de mudar de fornecedor, contrariando suas preferências auto-interessadas? A teoria da escolha racional não nos ajuda a responder essas questões.

Um raciocínio análogo ajuda-nos a analisar o processo de tomada de decisões em um jogo de promessa. Suponhamos que Bernardo (B) forneça algum serviço a Alice (A), baseado na promessa de que ela lhe retribuirá o serviço no futuro. Esta situação é representada de forma extensiva no diagrama da Figura 2.

Para entender o que está em jogo, podemos usar o método da indução retroativa. Primeiro, olhamos para o nó final e comparamos os *pay-offs* de Alice. Percebemos então que Alice prefere não fornecer o serviço, pois $U^e(4) > U^e(3)$, onde U^e denota utilidade esperada. Deslocamo-nos então para o segundo nó: Bernardo parece indiferente entre as opções de honrar sua promessa ou abandoná-la, pois seu *pay-off* é nulo nos dois casos. Mas, não dissemos antes que Ber-

Figura 2: jogo de promessa

A. faz uma promessa	B. fornece o serviço	A. honra sua promessa (3,3)
Não faz	Não fornece	Não honra
(0,0)	(0,0)	(4, -2)

Fonte: Hollis e Sugden 1993

Quando cumpriria seu plano baseado na promessa de Alice? Isso é intrigante, especialmente quando levamos em conta o pressuposto de racionalidade de conhecimento comum (CKR). Nesse caso, Alice será racional se mantiver sua promessa e fornecer o serviço?

A resposta às questões acima depende de algo que está por trás das escolhas individuais. Se a escolha de Alice é motivada apenas por suas preferências, ela não honrará sua promessa. Bernardo, por sua vez, sabe que Alice levará em conta o fato de que honrar sua promessa não será de seu interesse. Portanto, ele também não terá incentivo para prestar o serviço. Como resultado da ausência de confiança mútua, o jogo encerrar-se-á no primeiro nóculo da árvore retratada na Figura 2.

Ora, sabemos que as duas partes se sairiam melhor se confiassem uma na outra. O melhor resultado seria decorrente de uma confiança mútua genuína, possibilidade de que a teoria da escolha convencional é incapaz de dar conta.⁵ Despreza assim uma vasta literatura baseada em pesquisas empíricas que mostram que o comportamento cooperativo, típico de situações em que existe confiança entre os participantes, é parte da realidade econômica. Como argumentam Hollis e Sugden (1993: 13), alguns fatos corriqueiros da vida social — honrar promessas, cumprir ameaças, fazer favor a outrem — são “misteriosos” do ponto de vista da teoria dos jogos.⁶

Para decifrar os enigmas observados na teoria tributária do modelo de Savage, poderíamos invocar a existência de preferências morais, que seriam preferências como outras quaisquer. Nessa concepção, a escolha do indivíduo que se ex-

⁵ Hollis e Sugden (1993: 15) assinalam que a interpretação convencional desse tipo de situação não gera controvérsia entre teóricos de jogos, embora ameace minar a possibilidade de vida social entre agentes racionais, ao destruir a idéia de reciprocidade. Em suas próprias palavras, essa análise “... threatens to undermine social life among rational agents by destroying reciprocity and hence to undermine the claim of game theory to idealize the basis of everyday social intercourse also”.

⁶ Ver Bianchi (1998) e Muramatsu (1998) para alguns exemplos.

põe à perda de sua liberdade e de sua reputação ao denunciar uma fraude é perfeitamente análoga à escolha feita por esse mesmo indivíduo na véspera, ao comprar frutas no mercado da esquina; as duas escolhas revelam preferências que podem ser encaixadas em uma mesma função de utilidade, é tudo.

Mas invocar preferências morais pode ser um argumento *ad hoc*, que se prestaria a acomodar situações em que a teoria convencional é falseada, como nos adverte Sugden (1998: 75). Trata-se de um expediente, que não resolve o problema de acomodar no arcabouço teórico tradicional anomalias como a existência de compromissos morais, que constituem violações do padrão de comportamento auto-interessado. Concordamos com Sugden quanto à necessidade de investigar mais a fundo os determinantes das preferências. Assim, inspiradas pela noção de racionalidade instrumental de Hume, sugerimos que preferências morais decorrem de sentimentos morais, que constituem os alicerces das preferências não auto-centradas e do comportamento cooperativo ou prossocial. (v. Muramatsu no prelo; Frank 1988 e 1992; Katz 2000; Vromen 2000)

Voltaremos a esse ponto mais adiante. Por ora basta reiterar que, quando abrimos espaço no corpo da análise para razões para agir que vão além das preferências auto-interessadas de curto prazo (satisfação do auto-interesse material), o fornecimento de um serviço prometido pode ser racional. O exemplo ajuda a argumentar que é importante entender melhor compromissos e planos porque eles facilitam a coordenação social, assim beneficiando todos os envolvidos. Passaremos agora a discutir até que ponto a possibilidade de confiança mútua e coordenação interpessoal pode ser convenientemente explicada em termos de planos e pré-compromissos.

2.2. Planos, Compromissos, Moralidade e Coordenação Interpessoal

Nossa atenção se volta aqui para a relação entre planos, compromissos e moralidade. Quando o tomador de decisão resolutivo é incluído no corpo da análise, há espaço para explicar o comportamento em termos de razões para agir sustentadas por planos, em que a escolha decorre do desejo de agir de acordo com um curso de ação predeterminado. Em outros termos, no momento em que a pessoa reconsidera um curso de ação previamente escolhido, ela pode concluir que sua disposição para agir de acordo com o plano lhe dá razões suficientes para cumpri-lo. O que mantém a ação planejada? O texto de Gauthier não nos dá uma resposta satisfatória e direta, mesmo porque o autor não se propõe a tal. Em sua perspectiva, o que realmente importa para explicar compromissos genuínos é o fato de que o agente extrai benefícios do fato de ser um planejador resolutivo, bem como do fato de esperar a mesma disposição por parte daqueles com os quais interage. Mas isso explica satisfatoriamente as instâncias de coordenação interpessoal? Parece que não, como argumentaremos em seguida.

Sem dúvida, compromissos e planos são forças importantes atrás da coordenação social, como admite Gauthier: subjacente aos mesmos está a própria necessidade de coordenação social. É verdade que existem formas de coordenação

cuja emergência não depende crucialmente de forças morais, entre elas convenções como dirigir no lado direito (ou esquerdo) da rua.⁷ Mas o fato de o agente resoluto extrair benefícios de uma convenção, esperando a mesma disposição dos outros, não é suficiente para explicar sua emergência e persistência no tempo. A rigor, Gauthier parece esticar demais a idéia da ação planejada como elemento explicativo. Ele repele a hipótese de que a necessária “provisão de segurança” para uma coordenação interpessoal eficiente seja explicável a partir de um pano de fundo moral — em nosso exemplo, a segurança de Bernardo de que, quando chegar o tempo, Alice honrará sua promessa. Na verdade, alega o autor, não é necessário apelar para uma moralidade pré-existente (Gauthier 1996: 242).⁸ Isso soa exagerado. Embora concordemos que nem todo plano pressupõe um aparato ético específico, normas e sentimentos morais tendem a desempenhar um papel central na explicação das expectativas individuais.

Para uma explicação mais abrangente, podemos apelar aqui para o conceito de “expectativas normativas”, exposto por Sugden (1998). Estas equivalem a uma categoria de sentimentos morais que consistem em disposições internas para agir baseadas em normas que vão além do auto-interesse individual. A idéia básica dessa concepção é expressa pelo autor da seguinte forma: dizer que o indivíduo *i* tem uma expectativa normativa de que o indivíduo *j* fará *x* equivale a dizer que não apenas *i* espera que *j* faça de fato *x*, como também que *j* se sente movido por um tipo de pressão social ou obrigação moral a comportar-se como *i* espera, exatamente porque ele assim o espera. Sugden considera que essas expectativas normativas são responsáveis, juntamente com o auto-interesse, por criar e manter convenções e normas. Graças a fatores emocionais como ressentimento, culpa ou senso de dever, as pessoas dispõem-se a manter a palavra dada, cumprir planos, obedecer convenções e assim por diante. Frank (1992) também descreve como sentimentos morais atuam como mecanismos de compromisso, ajudando a pessoa a resolver problemas de inconsistência temporal.⁹

Com isso em mente, passaremos agora a uma abordagem alternativa à teoria da escolha racional. Essa abordagem é capaz de elucidar alguns enigmas dessa teoria, fornecendo *insights* sobre razões para agir que vão além das preferências auto-interessadas dos agentes.

⁷ V., por exemplo, Young (1993).

⁸ Curiosamente, Gauthier (1996: 242-3) admite que os agentes resolutos podem invocar razões éticas e cultivar sentimentos morais com o objetivo de executar seus planos.

⁹ Frank (1992) nos convida a considerar o caso de uma pessoa que se sente mal quando comete uma fraude. Esses sentimentos podem levá-la a comportar-se honestamente mesmo quando sabe que poderia cometer fraude sem ser descoberta. Ora, se as outras pessoas com as quais convive percebem que ela tem sentimentos desse tipo, elas a procurarão como parceira em empreendimentos que requerem confiança. Por outro lado, os mesmos sentimentos morais poderão ajudar a pessoa a agir com prudência em casos em que os incentivos materiais de longo prazo favorecem a conduta honesta.

3. A ABORDAGEM DE METAPREFERÊNCIAS COMO TENTATIVA DE ACOMODAR A *RATIONALE* DO COMPROMISSO MORAL

Na sessão anterior, mostramos que Gauthier endossa a visão de que a característica distintiva dos seres humanos é sua capacidade de fazer planos e assumir compromissos, a seu turno o fator mais importante para explicar a coordenação social. Somos céticas em relação a essa posição, embora aceitemos a idéia de que preferências auto-interessadas não são o único determinante do comportamento de escolha.

Na conclusão de seu artigo, Gauthier (1996: 242) faz uma caricatura do debate sobre o que vem primeiro: um plano ou uma estrutura moral? Talvez ele o faça porque tem em mente uma noção particular de compromisso, ou seja, mecanismos de pré-compromisso. (Ainslie e Haslam 1992b; Shapiro 2002).¹⁰ Tais mecanismos são úteis para explicar várias instâncias de escolha intertemporal, referidas a problemas de impulsividade e auto-controle. Contudo, uma noção mais abrangente de compromisso pode ser encontrada no artigo seminal de Amartya Sen sobre “Rational Fools”. Vejamos como Sen coloca a questão:

One way of defining commitment is in terms of a person choosing an act that he believes will yield a lower level of personal welfare to him than an alternative ... This broader sense may have particular relevance when one acts on the basis of a concern for duty, which, if violated, could cause remorse, but the action is really chosen out of the sense of duty rather than just to avoid the ill-fare resulting from the remorse that would occur if one were to act otherwise. ([1977] 1982: 92)

Ao apoiar-se num conceito de compromisso que coloca uma cunha entre a escolha e o bem estar pessoais, Sen desvia-se de boa parte da teoria da utilidade tradicional. De fato, o autor elabora sua noção de compromisso a partir da discussão de Frankfurt (1971) sobre autonomia de escolha e preferências de ordem mais elevada. O artigo pioneiro de Frankfurt oferece *insights* importantes sobre a possibilidade de acomodar preferências e desejos morais no arcabouço da escolha racional.

Na verdade, desde os tempos de Platão, os filósofos vêm discutindo como se organizam os desejos, preferências e crenças do ser humano, em termos dos diferentes esquemas de deliberação envolvidos nesse processo. Sob influência kantiana, Frankfurt (1971) sustenta que o aspecto característico dos seres humanos é a autonomia da vontade. Ao contrário dos outros animais, os seres humanos te-

¹⁰ Ainslie e Haslam (1992b: 179) creditam a Strotz o mérito de ter sido o primeiro economista moderno a postular explicitamente uma mudança de preferência esperada. Ele introduz a noção de pré-compromisso para designar mecanismos pelos quais a pessoa que prevê uma inconsistência possível entre o plano que traçou e seu comportamento no ponto futuro elimina a possibilidade de concretizar o desvio em relação a seu plano.

riam a capacidade de formar preferências de segunda ordem. Eles não apenas desejam um curso particular de ação mais do que outro, como são capazes de refletir acerca da propriedade de seus desejos e crenças:

They are capable of wanting to be different, in their preferences and purposes, from what they are ... No animal other than man, however, appears to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires. (Frankfurt 1971:7)

A perspectiva de Frankfurt inspira-se na concepção kantiana de imperativos categóricos. Esta dissocia a escolha moral da noção de racionalidade instrumental: a razão não seria um mero instrumento de coordenação das paixões e crenças humanas, mas, ao contrário, um fim em si mesmo, o verdadeiro motivo da ação humana. Na concepção de Kant, uma pessoa autônoma é governada por leis que impõe a si mesma, na forma de imperativos categóricos: “Faça x independentemente de seus desejos ou preferências pessoais”. Além disso, os imperativos categóricos são universalizáveis, no sentido de que se impõem a todos os atores humanos.¹¹

A idéia de racionalidade kantiana pode ser usada para explicar o que leva Ulisses de volta à Grécia. Nosso herói decide levar adiante seu compromisso com base em um imperativo categórico que pode ser assim formulado: “Honre sua obrigação de obter glória para o povo grego independentemente de seu desejo pessoal de navegar em direção às sereias”.

É importante notar que o modelo de racionalidade adotado por Kant distingue-se da perspectiva de Hume, em termos dos determinantes da ação racional. Para o primeiro, a escolha moral racional deve ser livre de qualquer estado psicológico e, como tal, independente de desejos e crenças humanas. Esse aspecto da perspectiva kantiana nos parece problemático, uma vez que é pouco aplicável aos modelos de escolha sobre os quais os economistas têm se debruçado para explicar o comportamento com base em preferências. Por esse motivo, endossamos a alternativa proposta por Hollis e Sugden (1993) com o intuito de refinar a teoria da ação moralmente orientada. Segundo os autores, a racionalidade da escolha moral deve acomodar diferentes esquemas de deliberação e ordenação de preferências e desejos, dispostos em uma estrutura que comporta diferentes níveis hierárquicos. Mais precisamente:

The lower tier can comprise a standard preference ordering, ranging over possible outcomes. The upper tier is usually taken to comprise a second order preference ordering whose domain is the first-order pref-

¹¹ Para uma abordagem mais completa da perspectiva kantiana, v. Kant (1992) e Hollis e Sugden (1993).

erences ... The idea is to let agents act sometimes *not on the preferences they have but on those which they prefer to have.*

(Hollis e Sugden 1993: 30)

A perspectiva adotada no trecho citado é compatível com as críticas de Frankfurt à teoria da escolha tradicional. Para o autor, o tomador de decisão retratado por essa teoria seria um libertino (*wanton*, no original inglês), no sentido de se resumir a um simples feixe de preferências dadas. Seu argumento é retomado por Sen ([1977]1982: 99), que lamenta a excessiva preocupação da teoria econômica com o quase “idiota social” (*social moron*) simbolizado pelo homem econômico puro, “alojado na glória de sua ordenação única de preferências”. O único traço verdadeiramente humano desse personagem fictício seria a consistência de escolha — no mais, seria um louco racional (*rational fool*).

Ao criticarem a psicologia da escolha empobrecida que está por trás da teoria tradicional, tanto Sen quanto Hirschman ([1985] 1992) evocam o conceito de meta-ordenações (*metarankings*), também designado por preferências de segunda ordem ou metapreferências. Para Hirschman, os economistas valorizam excessivamente a questão da parcimônia na construção de suas teorias. Como toda a virtude, diz ele (1992: 143), a parcimônia pode ser exagerada. De forma semelhante, Sen (1982: 99) argumenta que a teoria tradicional “tem muito pouca estrutura” e mantém-se irrazoavelmente silenciosa sobre a natureza das preferências individuais e outras razões para a ação. Parece-lhe absurdo que uma única ordenação de alternativas preferidas, retratada em uma função de utilidade, possa representar adequadamente a base de escolha do agente, seu bem estar e suas preferências auto-interessadas. Em função disso, o autor inclui o conceito de meta-ordenação (*metaranking*) no corpo da teoria, para explicar a possibilidade de compromisso genuíno. Sen destaca o valor instrumental dessa noção, que pode ser empregada em diferentes contextos: para descrever, por exemplo, um conjunto de prioridades políticas, um sistema de interesses de classe, conflitos associados à dependência química e assim por diante. Nesses vários contextos, a noção de meta-ordenações ajuda a reconstituir a complexidade da escolha individual, que abriga instâncias em que as razões que levam a pessoa a agir não podem ser reduzidas ao denominador comum da utilidade.¹²

É importante salientar que a abordagem de metapreferência enfrenta dificuldades associadas à tratabilidade formal. O próprio Sen ([1974] 1982: 80-83) tenta lidar com elas raciocinando em termos de relações binárias, mas sua tentativa não é inteiramente bem sucedida. Isso nos leva a sugerir um modelo alternativo, que acomoda a possibilidade de uma escolha determinada por meta-ordenações de preferências. O que temos em mente é o modelo lexicográfico de Moldau (1988, 1993), a seu turno um refinamento da crítica de Georgescu-Roegen à teoria da escolha tradicional. Muramatsu (1998) fez uma adaptação de tal tratamento analítico com o intuito de oferecer uma explicação mais rigorosa sobre a

¹² Sobre esquemas de utilidade múltiplos, v. também Brennan (1989) e Lutz (1993).

escolha moral determinada por objetivos múltiplos irreduzíveis, como veremos na próxima sessão.

4. UM TRATAMENTO SISTEMÁTICO DAS META-ORDENAÇÕES DE PREFERÊNCIAS

Vimos até aqui pavimentando o caminho para o argumento de que uma teoria capaz de descrever melhor os planos e compromissos racionais deve considerar razões para agir diferentes do desejo de satisfazer as preferências auto-orientadas do agente. Isso nos leva a aceitar o desafio de “complicar” a teoria da ação tradicional, nela incluindo o conceito de meta-ordenação. Argumentamos que Sen pouco diz sobre como se dá a escolha efetivamente determinada por meta-preferências. Como saída para esse impasse, propomos aqui um modelo de escolha com critérios múltiplos irreduzíveis, instrumento potencialmente útil para superar a dificuldade da tratabilidade formal e o risco de regressão infinita que ameaça a abordagem de metapreferências.

Para dar conta do recado, assumimos que o comportamento é determinado por uma hierarquia de necessidades, vontades, desejos e crenças. Tais categorias psicológicas são ordenadas em termos de seu grau relativo de importância. Elas constituem as razões para a ação ou os critérios de escolha. Outra importante pressuposição é que a escala de prioridades é variável e as razões para agir não podem ser reduzidas ao denominador comum da utilidade.¹³ Se assim é, a ordenação dos critérios pode mudar em função da situação enfrentada (ou percebida) pela pessoa. Voltemos à história de Ulisses. Com base no modelo a ser construído, podemos dizer que o desejo do herói grego de levar adiante suas intenções prévias — navegar diretamente para a Grécia — fornece um critério de escolha que é determinante na explicação de seu comportamento.

Antes de mais nada, é importante dizer que a noção de critério, tal como explicitada por Moldau (1993: 355) aplica-se a dois elementos: (a) um padrão de comparação entre alternativas e (b) a importância relativa do próprio critério. Os múltiplos critérios são tratados pelo princípio lexicográfico, ou seja, existe uma ordenação variável de critérios de escolha, em que esta é determinada pelo critério prioritário (“de primeira ordem”, na terminologia de Moldau). Essa pareceu-nos ser uma maneira razoável de dar um tratamento sistemático e explícito

¹³ Isso implica que não existe taxa marginal de substituição entre as diferentes razões para a ação. Aqui o axioma da continuidade é relaxado, com o objetivo de permitir a existência de critérios ou objetivos múltiplos irreduzíveis. Mostraremos que nossa versão de escolha lexicográfica não exclui totalmente a possibilidade de um certo *trade-off* entre alternativas no conjunto de escolha. A novidade é que as alternativas são ordenadas em termos de vários critérios específicos, os quais, a seu turno, são também ordenados em termos de sua importância relativa em cada ponto do conjunto de oportunidades do agente. Nesse caso, o axioma de continuidade só se sustenta para ordenações feitas em termos de critérios específicos e ordenações de critérios específicas.

a forças motivacionais tão diversas como preferências auto-interessadas, planos, compromissos, preocupações morais e assim por diante.

A estrutura desse modelo comportamental funda-se na idéia intuitiva de que quando existem duas combinações que geram o mesmo nível de satisfação em termos do critério prioritário, a tomada de decisão leva em conta o segundo critério mais importante da hierarquia.¹⁴ Este é o princípio da escolha lexicográfica, em que o critério de segunda ordem só se torna decisivo quando há empate entre as alternativas definidas pelo critério de primeira ordem.

Passemos agora a descrever brevemente os principais argumentos desse modelo de escolha, que assume que todas as pessoas têm competência cognitiva para fazer julgamentos acerca de importância relativa de suas razões para a ação.¹⁵ Consideremos que o agente enfrenta uma situação-problema hipotética cuja decisão é contingente em relação a dois critérios, i e j (por exemplo, auto-interesse material e senso de comunidade). Nesse caso, uma das seguintes possibilidades é observada:

- i é mais importante do que j ;
- i é menos importante do que j ; e
- i e j são igualmente importantes

Se assim é, para cada combinação de bens $x \in X$, existe uma ordenação de critérios $a, b, \dots, i, j, \dots, t$ tal que os critérios disponíveis são irreduzíveis uns aos outros. Vale lembrar que a ordenação de critérios pode variar. Existem, portanto, situações em que uma determinada combinação de bens x^m toma como prioritário o critério i , enquanto j é critério proeminente para x^n .

É necessário enfatizar que o tratamento sistemático desse modelo requer que definamos uma relação binária “pelo menos tão importante quanto” no espaço de produto (*space product*) $J \times X$, onde J é o conjunto de critérios e X o conjunto de escolhas de todos os pares (j, x) . Baseada nessa relação, uma relação de não preferência binária em X em termos do critério $j \in J$ denotado por Q_j pode ser formulada da seguinte forma:

$x^1 Q_j x^2$, se e somente se (j, x^1) é “pelo menos tão importante quanto” (j, x^2) , para todo $x \in X$ e $j \in J$.

Uma relação de preferência (P_j) e uma relação de indiferença (I_j) podem também ser definidas de acordo com $j \in J$. Neste caso, $x^1 P_j x^2$, se e somente se o critério j em x^2 é mais importante do que ele é em x^1 . Quando $x^1 I_j x^2$, o critério j é “igualmente importante” em x^1 e x^2 , ou seja, nos dois pontos do conjunto de escolha.

¹⁴ Como já foi dito, esse modelo geral também admite a possibilidade de dois ou mais critérios terem o mesmo grau de importância na ordenação. Esse é um ponto importante para a derivação de uma relação geral de preferências e para o estabelecimento de condições de ordenação de alternativas em termos de critérios específicos, cujos níveis de importância variam.

¹⁵ Detalhes sobre aspectos técnicos desse modelo podem ser encontrados em Moldau (1993).

Até aqui sugerimos que as alternativas podem ser comparadas em termos de um determinado critério irreduzível j . Há também espaço para a ordenação de vetores de bens em termos de critérios variáveis. Nesse modelo geral de escolha, alternativas diferentes podem ser completamente pré-ordenadas (elas satisfazem os critérios de completude e transitividade), mesmo quando tomadas em termos de diferentes critérios irreduzíveis.

Assim, para quaisquer duas combinações $x^1, x^2 \in X$ e quaisquer dois critérios $j, h \in J$ (conjunto de critérios), uma das três possibilidades é observada:

- O critério j em x^1 é mais importante do que o critério h em x^2
- O critério h em x^2 é mais importante do que o critério j em x^1
- O critério j em x^1 é tão importante quanto o critério h em x^2

Como podemos notar, o axioma da completude é satisfeito. Adicionalmente, se combinarmos a relação binária “é mais importante do que” com a relação “é tão importante quanto”, podemos derivar uma terceira relação binária capaz de acomodar situações nas quais um critério j em x^1 é “pelo menos tão importante quanto” o critério h em x^2 .

Mostraremos agora que o axioma da transitividade pode ser observado para quaisquer três combinações dadas $x^1, x^2, x^3 \in X$ e quaisquer três critérios $j, h, g \in J$, desde que duas condições sejam atendidas: (a) o critério j em x^1 é “pelo menos tão importante quanto” o critério h em x^2 ; (b) este último é “pelo menos tão importante quanto” o critério g em x^3 .

Para explicar como a escolha com múltiplas razões para agir é determinada, uma relação de preferências geral no conjunto de escolha X deve ser definida. Admitamos que qualquer combinação x está associada com um critério particular $j \in J$, de tal forma a podermos especificar uma relação de não-preferência geral Q . Assumamos ainda que J é um conjunto finito de critérios, constituído por m elementos. J pode ser assim representado por $\{a, b, c, \dots, j, \dots\}$, onde $m \geq 1$. Deve existir, portanto, um número ordenador designado por $k(j, x)$, que é um número inteiro no intervalo fechado entre 1 e m , o qual indica a ordem de cada critério em termos de sua importância relativa. Na terminologia de Moldau, quanto menor esse número, tanto mais importante é o critério na ordenação. Segue-se que:

Se o critério j em x é mais importante do que o critério h em x , então $k(j, x) < k(h, x)$.

Portanto, uma relação de não-preferência binária Q_k relacionada ao $k^{\text{ésimo}}$ critério pode ser definida no conjunto X . Assim sendo,

$x^1 Q_k x^2$, se e somente se $(j(k, x^1), x^1)$ é “pelo menos tão importante quanto” $(j(k, x^2), x^2)$ para cada $x \in X$, onde $k \leq m$.

Um raciocínio semelhante aplica-se às relações de preferência e indiferença P_k e I_k :

$x^1 P_k x^2$, se e somente se, $(j(k, x^2), x^2)$ é pelo menos tão importante quanto $(j(k, x^1), x^1)$

$x^1 I_k x^2$, se e somente se, $(j(k, x^1), x^1)$ é igualmente importante a $(j(k, x^2), x^2)$

Com isso em mente, podemos definir a relação geral de preferências (P), bem como as relações de indiferença (I) e de não-preferência (Q) em X:

$x^1 P x^2$, se e somente se $\exists G \geq 1$ e $x^1 I_k x^2$ para $k < G$ e $x^1 P_G x^2$, para qualquer $x \in X$

$x^1 I x^2$, se e somente se $x^1 I_k x^2$, onde $k = 1, \dots, m$ e $\forall x^1, x^2$

$x^1 Q x^2$, se e somente se $x^1 I x^2$ ou $x^2 P x^1$, $\forall x^1, x^2 \in X$

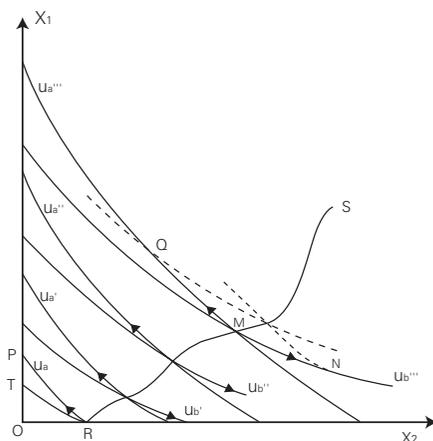
Embora não elaboremos todas as sutilezas subjacentes aos axiomas mencionados, é possível mostrar que as classes de não-preferência são pré-ordenações completas em X para cada critério $j \in J$ e para cada ordem $k \leq m$. (Moldau 1993: 359-61)

Vale destacar que o modelo se funda na hipótese de que o axioma de continuidade mantém-se em relação a cada critério específico $j \in J$. Assim sendo, existe uma função contínua real representativa da escolha com objetivos múltiplos irreduzíveis, função esta que pode ser designada por $U_j(x)$, onde $j = a, b, \dots$. É possível demonstrar que existe também uma função representativa de Q_k em X, para todo $k \leq m$. Essa última função é designada por $U_k(x)$, para iluminar o processo de determinação da escolha.

Devemos estar conscientes de que esse modelo de escolha é muito geral, o que lhe permite acomodar inclusive a instância de funções de utilidade com um único critério de escolha.

Tomemos agora o caso de uma pessoa que deve escolher entre dois cursos de ação, (x_1) e (x_2) , em termos de dois critérios, a, b $\in J$, tal como explicitado na Figura 3 abaixo. Por hipótese, o primeiro critério refere-se ao desejo do agente de buscar suas preferências auto-interessadas, ao passo que o segundo consiste em seu desejo de levar adiante um plano para contribuir para o bem estar social. Os dois critérios podem ser representados, respectivamente, por $U_a = U_a(x_1, x_2)$ e $U_b(x_1, x_2)$. Para esclarecer melhor como a escolha é determinada no modelo lexicográfico proposto, a Figura 3 fornece uma análise gráfica dos problemas de decisão.

Figura 3: Escolha entre dois cursos de ação, em termos de 2 critérios



Fonte: Adaptado a partir de Moldau (1998: 98)

Observando a Figura 3, vemos que o desejo de buscar preferências auto-interessadas é prioritário na região OPR, inclusive PR. No ponto R, ambos os critérios são igualmente importantes: o agente atribui a mesma importância às preferências auto-interessadas e àquelas baseadas no plano de contribuir para o bem estar social. Isso é assim porque $U_a = U_b$. Olhando novamente a curva PR (onde o critério predominante é a busca de auto-interesse), notaremos que qualquer combinação de bens com maior quantidade de x_1 (um bem privado puro, por exemplo) é preferida a R. Vale destacar que os critérios (a) e (b) são igualmente importantes para quaisquer combinações sobre a curva RS. À esquerda da mesma, as preferências auto-interessadas tornam-se o objetivo prioritário, ao passo que à sua direita a disposição de cumprir o plano de contribuir para o bem estar social é dominante, ou seja, o critério b torna-se absolutamente prioritário.

Devemos ter em mente que $U_{a'}$, $U_{a''}$, $U_{a'''}$, U_b , U_b'' e U_b''' (admitindo-se $U_{a'} < U_{a''} < U_{a'''} < U_{a''''}$ e $U_b < U_b' < U_b'' < U_b'''$) são níveis distintos de satisfação, associados, respectivamente, às razões para agir (a) e (b). Notamos que qualquer alocação abaixo da curva $U_{a'''}MU_b'''$ gera níveis de satisfações menores em termos do critério de primeira ordem. No ponto M, o critério de escolha (a) é tão importante quanto o critério (b). É importante notar que, apesar de M e N estarem sob a curva de ordenação dos objetivos ou critérios de primeira ordem ($U_{a'''}MU_b'''$), a cesta N é preferível a M. Isso ocorre porque N está associada a um nível de satisfação mais elevado do critério de segunda ordem, como mostram as curvas tracejadas no gráfico. Reiteramos aqui uma das idéias centrais do modelo lexicográfico proposto: quando a escolha não pode ser determinada pelo critério de primeira ordem, porque as alternativas são indiferentes com relação ao mesmo, o critério de segunda ordem torna-se decisivo.

O modelo representado na Figura 3 é geral o bastante para acomodar uma gama de motivos da vontade. Suponhamos agora que uma pessoa tenha três razões para agir, baseadas nas quais faz sua escolha. Os critérios de escolha seriam:

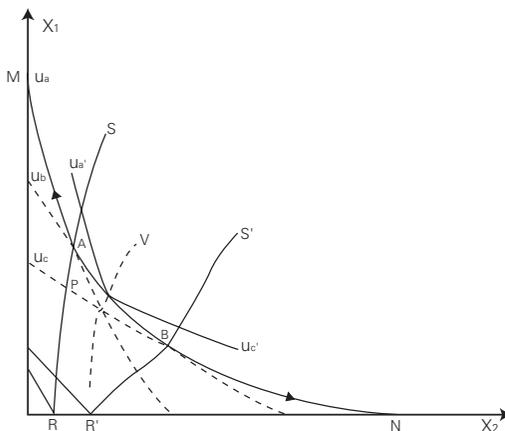
- desejo de satisfazer suas próprias necessidades e vontades;¹⁶
- desejo de conquistar uma auto-imagem positiva e aprovação social;
- desejo de agir de acordo com seu senso de dever.

Assim sendo, o tomador de decisão terá três funções representativas da satisfação de cada critério, $U_a(x_1, x_2)$, $U_b(x_1, x_2)$ e $U_c(x_1, x_2)$, tal como ilustrado pela Figura 4.

No gráfico da Figura 4, as curvas RS, R'S' e R''S'' retratam, respectivamente, as situações $U_a = U_b$ e $U_b = U_c$. Então, para qualquer alternativa à esquerda de RS, o critério (a) é absolutamente predominante. Entre RS e R'S', o critério (b) é visto como prioritário. À direita de R'S' o critério (c) passa a ser o mais importante. A análise da determinação da escolha em termos de razões irreduzíveis

¹⁶ Tais necessidades e vontades poderiam abrigar aquelas de seus familiares e amigos próximos, em relação aos quais o agente nutriria simpatia, vendo-os como uma extensão de si próprio.

Figura 4: Escolha entre dois cursos de ação, em termos de 3 critérios



Fonte: Adaptado a partir de Moldau (1988: 105)

para a ação começa na curva MABN, que corresponde ao nível mais alto de satisfação em termos do conjunto ordenado de critérios. Partindo de M em direção a A, o critério “desejo de satisfazer suas próprias necessidades e vontades” mantém-se constante. Contudo, outros critérios vão aos poucos adquirindo maior importância. No ponto A, o critério (b) passa a ser tão importante quanto (a), isto é, $U_a = U_b$. Entre os pontos A e B, o peso relativo do próximo critério (desejo de conquistar uma auto-imagem positiva e aprovação social) é constante, enquanto o critério (a) perde importância e o critério (c) sofre um aumento relativo de sua importância. No ponto B, os critérios (b) e (c) são igualmente importantes e coincidem com o critério de primeira ordem. De B para N, o desejo de honrar o próprio senso de dever converte-se em prioridade absoluta, e o agente racional terá incentivos para empenhar-se em satisfazê-lo.

Como podemos ver, esse modelo geral de escolha determinado por critérios múltiplos irreduzíveis nos permite explicar situações em que o critério prioritário de escolha varia. A escolha efetiva do indivíduo será definida pela curva MABN, pois esta exprime a ordenação de critérios mais elevada. Assim, qualquer ponto abaixo de MABN implica um nível de satisfação mais baixo do critério prioritário. No ponto P, por exemplo, o nível de satisfação gerado seria menor em termos do critério (b), enquanto Q associa-se a um menor nível de satisfação de (c) na hierarquia de necessidades.

Cabe agora refletir sobre as vantagens comparativas desse modelo em relação ao modelo de escolha tradicional. A resposta depende, certamente, das questões específicas que se quer responder. Como tentamos mostrar, explicações sobre a lógica de compromissos e planos têm necessariamente de incorporar razões para agir que estejam além das preferências auto-interessadas. Vemos com ceticismo a possibilidade do modelo convencional realizar essa tarefa. Além de permitir a incorporação dessas outras razões para agir, nosso modelo explicita as

fontes de preferências, ao invés de tomar a própria noção de preferência como um termo primitivo.

Uma segunda vantagem do modelo de escolha lexicográfica aqui sugerido é o fato de que possibilita dar um tratamento explícito às forças motrizes subjacentes à tomada de decisão racional de maneira relativamente parcimoniosa. Comparativamente à tentativa anterior de tratamento formal feita por Sen, ele mostra diretamente como a escolha é determinada quando vários critérios irreduzíveis estão em jogo, e o que ocorre quando a escolha não é determinada por um tipo de critério na hierarquia dos mesmos. A moral da história é que agir em conformidade com planos e com o senso de dever pode constituir um critério prioritário, em condições específicas. Cremos que inúmeras instâncias de escolha na vida prática iluminam a plausibilidade de nossa perspectiva teórica. (Hirschman [1985] 1992; Mansbridge 1990)

Contudo, cabe fazer certas qualificações. O modelo de escolha aqui construído não está totalmente a salvo dos problemas que afetam a abordagem de metapreferências, inclusive o risco de uma regressão infinita. Nosso modelo minimiza esse problema, ao descrever a escolha como sendo determinada pelo critério de primeira ordem, sendo que critérios menos prioritários na hierarquia só se tornam decisivos quando as alternativas são indiferentes em relação ao critério de primeira ordem.

OBSERVAÇÕES FINAIS

Esse artigo engajou-se em uma análise crítica da teoria de escolha tradicional, que não se pronuncia acerca dos fundamentos psicológicos do comportamento racional. Argumentamos aqui que a omissão apontada restringe o poder explicativo dessa teoria. Mais especificamente, buscamos mostrar que a abordagem econômica convencional é incapaz de dar conta, sem ajustes *ad hoc*, de todo um conjunto de fenômenos que podem ser rotulados de compromissos genuínos e planos.

Para tratar desse tema discutimos na seção 2 a racionalidade do comportamento planejado à luz das contribuições de Gauthier. Concordamos com este autor que, para entender a lógica de planos e compromissos, precisamos acomodar no corpo da análise razões para agir que não se restringem às preferências auto-interessadas. Contudo, discordamos de Gauthier em função da pequena importância que atribui às expectativas normativas na coordenação social. Não nos parece possível explicar a coordenação interpessoal pelo mero reconhecimento da habilidade humana de fazer e executar planos, ou pela emergência das convenções. Ao contrário, defendemos aqui o significado causal de sentimentos morais tais como senso de dever, culpa, vergonha, ressentimento, raiva e empatia no processo de tomada de decisão.

Com essa preocupação em mente, escrutinamos brevemente na seção 3 a abordagem kantiana do comportamento racional, que também focaliza a racio-

nalidade dos compromissos como escolhas contrapreferenciais. A intenção foi lançar nova luz sobre os *insights* decorrentes dessa perspectiva. Acatamos a idéia kantiana de imperativos categóricos, ainda que discordando de sua concepção de racionalidade livre de qualquer estado psicológico e, como tal, independente das preferências e crenças humanas. Esse aspecto da perspectiva kantiana nos pareceu pouco aplicável aos modelos de escolha sobre os quais os economistas têm se debruçado, que explicam o comportamento com base em preferências, ou seja, em estados psicológicos.

Ainda na seção 3 apresentamos uma posição mais abrangente, baseada em Sen e Hirschman, que por sua vez se inspiraram no conceito de Frankfurt sobre desejos de ordem superior. Endossamos seu pleito no sentido de “complicar” o discurso econômico, discutindo o papel das meta-ordenações de preferências ou metapreferências.

Admitimos que essa estratégia está sujeita a problemas de tratabilidade. Para superar essa dificuldade, sugerimos na seção 4 um modelo segundo o qual a escolha é determinada por um conjunto de objetivos múltiplos irreduzíveis. Esse modelo lexicográfico construído originalmente por Moldau abre espaço para a explicação do comportamento movido por diferentes critérios e ordenações de critérios, dispostos em uma estrutura hierárquica. Além disso, ele está melhor sintonizado com a noção humeana dos estados psicológicos como indutores da ação.

Para encerrar o artigo, reiteramos nossa crítica à teoria da escolha tradicional, por sua incapacidade de captar os mecanismos e processos causais envolvidos no fenômeno da escolha. Boas teorias fornecem descrições adequadas daquilo que se passa na realidade econômica, a despeito das necessárias simplificações, abstrações e omissões. No que concerne à nossa história particular, a mensagem a ser extraída é de que não há nada de errado em excluir elementos da psicologia da escolha que não são essenciais, desde que as principais forças motrizes que determinam o comportamento racional sejam adequadamente levadas em conta. A parcimônia é um valor importante na construção teórica, mas sua importância tem sido exagerada.

REFERÊNCIAS BIBLIOGRÁFICAS

- AINSLIE, G. e N. HASLAM (1992a) “Hyperbolic discounting”, in: Lowenstein, G. e J. Elster (eds.) *Choice over time*. N. York: Russell Sage Foundation.
- AINSLIE, G. e N. HASLAM (1992b) “Self-control”, in: Lowenstein, G. e J. Elster (eds.) *Choice over time*. N. York: Russell Sage Foundation.
- BIANCHI, A. M. (1998) “Are Brazilian economists different?” *Revista Brasileira de Economia* 52 (3), jul/set: 427-39.
- BRENNAN, T. (1989) “A methodological assessment of multiple utility frameworks.” *Economics and Philosophy* 5: 189-208.
- ELSTER, J. (2000) *Ulisses unbound*. Cambridge University Press.
- FRANK, Robert H. (1992) “The role of moral sentiments in the theory of intertemporal choice”, in: Lowenstein, G. e J. Elster (eds.) *Choice over time*. N. York: Russell Sage Foundation.

- FRANK, Robert. (1988) *Passions within Reason: The strategic role of the emotions*. New York: W. Norton & Company Inc.
- FRANK, Robert. (1992) "The role of moral sentiments in the theory of intertemporal choice", in G. Lowenstein e Jon Elster (eds.), *Choice over time*. N. York: Russell Sage Foundation.
- FRANKFURT, H. (1971) "Freedom of will and the concept of a person" *Journal of Philosophy* 68: 5-20.
- FRIEDMAN, M. e L. SAVAGE (1948) "The Utility Analysis of Choice Involving Risk" *The Journal of Political Economy* 56 (4): 279-304.
- GAUTHIER, David. (1996) "Commitment and choice: an essay on the rationality of plans", in: Farina, F.; F. Hahn e S. Vanucci (eds.) *Ethics, Rationality and Economic Behavior*. Clarendon Press.
- HIRSCHMAN, A. ([1985] 1992) *Rival Views of Market Society*. Cambridge: Harvard University Press, cap. 6: "Against parsimony: three easy ways of complicating some categories of economic discourse".
- HOLLIS, Martin e R. SUGDEN (1993) "Rationality in action". *Mind* 102:1-33.
- KANT, I. (1992) *Immanuel Kant: critical assessments*, vol III. Editado por R. Chadwick. Routledge.
- KATZ, L. (2000) *Evolutionary origins of morality: cross disciplinary perspectives*. Imprint Academic.
- LUTZ, Mark A. (1993) "The utility of multiple utility: a comment on Brennan". *Economics and Philosophy* 9: 145-54.
- MAKI, U. (1994) "Isolation, idealization and truth in economics", in B. Hamming e N. de Marchi, *Idealization in Economics*. Amsterdam-Atlanta, GA: Rodopi.
- MAKI, U. (2001) "Theoretical isolation and explanatory progress: transaction cost economics and the dynamics of dispute". Seminário do EIPE, Rotterdam, jan.
- MANSBRIDGE, Jane (1990) "The rise and fall of self-interest in the explanation of political life", in: J. J. Mansbridge (ed.), *Beyond self-interest*. The University of Chicago Press.
- MOLDAU, J. H. (1988) *A teoria da escolha com objetivos irreduzíveis e suas implicações*. Série Ensaaios Econômicos n° 69, Instituto de Pesquisas Econômicas, Universidade de São Paulo.
- MOLDAU, J. H. (1993) "A model of choice where choice is determined by an ordered set of irreducible criteria". *Journal of Economic Theory* 60 (2): 354-77.
- MURAMATSU, Roberta (1998) "Sobre a possibilidade do altruísmo: uma incursão pela análise econômica do comportamento não-egoísta". Dissertação de mestrado, Depto. de Economia, Universidade de São Paulo.
- MURAMATSU, Roberta (forthcoming). "Emotions in action: an inquiry into the psychobiological foundations of decision making". Tese de doutorado, Universidade Erasmus de Rotterdam.
- SAVAGE, L. J. (1954) *The foundations of statistics*. New York: Wiley.
- SHELLING, T. C. (1992) "Self-command: A new discipline", in: Lowenstein, G. e J. Elster (eds.) *Choice over time*. N. York: Russell Sage Foundation.
- SEN, A. (1982) *Choice, welfare and measurement*. Oxford: Basis Blackwell.
- SHAPIRO, Scott J. (2002). "Ulysses rebound". *Economics and Philosophy* 18: 157-182.
- SUGDEN, R. (1991) "Rational Choice: a survey of contributions from economics and philosophy". *The Economic Journal* 101: 751-85.
- SUGDEN, R. (1998) "Normative expectations", in: Ben-Ner, A. e L. Putterman (eds.), *Economics, values and organization*. Cambridge University Press.
- THALER, R. (1992) *The Winner's Curse: anomalies and paradoxes of economic life*. The Free Press.
- VROMEN, J. (2000) "Moral behaviour in evolutionary theorizing: variations on a theme?". Preparado para a Erasmus Summer School on Social Ontology, Rotterdam.
- YOUNG, P. (1993) "The evolution of conventions". *Econometrica* 61: 57-64.