

Carlos Roberto Bueno Ferreira

# QUEM DECIDE?

Para Além da  
Escolha Individual

UMA ANÁLISE  
DA ESCOLHA SOCIAL  
PELA  
COOPERAÇÃO JUSTIFICADA



Neste estudo seminal sobre a racionalidade e justificativa da escolha social, Carlos Roberto Bueno Ferreira nos oferece um modelo capaz de manter a potência formal da análise lógica dos mecanismos de escolha sem mitigar ou eclipsar os argumentos filosóficos fundamentais da vida social justa. *Para Além da Escolha Individual* é, de resto, uma excelente introdução aos problemas filosóficos, metodológicos e econômicos de teorias da escolha social, da cooperação social e da justiça procedimental, explorando diferentes tipos de paradoxos, dilemas e aporias decorrentes de possíveis abordagens teóricas da racionalidade coletiva. Trata-se de importante e original contribuição às pesquisas interdisciplinares em epistemologia social, filosofia da economia e teoria política em nossos dias de incertezas e de cenários cada vez mais complexos, onde agentes e legisladores precisam justificar processos decisórios para implementar políticas públicas e ações coletivas que promovam efetivamente o bem comum em nossas democracias. Além de reconstituir os argumentos e os modelos mais relevantes para uma teoria da escolha social, o Autor logra mostrar por que uma teoria da evolução da cooperação humana ajuda-nos a compreender o que leva o indivíduo a abrir mão da posição puramente egoísta sem perder de vista a racionalidade inerente à escolha individual.

**PhD. Nythamar de Oliveira** |



**Quem Decide?**  
**Para Além da Escolha Individual**

## *Direção Editorial*

---

Lucas Fontella Margoni

## *Comitê Científico*

---

**Prof. Dr. Nythamar Hilário de Oliveira**

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

**Prof. Dr. Fabrício Pontin**

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

**Prof. Dr. Thadeu Weber**

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

**Prof.<sup>a</sup> Dr.<sup>a</sup> Izete Pengo Bagolin**

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

**Prof. Dr. Nikolay Steffens Martins**

Universidade Federal do Rio Grande do Sul (UFRGS)

# **Quem Decide? Para Além da Escolha Individual**

Uma análise da escolha social pela cooperação justificada

**Carlos Roberto Bueno Ferreira**



**Diagramação:** Marcelo A. S. Alves

**Capa:** Carole Kümmecke - <https://www.conceptualeditora.com/>

**Ilustração de Capa:** Emma Atterholm

**O padrão ortográfico e o sistema de citações e referências bibliográficas são prerrogativas de cada autor. Da mesma forma, o conteúdo de cada capítulo é de inteira e exclusiva responsabilidade de seu respectivo autor.**



Todos os livros publicados pela Editora Fi estão sob os direitos da [Creative Commons 4.0](https://creativecommons.org/licenses/by/4.0/deed.pt_BR) [https://creativecommons.org/licenses/by/4.0/deed.pt\\_BR](https://creativecommons.org/licenses/by/4.0/deed.pt_BR)



Dados Internacionais de Catalogação na Publicação (CIP)

---

FERREIRA, Carlos Roberto Bueno

Quem Decide? Para Além da Escolha Individual: uma análise da escolha social pela cooperação justificada [recurso eletrônico] / Carlos Roberto Bueno Ferreira -- Porto Alegre, RS: Editora Fi, 2021.

133 p.

ISBN - 978-65-5917-327-3

DOI - 10.22350/9786559173273

Disponível em: <http://www.editorafi.org>

1. Escolha Social; 2. Cooperação; 3. Justificação; 4. Estado; 5. Brasil; I. Título.

CDD: 100

---

Índices para catálogo sistemático:

1. Filosofia 100

# Sumário

---

## **Apresentação** **9**

Nythamar de Oliveira

---

## **Prefácio à 1ª Edição** **12**

---

## **Introdução** **14**

---

## **1** **19**

### **Escolha social e a eleição de preferências em regimes democráticos**

1.1 A ideia da escolha social.....	20
1.2 Escolhas, ordens de preferências e valores individuais .....	32
1.3 A natureza formal da escolha .....	35
1.4 Condições para a função de bem-estar social .....	38
1.5 O paradoxo de arrow .....	42
1.6 Justiça social e alternativas ao paradoxo da escolha.....	46
1.6.1 (Re)Introduzindo comparação interpessoal de utilidade.....	47
1.6.2 O Paradoxo Liberal.....	51
1.6.3 Justiça, Bem-Estar e Bens Primários .....	55
1.7 Disposições finais ao primeiro capítulo.....	58

---

## **2** **63**

### **Cooperação no mundo dos egoístas**

2.1 O dilema do prisioneiro .....	64
2.2 Condições para cooperação .....	69
2.3 Um torneio de estratégias .....	72
2.3.1 Não Seja Ganancioso .....	77
2.3.2 Não seja o primeiro a quebrar a confiança.....	78
2.3.3 Reciprocidade .....	81
2.3.4 Clareza.....	82
2.4 A evolução da cooperação .....	84

---

**Escolha social pela cooperação justificada**

3.1 <i>First things first</i> : estabelecendo o equilíbrio cooperativo.....	90
3.2 Mantendo o equilíbrio cooperativo.....	95
3.3 Justiça e justificação.....	103
3.3.1 O caráter relacional da justiça.....	105
3.3.2 Princípio da Justificação Geral e Recíproca.....	108
3.3.3 Um modelo de Justificação.....	110
3.4 A questão da viabilidade.....	111
3.4.1 Auto-interesse e Racionalidade Econômica.....	113
3.4.2 Bloqueios de Viabilidade.....	115

---

**Conclusão**

---

**Projetando um possível modelo de justiça**

---

**Referências**

## **Apresentação**

*Nythamar de Oliveira*

Neste estudo seminal sobre a racionalidade e justificativa da escolha social, Carlos Roberto Bueno Ferreira nos oferece um modelo capaz de manter a potência formal da análise lógica dos mecanismos de escolha sem mitigar ou eclipsar os argumentos filosóficos fundamentais da vida social justa. *Quem Decide? Para Além da Escolha Individual* é, de resto, uma excelente introdução aos problemas filosóficos, metodológicos e econômicos de teorias da escolha social, da cooperação social e da justiça procedimental, explorando diferentes tipos de paradoxos, dilemas e aporias decorrentes de possíveis abordagens teóricas da racionalidade coletiva.

Trata-se de importante e original contribuição às pesquisas interdisciplinares em epistemologia social, filosofia da economia e teoria política em nossos dias de incertezas e de cenários cada vez mais complexos, onde agentes e legisladores precisam justificar processos decisórios para implementar políticas públicas e ações coletivas que promovam efetivamente o bem comum em nossas democracias. Além de reconstituir os argumentos e os modelos mais relevantes para uma teoria da escolha social, o Autor logra mostrar por que uma teoria da evolução da cooperação humana ajuda-nos a compreender o que leva o indivíduo a abrir mão da posição puramente egoísta sem perder de vista a racionalidade inerente à escolha individual.

Evita-se, destarte, os impasses e imperfeições de modelos atomistas (o chamado individualismo metodológico) e de seus antagônicos modelos coletivistas ou holistas (hodiernamente classificados sob o mesmo guarda-

chuva conceitual do comunitarismo), contribuindo para superar as viciadas querelas entre weberianos e durkheimianos, em sociologia, ou entre utilitaristas liberais e igualitaristas sociais, em teoria ético-política. Com efeito, o enfoque num modelo discursivo de justificação recíproca busca se firmar, segundo o Autor, como uma alternativa defensável para imprimir decisões autônomas, cooperativas e não autoritárias, fomentando a deliberação e a busca de uma racionalidade justificada capaz de assegurar a autonomia, sem arbitrariedade e como estratégia majoritária para coibir a atuação não-cooperativa.

A fim de traduzir a vontade individual em termos intersubjetivos, o Autor analisa e procura reconstituir os motivos que levam o indivíduo a cooperar, oferecendo uma solução procedimental em duas etapas, a saber: o estabelecimento do equilíbrio cooperativo e a manutenção e reequilíbrio da cooperação estabelecida. A justificação é agora considerada não apenas como um direito justificatório (*Recht auf Rechtfertigung*), mas como um dever recíproco, visando suprir o déficit informacional na tomada de decisão. A abordagem da escolha justificada e recíproca passa, portanto, a funcionar como um planejador central que, munido das declarações dos próprios agentes, é capaz de avaliar melhor a racionalidade que informa a decisão social podendo, inclusive, prever e detectar eventuais erros e anomalias, tanto no processo de seleção como nos resultados da escolha.

Trata-se, em última análise, de uma aposta pragmática na busca por uma solução que melhor se adequa à nossa realidade social em face da crise de representatividade e da falência aparente dos atuais mecanismos de eleição nas sociedades democráticas contemporâneas. Para tanto, é mister que se avalie atentamente as condições de racionalidade estabelecidas, considerando os requisitos formais de desempenho lógico, tais como foram historicamente desenvolvidos na teoria da escolha coletiva, visando uma resposta viável que não apenas represente um somatório de escolhas

individuais, mas possa de fato traduzir uma verdadeira escolha social. O Autor mostra-nos de forma convincente por que se faz necessária, hoje mais do que nunca, uma reavaliação da possibilidade de cooperação, haja vista as concessões individuais que devem tomar lugar num modelo de escolha social. Numa democracia tão dividida por polarizações e agendas ideológicas radicalizadas como a nossa, os desafios normativos da escolha social recíproca são ainda a única esperança que nos resta para resgatar o nosso compromisso ontológico com o bem comum que subjaz a nossas diferentes concepções, projetos e práticas de liberdade, igualdade e solidariedade cívica.

## Prefácio à 1ª Edição

A situação é a mais simples possível: eu, você e mais alguém, temos que definir o que é preferível entre  $x$ ,  $y$  ou  $z$ . Do ponto de vista da racionalidade econômica, avaliando de maneira simplista, a utilidade que se pode extrair de tais assertivas depende da capacidade dessa informação de servir adequadamente à relação entre a demanda dos agentes e a oferta daquilo que desejam ( $x$ ,  $y$  ou  $z$ ). Para a escolha social válida, contudo, não basta suprir interesses meramente individuais. Busca-se, portanto, uma função final de escolhas tal que o resultado represente um *juízo de preferência legítimo* que reflita a vontade de todos envolvidos.

Quando escrevi o presente trabalho o cenário político-social já indicava uma conjuntura de crescente polarização política, na qual escolhas coletivas muitas vezes aparentam uma completa desconexão com a vontade dos agentes interessados. Este quadro, ao que tudo leva a crer, tende a se potencializar. O sentimento de hipossuficiência representativa nos faz questionar a própria legitimidade dos *sistemas de votação* e de *escolha social* que temos a disposição.

“Quem escolheu isso?” ou “Isso não representa minha vontade”, infelizmente, são impressões mais comuns do que gostaríamos de admitir nos modernos Estados democráticos, os quais, paradoxalmente, proclamam ser alicerçados na soberania das escolhas populares.

Antes, entretanto, de atribuímos fatores subjetivos complexos para justificar tal fenômeno, como a crescente desinformação decorrente da era da *internet* e manipulação dos valores do debate público (que são tão antigos quanto a própria socialização humana), cabe o questionamento: *é mesmo possível a escolha social?*

Digo aqui *possível* com sentido de capacidade de existência. A pergunta que pretendi avaliar primeiramente, portanto, é ainda quanto ao *aspecto formal* da escolha, sem sequer adentrar ao estudo das motivações subjetivas. Trata-se de apreciação sobre a viabilidade de um modelo de escolhas que seja possível desde o ponto de vista lógico-matemático.

Ora, se, mesmo em abstrato e considerando os cenários os mais simples, tais que sejam exprimíveis em linguagem lógico-matemática, a escolha social não se demonstrar capaz de refletir a vontade final de um grupo de agentes, então todo o trabalho intelectual subsequente de tentar justificar tais escolhas como legítimas preferências sociais, demonstrar-se-á em vão.

Seria, portanto, impossível a escolha social? Quais os ajustes necessários para suprir as limitações e paradoxos dos métodos de escolha coletivos?

Convido o leitor, em um primeiro momento, a lembrar as origens dos métodos de votação, bem como entender e analisar os requisitos que levam a uma legítima função de escolha social.

Tentando uma alternativa ao problema de legitimidade de escolha social acima levantado, a presente obra traz a ideia da *cooperação justificada*, que se trata, em linhas gerais, de uma proposta de compatibilização da *teoria da cooperação* com a *justificação normativa* (no caso desta obra particularmente oriunda de sistemas de deliberação pública). A avaliação que se segue busca averiguar se é possível que pessoas, mesmo com posicionamentos políticos opostos, possam cooperar em função de um legítimo resultado social comum, sem deixar de lado seu interesse egoístico.

## Introdução

Esta obra é uma proposição otimista. A opção metodológica pela abordagem a partir da teoria da escolha social traz consigo os benefícios da objetividade econômico-matemática, de caráter formal e analítico, mas implica as desvantagens de uma análise que deixa de considerar as variáveis sociais e humanas inerentes ao processo de eleição. A solução, penso, deve estar em algum lugar entre a clareza da formalidade interpretativa dos axiomas lógicos e a sua debilidade natural em avaliar questões práticas e subjetivas.

Qual a relação de dependência entre a vontade individual e a legítima vontade social? A tradução da racionalidade que informa a escolha individual em um arranjo de escolha social pressupõe uma criteriosa série de condições formais. A teoria da escolha social, enquanto área do conhecimento afeta à evolução dos métodos de decisão coletiva, ocupou-se de estabelecer estes critérios de desempenho lógico aplicáveis aos mecanismos de escolha.

O tema da escolha social, contudo, lançou-se com maior repercussão desde a contribuição decisiva de Arrow em seu livro “*Escolha Social e Valores Individuais*”, em 1951, no qual o autor postula o *teorema geral da possibilidade*, reformulando as condições lógicas da escolha coletiva sob elegante asserção axiomática. Apesar do nome altivo, tal teorema ficou mais conhecido por sua versão pessimista, o *paradoxo de Arrow*, uma vez que concluiu pela impossibilidade de se alcançar a escolha social (de forma não autoritária e não impositiva).

Na esteira da ebulição intelectual provocada pela obra de Arrow foram sugeridas alternativas econômicas, políticas e filosóficas ao impasse

gerado pelo referido paradoxo. Foram traçados inevitáveis paralelos com as teorias clássicas e modernas da justiça, porquanto lidam, primordialmente, com o eterno conflito entre a solução utilitarista informada pela soma dos interesses particulares e os arranjos de direitos sociais, informados por argumentos filosóficos sobre a boa vida social.

O primeiro capítulo ocupa-se com os requisitos lógico-rationais para a legítima escolha coletiva e as possíveis alternativas ao impasse da escolha social. Para propor uma alternativa razoável ao paradoxo da escolha é preciso, antes de tudo, que se avalie atentamente as condições de racionalidade estabelecidas e os tipos respectivos de flexibilização que elas permitem.

Pressupondo-se um cenário no qual a interação contínua dos participantes gere um senso de extensão dos efeitos de cada escolha particular, o problema da *escolha social* se conecta com a temática da possibilidade de *cooperação social*. Neste sentido, a escolha social vai além da escolha individual, sem recorrer a uma função meramente agregativa, nem a uma autoridade impositiva externa. Também nessa acepção, a escolha pela cooperação assume um tom intersubjetivo e se contrapõe a atuação meramente auto-interessada e egoísta, assumida como conduta racional padrão na teoria econômica em geral.

A cooperação é o tema do segundo capítulo. As condições para o surgimento e manutenção da cooperação, bem como as estratégias de ação dos indivíduos visando o melhor resultado geral, são assuntos comuns a diversas áreas do conhecimento, dentre elas a teoria dos jogos, a economia de bem-estar, a filosofia política e as relações internacionais.

A obra pivô escolhida para a discussão de tais temas é “*A Evolução da Cooperação*” de Robert Axelrod, publicada pela primeira vez em 1984. Nela o autor sugere tanto causas sociais como biológicas para a ocorrência natural da cooperação.

Se pensarmos nas relações humanas como um jogo continuado de possíveis escolhas e resultados, estamos diante de problema que pede uma solução intersubjetiva. A possibilidade de cooperação passa inevitavelmente pelas respostas que esperamos dos demais integrantes do jogo social. O que leva o indivíduo a cooperar num mundo egoísta sem que haja uma imposição autoritária? Como inibir a atitude auto-interessada, mesmo quando ela se mostra vantajosa?

A resposta pessimista tem como melhor expoente a proposição hobbesiana de que o homem é dominado pelo egoísmo e pela competitividade, cenário no qual a cooperação não pode se desenvolver senão por meio de uma forte autoridade central.

Do outro lado, uma abordagem otimista busca avaliar a possibilidade da cooperação com base em normas de reciprocidade e estratégias de ação em jogos continuados. Trata-se de uma *Teoria da Cooperação* fundada na investigação sobre indivíduos que buscam o próprio interesse sem ingerência de uma autoridade impositiva. Estratégias cooperativas como o *tit-for-tat* (analisada no segundo capítulo) se mostraram potencialmente mais vantajosas do que propostas puramente competitivas ou retaliativas.

Trabalhando tanto os elementos da *teoria da escolha social* como os requisitos para a cooperação social, a presente obra propõe um procedimento em duas etapas que busca a implementação e a manutenção do equilíbrio cooperativo, objeto de análise no terceiro capítulo.

A primeira fase do procedimento visa à instalação da cooperação como estratégia estável. Trata-se, em última análise, da *escolha social pela cooperação*. Demonstrar-se-á que a cooperação pode emergir mesmo entre egoístas, sem depender necessariamente de uma mudança ideológica e sem a intervenção de uma autoridade externa. A cooperação, nesta primeira etapa, é entendida como o resultado mais vantajoso desde um ponto

de vista da soma de utilidade, que deve refletir uma função social de associação positiva.

Na fase inicial é crucial a liberdade dos participantes em escolher integrar o sistema cooperativo. Neste momento não são permitidos constrangimentos, comprometimentos e barganhas. A possibilidade da emergência da cooperação deve ser testada em termos objetivos.

Uma vez que estabelecido um equilíbrio cooperativo, a postura puramente egoísta não tem chances de prevalecer. Num ambiente ótimo, onde não há quebra de confiança, os resultados coletivos decorrentes da cooperação são os mais favoráveis e, portanto, os mais desejados.

Contudo, o critério de utilidade adotado na etapa de implementação do equilíbrio cooperativo não é suficiente para a manutenção de um sistema de cooperação social livre e justo. É preciso um mecanismo mais complexo que dê conta, por exemplo, de avaliações interpessoais de utilidade e de soluções que desafiem a maximização paretiana.

Daí a necessidade de agregar uma segunda fase ao procedimento, na qual o foco deve voltar-se para a questão da cooperação social entre cidadãos racionais livres em sociedades minimamente razoáveis. Neste segundo estágio a escolha deve ser ampla com um domínio irrestrito. Trata-se de um cenário de embate discursivo no qual indivíduos são dotados de autonomia para avaliar, inclusive, os fatores que levaram ao equilíbrio cooperativo em primeiro lugar.

Ponto comum que se pode notar é que a conduta auto-interessada do indivíduo não desaparece, mas deve ser traduzida de forma intersubjetiva, de preferência sob a interface de uma estratégia que iniba a conduta não-cooperativa dos demais participantes.

Assim, propõe-se um modelo discursivo de justificação reflexiva como alternativa para imprimir decisões autônomas, cooperativas e não autoritárias. Nesse modelo deve haver espaço sempre aberto para a

deliberação e uma racionalidade justificada que garanta a autonomia, a não arbitrariedade e sirva, analogamente, como estratégia majoritária para coibir a atuação não-cooperativa.

A justificação, nestes termos, passa a ser considerada não só como direito, mas como dever, visando suprir o déficit informacional na tomada de decisão. Na *teoria da escolha social* as preferências dos agentes são tomadas como informações meramente individuais, a partir das quais se buscam estimativas que deem conta de prever os resultados das interações entre os participantes do jogo contínuo que representa a vida em sociedade. Em outras palavras, procura-se projetar mecanismos tais que, quaisquer que sejam as preferências dos agentes, suas escolhas estratégicas levem aos mesmos resultados que a *função de bem-estar social* pré-determinou. Tal postura invariavelmente leva a resultados contraditórios e aparentemente incompatíveis com a verdadeira intenção do agente no ato da escolha. Isso representa um típico caso de eleição social mal informada.

Em face disso, a abordagem justificativa aqui proposta passa a funcionar como um planejador central que, munido das declarações dos próprios agentes, é capaz de avaliar melhor a racionalidade que informa a decisão social podendo, inclusive, prever e detectar eventuais erros e anomalias, tanto no processo de seleção como no resultado deliberado.

## **Escolha social e a eleição de preferências em regimes democráticos**

*“Uma compreensão do governo da maioria, da democracia, do liberalismo que faz sem o utilitarismo, e que faz mais do que afirmar que direitos estão certos, deve viajar por espaços mais misteriosos, deve subir escadas mais sinuosas, e deve empregar um altímetro mais complexo do que consistência transitiva.”*  
(Rae, 1980).

Regimes democráticos<sup>1</sup> presumem um grau de livre escolha de seus membros integrantes. As escolhas de cada indivíduo, por sua vez, devem de algum modo determinar a vontade social resultante do procedimento de eleição. O primeiro capítulo deste trabalho será dedicado ao estudo desta relação de interdependência entre as preferências individuais de cada membro da sociedade e os resultados sociais decorrentes dessas escolhas.

Quando cidadãos elegem representantes políticos está-se diante de um somatório de escolhas individuais ou de uma função que reflete uma escolha social? Um resultado eleitoral pode ser considerado como uma legítima escolha tomada socialmente ou consiste apenas no conjunto de todas as escolhas individuais de voto? E, mais além, a formação de uma preferência política representa, de fato, um acordo social? Ainda, pode-se dizer que o procedimento de escolha decorre de um conjunto de critérios lógico-matemáticos que garantem a coerência entre a vontade individual dos eleitores e a vontade social legitimamente representada?

---

<sup>1</sup> Em oposição a regimes autoritários, nos quais a vontade individual não é capaz de influenciar a decisão social resultante. Aqui considerando tanto um modelo ditatorial, em que a vontade social é determinada pela escolha de um ou alguns membros, como a noção de um ente social superior que define o resultado social independentemente da escolha dos indivíduos.

Essas são questões que se destacam em um modelo democrático de governo. Trata-se de um problema estrutural situado no próprio alicerce da legitimação da escolha em sociedade. A formação de preferências sociais depende de regras e critérios lógicos.

A abordagem por meio da teoria da escolha social permite uma análise formal detalhada e consistente dos requisitos que conformam uma verdadeira decisão coletiva.

### **1.1 A ideia da escolha social**

A teoria da escolha social é o ramo do conhecimento afeto à evolução dos métodos de decisão coletiva. Trata-se, em sua natureza, de conteúdo multidisciplinar que tem sido historicamente abordado a partir de variados pontos de vista, desde os puramente econômicos até perspectivas éticas e sociais mais abrangentes.

O problema surge com o mais simples dos cenários políticos: a necessidade de se definir como agir em busca do bem comum quando envolvidos dois ou mais indivíduos. Portanto, a busca pelo critério mais eficaz de escolha coletiva é assunto tão antigo quanto a própria socialização. Do outro lado da mesma moeda estão as políticas públicas e o bem-estar, que, segundo Paul Samuelson (1947, p. 203), são *os primordiais assuntos em comum entre economistas, filósofos, teólogos, panfletários, líderes sociais e reformistas*.

Com efeito, a filosofia sempre esteve ocupada em apontar métodos de decisão coletiva que fossem ao mesmo tempo capazes de evocar justiça social e satisfação do interesse individual, fosse pela soma utilitarista de vontades particulares visando à maximização de resultados (em moldes benthanianos, por exemplo) ou pela assunção de uma entidade social superior capaz de ditar a vontade geral (como na visão rousseauniana).

O que subjaz é a análise da relação entre as preferências individuais e o corpo de decisões coletivas, buscando-se verificar em que nível essas formas de escolha (individual e social) dependem uma da outra.

Neste ponto cumpre destacar a diferença instrumental entre os métodos concretos de decisão coletiva e a apreciação teórica acerca do *desempenho lógico* desses mesmos métodos.

Segundo Kotaro Suzumura (2002, p.2), enquanto os métodos de decisão coletiva em si possuem origem tão antiga quanto a sociedade humana, o estudo dos seus componentes lógicos parece possuir origem bem mais recente, podendo ser atribuída às obras de Marie Jean Antoine Nicolas Caritat (o Marquês de Condorcet) e Jean Charles de Borda, pensadores franceses que antecederam à revolução de 1789.

Começar-se-á por apreciar alguns dos argumentos formais e lógicos mais eficazes e contundentes afetos à teoria da escolha social. Para isso, será traçado um panorama das grandes contribuições que teceram a disciplina até o estado atual de conhecimento. O que se segue é uma breve exposição da evolução histórica acerca do tema da escolha social tomando por base a análise lógico-racional dos métodos de escolha coletiva.

Foi no contexto do iluminismo europeu e no ensejo das discussões sobre as implicações do sufrágio e dos direitos humanos na ordem social que Condorcet (1785) apresentou sua contribuição à teoria de tomada de decisão coletiva. Em seu ensaio denominado *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Condorcet verificou uma inconsistência no processo de tomada de decisão quando considerada a eleição entre preferências com um universo de três ou mais alternativas.

O autor notou a potencial ciclicidade que reside em qualquer votação tomada pelo critério da maioria simples, se analisado de forma par a par.

Este ciclo intransitivo<sup>2</sup> foi posteriormente denominado *paradoxo do voto*, ou o *paradoxo de Condorcet*. Em que pese o exemplo original tratar de uma amostragem com sessenta eleitores<sup>3</sup> e três alternativas, o referido paradoxo pode ser verificado em uma matriz tão simples quanto três eleitores com três alternativas de candidatos (A, B, C).

Eleitor 1	Eleitor 2	Eleitor 3
A	B	C
B	C	A
C	A	B

Tabela 1: Exemplo de paradoxo do voto

Como podemos verificar no exemplo acima, adotando-se o método de votação por maioria simples de pares chega-se a um ciclo de preferência social que não permite apontar um vencedor legítimo e democraticamente eleito. Não se pode declarar A vencedor, considerando que, transitivamente, a maioria (dois terços dos eleitores, no caso, eleitores 2 e 3) prefere

<sup>2</sup> A transitividade é conceito que será mais aprofundadamente analisado no próximo tópico referente ao paradoxo de Arrow. Trata-se da conclusão lógica de que se  $x P y$  e  $y P z$ , então  $x P z$ . A transitividade na escolha seja no modelo de escolha racional ou nos modelos de escolha coletiva, é considerada equivalente à racionalidade (e consistência) da decisão.

<sup>3</sup> Condorcet (1785b) aponta em seu *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* um famoso exemplo de uma situação de votação com 60 eleitores em três candidatos, mostrando um ciclo PMR (*pairwise majority rule*) como demonstrado no modelo abaixo:

A	B	B	C	C
B	A	C	A	B
C	C	A	B	A

n1: 23   n3: 2   n4: 17   n5: 10   n6: 8

Aqui, Condorcet observa que temos um “sistema contraditório” que representa o que veio a ser amplamente conhecido como o paradoxo de Condorcet. Em particular, descobrimos que a comparação PMR leva a: AMB (33-27), BMC (42-18) e CMA (35- 25). Há um ciclo na relação PMR nos três candidatos, de modo que nenhum candidato surge como sendo o vencedor em relação a cada um dos candidatos restantes (Gehrlein, 2006, p. 16-17).

C à A; também não se pode declarar vencedor B, pois a maioria (eleitores 1 e 3) prefere A à B; por fim, não se pode declarar vencedor C, porquanto a maioria (eleitores 1 e 2) prefere B à C.

Conforme aponta Gehrlein (2006, p. 17) o impasse resta evidente quando contraposto à clara necessidade, identificada por Condorcet, de se apontar um vencedor (PMRW - *Pairwise Majority Rule Winner*), uma vez que se trata de uma eleição que usa como critério a regra da maioria.

Qual candidato deve ser declarado vencedor? Na matriz da tabela 1 não existe um “vencedor Condorcet”, ou seja, uma alternativa viável não derrotada por qualquer outra em um embate direto. Ficava excluída, assim, a possibilidade de se basear consistentemente a escolha social no método aparentemente democrático de tomada de decisão coletiva por maioria (*majority rule*). O paradoxo apontou sinais inequívocos de que o desempenho lógico de votação e procedimentos relacionados à tomada coletiva de decisões deveria ser objeto de escrutínio teórico mais aprofundado.

A transitividade da escolha coletiva já era considerada um padrão de comportamento racional para decisões em grupo. A falta de transitividade, por sua vez, era vista como uma anomalia<sup>4</sup>.

Condorcet<sup>5</sup> era bastante inflexível em seu argumento e sustentava que a falta de transitividade na preferência para eleitores individuais era situação tão contraditória, que um sistema auxiliar deveria ser usado para eliminar “tais absurdos”. No entanto, mesmo depois de se eliminar a intransitividade das preferências dos eleitores individuais, descobriu-se que

---

<sup>4</sup> Em contraposição a esse entendimento Douglas Rae, a exemplo da abordagem assumida nos trabalhos de Arrow e Sen, afirmou sobre o peso da transitividade na formação de uma teoria de escolhas sociais: “*Uma compreensão do governo da maioria, da democracia, do liberalismo que faz sem o utilitarismo, e que faz mais do que afirmar que direitos estão certos, deve viajar por espaços mais misteriosos, deve subir escadas mais sinuosas, e deve empregar um altímetro mais complexo do que consistência transitiva*” (Rae, 1980).

<sup>5</sup> Condorcet (1788a, pg.156) mais tarde demonstra sua preocupação com a irracionalidade da intransitividade individual da preferência, afirmando: “*Claramente, se o voto de alguém fosse autocontraditório (intransitivo), teria que ser descontado, e devemos, portanto, estabelecer uma forma de votação que faça tais absurdos impossível.*”

a escolha coletiva dos eleitores ainda poderia produzir resultados intransitivos, sugerindo uma *resposta irracional na escolha coletiva de eleitores racionais* (Gehrlein 2006, p. 17).

Para contornar esse problema, Condorcet desenvolveu posteriormente métodos de eleição que definem como vencedor somente o candidato que obtivesse a maioria dos votos em todas as eleições “cara a cara” contra cada um dos outros candidatos. Tais métodos vieram inspirar os sistemas de múltiplos turnos de votação, comuns em ordenamentos jurídicos atuais.

Enfim, considerando a potência formal do argumento, o ponto levantado pelo exemplo de Condorcet é que a transitividade que ocorre naturalmente no ranqueamento de preferências individuais, não apresenta a mesma consistência quando considerada em modelos de escolha social<sup>6</sup>.

Contemporâneo de Condorcet, Jean Charles de Borda (1781) propôs método de votação que não apenas ranqueava, mas atribuía uma pontuação a cada alternativa ranqueada. Ao final, essa pontuação individual era somada e o candidato com a maior soma total era declarado o vencedor. O argumento é que qualquer critério puramente ordinal, tal qual o proposto por Condorcet, poderia resultar em uma falta de consistência interna na forma dos ciclos do método de votação por maioria simples de pares que acabamos de observar, pois esse critério ignora algumas dimensões de preferência.

Suponha que um determinado eleitor tenha a classificação de preferência linear nos candidatos A, B e C. Condorcet só responde pelo fato de que A é preferível em relação a C em uma comparação de maioria simples

---

<sup>6</sup> Essa dificuldade formal, que será abordada mais aprofundadamente na sequência deste trabalho, fundou-se como base comum para os diversos paradoxos que se sucederam na teoria da escolha social (como se pode extrair dos paradoxos de Condorcet, Borda, Arrow e Sen).

por pares entre A e C, ignorando, assim, a “intensidade” de preferência entre A e C, já que B, por exemplo, aparece entre eles no ranking. Para Borda, os ciclos que geram o paradoxo do voto somente surgem porque certas informações no ranking de preferência estão sendo ignoradas (Gehrlein 2006, p. 17).

Com efeito, a solução de Borda, parecia suavizar o problema da transitividade ao introduzir um elemento cardinal que agrega uma dimensão informacional à escolha. Entretanto, o acréscimo de informação eleva a complexidade do modelo matemático, mas não o livra de contradições.

O sistema cardinal proposto por Borda apresentou inconsistências que vieram a ser conhecidas como o Paradoxo de Borda, em que a legitimidade da escolha coletiva ficava comprometida se apreciada em relação à intensidade das preferências declaradas por cada candidato.

Uma explicação mais clara do fenômeno revela-se em casos em que um candidato A, por exemplo, ganha vantagem apenas porque os candidatos B e C (competidores mais fortes) dividiram os votos de intensidade da rejeição entre eles. Podemos compará-los a dois atletas que, tendo se exaurido competindo entre si, são vencidos por um terceiro que é mais fraco que qualquer um deles (Gehrlein 2006, p.7).

A solução proposta por Borda era uma tentativa de suprir a insuficiência da regra da pluralidade, a qual não permitia que os eleitores relatassem seus rankings de preferência completos em todos dos possíveis candidatos. Contudo, ao permitir uma dimensão de intensidade no voto, o sistema mostrou-se vulnerável a manipulações. Conforme afirma Black (1958) o funcionamento de tal método pode ser facilmente frustrado por eleitores que coloquem os adversários mais fortes que seus candidatos favoritos ao pé da sua lista (ou seja, atribuir pontuação propositalmente mais baixa). Isso daria uma grande vantagem aos candidatos intermediários. Manobras como esta fizeram com que Borda posteriormente

admitisse que *seu esquema somente fosse destinado a homens honestos* (Black 1958, p.182).

Como se pode notar, historicamente, no desenvolvimento das teorias da escolha social, onde houvesse um modelo de votação haveria concomitantemente investidas para desviar o método coletivo de escolha em favor de interesses individuais. Quanto mais simples e consistente o método, menos ele espelhava legitimidade. Quanto mais complexo e informado, mais propenso a manipulações.

Este binômio entre a clareza lógico-formal e a maior informação dentro dos modelos de eleição de preferências é assunto ainda hoje frutífero e que será abordado com a devida atenção nos capítulos que se seguem deste trabalho.

A próxima contribuição digna de nota no tocante à evolução da teoria de escolha social e os métodos de votação, já em meados do século XX, é a solução de *single-peaked preferences*<sup>7</sup> proposta por Duncan Black (1948). Black reavalia as contribuições anteriores de autores como o Conde de Borda e o Marquês de Condorcet, além de recuperar alguns trabalhos sobre esquemas de votação desenvolvidos por Charles Lutwidge Dodgson (mais conhecido pelo seu pseudônimo literário, Lewis Carroll). Tratava-se de uma condição simples e suficiente dentre as preferências dos eleitores na qual a regra de maioria determinava um resultado social, uma vez que somente existiria uma alternativa que receberia a maioria simples sobre qualquer outra (vencedor Condorcet), considerando-se que houvesse um número ímpar de eleitores (ou seja, excluindo-se a possibilidade de empate) e que a suposição do pico único fosse satisfeita.

A suposição de Black se vale de uma apresentação geométrica (gráficos de preferência) que permite que as alternativas sociais sejam

---

<sup>7</sup> Em tradução livre pode-se dizer “preferências de pico único”. Pico faz referência ao topo do gráfico geométrico representativo da função de preferências individuais.

representadas por uma variável unidimensional, na qual cada gráfico de indicadores de utilidade dos eleitores reflita um único pico. Teorema de Black<sup>8</sup> é o primeiro a apontar um resultado de possibilidade de preferência coletiva, e, com isso, abriu o caminho para o desenvolvimento moderno da teoria da escolha social. Com isso, as perguntas fundamentais sobre a possibilidade lógica da escolha coletiva estavam lançadas.

No lado da economia do bem-estar a análise crítica e sistemática quanto aos mecanismos econômicos também vinha em evolução constante. O paradigma utilitarista benthaniano dominou a política econômica do século XIX, permeando os trabalhos de John Stuart Mill, Alfred Marshall, Francis Edgeworth e Henry Sidgwick.

Jeremy Bentham foi contemporâneo de Condorcet e Borda, muito embora não compartilhasse o mesmo entusiasmo pela discussão acerca dos direitos humanos. Condorcet era um otimista do iluminismo europeu, preocupado com direitos individuais e a liberdade individual. O pensador francês tinha uma visão holística do conflito de valores humanos. Foi um porta-voz da diversidade do indivíduo e da liberdade de opiniões e *criticou as deficiências das teorias proto-utilitárias da felicidade* (Rothschild 2001, p.196).

Bentham, ao seu turno, era versado em direito e jurisprudência, um austero crítico do conceito de direitos naturais invioláveis<sup>9</sup>. Ao invés de basear suas políticas econômicas nos conceitos de direitos humanos universais, Bentham (1846) preferia o *princípio da maximização da felicidade*, de tal forma que o critério final para se julgar a adequação dos mecanismos econômicos era sua capacidade de trazer maior *felicidade* ao maior

---

<sup>8</sup> Theorem: *If all voters have single-peaked curves as preferences, then the median motion will be adopted by the committee* (Black 1948, p. 27)

<sup>9</sup> De fato, foi em seu comentário áspero sobre a "Declaração dos Direitos do Homem e do Cidadão", incorporada na Constituição francesa de 1791, que ele escreveu a famosa passagem: "*Direitos naturais são um simples absurdo: direitos naturais e imprescritíveis, retórica sem sentido - absurdo sobre pernas de pau*" (Bentham 1843, p. 500).

número de pessoas. De acordo com esta visão utilitarista, a tarefa do legislador devia ser a elaboração de leis e a instauração de instituições sociais e econômicas tal que permitissem que cada pessoa, agindo em busca de seu próprio interesse, pudesse trazer maior felicidade ao maior número de pessoas envolvidas.

Nos anos que se seguiram foram desenvolvidas diversas teorias com base utilitarista, dentre as quais a obra magna de John Stuart Mill, *Utilitarianism* (1863), que, a exemplo da contribuição de Condorcet, lidava com questões complexas como os sentimentos morais e a liberdade de discurso. Mill sugere que o agente moral utilitarista, ao classificar a utilidade de diferentes ações, deve considerar as experiências gerais das pessoas. Isto é, se as pessoas geralmente experimentam maior prazer em determinada conduta, pode-se concluir que agir conforme ela produz mais utilidade do que agir de outra forma que não traga o mesmo resultado. Assim, havia um esboço de avaliação interpessoal no que tange a formação dos valores sociais mais úteis, porquanto os prazeres<sup>10</sup> deviam ser comparados de acordo com o histórico de experiências (individuais e sociais). Não se trata, contudo, de um ranqueamento definitivo de utilidade, mas sim de um quadro relativo de avaliação da ação mais desejada para todos. A proposta de Mill, portanto, pretendia fazer uso da regra utilitarista de maximização para desvelar valores dentro de um cenário de experiências e liberdades individuais. Trata-se de tese que não impõe uma separação abrupta entre o agente moral de sua conduta objetivamente econômica.

---

<sup>10</sup> "If I am asked, what I mean by difference of quality in pleasures, or what makes one pleasure more valuable than another, merely as a pleasure, except its being greater in amount, there is but one possible answer. Of two pleasures, if there be one to which all or almost all who have experience of both give a decided preference, irrespective of any feeling of moral obligation to prefer it, that is the more desirable pleasure. If one of the two is, by those who are competently acquainted with both, placed so far above the other that they prefer it, even though knowing it to be attended with a greater amount of discontent, and would not resign it for any quantity of the other pleasure which their nature is capable of, we are justified in ascribing to the preferred enjoyment a superiority in quality, so far outweighing quantity as to render it, in comparison, of small account" (Stuart-Mill 1863, p.11).

O esquema de utilidade milliano, contudo, não foi o que prevaleceu. Também o antigo modelo de economia de bem-estar, desenvolvido por Arthur Pigou em *The Economics Of Welfare* (1920), que pressupunha a possibilidade da simples soma de utilidade entre diferentes indivíduos visando definir o objetivo social de utilidade máxima, foi duramente criticado por Lionel Robbins (1935). A posição de Robbins foi a que dominou nos anos que se seguiram. Consistia na negação categórica da possibilidade de comparações interpessoais de utilidade, uma vez que, na opinião do autor, tais comparações não poderiam reivindicar validade objetiva<sup>11</sup>.

A avaliação da opinião subjetiva dos agentes e a consequente comparação de utilidade agregada coletivamente mostraram-se inconsistentes demais em face da investida formal dos economistas de sua época.

No final da década de 1930, tornou-se amplamente reconhecido que as fundações da antiga economia do bem-estar estavam inadequadas ao paradigma econômico predominante. As novas fundações para a economia de bem-estar deveriam ser reconstruídas com informações de utilidade interpessoais não comparáveis.

O que se seguiu foi a retomada da tradição ordinalista, que tinha por expoente a obra de Vilfredo Pareto (1906). O *Princípio de Pareto* apontava que a mudança de um estado social para outro era considerada como socialmente benéfica se pelo menos um indivíduo ascendesse em sua posição de bem-estar sem que ninguém ficasse em posição inferior que a que originalmente se encontrava. Esta máxima passou a ser o critério padrão na avaliação da justiça na economia. *A implementação da eficiência de Pareto*

---

<sup>11</sup> A posição de Robbins (1981, p. 5), para ser exato, não rejeitou a possibilidade de fazer comparações interpessoais subjetivas de utilidade. O que ele sustenta é que comparações interpessoais subjetivas não podem reivindicar qualquer validade interpessoal "objetiva".

na alocação de recursos tornou-se o modelo central nesta fase da "nova" economia do bem-estar (Suzumura 2002, p. 7).

Não obstante, se por um lado o critério de Pareto satisfazia com eficiência e simplicidade os requisitos do modelo econômico de sua época (com uma noção ordinalista e sem comparações interpessoais de bem-estar), por outro, no campo da ética, ainda restava uma desconfortável sensação de insatisfação. Isso porque na maioria das políticas de distribuição econômica, não se pode evitar o favorecimento de alguns em detrimento de outros. O princípio paretiano, apesar de consistente, não era aplicável de maneira isolada, e a sua suposta neutralidade acabava por corroborar desigualdades de tratamento.

Na tentativa de suprir esse déficit da economia de bem-estar desenvolveram-se duas abordagens distintas. A primeira foi a introdução de *critérios de compensação*, proposta por autores como Nicholas Kaldor e John Hicks, que propunham a ampliação do princípio paretiano para que comportasse compensações hipotéticas entre os “ganhadores” e “perdedores”, do ponto de vista da política econômica. Tratava-se de uma tentativa de estender os limites da aplicabilidade do modelo de Pareto para situações em que o aumento do bem estar social pudesse abarcar perdas para alguns indivíduos. Este modelo de compensações, contudo, enfrentou fortes críticas em decorrência de suas contradições lógicas e da falta de simetria<sup>12</sup> e transitividade.

A segunda abordagem, desenvolvida por Abram Bergson (1938) e Paul Samuelson (1947), trazia a inédita noção de *função de bem-estar social*. O modelo de Bergson-Samuelson<sup>13</sup> oferecia a fundamentação contundente para justificar que o papel legítimo da economia de bem-estar deveria ser a busca

---

<sup>12</sup> Simetria é requisito lógico decorrente da noção de que  $\forall x, y \in A, se\ x\ R\ y, \text{ então } y\ R\ x$ . A assimetria por sua vez reflete a falta de simetria:  $\forall x, y \in A, se\ x\ R\ y, \text{ então } \neg(y\ R\ x)$ .

<sup>13</sup> Samuelson (1947, p. 220-221), "[it] is a legitimate exercise of economic analysis to examine the consequences of various value judgments, whether or not they are shared by the theorist, just as the study of comparative ethics is itself a science like any other branch of anthropology".

da repercussão lógica dos julgamentos de valores, independentemente das crenças éticas que eles representam. Trata-se, portanto de uma solução puramente formal. O que se propunha pelo conceito de *função de bem-estar social* era a tentativa de separar o que pertencia ao campo da ética e aquilo que era objeto de análise da economia de bem estar, enquanto áreas distintas de conhecimento. Esta nova escola de pensamento foi amplamente *reconhecida como o auge da abordagem ordinalista e científica sobre a economia de bem-estar* (Suzumura, 2002, p.8).

A base formal e lógica fornecida pela função de bem-estar social garantia a atuação consistente do modelo econômico social, sem que se precisasse recorrer ao conteúdo dos julgamentos de valores, pois tal tarefa ficaria reservada a filosofia ou à ciência social e política.

Da mesma maneira em que nas soluções que prevaleceram anteriormente (como no modelo utilitarista e no princípio paretiano), a proposta da função de bem estar-social representa solução formal que prioriza a consistência e eficiência, em detrimento da informação da escolha e da justiça moral e ética.

Como historicamente se pôde constatar, a tentativa de desvinculação entre a economia formal e a ética decorrente da política de distribuição gera um desconforto em termos de legitimação democrática das escolhas coletivas. Tal estado reativa a ebulição de questionamentos filosóficos fundamentais sobre a possibilidade de modelos de distribuição e de justiça que possam suprir o déficit ético e moral deixados pelos métodos tradicionais de economia de bem-estar social.

A questão fundamental, contudo, ainda permanece: dados os rígidos parâmetros formais para a tomada coletiva de decisões, estabelecidos pelas teorias de votação e pela economia de bem estar, seria possível a escolha social que resguarde os requisitos lógicos presentes na escolha individual?

De um modo geral, esta era a atmosfera intelectual em torno da teoria da escolha social quando Kenneth Arrow publicou sua tese de doutoramento *Escolha Social e Valores Individuais*, em 1951, tentando uma solução para a questão supramencionada. Em vista de sua natureza inovadora, bem como da sua influência sobre todo o campo da teoria da escolha social, dedicar-se-á a próxima seção a apreciação dos argumentos chave deste trabalho.

## 1.2 Escolhas, ordens de preferências e valores individuais

Primeiramente, cumpre analisar a possibilidade formal de se obter uma escolha legitimamente social, partindo-se da racionalidade que garanta a mesma consistência presente em ordens de preferências individuais.

A obra “*Social Choice and Individual Values*” de Kenneth Arrow de 1951 elevou qualitativamente o debate sobre a teoria da escolha social. Sem demérito às contribuições de seus precursores, Arrow propôs um método analítico capaz de avaliar não apenas um modelo de votação singular, mas todo processo de escolha (aplicável a qualquer esquema de eleição), por meio de um conjunto unificado de requisitos lógicos.

Ao revisitar o conceito de função de bem-estar social, apresentado no modelo de Bergson-Samuelson, Arrow propôs uma função<sup>14</sup> que fosse capaz de mapear os perfis de ordenação de preferências individuais em um único ordenamento de preferência social. Com efeito, se trata de solução formal e elegante que se destinava, em última análise, a verificar as condições necessárias para se alcançar uma escolha coletiva que fosse dotada de eficiência informacional e, ao mesmo tempo, democraticamente legítima.

Contudo, para melhor entender as implicações dessa premissa deve-se começar por definir “*escolhas*”, como elas podem ser ranqueadas e que

---

<sup>14</sup> Na segunda edição de *Social Choice And Individual Values*, em 1964, Arrow passa a adotar a terminologia *constituição* para referir-se à função de bem-estar social resultante do agregamento de preferências individuais (Arrow 1963, p. 105), para diferenciar do conceito inicialmente desenvolvido por Bergson.

tipo de racionalidade informa as preferências individuais, dado um conjunto particular de alternativas.

Arrow destaca que existem mais de um contexto e inúmeros modos de se fazer escolhas. Nas sociedades capitalistas democráticas, por exemplo, escolhas políticas são normalmente feitas por voto, enquanto que as questões econômicas são resolvidas pelos mecanismos de mercado (Arrow 1963, p.1). Não são incomuns, mesmo nas democracias, decisões motivadas por simples juízos individuais, decisões oligárquicas ou mesmo escolhas tomadas por influência de doutrinas abrangentes (informadas por regras tradicionais), tais como códigos religiosos.

Se as decisões sociais são feitas por um único indivíduo, estamos lidando com uma *ditadura*. Por sua vez, as decisões que seguem doutrinas tradicionais abrangentes indicam algum nível de *convenção* que transcende a esfera individual. Em ambos os casos, entretanto, parece estar presente uma definitividade que não encontramos em métodos como o mecanismo de mercado e a votação, os quais representam maneiras de fundir diversas preferências individuais no processo de eleição.

Uma ditadura ideal pressupõe uma só vontade envolvida no processo de escolha. Em uma sociedade governada por uma convenção, existe uma vontade comum. Em qualquer dos casos, não há conflito de interesses. Não há inconsistências, pois cada indivíduo pode ser considerado racional em relação às suas próprias opções. Na linguagem da economia existe consistência.

Seria, contudo, *possível atribuir tal consistência* (e, portanto, esta mesma racionalidade presente na escolha individual) *a modos coletivos de escolha, onde as vontades de muitos indivíduos estão envolvidas?* (Arrow 1963, p.2)

Arrow está interessado em saber se é formalmente possível construir um procedimento que permita a passagem de um conjunto conhecido de preferências individuais a um padrão de tomada social de escolhas.

A passagem de preferências individuais à escolha social é algo que já havia se mostrado um processo complicado. A consistência tão naturalmente presente na preferência racional individual não se repete ao considerar processos coletivos de escolha. Mesmo no mais simples dos cenários, que considere três alternativas e três eleitores, podem ser notadas inconsistências. Conforme já apontado anteriormente neste trabalho, no seio de qualquer votação tomada pelo critério da maioria simples, analisando-se de forma par a par, reside a potencial ameaça de intransitividade cíclica do denominado *paradoxo de Condorcet*<sup>15</sup>.

O impasse gerado pelo critério de Condorcet já havia apontado a fragilidade no desempenho lógico do modelo de escolha social por maioria simples de pares (*majority rule*), até então sinônimo de legitimidade democrática.

O método cardinal, por sua vez, como a solução proposta por Borda, conferia maior legitimidade para a escolha coletiva, porquanto adicionava um elemento de intensidade de escolha, ao seu turno traduzido em acréscimo na base de informação da preferência. Contudo o modelo provou-se manipulável e ineficiente, pois o incremento da complexidade formal não garantia a consistência lógica.

Do lado da economia de bem-estar, o paradigma paretiano e o modelo econômico predominante estabeleciam critérios lógicos que não poderiam ser ignorados.

A solução de Arrow, entretanto, mostrou-se um passo a frente em muitos desses sentidos, porquanto não era restrita ao critério da maioria,

---

<sup>15</sup> Para relembrar o exemplo de paradoxo de Condorcet, olhar o esquema da Figura 1.

ou mesmo qualquer método de votação tomado particularmente. O autor desenvolveu sua própria abordagem axiomática sobre a teoria da escolha social, o que lhe permitiu analisar todas as possíveis funções de bem-estar social de uma só vez<sup>16</sup>, via conjunto de lemas e teoremas que determinariam as condições necessárias e razoáveis para a eleição coletiva de preferências.

Arrow escolheu provar seu ponto assumindo a racionalidade como uma espécie de maximização, definindo as funções de utilidade em termos de utilidades individuais (de forma ordinal e sem comparações interpessoais), visando, dessa forma, a solução mais eficiente, mesmo que eminentemente formal.

### 1.3 A natureza formal da escolha

O objetivo fixado por Arrow é a verificação, em termos de desempenho lógico, da possibilidade de uma constituição<sup>17</sup> social (função de bem-estar social) capaz de garantir, simultaneamente, legitimidade democrática e eficiência informacional. Para tanto, escolhas individuais devem ser reduzidas a dados quantificáveis (relações de preferência). A comparabilidade interpessoal de utilidade deve ser impedida. Fatores como sentimentos, motivações éticas ou quaisquer outras que pudessem ser percebidas como não racionais em termos econômicos, também devem ser deixados de fora da equação.

Arrow desenvolve uma série de lemas e teoremas de forma a destacar um conjunto particular de axiomas lógico-matemáticos baseados no

---

<sup>16</sup> Com efeito, mesmo na mais simples das sociedades, composta de 2 indivíduos e 3 escolhas possíveis de estados sociais podem existir 6 diferentes ordenamentos de preferências ( $\alpha: x,y,z$   $\beta: x,z,y$   $\gamma: y,x,z$ ,  $\delta: y,z,x$   $\epsilon: z,x,y$  e  $\zeta: z,y,x$ ). Conforme sugere Suzumura (2002, p.11): “A função de bem-estar social funcionaria como um mapeamento definido pelo produto cartesiano  $\Delta \times \Delta$ , onde  $\Delta = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\}$  e toma seus valores em  $\Delta$ . Assim, mesmo na mais simples sociedade concebível, existem  $6^{36}$  funções de bem-estar social no sentido de Arrowiano, o que é um número astronômicamente grande (cerca de  $10^{28}$ )”.

<sup>17</sup> Ver nota 14.

modelo racional de escolhas binárias<sup>18</sup>, que ajudam a descrever as relações entre alternativas e indivíduos. Para um dado par de alternativas os indivíduos podem estabelecer uma relação (R) que pode ser de preferência (P) ou indiferença (I).

O primeiro axioma<sup>19</sup> aponta a conectividade que decorre de que  $\forall x$  e  $y$ , ou  $x R y$  ou  $y R x$  (lê-se “para todo  $x$  e  $y$ , ou  $x$  se relaciona com  $y$  ou  $y$  se relaciona com  $x$ ”, Arrow 1963, p. 13). O segundo axioma<sup>20</sup> afirma que,  $\forall x, y$  e  $z$ , se  $x R y$  e  $y R z$ , então  $x R z$  (lê-se “para todo  $x, y$  e  $z$ , se  $x$  se relaciona com  $y$  e  $y$  se relaciona com  $z$ , então  $x$  se relaciona com  $z$ ”, Arrow 1951, p. 13). Isso representa uma relação *transitiva*. Toda relação que

---

<sup>18</sup> Abaixo as referências axiomáticas comumente utilizadas. Para uma mais profunda compreensão da discussão que se segue aconselho ao leitor algum dos manuais de introdução à lógica matemática como Carnap (1958), Hubert e Ackermann (1960) ou Quine (1951).

- ∃ quantificador existencial ("para algum")
- ∀ quantificador universal ("para todo")
- condicional ("se, então")
- ↔ equivalência ("se, e somente se")
- ~ negação ("não")
- ∨ alternativa (inclusivo; "ou")
- & conjunção ("e")
- = identidade ("o mesmo que")
- ∈ element de ("pertencente a")
- ⊂ subconjunto de ("está contido em")
- ∩ intersecção de ("elementos pertencentes a dois conjuntos simultaneamente")
- ∪ união ("elementos pertencentes a algum dos conjuntos")

Uma variedade de propriedades que podem (ou não) estar presentes em relações binárias ajuda a definir o que o autor buscar avaliar quando propõe sua teoria. Dentre elas podemos apontar as que seguem:

- (1) Reflexividade:  $\forall x \in S: x R x$
- (2) Completude:  $\forall x, y \in S: (x \neq y) \rightarrow (x R y \text{ or } y R x)$
- (3) Transitividade:  $\forall x, y, z \in S: (x R y \ \& \ y R z) \rightarrow x R z$
- (4) Anti-Simetria:  $\forall x, y \in S: (x R y \ \& \ y R x) \rightarrow x = y$
- (5) Assimetria:  $\forall x, y \in S: x R y \rightarrow \sim (y R x)$
- (6) Simetria:  $\forall x, y \in S: x R y \rightarrow y R x$

Para Arrow, como será demonstrado, um “ordenamento” deve ser reflexivo, transitivo e completo (i.e., livre de anti-simetrias).

<sup>19</sup> AXIOM I: For all  $x$  and  $y$ , either  $x R y$  or  $y R x$  (Arrow, 1951, 1963, p. 13). Conectividade.

<sup>20</sup> AXIOM II: For all  $x, y$ , and  $z$ ,  $x R y$  and  $y R z$  imply  $x R z$  (Arrow, 1951, 1963, p. 13). Transitividade.

atenda a esses dois quesitos pode ser considerada um ordenamento (mesmo que fraco<sup>21</sup>).

Ainda quanto à formulação dos axiomas acima, Arrow trabalha algumas assertivas complementares que expressam conclusões lógicas autoevidentes, mas que esclarecem propriedades importantes de reflexividade, simetria e transitividade, considerando a relação (tanto de preferência como de indiferença) entre as alternativas  $x$  e  $y$ <sup>22</sup>:

- (a) *Para todo  $x$ ,  $x R x$*
- (b) *Se  $x P y$ , então  $x R y$*
- (c) *Se  $x P y$  e  $y P z$ , então  $x P z$*
- (d) *Se  $x I y$  e  $y I z$ , então  $x I z$*
- (e) *Para todo  $x$  e  $y$ , ou  $x R y$  ou  $y P x$*
- (f) *Se  $x P y$  e  $y R z$ , então  $x P z$*

*Tabela 2: Lema 1(Arrow, 1951,1963, p. 14)*

No que tange ao universo de escolhas possíveis dentro de cada arranjo de alternativas, Arrow destaca a função de escolha  $C(S)$ . Trata-se de um subconjunto de  $S$  (conjunto de todas alternativas) que contém a relação funcional decorrente das escolhas tomadas para cada situação possível. Assim, se há um elemento  $x$  em  $C(S)$ , pode-se concluir que ele foi o escolhido dentre as alternativas em  $S$ . No caso, se há  $y$  em  $S$  e  $x$  pertence a  $C(S)$ , então fica subentendido que houve uma relação de preferência ( $x R$

---

<sup>21</sup> Um ordenamento “fraco” se refere ao fato de não excluir a relação de indiferença. Seria o equivalente a uma preferência “maior ou igual” em relação às demais alternativas. Em um ordenamento “forte”, não há possibilidade de “empate”, assim, a relação de indiferença fica excluída.

<sup>22</sup> O lema de Arrow, como se pode notar, deixa clara opção do autor por um modelo de ordenamento “fraco” (que permite a relação de indiferença), que cumpra com as propriedades da conectividade, transitividade e que seja completo (uma espécie de simetria que exclui a possibilidade de identidade entre as alternativas). Em outras palavras, Arrow trata a relação binária de escolha de tal sorte que, caso declarada uma relação de preferência por uma das alternativas ( $x P y$ ), não é mais possível uma relação subjacente da alternativa não preferida ( $\sim y R x$ ). Simplificando ainda mais, se  $x$  for preferível à  $y$ , significa dizer que  $y$  não pode ser nem igual, nem indiferente e, evidentemente, nem melhor que  $x$ . Isso fica claro na formulação (e) do lema acima referido, no qual  $\forall x$  e  $y$ , ou  $x R y$  ou  $y P x$ . Portanto, as alternativas são mutualmente excludentes, o que torna o processo de decisão mais próximo da ideia geral dos modelos de escolha racional, nos quais as alternativas funcionam como vetores.

$y$ ) e, portanto, houve um escolha (ordenamento de preferências). Da mesma forma, nenhum  $z$  pode existir tal que  $z P x$ , pois isso representaria uma contradição intransitiva. Arrow define que  $C(S)$  é o conjunto de todas alternativas  $x$  em  $S$  tal que, para cada  $y$  em  $S$ ,  $x R y$  (Arrow 1951, p. 15).

Dadas estas informações, pode-se definir  $C([x, y])$  como conjunto composto por um universo de duas alternativas,  $x$  e  $y$ . Se  $x P y$ , então  $x R y$ ,  $x R x$  e  $\neg y R x$  (conforme visto acima no Lema 1). Assim, pode-se verificar que se  $x P y$  e  $x$  está em  $C([x, y])$ , então  $y$  não pode estar em  $C([x, y])$ , e  $x$  é o elemento unitário do conjunto  $C$ . Como colorário resta o segundo lema de Arrow: é condição necessária e suficiente que  $x P y$  que  $x$  seja o único elemento de  $C([x, y])$  (Arrow 1951, p. 16).

Pode-se verificar, dentre os pressupostos básicos da escolha racional (como preferência individual), que a escolha em qualquer ambiente pode ser determinada pelo conhecimento de preferências em situações de apenas dois elementos. A partir disso, Arrow representa seus mecanismos de escolha por relações ordenamento (ranqueamento), algo que, em seu ponto de vista, possui vantagens sobre as representações convencionais em termos de mapas de indiferença e funções de utilidade.

#### **1.4 Condições para a função de bem-estar social**

O ranqueamento relativo de um par fixo de possíveis estados sociais deve variar, de um modo em geral, com as mudanças nos valores dos indivíduos (Arrow 1951, p.22). Se a classificação social de valores não se altera mesmo em face da alteração das disposições de valores individuais, há boa chance de se estar diante de uma ditadura ou de uma noção de bem comum fundamentado em doutrinas abrangentes.

A posição de Arrow pretende desviar do ideal de bem comum pré-determinado, mas, concomitantemente, procura uma solução que não seja puramente utilitarista ou hedonista:

“Para a noção nominalista do período moderno, a suposição da existência do ideal social em algum reino platônico era desprovida de significado. A filosofia utilitarista de Jeremy Bentham e seus seguidores buscava, em vez disso, fundamentar o bem social somente no bem dos indivíduos. A psicologia hedonista associada à filosofia utilitária foi usada ainda para sugerir que o bem de cada indivíduo era idêntico aos seus desejos (Arrow 1951, p.22)”.

O autor busca reformular o conceito de função social de bem-estar inicialmente desenvolvido por Bergson, agregando a possibilidade de um resultado socialmente válido e relevante que não dependa plenamente do conjunto prévio de valores individuais (como no modelo utilitarista benthaniano), nem seja completamente independente das preferências individuais (como na situação ideal metafísica que designa sempre o mesmo ordenamento social apesar das alterações de valores individuais).

Na definição de Arrow, a função de bem-estar social devia ser entendida como resultante de um processo no qual cada conjunto de ordenamentos individuais,  $R_1, \dots, R_n$  (um ordenamento para cada indivíduo), leva a um ordenamento social correspondente de estados sociais alternativos,  $R$  (Arrow 1951, p. 23).

A função resultante, portanto, deve resguardar a racionalidade encontrada na função de escolha individual, mas não necessariamente precisa ser um mapeamento completo de todas as escolhas. Com isso em mente, Arrow impõe uma série de condições lógicas que se provaram verdadeiras quando aplicadas às preferências individuais e que encapsulam a noção de uma função de bem-estar social capaz de gerar um verdadeiro ordenamento social (sem refletir um mero somatório de escolhas individuais). A descrição que segue é uma explicação detalhada sobre essas condições propostas por Arrow que podem ser encontradas

originalmente em sua obra *Social Choice and Individual Values*, publicada originalmente em 1951.

**Condição 1: a regra de *Universalidade*, ou do *Domínio Irrestrito***

A primeira condição é que cada conjunto lógico possível de ordens individuais de um determinado conjunto  $S$  pode ser obtido a partir de um conjunto admissível de ordenamentos individuais de todas as alternativas. Em outras palavras, para qualquer conjunto de preferências individuais, a função de bem-estar social deve render uma classificação única e completa de escolhas sociais.

**Condição 2: *Monotonicidade ou Associação Positiva (Pareto)*.**

A segunda exigência é a associação positiva de valores sociais e individuais. Se qualquer indivíduo altera a ordem de preferência, elevando uma alternativa específica em importância, consequentemente, a função social de preferência deve ou promover o mesmo valor social ou mantê-lo inalterado, mas nunca pode ser classificada como inferior ao inicialmente estabelecido.

**Condição 3: *Independência das Alternativas Irrelevantes (IIA)*.**

Isso significa que, ao se considerar uma função  $C(S)$  como sendo a escolha social válida, então, assim como é o caso para um único indivíduo, a escolha feita a partir de qualquer ambiente fixo  $S$  deve ser independente da existência de alternativas fora de  $S$ . Também conhecido como *independência por pares (pairwise independence)* esta condição afirma que mudanças em rankings individuais de alternativas fora de um determinado subconjunto (alternativas irrelevantes) não devem exercer nenhum impacto sobre a função final de escolha social. Isso significa, em termos práticos, que, em uma eleição com três candidatos, se um deles morre (após a votação, é claro), as relações finais de preferência entre os

candidatos restantes tem de permanecer inalteradas. Caso contrário, estaríamos assumindo que o resultado da eleição depende de circunstâncias acidentais (como seria o caso de um candidato morrendo antes da votação).

#### **Condição 4: a *Soberania Cidadã*.**

Esta condição é a garantia de uma sociedade "livre para escolher", onde não há nenhuma função de bem-estar social imposta. Esta é muitas vezes referida como a regra de *não imposição*. Caso houvesse imposição, a sociedade nunca poderia expressar preferências sobre alguns dos pares de alternativas, porquanto a função social já teria sido dada.

#### **Condição 5: *Não Ditadura*.**

Isto significa que a escolha social não pode ser baseada unicamente na preferência de um indivíduo. Neste sentido, a função de bem-estar social tem que refletir de alguma maneira um fundir de vontades, e nunca pode decorrer da imposição de um único conjunto individual de opções sobre todos os demais.

Somente se atendidas às cinco condições<sup>23</sup>, existiria a possibilidade de uma fórmula geral na qual os cidadãos, considerado o universo ilimitado de preferências, possam determinar uma função de bem-estar realmente social. Atender a todas essas exigências tornaria possível a construção de um ordenamento social resultante de todos os estados sociais alternativos possíveis, partindo-se de um determinado conjunto de ranqueamentos individuais desses estados sociais, tudo em conformidade

---

<sup>23</sup> Na segunda edição de *Social Choice and Individual Values*, em seu capítulo VII, intitulado *notes of the Theory of Social Choice*, Arrow (1963, p.97) propõe uma simplificação das condições acima elencadas, introduzindo uma versão atenuada do princípio de Pareto em substituição às condições 2 (associação positiva) e 4 (soberania cidadã). A nova lista de quatro condições (U, I, P e D) foi amplamente adotada e serviu de base para as críticas futuras dos demais teóricos da escolha social. Contudo, tendo em vista uma melhor avaliação da ideia original do autor sobre o teorema da possibilidade, buscou-se manter a terminologia da obra de 1951.

com a soberania dos cidadãos e da racionalidade econômica (Arrow, 1951, p. 31).

Seria realmente possível um método que satisfaça todas essas condições? Arrow parece ter deixado este problema em aberto.

### 1.5 O paradoxo de Arrow

No capítulo V de sua obra, Arrow mostra-se determinado a provar o que chamou o *teorema geral de possibilidade*. Começa por afirmar que se há apenas duas alternativas a serem classificadas, é possível construir uma função de bem-estar social que satisfaça os axiomas I e II (conectividade e transitividade).

Teorema 1 ( Teorema da Possibilidade de duas alternativas ) : Se o número total de alternativas é dois, o método de decisão da maioria é uma função de bem-estar social que satisfaz Condições 2-5 e produz um ordenamento social das duas alternativas para cada conjunto de ordenação individual (Arrow 1963, p.48).

Isso significa dizer que em um ambiente com apenas duas alternativas a regra da maioria por pares (*majority rule*) satisfaz as condições lógicas de 2 a 5 para a constituição de uma função de bem-estar social. Neste sentido a descoberta de Arrow serve, por exemplo, como fundamentação lógica para o sistema Anglo-Americano de eleições, porquanto este consiste em um modelo bipartidário de alternativas.

O primeiro teorema comprova a conformidade da função social de duas alternativas a quatro condições lógicas: associação positiva, independência das alternativas irrelevantes, não ditadura e não imposição. Há, evidentemente, uma clara violação da primeira condição, a regra domínio irrestrito, uma vez que o número de alternativas está limitado a duas alternativas.

Em termos econômicos, isso significa que se tomadas apenas duas alternativas por vez, então se pode manter a consistência. Esta é a essência do terceiro lema: *para qualquer espaço de alternativas, o método de decisão da maioria é uma função de bem-estar social que satisfaz as condições de 2-5* (Arrow 1951, p. 48).

Como corolário, contudo, o teorema reforça que, mesmo no mais simples dos cenários, as cinco condições são incompatíveis se consideradas mais de duas alternativas. O paradoxo de Condorcet com três pessoas e três alternativas já havia demonstrado que o método de decisão da maioria não satisfaz a condição do domínio irrestrito sem gerar um ciclo intransitivo.

Arrow batizou sua conclusão de Teorema Geral da Possibilidade, um nome altivo para um resultado não muito otimista. Tal formulação ficou posteriormente mais conhecida como o Teorema da Impossibilidade, ou o Paradoxo de Arrow:

Teorema 2 (Teorema Geral da Possibilidade): Se há pelo menos três alternativas entre as quais os membros da sociedade são livres para escolher de qualquer maneira, então qualquer função de bem-estar social capaz de satisfazer as Condições 2 e 3 e produzir um ordenamento social que satisfaça os Axiomas I e II somente pode ser imposta ou ditatorial. (Arrow 1951, p.59).

A conclusão de Arrow é, à primeira vista, desoladora, pois traz consigo não só uma nova base axiológica para verificação lógica de preferências, mas a comprovação consistente da impossibilidade de escolhas sociais. Isso implica dizer que não existe um método de eleição que possa solucionar o paradoxo do voto, seja por votação ou qualquer regime de representação proporcional, não importa o quão complexo e detalhado, sem que se recorra em alguma medida a uma imposição autoritária.

Conforme admite o próprio Arrow, as três primeiras condições (*domínio irrestrito, associação positiva e independência de alternativas irrelevantes*), se tomadas em conjunto, servem para excluir as comparações interpessoais de utilidade. Sabendo disso, pode-se reinterpretar o teorema 2 para concluir que:

Se excluirmos a possibilidade de comparações interpessoais de utilidade, então os únicos métodos de passar de uma escolha individual para preferências sociais que serão satisfatórios e definidos por uma grande variedade de conjuntos de ordenações individuais serão ou impostos ou ditatoriais (Arrow 1951, p.59).

*Contrario sensu*, pode-se dizer que excluído o caráter intersubjetivo potencial de cada escolha individual, também se exclui a possibilidade da escolha genuinamente social que não seja alcançada por meio imposto ou ditatorial.

A abordagem formal de Arrow, entretanto, deixa claro que as comparações interpessoais devem ser excluídas, mesmo que somente para garantir a simplicidade e integridade do modelo de escolha social baseado em suas cinco condições lógicas essenciais. Trata-se, de certa forma, de homenagem à tradição da economia de bem-estar que o precedeu.

O teorema da impossibilidade, assim, estabeleceu-se como novo paradigma do escrutínio sistemático direcionado à solução racional do problema da escolha social.

Percebe-se, contudo, que o que realmente estava em jogo era mais do que um resultado matemático de um pressuposto lógico. Arrow optou por uma abordagem formal, uma solução racional conforme o paradigma da economia de bem-estar de sua época, e, ao concluir pela impossibilidade da escolha social, trouxe de volta à tona a discussão político-social sobre questões éticas.

Se por um lado essa nova abordagem não deixava espaço para contradição lógica, por outro, o resultado de impossibilidade em si mostrou-se inaceitável, especialmente para aqueles que ainda mantinham convicção na noção filosófica de justiça e distribuição social de bens e valores.

Arrow estava ciente das implicações decorrentes da rigidez formal de seu modelo. Como tentativa de composição dedica considerável porção da segunda parte de seu trabalho à avaliação das possíveis flexibilizações das condições inicialmente impostas, visando a verificação de alternativas ao impasse gerado por seu teorema:

Se quisermos fazer julgamentos de bem-estar social que dependam dos valores de todos os indivíduos, ou seja, que não são impostos ou ditatoriais, então devemos relaxar algumas das condições impostas. Continuará a ser sustentado que não há comparação interpessoal de utilidades significativa e que as condições em volta da palavra "satisfatório" devem ser aceitas. (Arrow 1951, p 60)

O autor busca possíveis relaxamentos das suas condições iniciais e esboça algumas alternativas para uma escolha social que não fosse puramente formal e nem vise apenas resultados matemáticos. Ele aborda escolhas de grupos partindo do ponto de vista da teoria dos jogos e aponta alguns problemas que surgem ao se considerar a ética distributiva ligada ao individualismo. Nos capítulos finais de seu livro, Arrow conclui em mais de uma ocasião que a *“hipótese individualista é insuficiente para produzir uma função de bem-estar social satisfatória”* (Arrow, 1951 p. 70), e que *“a distribuição real do bem-estar não se pode afirmar simplesmente em termos de dinheiro”* (Arrow 1951, p. 72).

Com efeito, o teorema da impossibilidade (assim como o paradoxo de Condorcet, antes o fizera) sugeriu muitas direções para pesquisa, dentre

elas: a melhor caracterização matemática de estruturas de votação, a ampliação dos insumos informacionais das escolhas sociais, a restrição da escolha a certos domínios, as compensações parciais de algumas das condições lógicas e a reavaliação da relação entre o equilíbrio competitivo e seu núcleo ideológico.

Nos anos que se seguiram, pensadores como Amartya Sen e John Rawls também aproveitaram a estrutura formal do pensamento de Arrow para desenvolverem suas próprias teorias que representavam, às suas maneiras, alternativas ao paradoxo.

A essa altura era inevitável a vinculação do assunto da escolha social às teorias clássicas e modernas da justiça, em particular no que tange ao longo conflito de ideias entre direitos universais e utilitarismo. Restou demonstrado com isso, o poder informativo da análise formal de escolha social como maneira de destacar os argumentos filosóficos mais básicos sobre a vida social boa.

### **1.6 Justiça social e alternativas ao paradoxo da escolha**

Diante do cenário imposto pelo paradoxo, restam duas principais alternativas: a) afastar qualquer possibilidade de uma racionalidade coletiva, tratando a escolha social como um mero somatório par a par de vontades individuais que determinam, por regra de maioria, a vontade de todos; ou b) pensar outros modos mais complexos de racionalidade que sejam capazes de considerar a agregação de preferências individuais de maneira a produzir uma legítima função de bem-estar social.

Economistas, juristas e filósofos que ainda pensavam em discutir temas tais como economia social e teorias da justiça escolheram a segunda opção. Não é exagero afirmar que foi o próprio Arrow quem, direta ou indiretamente, conduziu a discussão para esse desfecho.

Ao analisar o papel de Arrow na grande discussão sobre a possibilidade da Economia de Bem-Estar Social que se seguiria nos próximos anos (ou seja, a segunda metade do século 20) cumpre esclarecer alguns pontos.

O autor deixa clara sua opção metodológica pelo caminho formal (Arrow 1951, p. 6). Deve-se observar, contudo, que ele identifica claramente cada valor que está deixando de fora da equação<sup>24</sup>. Por exemplo, a comparação interpessoal de utilidade, o valor intrínseco do próprio processo de votação e até mesmo a ética distributiva são temas que Arrow deixa propositalmente de fora de suas formulações lógico-matemáticas. Somente após a declaração do segundo teorema (possibilidade geral) ele retorna para enfrentar estes tópicos, em uma abordagem muito menos rigorosa.

O legado do paradoxo, portanto, é mais do que simplesmente provar a impossibilidade matemática de decisões sociais que sejam racionais em sentido econômico. De um modo indireto, o teorema chama a atenção para a impossibilidade de uma função de bem-estar social com base na racionalidade formal e individualista que prevaleceu nas ciências econômicas até aquele ponto. Trata-se de uma crítica ao modelo economicista de prever o comportamento social, e, além disso, uma exposição da insuficiência de uma forma puramente matemática de fazer economia.

### **1.6.1 (Re)Introduzindo comparação interpessoal de utilidade**

Pode parecer contraditório que se queira reintroduzir a comparação interpessoal de utilidade no modelo de escolha social, depois de todo o esforço histórico em sentido contrário. Contudo, trata-se de um legítimo meio de desviar do paradoxo de Arrow.

---

<sup>24</sup> *It has been stated above that the present study confines itself to the formal aspects of collective social choice. The aspects not discussed may be conveniently described as the game aspects, especially since that term has acquired a double meaning. In the first place, no consideration is given to the enjoyment of the decision process as a form of play. There is no need to stress the obvious importance of the desire to play and win the game as such in both economic behavior and political." That such considerations are real and should be considered in determining the mechanics of social choice is not to be doubted but this is beyond the scope of the present study (Arrow 1963 p. 6-7).*

Com efeito, as razões para a exclusão de tais informações nos modelos de escolha permeia o binômio já levantado anteriormente que se estabelece entre a complexidade do esquema de voto e a sua eficiência em determinar resultados racionais e consistentes<sup>25</sup>.

A escolha pela abordagem mais consistente foi o que levou a Robbins (1936) a desconsiderar a comparação interpessoal de utilidade, porquanto tratava tal informação como desprovida de significado objetivo. A exclusão categórica proposta por Robbins, por sua vez, fundamentava-se na distinção entre comparações interpessoais descritivas e julgamentos de valor puramente normativos<sup>26</sup>.

Tal posição positivista foi a que veio ditar o parâmetro da economia de bem-estar, e não via problemas em descartar informações inconsistentes, desde que a integridade formal do procedimento permanecesse intacta.

De uma maneira inusitada, entretanto, a impossibilidade apontada pelo paradoxo arrowiano acabou por expor a fragilidade da abordagem puramente formal de se pensar a economia de bem-estar. Sem dúvida o modelo de Arrow reveste-se de argumentos lógicos rigorosos, todavia, a estratégia de utilizar uma base utilitarista e ordinalista, pode ter esvaziado demais o conteúdo da teoria da escolha social.

A posição de Arrow não admite tais comparações interpessoais. Contudo, não permitir não é o mesmo que negar a existência<sup>27</sup>. É incontestável, mesmo para o mais ávido positivista, que no processo social exista algum grau de comparação interpessoal de utilidade.

---

<sup>25</sup> Como vimos no modelo cardinal de Borda, um sistema complexo de eleição que permita cálculo comparativo de utilidade interpessoal fica mais suscetível a inconsistências lógico-racionais.

<sup>26</sup> Quanto ao assunto da dicotomia fato-valor ver Ferreira (2016).

<sup>27</sup> Mesmo Arrow reconhece que considerações (como o valor do processo de escolha em si e a comparação interpessoal) sejam reais e que devam ser consideradas na determinação da mecânica da escolha social, mas admite que isso está além do escopo do trabalho desenvolvido em *Social Choice and Individual Values*. (Arrow 1951 p. 6-7).

Sen (1970a) afirma que o formato da função de bem-estar social proposto por Arrow pode muito bem ser expandido para que admita as comparações interpessoais (conforme o que chamou de *funcionalidades de bem-estar*). Tal expansão certamente vem acompanhada da dificuldade de se criar um procedimento formal que possa abarcar essa informação de uma maneira satisfatória. Conforme Sen:

É certo que é difícil conceber um procedimento que tome nota das comparações interpessoais (ou da cardinalidade) em procedimentos de escolha social institucionalizados, mas isso é uma característica do mundo, não da teoria da escolha social como tal. (Sen, em Elster 1986, p. 220).

De fato, a complexidade das relações sociais às quais somos submetidos na prática excede em muito o nível de descrição possível no modelo da teoria da escolha social.

Entretanto, dentre as diversas justificativas para a criação de estruturas de bem-estar social informadas com comparações interpessoais de utilidade, nenhuma deve ser tomada como solução definitiva. Sen (1982) adverte para o fato de que mesmo com o aumento de informação da escolha não há garantia de que princípios como igualdade, equidade e justiça serão devidamente contemplados. Em outras palavras, pode-se dizer que o aumento da base informacional não necessariamente se traduz em mais legitimidade de escolha.

Dessa forma, a relevância moral da comparação interpessoal depende fortemente do critério de maximização a ser adotado. No caso do modelo utilitarista estaríamos falando de soma de bem-estar e, conseqüentemente, a comparação seria feita em termos de diferenças de bem-estar individuais. No modelo ralswiano de maximização segundo o princípio da diferença, por exemplo, o paradigma passa a ser o bem-estar do pior

indivíduo, e a comparação implicaria a noção de *níveis de bem estar* (Sen 1982, p.270).

O critério de maximização define o vetor que direciona a interpretação das novas informações acerca da escolha individual. Trata-se de alternativa ao impasse gerado pelo teorema de Arrow, na qual um observador ético forma suas próprias comparações interpessoais de utilidades e faz uso desta base informativa ampliada para definir uma função de bem-estar social.

Sem o critério interpretativo, contudo, corre-se o risco de não restarem correspondências (intersubjetivas) entre as comparações interpessoais. A questão da *arbitrariedade das unidades de utilidade é reflexo direto do problema da comparação interpessoal* (Sen 1970 p.92). Se existirem vários observadores éticos que formam suas respectivas comparações interpessoais subjetivas de utilidade, *seus julgamentos de bem-estar social podem conflitar de tal modo que algumas variantes dos teoremas de impossibilidade possam muito bem voltar fortalecidas* (Suzumura 2002, p.20). Tal situação já havia sido prevista por Arrow:

As pessoas parecem preparadas para fazer comparações da forma: Estado x é melhor (ou pior) para mim do que o estado y é para você... Comparações interpessoais do tipo “simpatia estendida” podem ser postas em forma operacional; o julgamento assume a forma: é melhor (no meu julgamento) ser eu mesmo no estado x do que ser você no estado y (Arrow 1963, p. 114-115).

Nestes casos as características que definem o indivíduo (como sua riqueza, suas habilidades e sua posição social) se confundem na comparação interpessoal.

Isso reforça que a teoria de escolha social de Arrow, a exemplo dos modelos que o antecederam, está enraizada na concepção filosófica que podemos chamar de consequencialismo de bem-estar, na medida em que

se baseiam na avaliação dos estados de coisas em termos de utilidades individuais obtidas desses mesmos estados de coisas.

### 1.6.2 O Paradoxo Liberal

Sen (1970) notou que o requisito paretiano, fundamental na abordagem de bem-estar da teoria da escolha social, é contraditório com o princípio da liberdade individual.

O paradigma liberal parece exigir tanto escolhas individualmente relevantes como liberdade para escolher o que se prefere. O princípio paretiano, por sua vez, exige que: se todos os indivíduos preferem  $x$  a  $y$ , então a sociedade também deve preferir  $x$  a  $y$ . Atendendo a esses dois requisitos é melhor que se permita ao indivíduo fazer o que ele quer, contando que se mantenha inalterado o resultado para os demais.

Sen propôs a definição de liberalismo através da condição L: *Para cada pessoa "i" existe ao menos um par de alternativas (x,y) tal que "i" seja decisivo na escolha social entres as alternativas, seja em que ordem for, isto é,  $x P_i y \rightarrow x P y$ , e  $y P_i x \rightarrow y P x$*  (Sen 1970, p. 87).

Em uma versão atenuada pode-se reduzir a abrangência do princípio para que não retrate todas as pessoas, mas ao menos o embate entre as preferências de dois indivíduos e suas respectivas influências na escolha social resultante. Sen chamou essa condição de  $L^*$  (liberalismo mínimo)<sup>28</sup>.

Contudo, mesmo considerando essa condição mais branda, o princípio liberal  $L^*$  é inconsistente com as condições U (domínio irrestrito) e P (Princípio de Pareto Atenuado) quando imposta a uma *função de decisão*

---

<sup>28</sup> Condição  $L^*$  (liberalismo mínimo) - Existem pelo menos duas pessoas  $k$  e  $j$  e dois pares de alternativas distintas  $(x, y)$  e  $(z, w)$  tais que  $k$  e  $j$  são decisivos sobre  $(x, y)$  e  $(z, w)$ , respectivamente, cada par tomado em qualquer ordem.

social (SDF), como mostra Sen<sup>29</sup> em seu Teorema 6\*.1: *Não há função de decisão social que satisfaça as condições U, P e L\** ( Sen 1970, p. 87).

Sen (1970, p. 80) oferece uma ilustração para esclarecer a natureza do paradoxo liberal. Considerando o cenário com duas pessoas (Senhores A e B) e três alternativas, sendo: a) Senhor A ler uma cópia do livro *O Amante de Lady Chatterly*; b) Senhor B ler o referido livro; ou c) nem um dos dois senhores ler o livro. O Senhor A, o puritano, prefere que ninguém leia o livro, seguido da alternativa que ele próprio leia e por último a opção em que o influenciável Senhor B seja exposto ao conteúdo do livro. Portanto temos que ele prefere (c, a, b). Já o Senhor B, o lascivo, prefere que qualquer um deles deva lê-lo em vez de nenhum, mas ainda prefere que o Senhor A leia em vez de ele próprio, porque gostaria que o Senhor A fosse exposto à obra provocante de D. H. Lawrence. Consequentemente, o Senhor B prefere (a, b, c).

Segundo Sen, argumento liberal pode ser feito para o caso que, dada a escolha entre o Senhor A ler o livro e ninguém lê-lo, sua própria preferência deveria ser refletida por preferência social. Então essa sociedade deveria preferir que ninguém leia o livro, ao invés de ter o Senhor A de ler o que ele claramente considera como um terrível livro. Por isso, a alternativa (c) seria socialmente preferível a (a).

Da mesma forma, existe um argumento liberal em favor de refletir a preferência do Senhor B na escolha social entre o Senhor B ler o livro e ninguém o ler. Assim, (b) seria preferível a (c). A sociedade, portanto, deveria preferir o Senhor B lendo o livro a ninguém o lendo, e ninguém lendo o livro ao Senhor A o lendo. No entanto, a alternativa de o Senhor B ler o livro é Pareto-inferior que a alternativa de o Senhor A lê-lo mesmo em

---

<sup>29</sup> A prova é a que segue: assumindo que  $x P_k y$ ,  $z P_i w$  e  $\forall i: (w P_i x \& y P_i z)$ . Pela condição  $L^*$ ,  $x P y$  e  $z P w$ . Pela condição  $P$ ,  $w P x$  e  $y P z$ . Isso viola a aciclicidade. Portanto, não há SDF que satisfaça as condições  $L^*$  e  $P$ , dada a condição  $U$ .

termos de um critério de Pareto atenuado (*weak Pareto*), e se a preferência social honra essa classificação, então é preferível (a) a (b).

Ao final, cada alternativa pode ser vista como pior que outra, e, desta forma, não existe melhor alternativa neste cenário, bem como não há escolha considerada satisfatória.

O paradoxo de Sen levanta a suspeita de que a solução paretiana, tão aclamada pelos economistas, não guarda compatibilidade com o princípio liberal quando considerado um domínio irrestrito.

Em termos de escolhas, as repercussões são consideráveis, porquanto o resultado de Sen implica reconhecer que não é permitido a mais de um indivíduo por vez a liberdade para escolher conforme aquilo que deseja isoladamente da preferência dos demais indivíduos da comunidade.

A garantia definitiva para a liberdade individual não deveria, portanto, ser procurada em regras de escolha social. Seria mais coerente empenhar-se no desenvolvimento de valores individuais que respeitem as escolhas individuais dos outros. Como adverte Sen (1982, p.289), o conflito apresentado pelo paradoxo liberal se aplica às sociedades que não possuem mecanismos de reciprocidade e nas quais a escolha par a par, baseada em valores liberais, pode conflitar com as escolhas fundadas no princípio paretiano.

Pode-se notar que não há conflito entre o princípio de Pareto e a condição de liberalismo mínimo se considerados qualquer par de alternativas separadamente. O conflito surge quando reunimos mais de um par de alternativas. Neste aspecto pode-se comparar o resultado de impossibilidade alcançado por Sen ao paradoxo levantado por Arrow, pois em ambos o requisito do domínio irrestrito é definitivo.

Cumprir destacar, ainda, que, ao contrário do teorema de Arrow, a condição da independência de alternativas irrelevantes (IIA) não foi exigida e tampouco foi requerida preferência social transitiva. Trata-se de

mera apreciação de inconsistência por ciclicidade intransitiva, em moldes semelhantes ao paradoxo do voto de Condorcet. O teorema é perturbador, especialmente por manter o corolário de impossibilidade mesmo em cenários atenuados (como o mínimo liberal).

O paradoxo serve como uma dura crítica ao requisito paretiano que, ao fim, não se presta como garantia de um resultado social que reflita as vontades individuais, mas concentra sua eficácia na maximização da vontade do maior número de pessoas:

Se alguém leva a sério o princípio de Pareto, como os economistas parecem fazer, então se tem que enfrentar problemas de consistência em estimar os valores liberais, mesmo os mais brandos. Colocando de outra maneira, se alguém tem certos valores liberais deve considerar esquivar-se da otimização paretiana. Muito embora o critério de Pareto tenha sido enaltecido como uma expressão da liberdade individual, parece que nas escolhas que envolvem mais de duas alternativas podem haver consequências que são, de fato, profundamente iliberais (Sen, 1982 p. 290).

A análise de Sen remonta ao problema que John Stuart Mill (1861) teve de enfrentar em sua simultânea crença no resultado de moralidade com bases utilitaristas, por um lado, e na inexorabilidade dos direitos liberais, por outro.

Com efeito, dada a notável difusão do consequencialismo de bem-estar, muitos economistas se propuseram a encontrar meios de desviar do teorema de Sen, examinando criticamente a articulação original de liberdade individual e as implicações sobre a impossibilidade de um critério paretiano liberal.

Entretanto, uma solução capaz de evitar a moralidade puramente utilitarista e, ao mesmo tempo, garantir a relação entre valores individuais de liberdade e o resultado social desejado, parece estar mais próxima de

um argumento político filosófico do que de uma descrição econômica com inspiração na escolha individual e a regra de mercado.

### 1.6.3 Justiça, Bem-Estar e Bens Primários

Destacam-se três questões distintas na análise teórica da escolha social relacionada ao bem-estar e direitos fundamentais: a questão da articulação analítica dos direitos, a questão da realização dos direitos, e a questão da concessão inicial dos direitos. (Suzumura 2002 p. 24).

Para tratar do assunto tomar-se-á como exemplo a obra de John Rawls, porquanto responde ao problema da escolha coletiva enfrentando cada uma das questões acima mencionadas. *Uma Teoria da Justiça* (Rawls, 1971) foi, muito provavelmente, a obra mais influente sobre a filosofia política da segunda metade do século 20. O autor propôs um procedimento ideal<sup>30</sup> para a seleção dos princípios de justiça. Pode-se argumentar, contudo, que quando se busca a seleção social de princípios de distribuição de bens e valores, mesmo que relativamente a uma ordenação hipotética, estamos discutindo uma função de bem-estar social. No caso da *posição original* de John Rawls<sup>31</sup>, a escolha deveria ser aquela que refletisse princípios morais universais, buscando desviar do critério utilitarista e da postura meramente auto-interessada.

Todavia, foi a noção welferista contida no princípio rawlsiano da diferença que mais chamou a atenção dos economistas de sua época. A formulação do conceito de equidade trazia consigo uma nova maneira de corrigir o problema do bem-estar dos menos favorecidos. Isto redefiniria aos níveis mínimos de disponibilidade de bens primários (*primary goods*)

---

<sup>30</sup> A abordagem procedimentalista de Rawls faz uso da situação hipotética chamada de *Posição Original*, onde os indivíduos escolhem os princípios básicos da sociedade por de trás do *véu da ignorância*, sem conhecer sua própria posição na ordem social resultante (ignorantes de suas identidades pessoais). Em tal situação de igualdade primordial, Rawls alega que seus princípios de justiça seriam geralmente aceitos como um acordo justo.

<sup>31</sup> Rawls faz referências diretas ao trabalho de Arrow em *Uma Teoria da Justiça*. Em particular, quando se refere ao procedimento da posição original e da distribuição de bens primários (ver em Rawls 1971, p. 114).

aceitáveis para a sociedade de uma maneira geral, o que, não obstante, implicaria em uma forma especial de comparação interpessoal de utilidade.

É fácil observar que a obra de Rawls beneficiou-se do estado da discussão sobre a questão da escolha social (em especial a evolução sobre os requisitos de desempenho lógico dos modelos de eleição de preferências), contudo, a contribuição do autor elevou o tema a um novo patamar como um todo.

Aparentemente, a alternativa rawlsiana dava conta de problemas antigos que a economia de bem-estar, como ramo da ciência econômica, não tinha capacidade de sobrepujar desde um ponto de vista puramente técnico.

Rawls criticou a posição welferista e propôs uma base informacional alternativa para os bens primários sociais, definíveis como "*coisas que se presume que todo homem racional queira*" e cujo uso fosse independente da concepção subjetiva sobre a utilidade de tais bens (Rawls 1971, p. 54). Elevou-se, com isso, a complexidade informacional da escolha, sem deixar de lado a racionalidade inerente à ordenação de preferências.

De forma similar, o autor criticou a regra utilitarista de agregação por ser "*indiferente quanto à maneira com que a soma constante de benefícios é distribuída*" (Rawls 1971, p. 71). Sua alternativa ao utilitarismo benthamiano foi refletida em seu sistema de justiça no qual todos os bens sociais primários (liberdade, oportunidade, renda, riqueza, e as bases do auto-respeito) deveriam "*ser distribuídos igualmente, a menos que uma distribuição desigual de qualquer um ou de todos esses bens seja vantajosa aos menos favorecidos*" (Rawls 1971, p. 297) ".

Para Rawls princípio da eficiência de Pareto, por si só, não era capaz de selecionar a melhor distribuição de bens primários<sup>32</sup>. Os princípios de justiça rawlsianos, desta forma, servem de vetores que complementam o critério paretiano, dirigindo a distribuição desses bens na busca da função de bem-estar social ideal, bem como determinando quando e como podem ocorrer exceções justificáveis à regra de maximização.

Trata-se de forma de comparação interpessoal de utilidade guiada por princípios que pudessem melhor aproveitar esse acréscimo de informação ao modelo de escolha.

Em suma, pode-se dizer que o sistema de Rawls resolvia a questão da escolha social: a) sem abandonar a regra de eficiência paretiana; b) sem recorrer a ditadura de doutrinas abrangentes; c) utilizando critérios de universalização; d) refletindo uma função social de domínio irrestrito; e e) apontando mecanismos adicionais para coibir a imposição autoritária e incentivar a cooperação em um regime de equilíbrio duradouro e sustentável.

É verdade que *Uma Teoria da Justiça* buscava uma concepção mais ampla, ou seja, um modelo procedimental capaz de deduzir princípios universais. A questão da escolha coletiva, contudo, é parte integrante de qualquer modelo de distribuição que almeje legitimidade social. O que a teoria de Rawls deixa claro é que, antes de adentrar em razões éticas e princípios universais, é preciso avaliar em que condições a escolha social de bens e valores vai ocorrer. Quais os requisitos lógicos e estruturas de eleição serão utilizados? Quais serão as concessões em termos de consistência formal e de racionalidade individual auto-interessada? O problema

---

<sup>32</sup> Pode-se notar a posição de Rawls das seguintes passagens: “Now these reflections show only what we knew all along, that is, that the principle of efficiency cannot serve alone as a conception of justice” (Rawls 1971, 1999, p.61) e “The principle of efficiency does not by itself select one particular distribution of commodities as the efficient one. To select among the efficient distributions some other principle, a principle of justice, say, is necessary” (Rawls 1971, 1999, p.59)

levantado pelo teorema de Arrow deve ser enfrentado, seja para confirmar suas conclusões ou apontar um meio de desviá-lo.

Embora não se trate de palavra final sobre o assunto, a concepção de justiça como equidade de Rawls exerceu forte influência no cenário da economia de bem-estar abrindo caminho para as contribuições que buscassem novas e possíveis maneiras de evitar os paradoxos da escolha social.

De fato, articulações alternativas sobre vantagens individuais foram propostas, apontando instigantes possibilidades na economia de bem-estar em geral, e na teoria da escolha social em particular. Propostas representativas para este efeito incluem, os *recursos* na teoria da equidade de Dworkin (1981) e a *abordagem das capacidades* em Sen (1985).

O caminho, portanto, está aberto para quem mais esteja disposto a testar os limites lógicos da escolha social. Creio, contudo, que a solução, seja qual for, deve ser capaz de manter a potência formal da análise lógica dos mecanismos de escolha sem perder de vista os argumentos filosóficos mais básicos sobre a vida social justa.

### 1.7 Disposições finais ao primeiro capítulo

O presente capítulo se propôs a levantar e analisar os requisitos formais para a escolha social. Foram avaliados métodos de votação e o desempenho lógico de modelos de escolha coletiva. Foram, ainda, destacadas as possíveis inconsistências e paradoxos potencialmente presentes em qualquer esquema de votação que considere mais de duas alternativas.

Uma solução sobre preferências coletivas que esteja para além da escolha individual, precisa ser capaz de apontar formas de compatibilização entre os critérios U, P, I e D<sup>33</sup>.

---

<sup>33</sup> Domínio irrestrito (*Unrestricted Domain*), princípio de Pareto, IIA (independência das alternativas irrelevantes) e não Ditadura.

O resultado de impossibilidade sugerido pelo paradoxo de Arrow é problema perturbador que acaba por servir como crítica à utilização de um modelo puramente formal na busca pela melhor política econômica de distribuição de bens e valores.

A avaliação da escolha desde um viés formalista mostra-se mais ajustada à predição do comportamento social de consumo conforme um esquema de eleição de preferências afeto à regra de mercado. A noção político-social de bem-estar, contudo, parece exigir solução mais complexa.

A compreensão do governo da maioria e da democracia, *“deve viajar por espaços mais misteriosos, deve subir escadas mais sinuosas, e deve empregar um altímetro mais complexo do que consistência transitiva”* (Rae 1980, p.544).

A solução que será proposta neste trabalho pretende seguir o caminho em busca de uma alternativa viável, consistente e duradoura para o problema da escolha social.

Como alternativa ao teorema de impossibilidade da escolha social, será necessário um procedimento que se acople no sistema de eleição de preferências individuais e ajuste (por meio de princípios vetores) a determinação da função de bem-estar social, visando um resultado que seja ao mesmo tempo livre, eficiente, cooperativo e não ditatorial.

Conforme já visto, neste mesmo sentido é o conceito de justiça como equidade de Rawls (2001) no qual é fundamental a noção de sociedade como um sistema justo de cooperação social prolongado no tempo, visando desenvolver uma concepção política de justiça para um regime democrático.

Trata-se de um exemplo de arranjo que não abandona os critérios lógicos de avaliação dos métodos de escolha, mas expande a possibilidade de se chegar a uma legítima escolha social por meio de princípios universais que servem como vetores para interpretação de informações que, na

abordagem puramente formal, não teriam espaço de apreciação no processo de decisão.

A posição de Rawls não afronta o critério de Pareto, porquanto um arranjo de estrutura social básica somente é considerado eficiente quando não há maneira de mudar a distribuição de bens e valores de modo a aumentar as perspectivas de alguns sem diminuir com isso as perspectivas dos outros. Existiriam, portanto, muitos arranjos eficientes deste tipo, cada um deles especificando uma divisão de vantagens para cooperação social. O problema seria escolher dentre estes arranjos de distribuições eficientes, uma função que fosse, também, justa. Segundo o autor, se alcançado isso, *teremos ido além da mera eficiência e, ainda assim, de uma maneira compatível com ela* (Rawls 1971, 1999, p 61).

A nova pergunta que clama iluminação passa a ser: que tipo de *arranjo eficiente* reflete melhor a distribuição de bens e valores, de forma a ser capaz de atingir e manter um *sistema de cooperação social equilibrado*?

Neste ponto o presente trabalho passa a se dedicar mais especificamente a uma solução para o problema da escolha social enquanto decisão política por um arranjo resultante de bem-estar social equilibrado e sustentável. Fica em segundo plano, portanto, a concepção de escolha coletiva considerada em sua aplicação meramente racional e agregativa de vontades individuais. Pode-se argumentar, entretanto, que o sistema de cooperação se aplica a ambos os aspectos da escolha sob as mesmas normas de desempenho lógico.

A escolha social a que se direcionará o presente esforço será, em seu âmago, a escolha pela cooperação social, pressupondo-se a um cenário otimista no qual a interação contínua dos participantes gere um senso de extensão dos efeitos de cada escolha particular. Neste sentido, a escolha social vai além da escolha individual, sem recorrer a uma função meramente agregativa, nem a uma autoridade impositiva externa. Também

nessa acepção, a escolha pela cooperação assume um tom intersubjetivo e se contrapõe a atuação meramente auto-interessada e egoísta, assumida como conduta racional padrão na teoria econômica em geral.

Cabe, contudo, um especial esclarecimento metodológico preliminar. Certamente não é o caso de que a teoria da escolha social seja aplicável somente em busca da cooperação. Escolha coletiva, em sua concepção ampla, pode muito bem ser aplicada em cenários puramente competitivos. Com efeito, esta é suposição inicial de Arrow<sup>34</sup> que aponta como fundamental o desejo de se vencer o “jogo<sup>35</sup>” da escolha social tanto no comportamento econômico como nas implicações políticas resultantes.

De fato, a “vitória” no jogo social tem relação direta com a legitimidade da função social resultante. Quanto mais pessoas tiverem suas preferências individuais traduzidas na função de bem-estar social, mais legítima a escolha social.

Entretanto, especialmente no que tange às implicações políticas, o jogo social nem sempre é competitivo. Pelo contrário, nestas condições, o jogo social, muitas das vezes, apresenta solução cooperativa eficiente e autônoma. Neste sentido a lição de Elster:

A teoria da escolha social tem ligações óbvias com a teoria dos jogos cooperativos e não cooperativos. Além disso, a conexão com teorias da justiça distributiva é ainda mais íntima. A estrutura da teoria foi consideravelmente clarificada pela consideração explícita destas conexões. A teoria das escolhas sociais e a teoria dos jogos cooperativos têm muitas características comuns. Em particular, a teoria de barganha e a teoria da escolha social são ambas preocupadas com a

---

<sup>34</sup> *There is no need to stress the obvious importance of the desire to play and win the game as such in both economic behavior and political* (Arrow 1951 p. 7).

<sup>35</sup> Arrow admitiu que existissem fatores reais além dos aspectos formais da preferência que deveriam ser considerados na determinação da mecânica da escolha social, mas alegou que estavam fora do escopo do seu estudo. Os aspectos não abordados pela sua teoria poderiam convenientemente ser descritos como aspectos “do jogo” (Arrow, 1951, p.6).

obtenção de um resultado social de preferências individuais, sendo a otimização de Pareto uma das principais restrições. (Elster 1986, p.3)

A cooperação social, assim, é tema que intersecciona a filosofia política, a teoria da escolha social e a abordagem matemática da teoria dos jogos.

Na teoria de Rawls *o papel dos princípios de justiça (como parte de uma concepção política de justiça) é o de especificar termos justos de cooperação social* (Rawls 2001, p. 7). O princípio da liberdade somado à exceção razoável ao critério paretiano contida no princípio da diferença servem como agentes balizadores na busca dos termos mais justos de cooperação entre cidadãos considerados livres e iguais. A teoria de Rawls, todavia, é elaborada sobre argumentação fortemente deontológica e fundada na noção procedimental kantiana de universalização de máximas individuais.

Seria possível a cooperação entre indivíduos racionais (mesmo que puramente auto-interessados) sem a necessidade da imposição externa de um modelo moral ideal ou de uma autoridade política? Como atingir e manter o equilíbrio cooperativo?

O segundo capítulo será destinado a esta investigação. Partindo-se da abordagem da teoria dos jogos<sup>36</sup> buscar-se-á verificar as condições exatas em que a cooperação pode ocorrer.

Tratar-se-á, doravante, da escolha social como escolha pela solução cooperativa, livre, eficiente, duradoura e não imposta por uma autoridade externa.

---

<sup>36</sup> A teoria moderna dos jogos é o estudo de modelos matemáticos de interação estratégica entre agentes racionais. A origem dessa abordagem é creditada à ideia da existência de equilíbrios de estratégias mistas em jogos de soma zero com duas pessoas, conforme provou John Von Neumann em seu trabalho seminal *On the Theory of Games of Strategy*, de 1928.

## Cooperação no mundo dos egoístas

*“Chess is not a game. Chess is a well-defined form of computation. You may not be able to work out the answers, but in theory there must be a solution, a right procedure in any position. Now real games... are not like that at all. Real life is not like that. Real life consists of bluffing, of little tactics of deception, of asking yourself what is the other man going to think I mean to do. And that is what games are about in my theory”. (Von Neumann – em diálogo com Jacob Bronowski em *The Ascent of Man*, 1973)*

A busca pelo auto-interesse é, talvez, o pilar mais resiliente da racionalidade econômica clássica. O modelo racional de escolha individual parece impor que cada indivíduo aja sempre egoisticamente visando o resultado mais vantajoso.

Com efeito, quando consideramos situações de escolha estanques e resultados pré-definidos de acordo com cada opção dos agentes, a racionalidade de qualquer decisão que não a opção egoísta fica comprometida, pois não corresponde à maximização do interesse que estava em jogo.

O jogo social, contudo, não goza de critérios tão estritos. Não se trata de um jogo isolado e sem repercussão nas escolhas futuras dos participantes. Dificilmente estamos diante de um jogo de soma zero<sup>1</sup>. Ou seja, as escolhas sociais não funcionam como um eterno jogo de xadrez<sup>2</sup>, onde o objetivo de um participante é derrotar o outro.

---

<sup>1</sup> Jogos de soma zero são jogos competitivos por natureza, nos quais só pode haver um vencedor e, conseqüentemente, nos quais existe um (ou mais) derrotados. São jogos nos quais o ganho de um jogador implique necessariamente a perda para o outro jogador.

<sup>2</sup> Curiosamente, Von Neumann sequer considerava o xadrez como um “jogo” na concepção particular em que empregava em sua teoria. Para Neumann o xadrez seria uma forma computacional bem definida, incapaz de avaliar as informações que motivam os jogadores.

Se pensarmos nas relações humanas como um jogo continuado de possíveis escolhas e resultados, estamos claramente diante de um problema que extrapola a simples maximização do ganho individual. A possibilidade de cooperação, portanto, passa inevitavelmente pelas respostas que esperamos dos demais participantes.

A pergunta permanece: como esperar a cooperação quando exista incentivo individual à atuação egoísta? A resposta pessimista é a mais difundida no imaginário coletivo, baseada no argumento hobbesiano de que o estado de conflito constante domina o homem, impondo uma competitividade avassaladora que o reduz à existência “*em medo contínuo, perigo de morte violenta e uma vida solitária, pobre, má, bruta e curta*” (Hobbes, 1651, p. 78). Neste cenário, a cooperação social não é capaz de subsistir senão por imposição de uma forte autoridade central.

Robert Axelrod, em seu livro “A Evolução da Cooperação”, de 1984, propõe uma visão menos derrotista, mediante a análise de padrões de comportamento e estratégias de reciprocidade. Inspirado na conduta entre nações nas relações internacionais, onde não há influência de uma entidade superior dotada de poder cogente, o autor tenta uma resposta para a questão: o que, de fato, leva o indivíduo auto-interessado a cooperar sem que haja uma imposição de uma autoridade central?

O problema básico surge em situações nas quais cada indivíduo tenha motivos razoáveis para agir egoisticamente, mas que, consideradas as alternativas do jogo, a escolha egoísta gera um resultado total pior que a opção cooperativa.

Para isso Axelrod (1984) utiliza uma representação simples capaz de despir o problema da cooperação de suas variáveis mais particulares permitindo a avaliação mais objetiva, valendo-se de um exemplo particular de dilema do prisioneiro em sua forma iterada.

## 2.1 O dilema do prisioneiro

É questão fundamental à teoria dos jogos a análise da “maneira de se jogar o jogo”, buscando destacar pontos de equilíbrio e estratégias racionais confiáveis, capazes, por sua vez, de prever o comportamento dos agentes. Na concepção original de Von Neumann o termo “jogo” refletia uma situação de conflito na qual o participante deveria (obrigatoriamente) fazer uma escolha, sabendo que os demais também estão tomando suas posições (Poundstone 1993, p.6). Alguns jogos são simples. Outros são permeados por antecipações e desconfianças que acabam por influenciar a maneira de se jogar.

O dilema do prisioneiro<sup>3</sup>, em sua concepção originalmente formalizada por Albert W. Tucker, é um exemplo de jogo que visa analisar a possibilidade de cooperação entre participantes auto-interessados, considerando um quadro de resultados previamente determinados. Trata-se de um cenário que incorpora a tensão entre a racionalidade individual (refletida no incentivo de ambos os lados serem egoístas) e racionalidade coletiva (refletida no maior retorno para ambos os lados na cooperação mútua).

Dois indivíduos podem cooperar entre si ou denunciar um ao outro. A denúncia produz um retorno individual maior do que a cooperação. Mas se ambos denunciam, o resultado é pior do que se ambos cooperarem. A situação ilustrada nesse experimento provoca o leitor a colocar-se no lugar dos participantes hipotéticos do jogo e avaliar suas alternativas de ação. Como o próprio nome indica, o experimento acaba por gerar um impasse que desafia a racionalidade dos agentes em face das opções de escolha,

---

<sup>3</sup> O jogo do dilema do prisioneiro foi inventado em 1950 por Merrill Flood e Melvin Dresher, e formalizado por A. W. Tucker. Dilema do Prisioneiro, assim denominado por indicar, em sua versão mais tradicional, situação em que dois indivíduos são presos em celas diferentes e incommunicáveis. Eles podem cooperar entre si ou denunciar o outro jogador. Se ambos cooperam, ganham desconto em suas penas. Se um coopera e outro denuncia, o denunciante sai ileso enquanto o agente que cooperou recebe a pena máxima. Se ambos denunciam ambos recebem uma pena maior que a inicial, mas menor que a máxima.

porquanto o comportamento de maximização do interesse de cada indivíduo leva a um desfecho pior para todos.

O dilema do prisioneiro, portanto, desafia as próprias bases da teoria dos jogos, pois confronta a noção de que exista uma melhor “maneira de se jogar” (estratégia). Trata-se de um constructo matemático preciso e aplicável a situações reais de decisão podendo ser extensivamente aproveitado em diversas áreas do conhecimento como a biologia, a psicologia, a sociologia, a economia e o direito. Os impasses deste tipo, em última análise, podem surgir em qualquer lugar onde exista um conflito de interesses (Poundstone, 1993, p.6).

Nesse contexto, e em busca das reais razões que levam indivíduos a cooperarem, Robert Axelrod propôs uma variação do modelo tradicional de dilema do prisioneiro. O exemplo de jogo proposto por Axelrod (1980 e 1980a) é uma matriz simples com dois jogadores, duas escolhas e quatro possíveis resultados. A maneira como o jogo funciona é mostrada na figura 1. Um jogador escolhe uma linha, cooperando ou denunciando. O outro jogador escolhe simultaneamente uma coluna, seja cooperando ou denunciando. Juntas, essas escolhas resultam em um dos quatro resultados possíveis ( $[3,3]$ ,  $[5,0]$ ,  $[0,5]$  ou  $[1,1]$ ). Se ambos os jogadores cooperam, eles prosperam e recebem a recompensa para a cooperação ( $R=3$  - *Recompensa para Cooperação Mútua*). No jogo de Axelrod a recompensa é de três pontos. Se um jogador cooperar, mas outro não, o jogador denunciante recebe cinco pontos (situação  $T=5$  - *Tentação de Denunciar*), enquanto o jogador cooperante recebe zero ponto (situação  $S=0$  - *Sorte/Azar do Denunciado*). Se ambos denunciam, ambos obtêm um ponto como punição (situação  $P=1$  - *Punição para Denúnciação Mútua*).

Figura 1  
Dilema do Prisioneiro

		<i>Jogador</i>	
		<i>Cooperar</i>	<i>Denunciar</i>
<i>Jogador</i>	<i>Cooperar</i>	$R=3, R=3$ Recompensa para Cooperação Mútua	$S=0, T=5$ Sorte/Azar do Denunciado e Tentação de Denunciar
	<i>Denunciar</i>	$T=5, S=0$ Tentação de Denunciar e Sorte/Azar do Denunciado	$P=1, P=1$ Punição por Denúnciação Mútua

Fonte: Axelrod, (1984, 2006)

Nota-se que a postura puramente auto-interessada é estratégia legítima. Ademais, a postura egoísta é a única que permite a maximização do benefício individual considerando os resultados possíveis do jogo.

Outra constatação digna de nota é que o exemplo de Axelrod não prevê perda significativa para os não cooperantes (como um resultado punitivo negativo, por exemplo -1, que indicaria uma situação não paretiana) tampouco indica o estado inicial dos participantes em termos de risco ou relevância da punição (como por exemplo uma pessoa muito abastada para quem o prêmio cooperativo não seja um real incentivo).

Desta forma, é sempre melhor denunciar se o participante sabe que o outro irá cooperar e, também, é melhor denunciar se é sabido que o outro irá denunciar! Ora, em ambos os casos há vantagem individual. A estratégia egoísta, portanto, compensa independentemente da ação do seu oponente (Axelrod 1984, 2006 p. 9).

Mas a lógica também vale para o outro participante, e a decisão egoísta, se assumida por ambos, sempre levará a um desfecho pior que a cooperação. Em outras palavras, a estratégia em busca de maximização do resultado individual, se perseguida como estratégia simultânea pelos dois participantes reflete um resultado que não satisfaz nem um nem outro,

levando invariavelmente a situação P [1,1] (Punição para a Denúnciação Mútua).

Esse é justamente o ponto do exemplo de dilema do prisioneiro proposto por Axelrod. No cenário acima exposto a racionalidade individual leva a um resultado pior para ambos do que seria potencialmente possível. Trata-se de uma abstração para avaliar a situação, naturalmente ocorrente nas relações sociais, em que a racionalidade individual aponta para uma atitude auto-interessada, mas, ainda assim, ambos sair-se-iam melhor se tomassem a via cooperativa.

Trata-se de um jogo relativamente simples, mas que permite a análise formal dos diversos fatores que destacam a relação do participante com as alternativas possíveis.

Primeiramente, o participante especifica a ordenação das quatro jogadas possíveis de acordo com seu potencial resultado. A melhor jogada seria situação T [5,0], a pior jogada seria a situação oposta S [0,5]. Dentre as jogadas intermediárias, a situação cooperativa R [3,3] se mostra mais vantajosa que a alternativa egoísta P [1,1]. Temos, portanto, a ordem:  $T > R > P > S$ .

Em um segundo momento procede-se a análise de estratégias a serem tomadas. No jogo em questão pode-se assumir que a chance de explorar o outro participante é a mesma de ser explorado por ele. Isso reforça que, em termos de resultado, a estratégia de cooperação mútua prevalece sobre a estratégia de exploração mútua. Somando-se esta etapa ao momento inicial de ranqueamento das jogadas possíveis, temos o panorama deste Dilema do Prisioneiro.

A partir disso, pode-se concluir que, dados dois jogadores egoístas que se enfrentam em um jogo único, a ação racional e consistente será de que ambos denunciarão o adversário, ficando com o resultado P [1,1], e, portanto, saindo-se pior do que era potencialmente atingível.

Conclusão semelhante pode ser alcançada se tomarmos o jogo com número determinado e finito de repetições, pois o incentivo para cooperação continua sendo previsível. Isso fica claro no caso do penúltimo movimento, no qual ambos os participantes devem esperar a postura egoísta do seu competidor, já que no último movimento, assim como no caso do exemplo do jogo único, não haverá mais decisão futura a ser influenciada.

Axelrod (1984, 2006 p.10) avalia que esta racionalidade não se aplica se os jogadores se enfrentam *repetidamente por número indefinido de vezes*. Isso porque, num cenário realista, os participantes não podem ter certeza de qual será a última interação que terá com o oponente. O autor afirma que é nesse caso, com número indefinido de interações, que a cooperação pode emergir. O problema passa a centrar-se, então, na busca pelas exatas condições necessárias e suficientes para que se verifique a emergência da cooperação.

## **2.2 Condições para cooperação**

Para verificar a possibilidade de cooperação, Axelrod define alguns parâmetros mínimos, dentro dos quais o jogo de repetições indefinidas deve ser aplicado.

Antes de tudo, o jogador deve ser capaz de reconhecer o outro participante e deve lembrar-se de como interagiram nas iterações anteriores. Somente assim, contando com a habilidade de reconhecer e lembrar, é possível a formação de um histórico de interações particulares, que serve de informação a ser considerada para influenciar as decisões futuras.

Além desta capacidade de memória são necessárias outras quatro condições, que restam elencadas por Axelrod (1981, p. 308) para garantir o funcionamento elementar do jogo iterado, sem que sejam aplicadas soluções externas ao jogo como solução ao dilema do prisioneiro: a) cada

jogador deve sempre considerar a cada interação todas as estratégias possíveis para si e para o outro participante (não pode haver ameaças ou comprometimentos); b) não há como saber, de antemão, qual será a jogada do outro participante (o que impossibilita uma análise de informações sobre intenção dos jogadores); c) não há possibilidade de fuga ou eliminação (cada jogador só tem a opção de cooperar ou denunciar); e d) não há negociação sobre os resultados (a punição e a recompensa são vinculadas à opção de cada jogador).

Dessa forma, não há espaço para promessas não garantidas por ação consequente que as implemente. A comunicação entre os participantes somente é permitida através da sequência de suas próprias decisões. O que permite a cooperação é o fato de que os jogadores irão encontrar-se novamente no futuro. As escolhas presentes não determinam somente o jogo atual, mas influenciam as escolhas subsequentes.

Concomitantemente, o futuro lança uma sombra sobre o presente e passa a influenciar a escolha da estratégia a ser tomada no jogo corrente. Assim, a pauta passa a ser o peso que o futuro pode imprimir na tomada de decisão no caso de jogos continuados, e qual o papel que a cooperação pode adquirir em casos como este.

Com efeito, é natural que pensemos que a decisão futura é menos importante que a decisão atual, porquanto, a recompensa do jogo futuro, por se tratar de potencial interação que pode nem vir a ocorrer, passa a ser vista como menos valiosa.

Na contabilidade cumulativa dos resultados, dado um lapso de iterações, passa a ser necessário um meio de avaliar o peso que a próxima decisão terá sobre a decisão em andamento. A sombra lançada pela recompensa da próxima decisão deve, portanto, ser avaliada como porcentagem em relação à escolha atual (Shubik, 1970). Tal coeficiente funciona como

*parâmetro de desconto*<sup>4</sup> que deve ser levado em consideração em busca da *estratégia*<sup>5</sup> ideal. Quanto menor este parâmetro (aqui denominado  $w$ ), menos importantes são os movimentos posteriores em relação aos mais próximos.

Mas qual é a melhor estratégia? Se consideradas todas as condições acima dispostas, existe uma regra de decisão que permite a maior pontuação possível em longo prazo? Existe uma estratégia que seja capaz de maximizar a recompensa de um jogador independentemente da estratégia adotada pelo outro participante? Axelrod responde a essa pergunta propondo o primeiro lema formal:

Proposition 1 - If the discount parameter,  $w$ , is sufficiently high, there is no best strategy independent of the strategy used by the other player<sup>6</sup>. (Axelrod, 1984, 2006, p. 15)

Isso significa que, no cenário de dilema do prisioneiro iterado tal qual proposto pelo autor (assumindo que o parâmetro de desconto seja digno de influência) a estratégia que melhor funciona é a que considera a estratégia do outro jogador.

Em jogos de soma zero, dado que o objetivo é eliminar o adversário, o jogador pode pressupor, de forma racional e válida, que não há possibilidade de cooperação. O participante deve sempre antecipar a pior jogada possível por parte do oponente, e agir de forma a minimizar o dano ou maximizar o ganho em sua jogada corrente. No exemplo do dilema do

---

<sup>4</sup> Assim, a próxima recompensa vale apenas uma fração do resultado atual. Uma cadeia inteira de denunciação mútua teria então um "valor presente" de  $P + wP + w^2P + w^3P \dots = P / (1-w)$ . O parâmetro de desconto,  $w$ , pode receber uma das duas interpretações. A interpretação econômica padrão é que o consumo posterior não é valorizado tanto quanto o consumo mais próximo. Uma interpretação alternativa é que movimentos futuros podem não ocorrer de fato, já que a interação entre um par de jogadores tem apenas certa probabilidade de continuar.

<sup>5</sup> Em geral, uma estratégia (ou regra de decisão) é uma função da história do jogo até o momento, traduzida, no caso do jogo de soma não zero, em uma probabilidade de cooperação no próximo movimento.

<sup>6</sup> Proposição 1. Se o parâmetro de desconto  $w$  for suficientemente alto, não há melhor estratégia independente da estratégia usada pelo outro jogador.

prisioneiro, diferentemente de um jogo de soma zero, o objetivo não é eliminação. Pelo contrário, como já demonstrado, a opção cooperativa traz recompensa mais benéfica que a conduta puramente egoísta, especialmente se considerada em jogos iterados.

Assim, podemos extrair as seguintes conclusões: a) é condição necessária para que emergja a cooperação a interação contínua entre os participantes, de tal forma a gerar tanto um histórico de decisões anteriores como uma expectativa de conduta futura; b) não é, contudo, condição suficiente a interação continuada, pois os arranjos de estratégias e comportamentos dos participantes exerce constante influência nas decisões a serem tomadas; c) não há, nestes termos, uma estratégia ideal que possa ser considerada a melhor, sem que haja a avaliação concomitante da estratégia adotada pelo outro jogador.

A demonstração de que não existe uma única “melhor estratégia” levanta a questão sobre os possíveis padrões de comportamento que podem se apresentar, ou mesmo determinar o jogo contínuo.

### **2.3 Um torneio de estratégias**

O *Dilema do Prisioneiro* é uma representação relativamente simples e que permite a análise da racionalidade que informa a escolha a ser tomada, dados os resultados possíveis do jogo proposto. O jogo nesses moldes permite uma ampla aplicabilidade interdisciplinar explorada por cientistas políticos, cientistas sociais, economistas, biólogos, psicólogos, entre outros, haja vista a capacidade do exemplo do dilema para avaliar potenciais padrões de comportamento. A literatura sobre o tema é frutífera e, não se restringe somente ao jogo em si, mas as causas externas que possam influenciá-lo. Como exemplo podemos apontar soluções que buscam contornar o dilema como a competição oligopólica (Samuelson 1973, p. 503), o método de troca de votos (Riker e Brams 1973) e mesmo análises

que permitam escolhas interdependentes ou uma tributação sobre a conduta egoísta (Clarke 1980).

Para que seja possível uma avaliação objetiva, contudo, é necessário um recorte preciso. É exatamente esta a proposta de Axelrod. Assim como fez Arrow, ao escolher os paradigmas para a escolha social, Axelrod (1984, 2006) limita a análise ao seu substrato formal, sem que sejam possíveis interferências externas ao jogo. Trata-se, portanto, de uma escolha metodológica que paga o preço quanto à complexidade das relações avaliadas, mas garante a credibilidade e consistência formal do argumento.

Para Axelrod o foco deve ser como jogar melhor o jogo e não como evitar o dilema. A fim de verificar como escolher de forma mais eficaz, é necessária uma nova abordagem capaz de avaliar as possibilidades estratégicas inerentes a um cenário de soma não-zero, ou seja, uma situação na qual os interesses dos participantes não sejam excludentes, mas coincidam ao menos parcialmente.

A solução proposta pelo autor enriqueceu enormemente a discussão sobre as estratégias de escolha e a emergência da cooperação: um torneio entre programas de computador.

Os parâmetros são os do exemplo de Dilema do Prisioneiro iterado, no qual cada programa concorrente tem como possíveis entradas a escolha cooperativa ou a não cooperativa. Convenientemente para o autor, uma disputa entre programas mediante a linguagem lógico-computacional é uma abordagem ideal para impedir diálogo entre os participantes e mesmo entre estratégias, restando disponíveis somente as informações das decisões já tomadas no jogo até então.

Axelrod convidou profissionais respeitados por seus trabalhos no âmbito da teoria dos jogos, muitos dos quais já haviam publicado seus próprios modelos de Dilema do Prisioneiro, e os enviou os requisitos para

o torneio. Foram submetidas entradas de diversas disciplinas como psicologia, economia, ciência política, matemática e sociologia.

As regras da competição impunham o enfrentamento direto de cada estratégia enviada em etapas eliminatórias. Cada jogo consistia em exatos 200 movimentos, nos moldes da matriz de recompensas do exemplo de Dilema do prisioneiro já explicado na figura 1 ( $[3,3]$ ,  $[5,0]$ ,  $[0,5]$  ou  $[1,1]$ ). Em um primeiro turno, no qual foram confrontadas as 14 estratégias inicialmente submetidas, houve 120.000 movimentos de 240.000 escolhas possíveis, o que corrobora a adequação da análise computacional em detrimento da aplicação empírica de experimentos de campo com pessoas.

A estratégia vencedora foi o programa *TIT FOR TAT*, submetido por Anatol Rapoport da universidade de Toronto.

Figura 2 - Tabela de resultados do torneio de Axelrod

		TABLE 2 Tournament Scores															
Other Players		TIT FOR TAT	TIDEMAN AND CHIERUZZI	NYDEGGER	GROFMAN	SHUBIK	STEIN AND RAPOPORT	FRIEDMAN	DAVIS	GRAASKAMP	DOWNING	FELD	JOSS	TULLOCK	(Name Withheld)	RANDOM	Average Score
Player																	
1.	TIT FOR TAT (Anatol Rapoport)	600	595	600	600	600	595	600	600	597	597	280	225	279	359	441	504
2.	TIDEMAN AND CHIERUZZI	600	596	600	601	600	596	600	600	310	601	271	213	291	455	573	500
3.	NYDEGGER	600	595	600	600	600	595	600	600	433	158	354	374	347	368	464	486
4.	GROFMAN	600	595	600	600	600	594	600	600	376	309	280	236	305	426	507	482
5.	SHUBIK	600	595	600	600	600	595	600	600	348	271	274	272	265	448	543	481
6.	STEIN AND RAPOPORT	600	596	600	602	600	596	600	600	319	200	252	249	280	480	592	478
7.	FRIEDMAN	600	595	600	600	600	595	600	600	307	207	235	213	263	489	598	473
8.	DAVIS	600	595	600	600	600	595	600	600	307	194	238	247	253	450	598	472
9.	GRAASKAMP	597	305	462	375	348	314	302	302	588	625	268	238	274	466	548	401
10.	DOWNING	597	591	398	289	261	215	202	239	555	202	436	540	243	487	604	391
11.	FELD	285	272	426	286	297	255	235	239	274	704	246	236	272	420	467	328
12.	JOSS	230	214	409	237	286	254	213	252	244	634	236	224	273	390	469	304
13.	TULLOCK	284	287	415	293	318	271	243	229	278	193	271	260	273	416	478	301
14.	(Name Withheld)	362	231	397	273	230	149	133	173	187	133	317	366	345	413	526	282
15.	RANDOM	442	142	407	313	219	141	108	137	189	102	360	416	419	300	450	276

Fonte: Axelrod (1980a p.11)

*Tit-for-tat*<sup>7</sup> começa com uma escolha cooperativa, e depois repete o que o outro jogador fez no movimento anterior. Esta regra de decisão é

<sup>7</sup> O significado mais adequado para *tit-for-tat* seria “retribuição equivalente”. Contudo, a tradução livre revela a utilização coloquial do termo, como sinônimo da expressão “olho por olho”.

provavelmente a regra mais discutida e conhecida de se jogar o dilema do prisioneiro<sup>8</sup>. É facilmente compreendida e programável.

Com efeito, em um torneio de computadores, que não leva em consideração fatores externos, tal programa concentra as propriedades desejáveis de não ser muito explorável e de pontuar alto contra outras estratégias cooperativas (incluindo a sua própria regra gêmea). Mas o que faz de *tit-for-tat* uma estratégia tão bem sucedida tanto em face de estratégias cooperativas como de regras não cooperativas?

Uma primeira conclusão é evidência de uma propriedade comum entre as estratégias que melhor pontuaram: a propensão à cooperação. As oito primeiras colocadas no ranking do torneio de Axelrod, eram estratégias consideradas bem intencionadas (*nice*).

Muitos teóricos do Dilema do Prisioneiro que submeteram seus programas na expectativa de explorar a atitude cooperativa dos demais e maximizar seu resultado com sutis manobras de denúncia foram surpreendidos, porquanto suas estratégias, na maioria dos casos, pontuaram pouco.

Isso gerou um influxo de questionamentos sobre a amostragem dos programas submetidos e os critérios de confrontamento, o que culminou com a realização de um segundo torneio, este com a participação de 62 programas, cujos resultados foram publicados em setembro de 1980 no *The Journal of Conflict Resolution*, (Axelrod 1980b, pp. 379-403).

Os novos resultados são mais completos de duas maneiras. Primeiramente, todas as estratégias possíveis são levadas em consideração, não simplesmente alguns subconjuntos arbitrariamente selecionados. Em segundo lugar, não apenas as condições de equilíbrio são estabelecidas, mas

---

<sup>8</sup> Sabe-se que provoca um bom grau de cooperação quando jogado com humanos (Oskamp, 1971; Wilson, 1971).

também são permitidos mecanismos capazes de *invasão*<sup>9</sup> de uma estratégia não-cooperativa para um equilíbrio cooperativo (e vice-versa).

Novamente *tit-for-tat* foi a estratégia vitoriosa. E o padrão constatado no primeiro turno de enfrentamentos prevaleceu, porquanto, dentre as quinze primeiras estratégias com maior pontuação apenas a oitava (HARRINGTON) era uma estratégia maliciosa. Não somente isso, mas também dentre as quinze piores classificadas, todas menos uma eram maliciosas.

As estratégias maliciosas não conseguem pontuar bem quando competem com outras estratégias também maliciosas, porquanto, dado a matriz de recompensas do jogo em questão, somente conseguem o resultado [1,1] por entrada e, portanto, sua média de pontos por jogo fica próxima do mínimo: 200 pontos. A simplicidade e versatilidade de *tit-for-tat* a garantem como estratégia que pontua melhor em ambos os cenários, o cooperativo e o não cooperativo.

*Tit-for-tat* provou-se também ser resistente à invasão de outras estratégias predominantemente maliciosas como ALL D (que consiste em sempre denunciar).

As questões do parâmetro de desconto  $w$  e a sustentabilidade da estratégia cooperativa *tit-for-tat* serão melhor analisadas mais adiante, mas a conclusão que podemos tomar, desde já, é que se todos em uma população adotam a estratégia *tit-for-tat* e, portanto, estão em equilíbrio cooperativo, nenhuma outra estratégia, seja ela mista ou puramente maliciosa consegue invadir se o parâmetro  $w$  é alto o suficiente. Contudo, da mesma forma, se o parâmetro  $w$  cair a padrões críticos ( $w \leq 0,5$ ) a estratégia cooperativa deixa de ser estável, pois passa a pagar em turnos

---

<sup>9</sup> O conceito de invasão será mais bem explorado no tópico Evolução da Cooperação. Trata-se de termo emprestado da biologia evolucionista que se refere a capacidade de uma estratégia suplantarem outra estratégia utilizada predominantemente por uma população.

alternados o pior resultado possível. Isso reflete a conexão entre a sombra que o futuro projeta sobre a decisão atual e a possibilidade da cooperação estável.

Com efeito, quanto maior o parâmetro de desconto  $w$ , mais eficaz *tit-for-tat*. Isso significa que quanto maior a chance de o jogador encontrar o outro no futuro, em uma relação continuada, maior a estabilidade<sup>10</sup> geral de *tit-for-tat*. Podemos, contudo, avaliar os resultados e concluir que o sucesso de TFT se deve a alguns motivos específicos.

O sucesso robusto de *tit-for-tat* deve-se à combinação de ser uma estratégia benigna, retaliatória, rápida para perdoar e simples. Sua natureza cooperativa impede que ela entre em problemas desnecessários. Sua rápida retaliação desencoraja o outro lado de persistir quando a escolha egoísta é tentada. Sua capacidade de perdoar ajuda a restaurar cooperação mútua. E sua clareza faz-se inteligível para o outro jogador, para que possa, assim, optar pela cooperação em longo prazo (Axelrod 1984, 2006 p. 54).

Avaliemos então, os conselhos de Axelrod quanto à escolha efetiva em jogos de Dilema do Prisioneiro.

### 2.3.1 Não Seja Ganancioso

A maioria das interações sociais não reflete um jogo de soma zero. Isso significa que muitas das situações permitem uma atitude cooperativa que, por vezes, é negligenciada, seja por falta de informação no momento da escolha ou mesmo por falta de confiança na atitude dos demais participantes do jogo.

---

<sup>10</sup> TIT FOR TAT demonstra-se uma estratégia geral estável se  $w \geq \max(T - R/T - P, T - R/R - S)$ . Isso significa que não pode ser invadida. Assim, dizer que ALL D, por exemplo, não pode invadir TFT significa que  $V(\text{ALL D}/\text{TFT}) \leq V(\text{TFT}/\text{TFT})$ . Como vimos,  $V(\text{ALL D}/\text{TFT}) = T + wP/(1 - w)$ . Já que TFT sempre coopera com seu programa gêmeo,  $V(\text{TFT}/\text{TFT}) = R + wR/W^2R... = R/(1 - w)$ . Dessa forma, ALL D não consegue invadir TIT FOR TAT quando  $T + wP/(1 - w) \leq R/(1 - w)$ , ou  $T(1 - w) + wP \leq R$ , ou  $T - R \leq w(T - P)$  ou  $w \geq T - R/T - P$  (Axelrod 1984, 2006, p. 67).

Em uma situação de equilíbrio cooperativo a tentação pela traição (*Tentação de Denunciar*) é maior, pois a recompensa é alta. Cedo ou tarde alguém cede à tentação. Contudo, como já apontado anteriormente, se o jogo for iterado e o coeficiente  $w$  (parâmetro de desconto) for significativo, a conduta egoísta não se paga. A quebra da confiança mútua gera uma onda retaliativa que leva o resultado do jogo, ao menos em quanto o impasse durar, ao pior resultado coletivo possível (*P - Punição para Denunciação Mútua* [1,1]).

Ademais, esta situação de desconfiança mútua tende a despertar um padrão de comparação entre resultados obtidos até então, e conseqüentemente, leva a um sentimento de inveja. Isso, por sua vez, gera tentativas de equiparação de resultado, mas essa retificação somente pode ser alcançada mediante mais rodadas de denúnciação.

A ausência de ganância ou inveja é característica determinante no sucesso de *tit-for-tat*. Justamente por não fazer tentativas de comparação de resultados obtidos pelas demais estratégias, a regra de *tit-for-tat* mantém-se consistente e desencoraja a atitude egoísta em longo prazo. Em outras palavras, *tit-for-tat* não visa pontuar mais que a estratégia adversária (aliás, seu código de programação nem considera o histórico de pontuação, mas tão somente a conduta da jogada imediatamente anterior).

### **2.3.2 Não seja o primeiro a quebrar a confiança**

Confiança é um equilíbrio que não se sustenta unilateralmente. Trata-se de garantia mútua. Tanto em experimentos teóricos no campo econômico como na prática das relações sociais e internacionais pode-se notar que a cooperação entre os agentes tende a ser condicional. Isso significa que os indivíduos estão dispostos a cooperar, contando que os outros também cooperem com eles.

Isso pode ser traduzido de formas diversas em uma estratégia de interação continuada. A maneira conservadora, e um tanto pessimista, presume a posição não cooperativa, e se defende de uma possível postura auto-interessada. Do outro lado, a visão otimista confia livre de amarras e, conseqüentemente, fica exposta à exploração alheia. O desafio, no caso do Dilema do Prisioneiro, é uma estratégia que seja equilibrada.

O resultado do torneio de Axelrod é categórico, sendo a natureza benigna o traço característico recorrente das estratégias que melhor pontuaram. Mas o que se pode extrair disso? Por que as regras cooperativas são tão eficazes?

Certamente não é o caso de que a cooperação seja sempre a melhor escolha, especialmente se considerarmos uma gama de outras estratégias maliciosas com grande potencial de invasão<sup>11</sup>. Contudo, as populações que seguem regras cooperativas (havendo a garantia de uma relação contínua, ou seja um parâmetro de desconto alto) são as mais resistentes à invasão de estratégias maliciosas. Isso é facilmente comprovado, pois as estratégias cooperativas pontuam consistentemente entre si.

*Tit-for-tat* é a estratégia de cooperar no primeiro movimento e repetir o último movimento do adversário nos demais movimentos. Na interação entre duas pessoas regidas pelo programa *tit-for-tat*, por exemplo, elas cooperarão em todas as rodadas e o resultado será de 600 pontos para cada (200 vezes [3,3]).

Do outro lado, é plenamente racional que o indivíduo queira se precaver, e agir não cooperativamente até que o outro passe a cooperar. De fato, essa foi a aposta da maioria dos programadores que submeteu estratégias ao torneio de Axelrod. Contudo, tal estratégia mostrou-se arriscada, especialmente em se tratando de um jogo continuado de 200 repetições.

---

<sup>11</sup> Invasão aqui significa a colonização da estratégia alheia.

Isso porque a atuação inicial não cooperativa desencadeia uma série retaliativa por parte do outro participante. Isso coloca ambos em uma situação complexa de avaliação de padrões de exploração mútua. A cada rodada de retaliação a implicação futura para o agir não cooperativo aumenta, pois o movimento de perdão passa ser explorado como uma fraqueza.

Ora, se um indivíduo está utilizando *tit-for-tat* (estratégia de cooperar na primeira vez e depois copiar a ação do adversário nas demais vezes) e outro *ALL D* (estratégia de sempre denunciar) o resultado será de 195 pontos para um e 205 para outro (1 vez [0,5] e 199 vezes [1,1]). Comparando com os 600 pontos obtidos pelo enfrentamento de duas estratégias cooperativas, nota-se que denunciar primeiro não compensa.

Não podemos esquecer, entretanto, que a questão passa pelo coeficiente de valorização das interações futuras em relação aos ganhos imediatos no jogo atual. Nos casos em que o participante não consegue visualizar a continuidade da interação com outro participante, manter a conduta cooperativa pode não ser a melhor escolha inicial. Para indivíduos itinerantes<sup>12</sup>, por exemplo, que não tem a expectativa de encontrar o mesmo participante do jogo no futuro, a estratégia não cooperativa é uma alternativa atraente. Mas, se existe a expectativa de interações futuras, ser o primeiro a quebrar a confiança pode comprometer tanto a relação em si como a possibilidade de resultados cooperativos.

Como já visto, *tit-for-tat* só se mantém como estratégia estável se o parâmetro de desconto ( $w$ ) for alto. Contudo, uma vez estabelecida a cooperação com base na reciprocidade, a estratégia cooperativa não pode mais ser invadida, nem por uma estratégia individual, nem por um grupo de estratégias maliciosas. Uma população de estratégias cooperativas (onde

---

<sup>12</sup> Quanto a isso, ver Gropper 1975 p. 106-107 onde o autor avalia a conduta de grupos de ciganos e sua relação com as sociedades com as quais convivem por períodos não extensos.

não há quebra de confiança) desfruta de resultados tão favoráveis que acaba por afastar aglomerados ou indivíduos que estejam usando estratégias maliciosas, protegendo-se, assim, de invasões externas ao equilíbrio cooperativo<sup>13</sup>.

Em jogos tão continuados como, por exemplo, o convívio social e a relação com o meio ambiente o coeficiente  $w$  é o mais alto possível. A solução cooperativa, uma vez que estabilizada, faz com que o indivíduo se recinta de não ter cooperado desde o começo.

A lição que resta, portanto, é que agir egoisticamente pode parecer promissor no princípio, mas em longo prazo pode destruir o ambiente necessário para o seu próprio sucesso (Axelrod 1984, 2006 p. 117).

### 2.3.3 Reciprocidade

Reciprocidade é fundamental tanto para manutenção da confiança cooperativa como para coibir o comportamento malicioso. A programação de *tit-for-tat* é simples e consiste em duas diretrizes fundamentais: a) começa cooperando (não seja o primeiro a quebrar a confiança); e b) pratique a reciprocidade (tanto da cooperação como do egoísmo).

Esse conjunto de regras simples desbancou programas muito mais sofisticados, capazes de apreender com o histórico de rodadas anteriores e avaliar as probabilidades da próxima ação de seu oponente. Aparentemente, essas duas diretrizes foram capazes de lidar da maneira mais eficiente tanto com programas bem intencionados como com estratégias maliciosas.

Estratégias excessivamente cooperativas, para as quais o objetivo principal era evitar a punição por denúncia mútua, foram facilmente

---

<sup>13</sup> Este é o conteúdo da proposição 7 de Axelrod, que atesta que se uma estratégia individual não pode ser invadida por outra estratégia individual, ela também não pode ser invadida por um grupo de estratégias. É exatamente o caso de TFT, quando o coeficiente  $w$  é alto o suficiente (Axelrod 1984, 2006 p. 67).

exploradas. Na outra mão, estratégias maliciosas não souberam perdoar quando necessário e terminaram com resultados piores. Dentre os muitos modelos submetidos *tit-for-tat* foi o que demonstrou a melhor dosagem entre perdão e punição.

De outro lado, estratégias como a ‘Retaliação Permanente’, comum na esfera das relações internacionais, também sofrem por serem muito duras. Nestes casos, o primeiro movimento é de cooperação, o que classifica a estratégia como bem intencionada, mas os movimentos posteriores somente serão cooperativos enquanto o outro jogador cooperar. Uma vez quebrada a confiança nunca mais a opção cooperativa volta a ser possível. Em que pese ser estratégia clara e que representa o máximo desencorajamento da postura maliciosa, tal regra afeta seu próprio desempenho final por não ser capaz de retribuir uma futura tentativa de cooperação. Em outras palavras, é uma estratégia demasiadamente inflexível para o seu próprio bem.

#### **2.3.4 Clareza**

Tão importante quanto a reciprocidade na medida correta é fazer-se entender pelo outro participante. Quanto melhor o outro jogador entender a sua estratégia, melhor ele poderá se posicionar perante ela. Penso que essa máxima seja mais aplicável ainda quando se fala em um jogo discursivo em que ambos participantes possam comunicar-se e esclarecer de antemão sua estratégia. O caso de Axelrod, entretanto, era de um torneio de programas, que, ao menos em uma primeira análise, não tem problemas para “entender” o outro jogador e, ainda assim, a estratégia mais simples foi a que se mostrou mais eficaz.

Estratégias sofisticadas que utilizavam complexas regras para avaliar o histórico do jogo e calcular a maximização estatística do próximo movimento não conseguiram o mesmo sucesso de *tit-for-tat*. Essas regras que

tentaram inferir como o outro jogador iria se portar na jogada subsequente esbarraram em problemas como não levar em conta a mudança que seria desencadeada pela sua própria mudança de comportamento. Ora, em um jogo de soma não zero, não podemos presumir que o outro participante esteja agindo para derrotar o oponente.

O outro participante sempre vai avaliar o seu comportamento e tentar agir da forma que convém para sua própria estratégia, seja ela bem intencionada ou não. No caso de *tit-for-tat*, por exemplo, uma vez que o oponente entende que a reciprocidade vai ser imediata, ele passa agir de acordo, e, eventualmente, passa a preferir a postura cooperativa, se considerar que a recompensa (R) é mais vantajosa que a situação de mútua denúnciação (P).

*Tit-for-tat*, nesse sentido, beneficia-se da sua simplicidade, já que somente considera uma jogada anterior. Tal característica, ao mesmo tempo em que torna clara a estratégia para o oponente, permite que a retribuição da cooperação seja imediata, sem a avaliação estatística das jogadas anteriores. Restabelecer a confiança deve ser prioridade, já que o resultado cooperativo é o mais vantajoso.

Ao lembrar-se apenas da jogada passada, *tit-for-tat* não cai na armadilha de enredar-se em complexas avaliações sobre a potencialidade da próxima rodada. Tal previsão, caso não se concretize (e a probabilidade nestes casos é apenas potencial, pois não há como ter certeza do próximo movimento) leva a uma série de rodadas de desconfiança mútua, o que retarda o restabelecimento da cooperação. Aliás, considerando a simplicidade do programa, é irrelevante o que o outro participante fez no movimento anterior ao último, pois essa informação sequer será computada. Ao mesmo tempo, a conduta futura do outro jogador somente é relevante para definir a jogada imediatamente posterior, ou seja, não se projeta uma estatística do jogo nem para frente nem retrospectivamente.

## 2.4 A evolução da cooperação

Como visto, é fundamental a questão da emergência da cooperação e da sua manutenção. A abordagem de enfrentamento de estratégias proposta por Axelrod traz comprovação formal do sucesso de *tit-for-tat*. O torneio de estratégias, contudo, não é capaz de avaliar a resiliência de cada estratégia.

Imagine o cenário específico em que a cooperação floresceu e, com isso, todos os integrantes desta população aderiram à utilização da mesma estratégia. Estaria tal população protegida contra a investida de outros que adotam estratégias eminentemente egoístas (como *ALL D*<sup>14</sup>, por exemplo)?

Outra pergunta que surge com sentido semelhante é: poderia uma população que está em equilíbrio egoísta (ou seja, todos utilizando somente a estratégia *ALL D*) ser convertida em uma população cooperativa?

Uma abordagem biológica evolucionista pode muito bem emprestar terminologia adequada para se compreender o que está sendo discutido. John M. Smith (1976) exemplifica um cenário no qual surge, população de indivíduos que emprega uma estratégia padrão de conduta, um indivíduo mutante que emprega estratégia diferente. Nestes termos, a *estratégia mutante* somente prevalece se for capaz de produzir resultado melhor que ganho já estabilizado na interação estabelecida pela estratégia padrão. Diz-se, neste caso, que a nova estratégia *invadiu* a estratégia padrão. É consecutivo lógico dessa abordagem evolucionista que uma estratégia coletivamente estável não pode ser suscetível a invasões.

A estabilidade cooperativa, por sua vez, é conceito mais esquivo. Como visto, estratégias benignas tendem a não se saírem tão bem quando confrontadas por estratégias maliciosas. Seria possível encontrar uma

---

<sup>14</sup> Conforme referido anteriormente, *ALL D* consiste em sempre denunciar.

estratégia cooperativa capaz de resistir à investida de estratégias mutantes mal intencionadas?

Com efeito, Axelrod (1984, 2006 p. 58-70), dispôs-se a testar o desempenho de *tit-for-tat* também enquanto estratégia estável. Para isso, lançou pressupostos e condições que permitissem que tal estratégia fosse imune à invasão de qualquer outra.

Cumpra lembrar que *tit-for-tat* é estratégia benigna que começa por cooperar e em seguida atua de forma retributiva. A interação com outras estratégias bem intencionadas é a melhor possível, contudo, quando em face de estratégias maliciosas é inevitável se deparar com a eventual opção egoísta (D).

Uma primeira condição é que *tit-for-tat* é estável se e somente se o coeficiente  $w$  é alto. Isso significa que ninguém que utilize outra estratégia alcançará resultado melhor que *tit-for-tat* quando o futuro lança sombra suficiente sobre o presente, conforme ilustra Axelrod (1984, 2006, p. 60).

Na verdade, qualquer estratégia que comece por cooperar, depende de um alto fator  $w$  para que possa ser considerada estável, já que em interações que não dependem de encontros futuros a tentação de agir egoisticamente é maior.

A maior dificuldade resta exatamente no confronto com estratégias egoístas. Com efeito, somente existe uma estratégia que é sempre coletivamente estável, e ela é a mais egoísta das estratégias. *ALL D* é a regra na qual sempre se denuncia o oponente. Ora, se o outro irá sabidamente denunciar, não existe razão alguma para cooperar.

Eis um problema considerável para a evolução da cooperação. Se tomada uma população inicial de não cooperantes, nenhum indivíduo estrangeiro agindo isoladamente terá condições de mudar o cenário estabelecido. A estratégia *ALL D* é estável e impossível de ser invadida por qualquer agente singular.

Num mundo de não cooperantes o indivíduo disposto a cooperar não encontra um par para agir reciprocamente, o que o impele a assumir a posição egoísta, mesmo que contra sua vontade.

Axelrod nota, contudo, que se os estrangeiros chegarem em pequenos aglomerados a situação se transforma, e a chance de se iniciar a cooperação, mesmo em um mundo de egoístas, emerge.

Para que isso ocorra, contudo, é necessário que a estratégia cooperativa (que pretende invadir em pequenos aglomerados) seja retributiva ao máximo. Convenientemente, *tit-for-tat* atende a esses pressupostos. Se considerados os valores de recompensa conforme a tabela do Dilema do Prisioneiro<sup>15</sup> e um jogo com dez repetições (coeficiente  $w=0,9$ ) o máximo de pontuação de cada indivíduo na população que utiliza *ALL D* como estratégia estável será 10 (10 vezes  $P=1$ ). Se um único invasor utilizando *tit-for-tat* for confrontado com os integrantes da população referida, ele será explorado na primeira jogada ( $S=0$ ), mas depois agirá de forma a retribuir na nove jogadas seguintes (9 vezes  $P=1$ ). Contudo, se dois ou mais invasores usando *tit-for-tat* interagirem entre si, eles conseguem uma pontuação de 30 (10 vezes  $R=3$ ) cada.

Dessa forma, tem-se que os indivíduos utilizando *ALL D* somam 10 pontos quando interagem entre si. Os indivíduos usando *tit-for-tat* somam 9 pontos quando interagem com os demais indivíduos da população estabilizada que utiliza a estratégia maliciosa. E, por fim, os invasores usando *tit-for-tat* somam 30 pontos quando interagem entre si. Axelrod conclui que Invasores utilizando *tit-for-tat* terão melhores resultados que os nativos da população que utiliza *ALL D* se ao menos 5% das interações ocorrerem entre indivíduos cooperantes<sup>16</sup> (Axelrod 1984, 2006, p.64).

---

<sup>15</sup> Ver figura 1 na p. 67 ( $T=5$ ,  $R=3$ ,  $P=1$  e  $S=0$ )

<sup>16</sup>  $30p + 9(1 - p) > 10$  or  $21p + 9 > 10$  or  $21p > 1$  or  $p > 1/21$ .

Assim, mesmo um pequeno grupo de indivíduos cooperativos pode invadir uma população que utiliza *ALL D* como estratégia estabilizada.

Se considerarmos um cenário com ainda mais iterações do jogo, maiores as chances de a cooperação se estabelecer (com ainda menos interações diretas entre os cooperantes invasores).

Mais além, a última proposição<sup>17</sup> de Axelrod demonstra que, uma vez estabelecida a cooperação baseada na reciprocidade, ela se mantém estável mesmo em face de invasores maliciosos agrupados.

A conclusão de Axelrod é reconfortante e oferece elementos consistentes para basear uma nova aproximação à teoria da escolha social com vistas à cooperação. Conforme a prova supramencionada a cooperação pode ser alcançada mesmo no mundo dos egoístas, sem uma autoridade central impositiva, mas a partir de um grupo de indivíduos que possam contar com reciprocidade (Axelrod 1984, 2006, p.69).

Ao final, a aplicação da noção evolucionista biológica de estabilidade coletiva à análise do comportamento social permite que se revelem tipos de estratégias capazes de prevalecer em face de qualquer outra (mesmo que maliciosas).

Cumprе ressaltar, contudo, que o fato de se encontrar uma estratégia cooperativa estável não quer dizer que tal estratégia seja a única conduta racional possível. Se de fato existir uma estratégia alternativa mais bem-sucedida, ela será apontada pelo indivíduo *mutante* mais cedo ou mais tarde. A proposição dessa nova estratégia, contudo, dentro do contexto social (de confronto e interação de estratégias), pode ocorrer por meio de procedimentos de deliberação consciente, ou mesmo através de avaliações contingentes de tentativa e erro. Da mesma forma, nada impede que sejam

---

<sup>17</sup> Proposition 7. If a nice strategy cannot be invaded by a single individual, it cannot be invaded by any cluster of individuals either. (Axelrod 1984, 2006, p.67).

estabelecidas exceções à conduta cooperativa (mormente como nos casos em que a conduta egoísta seja em benefício dos mais necessitados<sup>18</sup>).

A prioridade, portanto, deve ser alcançar e manter o equilíbrio cooperativo. O segundo passo deve ser o estabelecimento de mecanismos de correção de curso munidos com exceções à estratégia estabilizada.

O presente trabalho parte do pressuposto otimista de que *tit-for-tat* é uma estratégia estável e capaz de invadir mesmo a mais egoísta das estratégias.

Considerando a natureza iterada do jogo social, panorama no qual são feitas as escolhas políticas sobre a justa divisão de bens e valores, não é exagero assumir que o futuro sempre lança sombra suficiente a influenciar a decisão presente.

Com isso em vista, pretende-se estender os princípios subjacentes ao programa *tit-for-tat* como vetores válidos para informação da escolha social, como alternativa ao paradoxo de Arrow, e visando uma função de bem-estar social eficiente e não imposta.

---

<sup>18</sup> Neste ponto pode-se referir algo como a exceção constante no princípio da Diferença da teoria de John Rawls.

## **Escolha social pela cooperação justificada**

*The most fundamental idea in this conception of justice is the idea of society as a fair system of social cooperation over time from one generation to the next. We use this idea as the central organizing idea in trying to develop a political conception of justice for a democratic regime. (Rawls 2001, p. 5)*

A presente obra consiste na proposta de um modelo de escolha social pela cooperação justificada que visa uma função de bem-estar social não imposta, equilibrada e duradoura.

Conforme visto até aqui, uma solução cooperativa que se apresente naturalmente e sem a necessidade da imposição de uma autoridade externa é uma proposição otimista. Contudo, não se trata de um modelo idealizado e ingênuo, tampouco de argumento deontológico que pressupõe uma noção moral transcendental. A cooperação, mesmo entre egoístas é plenamente possível, e os requisitos para isso foram extensivamente apreciados no capítulo 2.

Também se pôde comprovar que uma população de estratégias cooperativas se mostra imune a invasão de estratégias maliciosas singulares. Tal situação é justamente o caso de um jogo continuado em cuja sombra do futuro exerça influência considerável sobre as decisões a serem tomadas no presente.

Isso significa que, desde um ponto de vista da soma de utilidade, uma vez que estabelecido um equilíbrio cooperativo, a postura puramente egoísta não tem chances de prevalecer. Num ambiente ótimo, onde não há quebra de confiança, os resultados coletivos são os mais favoráveis

possíveis. Assim, a exemplo do que ocorre na biologia, o invasor mal intencionado é identificado, discriminado e afastado.

O primeiro passo, portanto, deve ser alcançar o equilíbrio cooperativo.

### **3.1 *First things first*: estabelecendo o equilíbrio cooperativo**

Toda escolha social é, em alguma medida, escolha por cooperação. A escolha social que tende ao equilíbrio duradouro, por sua vez, somente alcança seu máximo potencial por meio da cooperação.

A escolha individual egoísta é dada de antemão como conduta racional. Para que essa racionalidade seja traduzida em termos intersubjetivos, um fator de cooperação deve estar contido no procedimento de formação da decisão coletiva, caso contrário, o indivíduo sair-se-ia melhor em confiar em suas escolhas puramente auto interessadas.

O que a presente obra propõe, como etapa inicial, é a urgência de se estabelecer um equilíbrio cooperativo, pois uma vez estabelecido tal estado de coisas, a possibilidade de continuidade de estratégias e condutas mal intencionadas é drasticamente reduzido. Com isso surge a necessidade de se estabelecer um procedimento capaz de otimizar a conduta dos indivíduos em direção a este equilíbrio.

O cenário que será aproveitado é o mesmo desenvolvido na obra de Axelrod: a possibilidade de surgimento de cooperação mesmo entre agentes puramente egoístas. Cumpre destacar que o termo cooperação<sup>1</sup> aqui assumido não se refere, ao menos nesta primeira fase, a uma forma de avaliação qualitativa intersubjetiva do resultado da escolha social.

A concepção de cooperação, nesta etapa inicial, é direcionada para uma avaliação utilitarista de maximização do resultado social por meio da

---

<sup>1</sup> Vale estressar que a cooperação nos termos aqui expressos, não trata de valores subjetivos psicológicos como o prazer altruísta que algum participante possa experimentar em sacrifício do seu ganho de bem-estar.

ação cooperativa de indivíduos auto-interessados (isto é, sem considerar altruísmos e autoridades externas). Com efeito, a cooperação somente se sustenta em longo prazo se proporcionar resultado coletivo melhor que a conduta meramente egoísta.

As noções de *estratégias mutantes* e de *invasão* (Smith 1976), emprestadas da abordagem biológica evolucionista, são de fundamental importância na definição da melhor maneira de se jogar o jogo continuado da escolha social.

Segundo o jogo proposto por Axelrod (1984, 2006, p. 59) existe uma função dentre os parâmetros de resultados (T, R, P e S) que permite que uma estratégia cooperativa seja coletivamente estável. Tal situação é justamente o caso de um jogo continuado em que a sombra do futuro<sup>2</sup> exerça influência considerável sobre as decisões a serem tomadas no presente.

As conclusões lógico-axiomáticas, dispostas por Axelrod (1984) e discutidas no segundo capítulo desta obra, atestam a estabilidade de uma população de estratégias cooperativas do tipo *tit-for-tat* e sua consequente imunidade às estratégias invasoras.

Mais além, restou comprovado (Axelrod 1984, 2006, p. 64) que mesmo um pequeno grupo de indivíduos dispostos a cooperar nos parâmetros de *tit-for-tat* pode invadir uma população que utiliza estratégia eminentemente egoísta, fazendo emergir a cooperação, sem necessidade de uma autoridade impositiva externa. Isso ocorre por que a interação entre os indivíduos cooperantes, mesmo que ainda não predominante naquela população, eleva o resultado coletivo para níveis maiores que os alcançados pela estratégia maliciosa nativa.

Tal conclusão oferece argumentos sólidos para uma proposta otimista em relação à noção de cooperação social. O jogo social, considerando

---

<sup>2</sup> O coeficiente  $w$  (Shubick 1970) deve ser alto o suficiente.

seu aspecto intersubjetivo, não é jogo de soma zero. Trata-se de jogo indefinidamente continuado no qual a opção cooperativa é constante.

O aspecto da continuidade implica dois corolários: a) que os jogadores serão capazes de memorizar o histórico de jogadas passadas; e b) que a decisão futura deve lançar uma sombra sobre o presente, influenciando a escolha da estratégia a ser tomada no jogo corrente. Assim, escolhas presentes não determinam somente o jogo atual, mas influenciam as escolhas subsequentes dos jogos futuros.

A liberdade de escolha<sup>3</sup> também deve ser considerada como aspecto intrínseco a esta fase de estabelecimento da cooperação. Assim, cada jogador deve sempre considerar a cada interação todas as estratégias possíveis para si e para o outro participante.

Neste momento não são permitidos constrangimentos<sup>4</sup>, compromissos e barganhas. Também, não há como saber, de antemão, qual será a jogada do outro participante<sup>5</sup> (o que impossibilita uma meta análise de informações sobre intenção dos jogadores). Não há espaço, portanto, para promessas não garantidas por ação. A comunicação entre os participantes somente se extrai da sequência de escolhas dos participantes. O que permite a cooperação é o fator de continuidade, pois os jogadores sabem que irão encontrar-se novamente no futuro.

Nota-se que não cabe proposição discursiva ou deliberação, porquanto o que está sendo avaliado é primordialmente a escolha inicial de se participar de (e permanecer em) um sistema de cooperação.

Um último e fundamental critério do jogo social é que o resultado (seja cooperativo ou não) deve refletir uma função social de associação

---

<sup>3</sup> Pode-se fazer o paralelo com a condição do (U) Domínio Irrestrito de Arrow (1951, p. 24), no que tange à escolha social.

<sup>4</sup> Aqui se trata da mesma condição de não-Ditadura (D) da teoria de Arrow (1951, p. 30).

<sup>5</sup> Pode-se argumentar a relação tanto ao critério IAI (Independência das alternativas irrelevantes) de Arrow (1951, p. 26) como à natureza individual da escolha que não considera comparação interpessoal de utilidade.

positiva<sup>6</sup>. Ou seja, o resultado do jogo somente pode melhorar a condição inicial dos participantes. De fato, se o indivíduo pudesse considerar aderir a um jogo continuado futuro que o deixasse em condição pior que a atual, não haveria razão para participar.

Para aderir a um sistema de cooperação, contudo, não só deve haver real vantagem no resultado refletido pela função de bem-estar, mas a vantagem deve ser maior que o resultado da conduta egoísta.

Cumprе relembrar que no exemplo de Axelrod (1984, 2006, p. 59), mesmo a conduta egoísta era eficiente em termos paretianos, pois o participante que agisse egoisticamente sempre conseguiria resultado melhor que o inicial<sup>7</sup>. Isso fazia da escolha egoísta não só uma escolha válida e eficiente, mas a melhor escolha em jogos singulares (nos quais os participantes não voltarão a interagir novamente no futuro), pois a tentação da denúncia é mais alta.

Assim, para que emerja a cooperação, sem imposição externa, deve-se oferecer vantagem mais atraente que a proporcionada pela mera postura auto-interessada.

Dadas essas características do jogo social, a cooperação pode emergir mesmo entre egoístas, sem necessitar de imposição externa. Para isso é necessário que ao menos um grupo de participantes inicie o ciclo de cooperação, valendo-se de uma estratégia retributiva do tipo *tit-for-tat*.

A programação de *tit-for-tat* consiste em apenas duas diretrizes fundamentais: a) comece cooperando (não seja o primeiro a quebrar a confiança); e b) pratique a reciprocidade (tanto da cooperação como do egoísmo).

---

<sup>6</sup> Referência ao critério (P) na teoria de Arrow (1951, p.25).

<sup>7</sup> Mesmo o pior desfecho, que consistia em cooperar quando o outro participante não cooperasse gerava uma situação paretianamente eficiente (Axelrod, 1984 p. 59).

A reciprocidade é fator de destaque nessa fase de instalação do equilíbrio cooperativo, pois é através desta diretriz que os demais participantes sentir-se-ão desencorajados a tentar a vantagem pessoal em detrimento da cooperação.

A aposta otimista gira em torno do surgimento de um momento de virada em que os participantes compreendam que, em longo prazo, a alternativa cooperativa gera mais ganho que a conduta egoísta. Contudo, como já apodado anteriormente, trata-se de otimismo fundado em boas probabilidades, uma vez que a continuidade indeterminada<sup>8</sup> imanente ao jogo social garante que essa compreensão acerca do benefício da cooperação se desvelará no decorrer do jogo.

Não há como negar que esta primeira etapa de busca pelo equilíbrio cooperativo é utilitarista, mormente se considerada a concepção de cooperação no mundo de egoístas, conforme o tema desenvolvido no segundo capítulo. A escolha por essa metodologia, no entanto, busca justamente corroborar que mesmo no cenário improvável de puro auto-interesse, e sob as condições mais objetivas, a cooperação pode emergir.

A prova formal, aqui apresentada, da possibilidade de emergência da cooperação é argumento contundente que afasta a desconfiança já arraigada no senso comum de que a cooperação é algo intangível e efêmero. De fato, essa desconfiança é fruto da suposição que confunde cooperação com formas de altruísmo, solidariedade e sacrifício<sup>9</sup>. Se tomada a noção utilitarista, despida dessas suposições, a cooperação assume contornos bem tangíveis. Se, além disso, estiverem presentes os critérios da reciprocidade

---

<sup>8</sup> O coeficiente  $w$  deve ser maior que 0,5. Axelrod (1984, p. 65) aponta que interação social de um jogo com 10 iterações corresponde a um coeficiente  $w = 0,9$ . Por sua vez 200 iterações corresponde a um coeficiente  $w = 0,99654$ . Ou seja, somente jogos com uma expectativa de interação menor que 5 vezes não projetam uma sombra capaz de induzir a cooperação.

<sup>9</sup> Pode-se arguir quanto a isso que todas as formas de altruísmo, solidariedade e sacrifício são formas sutis de interesse próprio (Elster 2016, p.10).

e continuidade das relações sociais, a cooperação deixa de ser mera probabilidade, e torna-se a estratégia dominante.

Penso que a solução viável sobre escolha social deve se acoplar ao modelo utilitarista, e não tentar suplantá-lo ou desviá-lo por completo<sup>10</sup>. Esse acoplamento deve adicionar parâmetros que conduzam a atuação coletivamente desejada.

Em outras palavras, a conduta auto-interessada do indivíduo não desaparece, mas deve ser traduzida de forma intersubjetiva, de preferência sob a interface de uma estratégia que iniba a conduta egoísta dos demais participantes. Trata-se, portanto, de procedimento que direciona a escolha individual inicial rumo à cooperação social.

### 3.2 Mantendo o equilíbrio cooperativo

A prioridade, conforme demonstrado, deve ser alcançar o equilíbrio cooperativo. O segundo passo deve ser a manutenção do equilíbrio e o estabelecimento de mecanismos de correção de curso munidos com exceções à estratégia estabilizada.

Cumpra ressaltar, antes de tudo, que o fato de se encontrar uma estratégia cooperativa estável não quer dizer que tal estratégia seja a única conduta racional possível daí em diante. Se de fato existir uma estratégia alternativa mais bem-sucedida, ela será apontada pelo indivíduo *mutante* mais cedo ou mais tarde.

A proposição de uma nova estratégia, contudo, dentro do contexto social (confronto e interação de estratégias), pode ocorrer: a) por meio de procedimentos de deliberação consciente; b) através de avaliações contínuas de tentativa e erro ou c) pela combinação dessas duas formas.

---

<sup>10</sup> Nesse sentido a lição de Honneth: “It is important to stress, however, that this second model of conflict, based on a theory of recognition, should not try to replace the first, utilitarian model but only extend it. It will always be an empirical question as to the extent to which a social conflict follows the logic of the pursuit of interests or logic of the formation of moral reactions.” (Honneth 1995, p.165)

Ademais, nada impede que sejam estabelecidas exceções à estratégia estabilizada (mormente como nos casos em que a conduta egoísta seja em benefício dos mais necessitados<sup>11</sup>), sem que haja um abandono completo da estratégia cooperativa.

O critério utilitarista adotado na etapa de implementação do equilíbrio cooperativo não é suficiente para a manutenção de um sistema de cooperação social, porquanto estritamente focado no aspecto formal da cooperação, isto é, no somatório mais eficiente de resultados gerais de bem-estar. É preciso um mecanismo mais complexo que dê conta, por exemplo, de avaliações interpessoais de utilidade e de soluções que desafiem a maximização paretiana.

Daí a necessidade de agregar uma segunda fase ao procedimento, na qual o foco deve voltar-se para a questão da cooperação social entre cidadãos racionais livres em sociedades minimamente razoáveis<sup>12</sup>. Neste segundo estágio a escolha deve ser ampla com um domínio irrestrito. Não se está mais diante da decisão binária entre cooperar ou não, mas de um cenário de embate discursivo no qual indivíduos são dotados de autonomia para avaliar, inclusive, os fatores que levaram ao equilíbrio cooperativo em primeiro lugar.

Com efeito, o que sugere que uma sociedade democrática seja entendida como um sistema de cooperação social é o fato de que, do ponto de vista político, e no contexto da discussão pública de questões básicas de direito político, *seus cidadãos não consideram sua ordem social como uma ordem natural fixa, ou como uma estrutura justificada por doutrinas religiosas ou princípios hierárquicos expressando valores aristocráticos* (Rawls 2001, p.6).

---

<sup>11</sup> Neste ponto pode-se referir algo como a exceção do princípio da Diferença da teoria de John Rawls.

<sup>12</sup> Também em referência a obra de John Rawls, é necessária uma *sociedade bem ordenada*, minimamente capaz de considerar a importância da autonomia das escolhas individuais e as vantagens da cooperação social.

Assim, para desviar da imposição social autoritária e do utilitarismo clássico são necessários vetores morais que direcionem a atuação política em torno de uma concepção de justiça. Neste sentido, existem brilhantes e amplas soluções já propostas com diferentes bases: com foco em princípios deontológicos como em Rawls (1971), com viés deliberativo em Habermas (1999), e mesmo como forma mais contingente de combate às desigualdades como em Sen (2000) e Nussbaum (2006).

A proposta do presente trabalho, contudo, é mais singela. O que se está sugerindo é um procedimento justificativo complementar que seja capaz de legitimar a cooperação social e que possa ser acoplado ao modelo utilitarista de modo a conduzir a uma função de bem-estar que represente mais que o somatório de escolhas e resultados individuais.

Não se trata, portanto, de uma teoria sobre justiça propriamente dita. O esforço se concentra muito mais na identificação e coibição da atuação nociva ao equilíbrio social que na busca de valores universais (que *a priori* possam determinar a atuação política).

É necessário, todavia, que se parta da noção de sociedade como um sistema justo de cooperação social. 'Justo', aqui, não se refere ao conteúdo da decisão política em si, mas ao procedimento que gerencia a relação intersubjetiva dos indivíduos em prol da manutenção cooperativa.

Para isso será aproveitada a definição de cooperação social como disposta por John Rawls (2001) em *Justiça como Equidade, uma Reformulação*. A teoria de Rawls é admitidamente baseada na ideia de cooperação, sendo *o papel dos princípios de justiça (como parte da concepção política de justiça) especificar os justos termos de cooperação social* (Rawls 2001, p.7).

Como características pode-se dizer que: a) a cooperação social é guiada por regras previamente reconhecidas e procedimentos que devem ser aceitos pelos agentes. Cooperação, neste aspecto, não pode ser confundida

com coordenação (na qual a atividade pode, por exemplo, ser determinada por uma autoridade central); b) a ideia de cooperação inclui a noção de justos termos de cooperação, com os quais os agentes devem concordar, contando que os demais agentes também atuarão conforme tais critérios. Estes termos refletem a ideia de reciprocidade; c) cooperação inclui a ideia de vantagem racional, aqui considerada como vantagem buscada desde o ponto de vista de cada agente perseguindo seu bem-estar.

O primeiro aspecto refere-se às claras “regras do jogo”. A cada “jogada” o indivíduo deve ser capaz de avaliar as consequências de suas escolhas de acordo com um padrão bem definido de resultados possíveis. É importante pontuar que a liberdade de escolha deve estar presente e a solução, portanto, não pode ser ditatorial, porquanto não mais se trataria de ação cooperativa, mas sim de coordenação.

O segundo ponto, trata dos justos termos de cooperação, o que remete ao procedimento instalado na ordem social para a manutenção da cooperação social. O procedimento justo, como já referido, é um critério que deve valer para todos de maneira a garantir a continuidade da cooperação por meio da confiança no sistema de reciprocidade.

Por fim, o terceiro atributo ressalta a racionalidade do agente cooperativo. Deve ficar claro que quando se está falando de cooperação social não se pode exigir conduta que esteja além da ação racional do indivíduo.

Neste sentido a redistribuição de bens e valores é exceção ao equilíbrio cooperativo. Devem ser tratadas como questões complementares, regidas por procedimentos específicos, que devem definir por meio de discurso público os paradigmas de distribuição equitativa (de bens, valores e oportunidades) a serem (re)incorporados nas “regras do jogo social”. Não se pode, portanto, confundir cooperação nem com altruísmos nem com princípios de redistribuição de bens e valores.

Um bom paralelo para explicar a dinâmica entre a cooperação e as exceções possíveis ao equilíbrio cooperativo é o princípio da diferença conforme desenvolvido em *Uma Teoria da Justiça* (Rawls 1971 p. 53).

A pergunta fundamental para Rawls é: qual a mais aceitável concepção política de justiça para especificar os termos justos de cooperação entre cidadãos considerados livres e iguais (e como razoáveis e racionais) como membros normais e plenamente cooperantes da sociedade ao longo de uma vida completa, de uma geração para a outra? (Rawls 2001, p.8). A sua resposta são os princípios de justiça. O primeiro princípio, também referido como princípio da liberdade, estabelece as bases da cooperação social: direitos iguais ao maior esquema de liberdade que seja compatível com esquema similar de liberdades dos outros (Rawls 1971 p. 53). O segundo princípio, denota uma exceção ao primeiro, porquanto trata dos casos em que desigualdades sociais e econômicas seriam toleradas, sem colocar em risco a estabilidade da cooperação garantida pelo primeiro princípio.

Conforme já ressaltado, o presente trabalho não tem a pretensão de definir a liberdade, equidade ou qualquer outro valor moral *a priori* como primordial ao estabelecimento do equilíbrio cooperativo. Contudo, guardadas as proporções, é essa a relação entre os procedimentos que se pretende desenvolver.

A exceção ao princípio cooperativo (estabelecido pela confiança projetada pela sombra do futuro na decisão presente e pelo critério de reciprocidade) não é capaz de colocar em risco o equilíbrio cooperativo. Uma exceção em benefício dos menos favorecidos, por exemplo, deve ser considerada uma “bolha” que excepciona o critério de reciprocidade. Tal exceção deve ser autorizada por todos os agentes que devem concordar em manter esse estado de coisas alternativo em paralelo.

Duas questões são dignas de nota quanto ao que foi referido acima. Primeiramente, fala-se em “bolha”, pois é situação frágil que deve ser

tratada com especial cuidado. Como visto no capítulo segundo deste trabalho, uma população em equilíbrio cooperativo não respeita a exceção maliciosa. Num ambiente ótimo, onde não há quebra de confiança, os resultados coletivos são os mais favoráveis possíveis. Assim, a exemplo do que ocorre na biologia, o invasor mal intencionado é identificado, discriminado e afastado. Daí a necessidade de se identificar a política de benefício aos menos favorecidos como uma exceção tolerável, para que não seja expelida do sistema de cooperação.

O segundo ponto refere-se novamente a noção de cooperação. Cumprir destacar que a cooperação, em seu aspecto racional, mesmo nesta segunda etapa, é considerada em função do bem-estar individual dos agentes. Trata-se de uma avaliação objetiva acerca do que significa cooperar. Assim, desde o ponto de vista racional, não há cooperação em uma medida de redistribuição, mas sim uma exceção, porquanto o agente beneficiado não está necessariamente retribuindo o benefício de maneira eficiente (paretianamente) aos demais agentes.

Essas exceções, contudo, são fundamentais para o bom funcionamento do equilíbrio cooperativo. Apesar de não estarem sendo apreciados os requisitos de justiça em si, é indispensável que um equilíbrio duradouro guarde critérios justos de restabelecimento de equidade. Deve haver mecanismos capazes de distinguir entre a conduta puramente maliciosa e a ação permitida pelos demais agentes em prol dos que foram preteridos em alguma circunstância do jogo social.

Exceções, cumprir ressaltar, são tão importantes quanto a cooperação em si e, diga-se mais, considerando o modelo aqui proposto, é nelas que o elemento moral se faz presente, corrigindo e direcionando o curso da cooperação social para refletir um sistema legítimo que faça mais que somente a agregação maximizada de preferências.

Isso a racionalidade econômica não consegue, por si só, dar conta. Se tomado o critério utilitarista de eficiência paretiana, por exemplo, não há como discriminar o que motivou a conduta não cooperativa. Apenas é possível a análise do resultado da conduta em relação a maximização de bem-estar. Uma ação de exceção, nestes termos, não poderia ser diferenciada de uma conduta maliciosa.

Essa segunda etapa do procedimento deve, portanto: a) permitir escolha com um domínio irrestrito, por meio de um cenário de embate discursivo no qual indivíduos são dotados de autonomia para avaliar, inclusive, os fatores que levaram ao equilíbrio cooperativo; b) servir de mecanismo capaz de definir as exceções razoáveis ao sistema de cooperação; c) ser capaz de discernir entre condutas mal intencionadas e as exceções devidamente autorizadas ao equilíbrio cooperativo.

Tais quesitos representam sérios desafios à escolha social, porquanto implicam, de uma só vez, a relativização da eficiência paretiana e a inclusão de comparações interpessoais de utilidade. Se, como anunciado anteriormente, se pretende desviar de uma solução que determine os parâmetros de justiça e vida boa *a priori*, o modelo que melhor se adequa à essa segunda fase do procedimento é o da justificação discursiva.

A justificação imprime transparência ao procedimento, agindo na identificação de estratégias verdadeiramente malignas, e condutas que comprometam o equilíbrio já alcançado.

Além disso, é necessário um modelo que represente uma solução intersubjetiva e que permita reciprocidade imediata e na medida certa. Tanto a recompensa pela cooperação como a retaliação da atuação maliciosa devem ser apreciadas de modo rápido e contínuo, de tal sorte que reste garantida a longevidade da relação interpessoal.

Para isso o modelo aqui proposto busca elementos no procedimento discursivo-deliberativo, tal qual proposto por Habermas e desenvolvido

por Forst. Trata-se de alternativa à racionalidade limitada que o modelo de escolha individual oferece.

A solução discursiva justificada se apresenta como uma possível resposta ao déficit de representação e legitimidade que maculam os mecanismos clássicos de escolha.

A concepção de *justificação reflexiva*<sup>13</sup> (Forst, 2014) desponta como alternativa para imprimir decisões autônomas, cooperativas e não autoritárias. Nesse modelo deve haver espaço sempre aberto para a deliberação e uma racionalidade justificada (para além do liberal paretiano) que garanta a autonomia, a não arbitrariedade e sirva, analogamente, como estratégia majoritária para coibir a atuação não-cooperativa.

A justificação, nestes termos, passa a ser considerada não só como direito (Recht auf Rechtfertigung), mas como dever, visando suprir o déficit informacional na tomada de decisão. Na teoria da escolha social as preferências dos agentes são tomadas como informações meramente individuais, a partir das quais se buscam estimativas que deem conta de prever os resultados das interações entre os participantes do jogo contínuo que representa a vida em sociedade.

Em outras palavras, procura-se projetar mecanismos tais que, quaisquer que sejam as preferências dos agentes, suas escolhas estratégicas levem aos mesmos resultados que a função de bem-estar social pré-determinou. Tal postura invariavelmente leva a resultados contraditórios e aparentemente incompatíveis com a verdadeira intenção do agente no ato da escolha. Isso representa um típico caso de eleição social mal informada.

Em face disso, a abordagem justificativa aqui proposta passa a funcionar como um planejador central que, munido das declarações dos próprios agentes, é capaz de avaliar melhor a racionalidade que informa a

---

<sup>13</sup> A questão da justificação é tema de destaque em qualquer proposição de teoria de justiça. O conceito de justificação reflexiva desenvolvido por Forst é o modelo sobre o qual o presente trabalho se inclinará.

decisão social podendo, inclusive, prever e detectar eventuais erros e anomalias, tanto no processo de seleção como no resultado deliberado.

Ao mesmo tempo, a escolha social justificada funciona como solução que traduz a vontade individual em vontade social legítima, sem perder o embasamento racional desde o ponto de vista de cada participante da sociedade.

O diálogo sempre aberto (domínio irrestrito) deve permitir ordenamentos coerentes sem que haja a imposição da vontade arbitrária externa (não ditadura).

### 3.3 Justiça e justificação

As *Instituições Justinianas*, uma espécie de introdução ao Direito Romano, trazem a seguinte passagem logo em suas linhas iniciais: *Iuris praecepta sunt haec: honeste vivere, alterum non laedere, suum cuique tribuere*<sup>14</sup>. Trata-se da enunciação exordial dos preceitos máximos da Lei Romana baseados na noção de justiça distributiva mais difundida desde a antiguidade: dar a cada (um) o que é devido<sup>15</sup>. Este antigo princípio, aqui referido como *suum cuique*, é um critério que possui inegável mérito, figurando no nosso imaginário de Justiça desde Platão, centrando-se na questão primária: *quem recebe o quê?*

Nota-se, desde já, que o que está em jogo é a justa distribuição de bens entre os indivíduos. Trata-se de uma visão distributiva e orientada aos destinatários desses bens, enquanto participantes de um sistema social que lhes deve garantir o suficiente para que gozem de uma vida boa.

Em que pese a inegável eficiência desse critério, considerando-se a justa repartição de bens materiais por via do Direito Civil, e.g., essa noção

---

<sup>14</sup> Os Preceitos da lei são os seguintes: viver honestamente, não lesar ninguém e dar a cada um o que é devido. (Inst. 1,1,3-4); Institutiones Justiniani, 1535.

<sup>15</sup> Também se pode traduzir como: *a cada (um) o que é seu*.

negligencia outros aspectos essenciais da Justiça. Um primeiro ponto que pode ser destacado é que uma visão puramente baseada em bens materiais não evidencia a origem dos bens a serem distribuídos, ignorando, muitas vezes, as relações de dominação subjacentes na etapa da produção dos bens em si. Da mesma forma, a figura de *quem distribui* não fica bem delineada. Surge a necessidade de um *distribuidor neutro*, um programador da métrica de justiça. Tal abordagem mecânica coloca em dúvida a própria justiça enquanto realização política, porquanto considera os indivíduos como meros recipientes passivos.

Ainda, pode-se destacar que uma visão de justiça puramente centrada na distribuição de bens deixa de enfrentar da maneira devida à questão da *injustiça*, já que não há diferenciação de contextos em que ocorrem as privações de bens. Assim, tratar-se-ia da mesma maneira um indivíduo que sofreu privação em razão de uma catástrofe natural e outro que foi privado de sua parte mercê de exploração econômica ou política. É necessário que o espectro semântico da justiça seja capaz de discernir entre um ato de solidariedade moral e outro de obrigação civil.

Para Forst (2014) justificadas reivindicações a bens não surgem espontaneamente, mas somente podem ser alcançadas por meio do discurso, nos contextos dos respetivos procedimentos de justificação, nos quais todos podem participar como indivíduos livres e iguais. Este é o requisito fundamental de justiça e, concomitantemente, é a base para a interpretação teórico-discursiva de uma noção alternativa de justiça. Essa outra figura da justiça, mais completa e apta, deve ser voltada para relações e estruturas intersubjetivas e não focada em estados subjetivos ou supostamente objetivos. Somente assim, buscando uma resposta satisfatória para questão da justificação das relações sociais, por meio da verificação do real “poder de justificação” que detém cada cidadão ou grupo, pode-se desenvolver uma concepção de justiça radical. Nas palavras de Forst: *a questão*

*básica da justiça não é uma questão de o que você tem, mas de como você é tratado* (Forst, 2014, p.19).

Na busca dessa noção mais adequada, que contemple algo a mais que simplesmente o princípio da *suum cuique*, deve-se voltar ao próprio conceito de justiça. Resta implícita na definição daquilo que é justo a necessidade de afastar a arbitrariedade, porquanto a existência de comandos arbitrários leva a formas de dominação injustificadas. Não se busca, contudo, eliminar todas as diferenças entre os indivíduos em prol de um conceito intangível de justiça metafísica. O contexto deve permanecer relevante. O fundamental para a questão da justiça é a supressão da arbitrariedade como forma de dominação, não a eliminação das contingências. *A Justiça deve permanecer uma tarefa humana que visa a não dominação, não uma tarefa para os deuses num mundo livre de contingências naturais e históricas* (Forst, 2014, p.21).

A dominação, neste cenário, significa controle arbitrário sobre outros. Note-se que, no caso, a injustiça surge quando o controle ocorre sem a devida justificação. Portanto, pode-se perceber que o impulso que afasta a injustiça não é necessariamente o de querer um bem, mas, antes de tudo, a vontade de não ser dominado de forma injustificada.

A noção alternativa de justiça que Forst propõe prima pelo direito básico à justificação. Assim, nenhuma relação política ou social deve ser imposta se não puder ser justificada adequadamente em face dos que estão envolvidos. Esta é a faceta que fica suprimida na abordagem de justiça centrada prioritariamente no indivíduo como recipiente de bens, como no princípio da *suum cuique*. A luta por justiça deve ser emancipatória. Em termos reflexivos, isso significa que cada um deve ser respeitado enquanto sujeito e destinatário, concomitantemente, das justificações normativas.

### **3.3.1 O caráter relacional da justiça**

O recente debate sobre equidade pode ajudar a reforçar a noção reflexiva e intersubjetiva da justiça em contraposição a um ideal de justiça subjetiva ou mesmo objetiva: justiça é sempre uma matéria relacional.

Afinal, equidade de quê? De recursos, de bem estar ou de capacidades?<sup>16</sup> Onde deve estar o foco da justiça, no resultado ou no procedimento? Martha Nussbaum (2006) critica a noção de *justiça procedimental pura*, assim desenvolvida por John Rawls (1971, p. 74), arguindo que os defensores de uma visão baseada no resultado consideram prioritariamente a qualidade de vida dos indivíduos. Assim, ao final, sempre rejeitarão a solução procedimental, por mais elegante que se apresente, se ela não conduzir ao resultado esperado pelas instituições no que toca à dignidade e justiça.

Com efeito, a concepção proposta por Amartya Sen (2000) prioriza o resultado ao procedimento puro. A justiça serve como um instrumento que deve produzir algo palpável, não podendo ser reduzida a mera engrenagem interna de uma máquina. Deste modo, serve como exemplo de teoria ambivalente que engloba ambas as noções de justiça aqui apresentadas.

A abordagem de Sen é sistema que analisa as reais capacidades para alcançar as funcionalidades (econômicas ou morais). Essas funcionalidades, por sua vez, podem ser valoradas de forma diferente e por motivos diversos em cada situação. Busca-se, com isso, afastar a ideia de que existe uma só racionalidade para a escolha, gerando uma maneira particular de avaliar igualdades entre as pessoas (e mesmo entre nações).

Justiça não é um critério universal de quantidades de bens, tampouco a erradicação completa de privação. Como já exposto, para Forst (2014, p. 27) o objetivo primário da justiça não é que os indivíduos possam obter

---

<sup>16</sup> Ver Cohen, "Equality of what? On welfare, goods and capabilities".

certos bens, mas que eles sejam agentes dotados de direitos iguais dentro de um contexto social, nacional ou internacional, e que possam levantar reivindicações específicas a bens sob esse fundamento.

Sen (2011) critica as teorias ideais sobre justiça, arguindo que a acessibilidade comparativa da qualidade de vida e da justeza de uma sociedade podem ser avaliadas mesmo em face de desacordos em relação ao que seria a justiça perfeita. Nessa esteira, a abordagem das capacidades serve tanto como fundamentação da justiça material como da razão pública como meio de justificação.

Forst (2014) destaca que a sua concepção de justiça, a exemplo da de Sen, preza pelo caráter relacional e estruturado, o qual não busca uma teoria ideal, mas questiona as relações sociais de controle ou dominação que existem e precisam ser transformadas em relações justificáveis.

Contudo, apesar desse paralelo, Forst destaca que a abordagem de Sen, se analisada de um ponto de vista relacional e voltado à prática, se omite em pontos importantes, principalmente nas questões da injustiça, das obrigações, dos princípios e das instituições da justiça.

A teoria de Sen (2011, p. 23) é direcionada à eliminação ou diminuição de injustiças concretas. A origem das assimetrias de capacidades, se tomada uma concepção de justiça material, é tema que faz grande diferença. O autor indiano é plenamente consciente de que há uma diferença entre pessoas famintas em decorrência de circunstâncias que estão além do seu controle e pessoas sucumbindo à fome por manobra de outras que planejaram esse resultado (ou ao menos o previram e não evitaram). Contudo, em sendo a teoria das capacidades primariamente centrada nos resultados, seu potencial de integrar essas distinções à estrutura básica da justiça é limitada.

Outra objeção apontada por Forst é relativa às obrigações. A justiça, de acordo com a visão relacional, questiona *as relações entre pessoas*.

Desta forma, a responsabilidade é variável, porquanto deve abarcar desde aqueles que deliberadamente cometeram uma injustiça, àqueles que apenas se beneficiaram, até os que somente se envolvem de maneira auto-interessada. Sen adota uma concepção consequencialista em relação à obrigação, baseada nas capacidades e nos poderes efetivos de ação, assumindo uma visão de justiça imperfeita e direcionada a certos resultados.

O argumento principal, segundo Forst (2014, p.30), deve presar pela prioridade dos princípios de igual respeito, participação e não dominação dentro das estruturas básicas de uma sociedade ou mesmo entre contextos transnacionais. O princípio fundamental de distribuição deve ser aquele que determina *quem* tem a autoridade para decidir.

### **3.3.2 Princípio da Justificação Geral e Recíproca**

Ao assumir a necessidade de uma justificação geral e recíproca, Forst reconhece a importância dos dois aspectos essenciais à justiça reflexiva, os quais consubstanciam sua teoria: a autonomia e a natureza discursiva da justificação.

A autonomia figura no núcleo do processo normativo. Ao tratar do tema, o autor vale-se da teoria de John Rawls, a qual considera um exemplo que melhor se enquadraria na segunda noção de justiça (a reflexiva). Adotando a concepção kantiana, baseada na noção de autodeterminação dos indivíduos, Rawls sugere que a autonomia de cidadãos livres e iguais não deve se restringir ao indivíduo enquanto mero recipiente passivo de bens necessários a uma vida boa, mas deve ser vista como força construtiva de sujeitos à justificação. O conceito essencial de autonomia é aquele ligado a determinação ativa da estrutura básica da sociedade. Portanto, o indivíduo somente é considerado autônomo se for, ao mesmo tempo, formador e destinatário das normas que lhe são impostas, em outras

palavras, somente deve se submeter às regras nas quais ele próprio participou do procedimento de justificação.

Forst (2014, p.32) destaca que o ponto crucial é a ausência de controle social injustificado, melhor dizendo, a confirmação de um estado de não dominação. Esta seria a forma mais justa de evitar arbitrariedade social e as relações de controle que dela decorrem. É nisso que consiste a noção de justiça como equidade, desenvolvida por Rawls, em contraponto à ideia de justiça alocativa do tipo *suum cuique*.

Sem embargo, é na concepção teórico-discursivo que a abordagem de Forst se diferencia da posição adotada por Rawls. Como vimos, é na natureza discursiva que são contemplados os princípios fundamentais de justiça referentes ao *direito de justificação*.

Forst define justiça como uma virtude humana e um imperativo político-moral capaz de se opor às relações de controle arbitrário ou de dominação. Dominação, por sua vez, nada mais é que o controle injustificado, o qual não deve ser permitido em uma sociedade justa, porquanto é pressuposto que os indivíduos sejam livres e iguais para assentir sobre as reivindicações normativas. Eis o *princípio supremo da justificação geral e recíproca* (Forst, 2014, p. 33), no qual não cabe a cada participante simplesmente projetar suas razões, mas sim, devem justificá-las discursivamente.

Aqui é inegável a influência da teoria habermasiana, em cima da qual se desenvolve o trabalho de Forst, que assume que cada sujeito de um determinado *contexto de justiça* possui o *direito fundamental à justificação*, ou seja, o direito de oferecer razões apropriadas para as normas de justiça que se aplicarão de maneira geral. Toda a norma posterior a isso é relacional, pois deve ser construída por meio de procedimento recíproco e geral de justificação. Tal critério serve como uma espécie de filtro para as reivindicações por privilégios.

Ao final, as duas noções de justiça apresentadas por Forst são complementares. Talvez a diferença consista em duas ideias morais de seres humanos: por um lado a garantia de certos bens necessários a uma vida boa, por outro a noção de dignidade que advém da postura de não se deixar dominar de forma injustificada (Forst, 2014 p. 37).

### **3.3.3 Um modelo de Justificação**

Em filosofia política a questão dos princípios que legitimam o justo exercício dos comandos políticos e das normas jurídicas é assunto frutífero, há muito debatido, e a respeito do qual foram desenvolvidas inúmeras teorias, desde concepções puramente ideais de justiça, até as mais voltadas aos resultados práticos, voltados a vida digna. Eis a dificuldade de se encontrar a correta abordagem metodológica.

Deve-se partir de uma teoria racional, que contenha princípios morais abstratos para, então, encontrar uma aplicação prática? Ou a fonte da justiça deve ser a realidade concreta, a partir da qual devem ser analisadas as aplicações mais justas, dados os contextos políticos em que se apresentam?

A resposta, passa pela virada reflexiva à questão da justificação político-normativa. Tomando o viés da Justiça, a questão da justificação é essencial para que se possa aceitar, como cidadãos livres e iguais, as normas que serão impostas de maneira geral. Essa justificação dinâmica é tanto concreta e historicamente situada como estrutural e teórica.

Isso, contudo, pressupõe um conceito de justiça discursiva, nos moldes habermasianos, como fundamentação procedimental e substancial de uma sociedade justa, na qual devem ser observados tanto os princípios da autonomia (como autodeterminação) e da justificação (como forma de não dominação), como os princípios de distribuição de bens, bem-estar e capacidades.

A proposta justificativa é uma abordagem que não começa nem a partir de um "ideal", nem de uma concepção empírica de política supostamente "realista", que exclui princípios de ordem superior, levando a um decisionismo.

A questão da legitimidade normativa deve ser enfrentada, em si, como uma questão de justiça, o que significa um comando justificado, não arbitrário, e no qual não haja dominação. O Justo passa a ser aquilo que permite, de uma maneira geral e recíproca, que os sujeitos se submetam apenas àquilo que autorizaram de forma livre e igual. Somente assim é possível a escolha social genuína.

Voltando-se à teoria da escolha social, a aplicação da noção de justificação reflexiva funciona como solução que traduz a vontade individual em vontade social legítima, sem perder o embasamento racional afeto à percepção de justiça desde o ponto de vista de cada participante da sociedade.

O diálogo sempre aberto (domínio irrestrito) deve permitir ordenamentos coerentes (transitivos ou quasi-transitivos) sem que haja a imposição da vontade arbitrária externa (não ditadura).

Da mesma forma no que tange o modelo de equilíbrio colaborativo, um sistema munido de um mecanismo de justificação reflexiva, se apresenta como solução sustentável para uma sociedade bem ordenada, haja vista sua tendência natural de projetar um futuro harmônico e afastar eventuais estratégias maliciosas por parte de seus integrantes.

Trata-se de um sistema que valoriza o maior resultado coletivo sem sacrificar os interesses individuais.

### **3.4 A questão da viabilidade**

As teorias com enfoque kantiano, conforme desenvolvidas por Rawls (1971) e Habermas (1999), por exemplo, focam nas condições ideais de um procedimento justo capaz de gerar resultados que reflitam os princípios

normativos selecionados na busca pela justiça. O que está em jogo é sempre o balanço entre a liberdade individual e a equidade substancial de uma sociedade. Em ambos os casos, contudo, as concepções de justiça foram modeladas para funcionarem com certa independência das questões contingenciais que implicariam em inviabilidade ou distanciamento entre a teoria e a realidade. Tal caráter ideal não é privilégio das teorias liberais, porquanto as ditas teorias comunitaristas também caem em um problema similar. Ao afirmar uma substancialidade ética que é anterior ao indivíduo, a posição comunitarista acaba por atrelar o justo ao que esteja alinhado com uma “vontade geral”, esvaziando o papel que a liberdade individual exerce na realização da justiça.

Existem algumas abordagens menos ambiciosas que buscam combater as injustiças de forma direta, partindo da viabilidade<sup>17</sup>. Tais abordagens costumam considerar os aspectos contingentes e contextuais, como a motivação e as preferências individuais dos indivíduos em determinado modelo econômico.

Conforme aponta Herzog (2016, p. 959) não se trata de apreciar se uma teoria mais idealista é menos viável que outra teoria menos idealista. Ambas as posições devem enfrentar a instância de consideração da viabilidade, não interessando se tratam os respectivos princípios de justiça como entidades independentes dos fatos contingentes ou se partem da realidade material para corrigir injustiças sociais já identificadas.

Tomadas essas delimitações, dois erros mais comuns podem ser identificados: a) considerar viáveis coisas que, de fato, são inviáveis. Este é o caso comumente atribuído às teorias de justiça ideais. Tal erro é apontado nas críticas ao idealismo exagerado, e acaba, dado o distanciamento com a realidade, relegando algumas teorias à irrelevância prática. b)

---

<sup>17</sup> Um exemplo disso é a abordagem de Amartya Sen (2011) no que tange sua concepção de Justiça. Vide “A ideia da Justiça”.

considerar inviáveis coisas que, na verdade são viáveis. Segundo Herzog este erro pode ser comparado com um resultado de exame médico “falso positivo”, pois uma doença foi identificada apesar de o paciente estar saudável (Herzog 2016 p. 960).

A análise que segue foca-se no segundo erro (considerar inviáveis coisas que não o são), porquanto os debates referentes à ideologia de mercado e aos modelos de individualismo racional econômico mostram uma deficiência justamente nesse campo. É o caso das presunções econômicas sobre aquilo que motiva as pessoas a escolherem ou agirem somente com base em seus interesses particulares.

### **3.4.1 Auto-interesse e Racionalidade Econômica**

Com efeito, a ideia de que somente valores auto-interessados seriam racionais é recorrente na teoria econômica, seja ela imputada à teoria clássica ou não. Essa aproximação do racional com o egoísmo individual mostrou-se uma barreira para que se pudessem introduzir elementos éticos na economia.

Na teoria econômica dominante existem dois métodos básicos para se determinar a racionalidade de uma conduta: (1) considerar a racionalidade como uma consistência *interna* da escolha e; (2) identificar a racionalidade com a maximização do auto-interesse (Putnam 2008, p. 73).

Se observarmos o primeiro método, de pronto podemos identificar sua fragilidade, porquanto a própria ideia de que a consistência *interna* de uma escolha possa determinar a racionalidade dela conseqüente mostra-se insuficiente. Sen observa que uma pessoa pode fazer exatamente o oposto do que deveria fazer para conseguir algo que deseja egoisticamente, tudo com perfeita coerência interna (Sen, 1992, p. 113).

Dessa forma, podemos notar que aquilo que consideramos coerente num conjunto de escolhas observadas depende, necessariamente, da

interpretação dessas escolhas e de algumas características externas a elas, como a natureza de nossas preferências, valores e motivações (Sen, 1992, p. 114).

Para Amartya Sen (1992, p. 114), o segundo enfoque, que se concentra na maximização do auto-interesse, é ainda mais difícil de defender. Ele estaria baseado em um critério de correspondência *externa* entre a escolha e o interesse pessoal. Desde já, podemos concluir que se trata de uma estratégia que nega o primeiro método de atribuição da racionalidade, o qual somente considerava a consistência *interna* das escolhas.

O problema aqui é outro. A posição de Sen (1992) é contrária a qualquer enfoque que simplifique demasiadamente as motivações para a conduta humana. Assim, o economista indiano perguntaria “por que a racionalidade residiria unicamente na busca do interesse próprio, com a exclusão de todo o resto?” E logo acrescentaria que pode até não ser absurdo considerar a maximização do interesse próprio como conduta racional, mas propor que todo o resto (que não seja a maximização do auto-interesse) seja irracional parece inaceitável (Sen 1992, p. 115).

Adotar esse critério implica numa postura que acaba por rejeitar a ética nas decisões reais dos indivíduos. A estratégia de utilizar a racionalidade como um “intermediário” que obrigue os atores econômicos a uma necessária maximização dos respectivos interesses próprios parece uma tentativa de padronização comportamental insuficiente. Nas palavras efusivas de Sen, “tentar usar as exigências da racionalidade na batalha a favor da suposição comportamental padrão da teoria econômica é como liderar uma carga de cavalaria montado num burro manco” (Sen, 1992, p. 115).

A questão, aqui, consiste em verificar se a busca da maximização do interesse individual é mesmo uma característica do comportamento real. “Será que o ‘homem econômico’, na busca de seus interesses, proporciona

um melhor entendimento do comportamento humano, ao menos no campo da economia?” (Sen, 1992, p.116).

Nessa mesma linha de argumentação Herzog (2016) aponta que a motivação para ação humana é uma questão muito mais complexa. Ela refere o paradoxo de Esterlin, o qual sustenta que depois de alcançado certo patamar de renda, o aumento de PIB não surte efeitos de aumento de felicidade (Herzog 2016, 964). A autora prossegue apontando que o incentivo monetário, apesar de ser eficaz em muitas circunstâncias, não serve de fator determinante para a tomada de decisão em todos os casos.

Enfim, existe suporte suficiente para sugerir que o auto-interesse não representa um limite de viabilidade tão fixo, porquanto podemos claramente notar que a motivação final da ação dos indivíduos é capaz de considerar o ganho de utilidade derivado do aumento de bem estar dos outros. Não se trata de esperar condutas sempre altruístas e desvinculadas de um retorno financeiro, mas sim de desfazer a presunção recorrente na ideologia de mercado de que o auto-interesse é algo dado, e que seria inerente à natureza humana. É justamente esse raciocínio que gera os “bloqueios de viabilidade”.

### **3.4.2 Bloqueios de Viabilidade**

O modelo de ideologia de mercado não assumiu as proporções que hoje observamos por mero acaso. A figura do “homem econômico não se desenvolveu e alcançou o imaginário social por ser uma mera estratégia metodológica, mas por ser utilizada como uma presunção da natureza humana” (Herzog 2016, p. 966).

A ideologia econômica pode influenciar diretamente a estrutura da motivação individual ao transformar a percepção daqueles interesses que são legítimos ou não de serem perseguidos. Tal ideologia também é capaz de afetar a auto-percepção dos indivíduos, reduzindo as opções reais

àquilo que traz um maior retorno financeiro. Conforme alega Herzog (2016, p. 966) “apontar que a ideologia econômica é enviesada, e os seres humanos tem motivações mais complexas, poderia mudar diretamente o horizonte de viabilidade de certas instituições econômicas”.

Existe, contudo, um efeito indireto dessa percepção dos indivíduos como seres auto-interessados. Tal efeito tem relação com a percepção que os indivíduos possuem sobre as motivações dos demais indivíduos. Conforme aduz Herzog (2016 p. 967), pouquíssimas pessoas veem a si mesmas como motivadas exclusivamente pelo auto-interesse. O problema se instaura quando essas mesmas pessoas passam a assumir que todos os demais indivíduos somente agem de maneira auto-interessada. Isso pode forçar que o agente se antecipe ao egoísmo alheio e passe a agir de forma a defender seu próprio interesse, por precaução.

Como observado no segundo capítulo deste trabalho, experimentos econômicos na teoria dos jogos, em especial os que preveem jogos continuados, apontam que a cooperação entre os agentes é condicional. Isso significa que os indivíduos estão dispostos a cooperar, contando que os outros também cooperem com eles. Considerando-se a disposição de um jogo de “dilema do prisioneiro iterado”, é possível concluir que o melhor resultado é sempre o cooperativo.

O problema surge com a quebra da confiança. Sabendo que a melhor estratégia para o outro participante, em termos de ganho pessoal, é sempre buscar o seu próprio interesse, o agente tende a se proteger e buscar defender seus próprios interesses.

Herzog (2016) aponta que se o indivíduo está influenciado pela ideologia de mercado ele está mais propenso a acreditar que os outros não estão dispostos a cooperar, a menos que haja também satisfação de seus interesses particulares. Tais problemas podem ser contornados em grupos menores, nos quais uma acareação ou uma declaração das reais

motivações pode afastar a desconfiança mútua instaurada. Entretanto, em se considerando sociedades como um todo, não é possível esse esclarecimento. Daí o malefício que a presunção de egoísmo imanente à teoria econômica dominante pode causar no imaginário coletivo.

Mesmo em jogos onde não haja perdas para os participantes a solução paretiana, dominante na ideologia de mercado, pode contrastar como mais básico princípio de liberdade de escolha (como apontado por Sen (1970a) no paradoxo desenvolvido em *The Impossibility of a Paretian Liberal*).

Sen (1992) busca inverter o paradigma liberal, colocando a vontade individual como condicionada, já no momento da escolha, a certas prioridades e valores sociais. Ao fazê-lo, ele também questiona o individualismo mecânico, na medida em que sustenta que indivíduos valorizam o ganho de utilidade da sociedade como um todo. A questão, para Sen é: se somos capazes de sacrificar nosso auto-interesse em nome de prioridades familiares ou sociais (ainda que no estilo utilitário de Mill (1863)), porque não fazemos isso com mais frequência? Ou, colocado de forma diferente, se somos capazes de fazer sacrifícios individuais em nível difuso, de acordo com nossa afiliação comunitária, partidária ou familiar, porque esses mesmos sacrifícios não são reproduzidos em um nível social?

Os desequilíbrios funcionais causados pela falta de confiança na motivação dos outros indivíduos acabam por limitar o conjunto de opções possíveis (domínio irrestrito), impedindo, por exemplo, que uma atitude cooperativa que geraria a maior utilidade mútua seja concretizada.

Estes desequilíbrios foram chamados por Herzog de “bloqueios de viabilidade”. Isso se verifica quando indivíduos estariam dispostos a contribuir mais, contudo deixam de fazê-lo, por considerarem que os outros não estariam dispostos a fazer o mesmo. Estamos em face de um impasse, uma situação de desconfiança recíproca.

A solução, segundo Herzog seria o foco em mecanismos que evitem a crença mútua no interesse individual. Isso manteria normas, como a justa distribuição de ônus sociais, no seu devido lugar, permitindo que sejam desfeitos os “bloqueios de viabilidade”.

Economistas como Sen (2011) e Wiens (2013) tem enfrentado o tema da insuficiência dos modelos econômicos exclusivamente centrados no interesse individual, apontando que é necessário um modelo mais complexo que leve em consideração motivações e comparações interpessoais de utilidade em busca de soluções que promovam maior justiça e bem estar social.

Segundo Herzog, com relação à ideologia econômica, os teóricos políticos e os ativistas sociais poderiam resolver dois problemas de uma só vez se dispensassem as crenças falsas e ajudassem a expandir o horizonte de viabilidade, gerando instituições mais equitativas, e reafirmando que todos estão dispostos a fazer suas respectivas partes na busca por justiça (Herzog 2016, p.968).

O objetivo aqui, contudo, era o de apontar a falha argumentativa da presunção geral que ocorre na “ideologia econômica” calcada no auto-interesse dos indivíduos. Penso que reestruturar o imaginário coletivo de como os indivíduos veem a si mesmo e aos outros é etapa crucial para que possamos implementar novas alternativas de justiça social e uma justa distribuição de riquezas, afastando as situações artificialmente criadas e erroneamente difundidas, que acabam por impedir a viabilidade de ações cooperativas antes mesmo de elas serem problematizadas.

## **Conclusão**

### **Projetando um possível modelo de justiça**

O objetivo do presente esforço foi a análise e proposição de um modelo de escolha que melhor desse conta de traduzir a vontade individual em uma escolha legitimamente social.

Buscou-se desenvolver um sistema equilibrado e cooperativo, que não fosse nem idealizado demais, por tomar emprestadas características puramente discursivas, nem uma solução casuística e completamente determinada pelo contexto.

É verdade que para se propor um modelo de justiça é necessário muito mais que somente definir uma estratégia de escolhas. Em uma teoria de justiça é preciso que sejam definidos os princípios balizadores da sociedade e a forma de equilíbrio desejado.

Contudo, vale destacar que, antes de tudo, se trata de uma proposição de mudança ética e moral, que ocorre no campo deontológico (do dever ser).

Não é a intenção do presente estudo, entretanto, discutir o que é mais inerente ao comportamento humano: a cooperação ou o egoísmo. A falibilidade, por sua vez, parece atributo mais facilmente imputável de maneira universal.

Daí a proposta de avaliar algo palpável, como as condições reais em que a cooperação emerge mesmo entre indivíduos egoístas. Penso que uma solução duradoura e consistente tenha que partir de avaliação concreta, nestes termos. Uma solução que menos exija que nos comportemos de forma infalível tem mais condições de se estabelecer e prevalecer.

Nesse sentido a evolução biológica é tão ou mais informadora da nossa trajetória social quanto a experiência histórica. A Moral Universal pode, por vezes, não traduzir um sistema de justiça humana, ou seja, uma ética contextualizada e falível, tal qual possamos enxergá-la.

Na tentativa de evitar o salto moral metafísico passa-se a buscar uma imbricação, avaliada de forma incidental, entre a conduta socialmente mais benéfica e a noção deontológica de justiça universal.

A estratégia *tit-for-tat* (que, como vimos, consiste em apenas duas diretrizes: começar cooperando e agir em reciprocidade), por exemplo, traz subjacente a noção milenar da regra de ouro<sup>1</sup>, fundadora de toda a ética ocidental judaico-cristã. Antes da motivação ética abrangente (teológica ou não), o que está em jogo é a possibilidade humana de interagir da forma mais eficiente possível. Isso não decorre do entendimento perfeito da Lei Moral, mas sim da disposição humana de, mesmo que em face da inclinação à atuação egoísta, escolher agir cooperativamente.

*Tit-for-tat* traz consigo essa condição. Não é por acaso que o torneio entre estratégias promovido por Axelrod teve como vencedor um programa cooperativo com mecanismos de reciprocidade. A metodologia utilizada, diga-se, sequer cogitou avaliar o entendimento perfeito da Lei Moral única.

Trata-se de uma engenharia reversa que permite entender uma engrenagem maior a partir dos atributos que nos são visíveis. Justiça Perfeita somente pode ser entendida de um ponto de vista de dentro da Lei Moral. A justiça humana, de outro lado, é um conceito que somente permite tentativas. O caminho mais justo deve resguardar sua posição intermediária entre o rigor da certeza moral e a compaixão com a falibilidade humana.

---

<sup>1</sup> A máxima da ética de reciprocidade: “cada um deve tratar os outros como gostaria que ele próprio fosse tratado”.

Conforme se verificou no segundo capítulo desta obra, uma população que age utilizando-se de estratégias cooperativas, em longo prazo, torna-se estável e imune a invasões de outras estratégias maliciosas. Isso deve funcionar como objetivo a ser alcançado. Da mesma forma, ao adotar estratégias cooperativas justificadas em longo prazo, é possível imprimir a mudança mesmo que dentro de populações que se valem de estratégias egoístas há séculos.

Assim, resta reduzida a discussão ideológica de “como” deve ocorrer a transição de um modelo colonizado e injusto, para um modelo melhor, com base em princípios mais justos (no caso, princípios de cooperação). Como provou Axelrod, mesmo a pior (mais egoísta) das estratégias egoístas (*ALL D*) pode ser invadida por grupos de estratégias cooperativas.

O modelo de justiça social cooperativa, em termos de resultados, sempre será o que melhor gera e sustenta o maior bem estar geral, seja ele traduzido em termos de estima (e valores conexos de afetividade social), de ecologia, e mesmo puramente em parâmetros de distribuição de bens primários. Trata-se de abordagem otimista, mas que se impõe como possível alternativa aos falidos sistemas de representação e de escolha social de que dispomos.

No que se refere à escolha social, é necessário deixar claro o que a abordagem justificativa aqui apresentada não busca solução eletiva par a par, mas sim um sistema aberto de domínio irrestrito. Isso significa dizer que não é possível restringir o ato de decisão a uma opção binária. Com isso, busca-se evitar que a escolha final tome somente em conta o somatório das escolhas individuais que a conformam. Conforme já demonstrado por Arrow (1951, p.48), em seu primeiro teorema<sup>2</sup> existe um ordenamento

---

<sup>2</sup> Teorema 1 (Teorema da Possibilidade de duas alternativas): Se o número total de alternativas é dois, o método de decisão da maioria é uma função de bem-estar social que satisfaz Condições 2-5 e produz um ordenamento social das duas alternativas para cada conjunto de ordenação individual.

social de duas alternativas para cada conjunto de ordenação individual. Dentre um espectro de possíveis opções (o universo de frutas, por exemplo) sempre há a solução de se escolher duas a duas (entre maçãs e laranjas, por exemplo, e posteriormente entre laranjas e bananas), sem que se incorra em inconsistência racional.

Por esse motivo, também a votação par a par resta descartada, já que traz consigo a mesma implicação. O modelo de plebiscitos, por exemplo, não seria suficiente para garantir um modelo de escolha social, porquanto, na realidade, trata-se de um somatório de vontades individuais que não traduzem sequer uma função de bem-estar digna da complexidade inerente à vida em sociedade.

A opção por um modelo que considere mais de duas opções que representem funções de bem-estar social, cobra o preço sobre o quesito da transitividade e da IAI (independência das alternativas irrelevantes). O paradoxo de Condorcet<sup>3</sup>, já demonstrado, adverte sobre o risco de abordar a questão da escolha quando o número de opções é maior que dois, porquanto o resultado final pode sequer representar a vontade da maioria.

O modelo que se busca é algo mais próximo da proposta de Sen (1999, p. 118-133) que pretende uma expansão dos mecanismos informacionais da escolha, visando suprir alguma eventual discrepância entre a ordem de preferências declaradas e o resultado final da eleição.

A abordagem justificativa aqui proposta tem a finalidade de suavizar o critério paretiano liberal, porquanto, em decisões justificadas, a possibilidade de contextualização resta alargada e, ao mesmo tempo, permite-se um balizamento moral no momento da escolha. Conforme já referido (Sen 2011, p. 23) deve ser possível a diferenciação entre pessoas famintas em decorrência de circunstâncias que estão além do seu controle e pessoas

---

<sup>3</sup> Ver capítulo 1, item 1.1.

sucumbindo à fome por manobra de outras que planejaram esse resultado. Por isso, apreciações puramente econométricas e escolhas sociais que não tragam consigo capacidade de ajuste e adequação ao contexto fático aos quais se aplicam, devem ser evitadas.

Escolha social deve trazer resultado que represente um *juízo de preferência legítimo* que reflita a vontade de todos envolvidos. Só assim, é possível o resgate mínimo da legitimidade nos procedimentos de eleição coletiva realmente alicerçados na soberania popular.

O esforço do presente trabalho, portanto, parte da noção fundamental de sociedade como um sistema justo de cooperação social prolongado no tempo, visando desenvolver uma concepção política de justiça para um regime democrático.

Trata-se, ao fim, da proposição de uma abordagem em prol da escolha livre e justificada pela cooperação social, a qual deve ser mantida de maneira equilibrada, contínua e não ditatorial.

## Referências

- ARROW, K. J. **A difficulty in the concept of social welfare.** Journal of Political Economy, n. 58, 1950.
- ARROW, K. J. **Social Choice and Individual Values.** New York: Wiley, (1951) (2ª ed) 1963.
- ARROW, K. J. **Some Ordinalist Utilitarian Notes on Rawls' Theory of Justice.** Journal of Philosophy, vol. 70, 1973. p. 245-263.
- ARROW, K. J. **General Economic Equilibrium: Purpose, Analytic Techniques, Collective Choice.** American Economic Review, vol. 64, 1974. p. 253-272.
- ARROW, K. J. **Extended Sympathy and the Possibility of Social Choice.** American Economic Review, Papers and Proceedings, vol. 67, pp. 219-25. 1977.
- ARROW, K. J.; SEN, A. K. e SUZUMURA, K. (eds.) **Social Choice Re-examined.** Vol 2. Proceedings of the IEA Conference held at Schloss Hernstein, Bernd. London. Macmillan Press LTD, 1996.
- ARROW, K. J.; SEN, A. K. e SUZUMURA, K. (eds.) **Social Choice Re-examined.** Vol 1. Proceedings of the IEA Conference held at Schloss Hernstein, Bernd. London. Macmillan Press LTD, 1997.
- ARROW, K. J.; SEN, A. K. e SUZUMURA, K. (eds.) **Handbook of Social Choice and Welfare.** Vol. I Oxford. North Holland, 2002.
- ARROW, K. J.; SEN, A. K. e SUZUMURA, K. (eds.) **Handbook of Social Choice and Welfare.** Vol. I Oxford. North Holland, 2011.
- ARROW, K. J. **Commentary in The Arrow Impossibility Theorem.** Volume 2. New York: Columbia University Press, 2014.

AXELROD, R. **Effective Choice in the Prisoner's Dilemma**. The Journal of Conflict Resolution, Vol. 24, No. 1 Mar. 1980. p. 3-25

AXELROD, R. **More Effective Choice in the Prisoner's Dilemma**. The Journal of Conflict Resolution, Vol. 24, No. 3 Sep. 1980. p. 379-403

AXELROD, R. **The Emergence of Cooperation among Egoists**. The American Political Science Review, Vol. 75, No. 2 Jun. 1981. pp. 306-318

AXELROD, R; HAMILTON W. D. **The Evolution of Cooperation**. Science, New Series, Vol. 211, No. 4489 Mar. 27, 1981. p. 1390-1396

AXELROD, R. **The Evolution of Cooperation**. Basic Books, Inc. Publishers New York, NY 1984.

AXELROD, R. **An Evolutionary Approach to Norms**. The American Political Science Review, Vol. 80, No. 4 Dec, 1986, p. 1095-1111

AXELROD, R. **The Evolution of Cooperation** – Revised Edition. Basic Books, Inc. Publishers New York, NY, 2006.

BENTHAM, J. **An Introduction to the Principles of Morals and Legislation**. Payne: London 1789.

BENTHAM, J., **Anarchical fallacies**. Primeiramente publicado em: J. Bowring, ed., The Works of Jeremy Bentham, Vol. II William Tait, Edinburgh (1843). Republicado em 1995. (Theommes Press, Bristol) p. 489-534.

BLACK, D. **On the rationale of group decision-making**, Journal of Political Economy 56, 1948. p. 23-34.

BLACK, D. **The Theory of Committees and Elections**. Cambridge University Press, Cambridge, 1958.

BORDA (J.C. de Borda) **Memoire sur les elections par scrutin**, Mermoires de l'Academie Royale des Sciences. 1781, p. 657-665. Republicado em 1907 por Clarendon Press, Oxford. 1781.

BRONOWSKI, Jacob. **The Ascent of Man**. London: British Broadcasting Corporation, 1973.

CLARKE. E. H. **Demand Revelation and the Provision of Public Goods**. Cambridge. Massachusetts: Ballinger, 1980.

COHEN M. D.; AXELROD, R. **The Adaptive Value of Changing Utility**. The American Economic Review, Vol. 74, No. 1 Mar., 1984. p. 30-42.

CONDORCET, Marquis de (Caritat, J. A. N.) **Essai sur l'Application de l'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix**. Paris, 1785.

CONDORCET M. de. **An essay on the application of probability theory to plurality decision making**: An election between three candidates. Em: Sommerlad F, McLean I 1989, eds. The political theory of Condorcet. University of Oxford Working Paper, Oxford, 1785a. pp 90-108.

DAHL, R. **A preface to democratic theory**. Chicago University Press, Chicago IL.1956.

DAHL, R. **Dilemmas of Pluralist Democracy: Autonomy vs. Control**. Yale University Press, 1983.

DAHL, R., RAE, D. **Who Governs: Democracy and Power in an American City**, Second Edition. Yale University Press, 2005.

DAWKINS, R. **The Selfish Gene**. Oxford: Oxford University Press, 1976.

DEWEY, J. **The Theory of Valuation**. Chicago: Chicago University Press, 1939.

DOWNING, L. **The Prisoner's Dilemma Game as a Problem-Solving Phenomenon**: An Outcome Maximizing Interpretation. Simulation and Games 6: 1975. p. 366-391.

- BERNI, D. A. **Teoria dos jogos**: jogos de estratégia, estratégias decisórias e teoria da decisão. Rio de Janeiro: Reichmann e Affonso eds. 2004.
- DWORKIN, R. **Sovereign Virtue**: The Theory and Practice of Equality. MA: Harvard University Press, 2000.
- ELSTER, J. **Sour Grapes**, Cambridge: Cambridge University Press, 1983.
- ELSTER, J. e HYLLELAND, A. eds. **Foundations of social choice theory**. Studies in rationality and social change, Cambridge University Press, 1986.
- ELSTER, J. (ed.) **Rational Choice: Readings in Social & Political Theory**. New York University Press, 1986.
- ELSTER, J. **The Cement of Society: A Study of Social Order** (Studies in Rationality and Social Change). Cambridge University Press, 1989.
- FERREIRA, C. R. B. **O entrelaçamento fatos-opiniões e a crítica a uma ciência sem valores**. Saarbrücken, Alemanha: Novas Edições Acadêmicas, 2016.
- FORST, R. **Contextos da justiça**. Filosofia Política para além de liberalismo e comunitarismo. São Paulo: Boitempo Editorial, 2010.
- FORST, R. **The Right to Justification**: Elements of a Constructivist Theory of Justice, trans. J. Flynn. Columbia University Press, New York. 2012.
- FORST, R. **Justification and critique**: towards a critical theory of politics. Trans. Ciaran Cronin. Polity Press, Cambridge, UK, 2014.
- FRIEDMAN, J. W. **A Non-Cooperative Equilibrium for Supergames**. Review of Economic Studies 38:1-12. 1971.
- GEHRLEIN W. V. **Condorcet's Paradox**. Theory and decision library General Editors: W. Leinfellner (Vienna) and G. Eberlein (Munich) series c: game theory, mathematical programming and operations research, Vol 40. Berlin: Springer-Verlag, 2006.

GILBERT, P.; LAWFORD-SMITH, H. **Political feasibility**. A conceptual exploration. *Polit. Stud.* 60(4): 2012 pp.809–825

GROPPER, R. **Gypsies in the City**: Cultural Patterns and Survival. Princeton University Press. 1975.

HABERMAS, J. **Teoria da ação comunicativa**. Vol. 1 e 2. São Paulo, Cortez Editora, 1999.

HABERMAS, J. **A inclusão do outro**: Estudos de Teoria e Política. Tradução de George Sperber e Paulo Astor Soethe. São Paulo: Loyola, 2002.

HABERMAS, J. **Ética da discussão e a questão da verdade**. São Paulo, Martins Fontes, 2007.

HERZOG, L. **Higher and Lower Virtues In Commercial Society**: Adam Smith And Motivation Crowding Out. *Renewal* vol. 20, n, 2/3. 2010. p. 107-115.

HERZOG, L. **Adam Smith´s Account of Justice Between Naturalness And Historicity**. *Journal of the history of philosophy*, vol 52, n. 4. 2014. p. 703-726.

HERZOG, L. **Distributive Justice, Feasibility Gridlocks and The Harmfulness Of Economic Ideology**. *Ethic Theory Moral Prac* 18, 2015. p. 957-969.

HOBBS, T. **Leviathan**. New York: Collier Books, (1651) 1962.

HONNETH, A. **The Struggle for recognition**. MIT Press. Cambridge. 1995.

HONNETH, A. **The Pathologies of Individual Freedom**. New Jersey. Princeton University Press, 2010.

MASKIN, E.; SEN, A.; ARROW, K.; DASGUPTA, P.; PATTANAIK, P., K.; STIGLITZ, J. E. **The Arrow Impossibility Theorem**. Columbia University Press, 2014.

MILL, J. S. **Utilitarianism** (1863). Ontário: Batoche Books, 2001.

NUSSBAUM, M. **Tragedy and human capabilities**: a response to Vivian Walsh. *Review of Political Economy*, 15, 3, p. 413-418, 2003.

NUSSBAUM, M. **Frontiers of Justice**: Disability, Nationality, Species Membership. Harvard University Press, Cambridge, MA. 2006.

OLIVEIRA, N. H. de. **Tractatus politico theologicus**: teoria crítica, libertação e justiça. Porto Alegre-RS: Editora Fi. 2016.

OLIVEIRA, N. H. de. **Tractatus practico theoreticus**: ontologia, intersubjetividade e linguagem. Porto Alegre-RS: Editora Fi. 2016

PARETO, V. **Manuale di Economia Politica**. Societa Editrice Libreria, Milan, (1906) English translation: *Manual of Political Economy*. A.M. Kelley, New York, 1927.

PIGOU, A.C. **The Economics of Welfare** (Macmillan, London). 1920.

PINK, D. **The Surprising Truth About What Motivates Us**. New York. Penguin. 2009.

POUNDSTONE, W. **Prisoner's dilemma**. Anchor Books. New York, 1993.

PUTNAM, H. **O Colapso da Verdade e outros ensaios**. Trad. Pablo Ruben Mariconda e Sylvia Gemignari Garcia. Aparecida, SP. Idéias e Letras, 2008.

RAE D. W. **An altimeter for Mr. Escher's stairway**: A comment on William H. Riker's 'Implications from the disequilibrium of majority rule for the study of institutions. *American Political Science Review* 74: 1980 p. 451-455.

RAWLS, J. **A Theory of Justice**. Cambridge, Mass. Harvard University Press, 1971;

RAWLS, J. **Political Liberalism**. New York: Columbia University Press 1996.

RAWLS, J. **The Law of Peoples**. Harvard University Press, Cambridge, MA. 1999

RAWLS, J. **The domain of the political and overlapping consensus**. In S. Freeman (ed.) *Collected Papers*. Harvard University Press, Cambridge, MA, 1999a. pp. 473–96.

RAWLS, J. **Justice as Fairness: A Restatement**, ed. E. Kelly. Harvard University Press, Cambridge, MA. 2001.

RIKER, W; BRAMS, S. J.. **The Paradox of Vote Trading**. *American Political Science Review* 67. 1973, p.1235-47.

ROBBINS, L. **An Essay on the Nature and Significance of Economic Science**, (1ª ed 1932) 2ª Edition, London: Macmillan, 1935.

ROBBINS, L. **Interpersonal comparisons of utility**, *Economic Journal* 48: 1938. p. 635-641.

SAMUELSON, P. A. (1947), **Foundations of Economic Analysis** (Harvard University Press, MA).

SAMUELSON, P. A. **Economics**, New York: McGraw-Hill. 1973.

SEN, A. **Collective Choice and Social Welfare**, San Francisco: Holden-Day; Amsterdam: North-Holland, 1970.

SEN, A. **The Impossibility of a Paretian Liberal**. *Journal of Political Economy*. 78: 1970a. p. 152–157 reprinted in Sen 1982a.

SEN, A. **Interpersonal aggregation and partial comparability**. *Econometrica* 38; 1970b. reprinted in Sen 1982a.

SEN, A. **Choice functions and revealed preference**. *Review of Economic Studies*, 38; 1971. reprinted in Sen 1982a.

SEN, A. **Behavior and the concept of preference**, *Econometrica* 40, 1973. p. 241- 259.

SEN, A. **Liberty, unanimity and rights**, *Econometrica* 43; 1976. reprinted in Sen (1982a).

SEN, A. **Social choice theory: a re-examination**, *Econometrica* 45; 1977a reprinted in Sen (1982a).

SEN, A. **On weights and measures**: informational constraints in social welfare analysis, *Econometrica* 45; 1977b reprinted in (1982a).

SEN, A. **Rational fools**: a critique of the behavioral foundations of economic theory, *Philosophy and Public Affairs* 6, 1977c. reprinted in Sen (1982a).

SEN, A. **Personal utilities and public judgment**: or what's wrong with welfare economics? *Economic Journal* 89; 1979b. reprinted in Sen (1982a).

SEN, A. **Interpersonal comparisons of welfare**. in M. Boskin (ed.), New York: Academic Press; 1979c. reprinted in Sen (1982a).

SEN, A. **Utilitarianism and welfarism**, *Journal of Philosophy* 16. 1979d.

SEN, A. **Equality of what?** in S. McMurrin (ed.), *Tanner Lectures on Human Values*, Cambridge: Cambridge University Press; 1980. reprinted in Sen (1982a).

SEN, A. **Resources, values and development**. Harvard University Press, Cambridge, Massachusetts. 1981.

SEN, A. **Choice, welfare and measurement**. Harvard University Press, Cambridge, Massachusetts. 1982a.

Sen, A. **Liberty as control**: an appraisal', *Midwest Studies in Philosophy* 8. 1982b

SEN, A. **Rights and agency**. *Philosophy and Public Affairs* 11; 1982c.

SEN, A. and Williams, B. (eds.), **Utilitarianism and Beyond**, Cambridge: Cambridge University Press. 1982d.

SEN, A. **Liberty and social choice**, *Journal of Philosophy* PJJ; 1983a

SEN, A. **Evaluator relativity and consequential evaluation'**, *Philosophy and Public Affairs* 12; 1983b.

SEN, A. **Poor, relatively speaking**. *Oxford Economic Papers* 35; 1983c

SEN, A. **Commodities and Capabilities**. North-Holland; Amsterdam 1985.

SEN, A. **On ethics and economics**. UK: Blackwell Publishing, 1988.

SEN, A. **Comportamento econômico e sentimentos morais**. *Lua Nova* no. 25. São Paulo, 1992.

SEN, A. **Capability and well-being**, in M. Nussbaum & A. Sen (Eds). *The quality of Life*, p 30-53 - Oxford. Clarendon Press, 1993.

SEN, A. **Desenvolvimento como liberdade**. Trad. Laura Teixeira Motta. Revisão técnica. Ricardo Doninelli Mendes. São Paulo: Companhia das Letras, 2000.

SEN, A. **Rationality and freedom**. Harvard University Press, Cambridge, Massachusetts. 2004.

SEN, A. **Adam Smith and the contemporary world**. *Erasmus Journal for Philosophy and Economics*, Volume 3, Issue 1, Spring 2010, pp. 50-67.

SEN, A. **A idéia de Justiça**. Trad. Denise Bottmann, Ricardo Mendes. São Paulo: Companhia das Letras, 2011a.

SEN, A. **Uses and abuses of Adam Smith**. *History of Political Economy* 43:3.. Duke University Press 2011b, p. 257- 271.

SEN, A. **Poverty and informed reasoning**. *Field actions science reports* [online], Special Issue 4, 2012.

- SEN, A. **Arrow and the impossibility theorem**: where do we go from here? in *The Arrow Impossibility Theorem*. New York: Columbia University Press, 2014a.
- SEN, A. **Economics and philosophy**. Volume 30/ Special Issue/ 1, march 2014b. pp. 1-10.
- SEN, A. **The reach of social choice theory**. *Soc. Choice Welf.* 39, 2014c. pp. 259-272.
- SCAZZIERI, R.; SEN, A.; ZAMAGNI, S. **Markets, Money and Capital**. Cambridge University Press, 2009.
- SHUBIK, M. **Game Theory, Behavior, and the Paradox of Prisoner's Dilemma**: Three Solutions. *Journal of Conflict Resolution* 14: 1970, p.181-94.
- SMITH, A. **The theory of moral sentiments**. Reprint. Originally published: Oxford: Clarendon Press, 1976. Liberty Fund. Indianápolis, 1984.
- SMITH, A. **A riqueza das nações**. São Paulo. Editora Nova Cultural Ltda, 1996.
- SMITH, J. M., and Parker, G. A. **The Logic of Asymmetric Contests**. *Animal Behavior* 24, 1976. p. 159-75.
- WEBER, T. **Ética e filosofia política**: Hegel e o formalismo kantiano. 2ª ed. Porto Alegre-RS: ediPUCRS, 2009.
- WIENS, D. **Prescribing institutions without ideal theory**. *Journal of Politic Philos* 20(1), 2012 p. 45-70.
- WIENS, D. **Demands of justice, feasible alternatives, and the need for causal analysis**. *Ethical Theory Moral Pract* 16(2): 2013. p. 325-338.

A Editora Fi é especializada na editoração, publicação e divulgação de pesquisa acadêmica/científica das humanidades, sob acesso aberto, produzida em parceria das mais diversas instituições de ensino superior no Brasil. Conheça nosso catálogo e siga as páginas oficiais nas principais redes sociais para acompanhar novos lançamentos e eventos.



**[www.editorafi.org](http://www.editorafi.org)**  
**[contato@editorafi.org](mailto:contato@editorafi.org)**