

A EXPLICAÇÃO CAUSAL (HUMEANA) DA AÇÃO E A TEORIA DOS JOGOS

André Guimarães Augusto*

Resumo

Nesse artigo é argumentado que a teoria dos jogos apresenta uma explicação causal para a ação humana. É adotado como hipótese do artigo que toda teoria social, pela própria natureza de seu objeto, formula ou adota, explícita ou implicitamente, uma explicação para a ação humana. O argumento central do artigo é de que a teoria dos jogos não-cooperativos adota a explicação causal de David Hume para a ação humana e que essa explicação elimina as características da posição de finalidades e da alternativa da ação.

Abstract

This article argue that game theory gives a causal explanation for human action. The hypotheses of article is that every social theory develop or adopt a explanation for human action given the nature of his subject. The main argument of the article is that non cooperative game theory adopt David Hume's causal explanation for human action and that this explanation eliminate the features of intentionality and alternative of action.

Sub-área: metodologia e caminhos da ciência

Sessões ordinárias

A teoria dos jogos tem crescido em importância dentro do pensamento econômico contemporâneo. Desde suas primeiras aplicações para as ciências sociais com o trabalho de Von Newmann e Morgenstein em 1944 e dos trabalhos de John Nash na década de 1950, a teoria dos jogos cresceu no número de adeptos e ganhou espaço com a criação de centros de pesquisa espalhados pelo mundo, revistas especializadas e um crescente número de artigos publicados em revistas de economia. O prêmio Nobel concedido a John Harsanyi, John Nash e Reinhard Selten pela análise do equilíbrio em jogos não-cooperativos em 1994 e a Thomas Schelling e Robert Aumann pela análise da cooperação e do conflito através da teoria dos jogos em 2005, deixa claro que a teoria dos jogos faz parte da “nova face do mainstream”, conforme a expressão de Collander, D. et alli (2004).

Nesse artigo é argumentado que a teoria dos jogos apresenta uma explicação causal para a ação humana. É adotado como hipótese do artigo que toda teoria social, pela própria natureza de seu objeto, formula ou adota, explícita ou implicitamente, uma explicação para a ação humana. O argumento central do artigo é de que a teoria dos jogos não-cooperativos adota a explicação causal de David Hume para a ação humana e que essa explicação elimina as características da posição de finalidades e da alternativa da ação.

Na primeira parte é apresentado um breve sumário da teoria dos jogos não cooperativos estáticos, considerando-se os casos de informação completa e incompleta. Na segunda parte do artigo é argumentado que a teoria dos jogos adota a explicação causal humeana da ação: na primeira subseção é apresentada a equivalência entre os desejos e as preferências como causa da ação em David Hume e na teoria dos jogos respectivamente; na subseção seguinte é argumentado que a noção humeana de causalidade é adotada pela teoria dos jogos. Na terceira parte se argumenta que a explicação causal para a ação não é capaz de apreender as características de posição de finalidade e escolha entre alternativas presentes na ação. O artigo termina com algumas considerações finais.

* Professor Adjunto do Departamento de Economia da Universidade Federal Fluminense.

1 – Jogos não cooperativos: um breve sumário

A teoria dos jogos tem suas origens nos trabalhos do matemático alemão David Hilbert sobre os fundamentos da matemática na virada do século XX. Propondo uma abordagem axiomática da matemática e sua aplicação a todos os campos do conhecimento, o programa de Hilbert é aplicado pela primeira vez na explicação dos jogos com o trabalho de Zermelo, “On application of Set Theory to the theory of Chess” de 1912. O trabalho de Zermelo inaugura o programa de pesquisa da matemática dos jogos no qual John Von Newmann também estava engajado. O programa de pesquisa da teoria dos jogos em economia tem suas origens na participação de Von Newmann e Morgenstein no Colóquio de Viena na década de 1930. Liderado por Karl Menger, filho de Carl Menger, o colóquio de Viena reunia economistas, filósofos e matemáticos influenciados pela matemática de Hilbert, pelo positivismo lógico do círculo de Viena e pela economia austríaca. É nesse ambiente que são elaborados os primeiros trabalhos de economia de Von Newmann. (Leonard, R.; 1995)

A teoria dos jogos só se torna um programa de pesquisa dentro das ciências sociais com a publicação da *Theory of Games and Economic Behavior* de John Von Newmann e Oscar Morgenstein em 1944. Nessa obra, Von Newmann e Morgenstein apontam a teoria dos jogos como um método de explicação para o comportamento individual em condições sociais. Em uma economia do tipo “Robson Crusoé”, isto é, do indivíduo isolado, as únicas variáveis que influenciariam a decisão dos agentes seriam aquelas que estariam diretamente sob seu controle e os “dados mortos” compostos pelo ambiente físico. Em uma economia social o comportamento dos agentes depende também dos dados que refletem a vontade e a intenção dos outros; dessa forma, o resultado da ação de cada um irá depender das ações dos outros. (Von Newmann, J. e Morgenstein, O. 1964).

Na teoria dos jogos a interação entre os agentes resulta em conseqüências que são valoradas pelos agentes, valoração esta que é representada pelo payoff ou função utilidade. Um payoff mais alto representa um resultado preferido a todos os outros e os payoffs possíveis representam a valoração dos agentes com relação aos resultados de sua interação de acordo com sua escala de preferências. Os agentes (jogadores) buscam obter o melhor resultado possível definindo um plano de ação (estratégia) considerando

a ação dos outros agentes (jogadores). A solução de um jogo é o resultado esperado da interação dos agentes (jogadores) e oferece um conjunto possível de regras de comportamento em cada situação possível.

Nesse quadro geral diferentes tipos de interação são formalizados pela teoria dos jogos. Os jogos cooperativos tratam de situações em que os compromissos são implementados, se focando na formação de grupos e na distribuição do payoff entre seus membros; nos jogos não cooperativos, em que compromissos não são obedecidos, a questão fundamental é a definição de estratégias dos indivíduos (Aumann, 1987a, p.54-55). Os jogos também podem ser dinâmicos formalizando situações de interação repetida ou estáticos quando a interação ocorre uma única vez. O argumento desse artigo se refere especialmente aos jogos não cooperativos estáticos.

A solução para jogos não-cooperativos com n participantes foi formulada por John Nash. A solução de Nash, ou equilíbrio estratégico, consiste em um perfil de estratégias em que cada jogador tem uma estratégia que maximiza o seu payoff se a estratégia dos outros se mantém fixa. Nash demonstrou que é possível definir pelo menos uma solução estratégica para jogos não cooperativos de n -pessoas. A solução estratégica, no entanto pode ser um composto das estratégias disponíveis para os jogadores, executadas com determinada probabilidade ao longo do jogos, isto é, estratégias mistas. Pode haver também casos de equilíbrios múltiplos, isto é, em que mais de uma solução estratégica é possível.

Em uma situação interativa o resultado final do jogo não depende apenas da ação escolhida por um jogador mas também das ações de todos os outros que participam do jogo. Isso implica que, ao decidir a sua ação, o jogador não deve considerar apenas as suas preferências quanto ao resultado do jogo mas também as possíveis decisões dos outros jogadores; em outras palavras as crenças com relação aos outros agentes – i.e.com relação às suas preferências e ao seu raciocínio deliberativo – são parte do determinante da escolha do plano de ação (estratégia) dos jogadores.

O equilíbrio de Nash assume que no início do jogo todos os jogadores têm a mesma informação e que essa informação é completa, incluindo as regras do jogo, as funções de payoff dos outros jogadores e as estratégias disponíveis para os outros jogadores, incluindo aquelas que eles realmente irão escolher; ou seja no equilíbrio de Nash os jogadores têm toda a informação sobre a estrutura matemática do jogo.

John Harsanyi elaborou uma solução para jogos em que os jogadores não têm todas as informações no início do jogo, isto é, jogos com informação incompleta. Harsanyi (1994) argumenta que a informação incompleta pode ser reduzida ao caso em

que os jogadores não têm conhecimento das funções de payoff uns dos outros. Nesse caso os jogadores formam expectativas sob a forma de uma distribuição de probabilidade subjetiva acerca dos payoffs possíveis dos outros jogadores, segundo uma regra bayesiana. Mas essas expectativas não são o suficiente para dar conta da informação incompleta em um jogo, pois nesse caso, as expectativas estão submetidas a uma lógica especular infinita: cada jogador deve formar expectativas sobre as expectativas dos outros acerca de seu payoff, sobre as expectativas das expectativas dos outros sobre sua expectativa acerca de seus payoffs e assim por diante. Essa lógica especular infinita tornaria impossível a operacionalização de um jogo com informação incompleta pois as distribuições subjetivas de probabilidade teriam que ser feitas em ordens cada vez mais elevada sem um ponto final.

A solução de Harsanyi para o problema da lógica especular infinita é a caracterização de cada jogador como formado por vários tipos, onde cada tipo tem sua própria função utilidade e distribuição de probabilidade subjetiva acerca do tipo dos outros jogadores. No início do jogo uma Loteria escolhe os tipos dos jogadores; cada jogador sabe seu tipo mas só pode formular uma probabilidade subjetiva sobre qual tipo é o tipo ativo dos outros jogadores. Essa probabilidade é condicionada à informação que o jogador tem – o seu tipo – e supõe uma probabilidade antecedente sobre as diversas combinações possíveis dos tipos de jogadores. Essa probabilidade anterior é de conhecimento comum de todos os jogadores. O equilíbrio estratégico é definido como um conjunto de estratégias para cada tipo que maximiza o seu payoff, dadas as estratégias dos outros tipos. (Harsanyi, J.; 1994)

O modelo de jogos bayesianos de Harsanyi introduz explicitamente a formação das crenças nas teorias dos jogos. As crenças são abordadas como um cálculo probabilístico e estão sujeitas a restrição das leis de coerência formal da probabilidade – aditividade finita, não-negatividade. Uma hipótese comum nos jogos bayesianos é de que as crenças são compartilhadas por todos os jogadores – hipótese do *common knowledge* -, já que todos tem a mesma distribuição da probabilidade antecedente acerca das combinações possíveis dos tipos; as diferenças nas probabilidades subjetivas resultam das diferenças de informação que os jogadores-tipos tem no início do jogo.

Em resumo, na teoria dos jogos os graus de crença são representadas por probabilidades bayesianas com relação à todos os elementos que podem estar sujeitos à incerteza – i.e., os estados do mundo -, incluindo às preferências dos outros jogadores. As crenças dependerão então das informações dos jogadores sobre a estrutura do jogo. Se houver informação completa – isto é se todos os jogadores têm todas as informações

sobre toda a estrutura do jogo, incluindo as preferências de todos os jogadores – as crenças corresponderão à verdadeira preferência dos jogadores e serão fixas. Com informação incompleta, a teoria dos jogos assume que os jogadores possuem as mesmas probabilidades iniciais com relação aos possíveis estados do mundo e formarão probabilidades subjetivas a respeito do estado do mundo efetivo a partir do diferente conjunto de informações que possuem.

2- A explicação causal (humeana) da ação na teoria dos jogos

Nessa seção será argumentado que a teoria dos jogos adota uma explicação causal para a ação humana, que remonta a David Hume. Primeiramente as causas da ação são identificadas com as preferências dos agentes; é argumentado que essa explicação é equivalente à explicação de David Hume que identifica o desejo como causa da ação. Um elemento adicional e ponto de controvérsia é o papel da racionalidade na determinação da ação em David Hume vis a vis os modelos racionalistas da teoria dos jogos. Esses argumentos serão desenvolvidos na subseção seguinte. Na segunda subseção é apresentado o argumento de que a explicação da teoria dos jogos para a ação atende aos critérios de causalidade de David Hume.

2.1 – Preferências, crenças e racionalidade: a causa da ação na teoria dos jogos.

Conforme apresentado na seção anterior, nos jogos não-cooperativos a estratégia dos agentes depende de suas preferências, dadas suas crenças. Essa relação é de dependência causal, no sentido da causalidade definida por David Hume, conforme será argumentado na subseção seguinte.

Se a ação depende das preferências, dadas as crenças, que sentido é atribuído a esses conceitos? Embora não haja muita discussão nem definição prévia das preferências na teoria dos jogos, seu sentido é o mesmo da teoria neoclássica padrão. Nesse ponto vale a pena dar a palavra para dois teóricos da teoria dos jogos; segundo Rubinstein (2006): “We view “preferences” as the mental attitude of an individual (economic agent) towards alternatives.” (p.1) e de acordo como Aumann (1987b): “(...) utilities directly express tastes, which are inherently personal. It would be silly to talk about “impersonal tastes”, tastes that are “objective” or “unbiased”” (p. 13). Assim, as preferências consistem em um estado mental, de natureza pessoal.

Sendo algo de inerentemente pessoal não é possível remeter a explicação de gostos ou preferências nem mesmo a uma conformação mental comum a todos os homens, uma vez que essa teria natureza “objetiva” ou pelos menos “impessoal”. Em outras palavras preferências não são passíveis de explicação racional. Mas além disso, não há também razão normativa para as preferências; porque um agente prefere X a Y é um assunto, pode se dizer, “de foro íntimo” para o qual nenhuma justificativa pode ser dada. Em resumo, as preferências são aracionais e portanto funcionam como um conceito primitivo na explicação da ação, não sendo passível de explicação ou justificativa.

Esse argumento com relação às preferências remete a teoria da ação de David Hume. Segundo David Hume as ações são resultados da vontade, definida como “ a impressão interna que sentimos e de que temos consciência quando deliberadamente geramos um novo movimento em nosso corpo ou uma nova percepção em nossa mente” (Hume, 2000, p.435). Para Hume uma impressão interna é uma percepção, um estado mental direcionado para a própria mente; de acordo com a definição anterior a vontade pode ser interpretada como uma disposição mental para agir.

Hume explica essa disposição para agir a partir da idéia de bem ou mal, identificadas por ele com o prazer e a dor, e do desejo gerado por essas idéias. Assim,

(...) quando temos a perspectiva de vir a sentir dor ou prazer por causa de um objeto, sentimos, em consequência disso, uma emoção de aversão ou propensão, e somos levados a evitar ou abraçar aquilo que nos proporcionará esse desprazer ou essa satisfação. (Hume, 2000, p.450).

o desejo é uma “propensão que nos une ao objeto” (Hume, 2000, p.474). Para Hume portanto, a disposição para agir (vontade) se segue à propensão a um objeto (desejo) despertado pela idéia de prazer.

É possível afirmar que o conceito de desejo de David Hume é equivalente ao de preferência do *mainstream* compartilhado pela teoria dos jogos não-cooperativos. Não se quer argumentar aqui que a teoria dos jogos compreende as preferências como um indicador de prazer ou dor, ao estilo da filosofia utilitarista¹. O que será argumentado é que o desejo em David Hume é um estado mental, aracional e capaz de gerar a disposição para agir da mesma forma que as preferências na teoria dos jogos.

¹ Aumann, em entrevista a Damme, oferece uma interpretação da ação nos moldes utilitaristas clássicos: “We have to be conscious for that, we can’t taste something without being conscious. The consciousness enables us to experience pleasure and pain.”(apud Damme, E. 1998, p.194)

O desejo segundo Hume é uma paixão direta, uma “impressão que decorre do bem e do mal de maneira mais natural e sem preparação.” (Hume, 2000, p.474); definido como uma paixão, e portanto como uma impressão o desejo é um estado mental da mesma forma que as preferências.

Da mesma forma que as preferências os desejos em David Hume são aracionais. É conhecida a afirmação de Hume :

Não é contrário à razão eu preferir a destruição do mundo inteiro a um arranhão em meu dedo. Não é contrário à razão que eu escolha minha total destruição só para evitar o menor desconforto de um índio ou de uma pessoa que me é inteiramente desconhecida. Tampouco é contrário a razão eu preferir aquilo que reconheço ser para mim um bem menor a um bem maior, ou sentir uma afeição mais forte pelo primeiro do que pelo segundo.(Hume, 2000, p.452).

É possível inferir dessa citação que para Hume a razão não é capaz de justificar um desejo, da mesma forma que não é capaz de justificar as preferências que, dessa forma são “inerentemente pessoais”.

Para Hume, uma paixão só pode ser contrária à razão quando está fundada em falsa suposições ou não escolhe os meios adequados (Hume, D.; 2000, p.452). Isso remete ao caráter instrumental da razão na ação (escolha dos meios) e às crenças (verdade ou falsidade das suposições); esse segundo ponto será tratado mais adiante. O ponto importante para o argumento desenvolvido aqui é que o caráter instrumental da razão, afirmado por Hume, implica que ela não pode ser capaz de gerar a disposição para agir e portanto não fornece uma explicação para a ação.

Assim, o caráter racionalista da teoria dos jogos se refere apenas à natureza instrumental da razão, a busca de melhores meios possíveis para satisfazer as preferências, e à formação das crenças. No entanto, esse ponto merece um pouco mais de atenção, uma vez que o requisito de consistência é um componente de racionalidade normalmente atribuído às preferências. Sugden (2006), por exemplo, argumenta que por esses requisitos de racionalidade atribuído às preferências, dentre outros motivos, David Hume não pode ser visto como precursor da teoria da escolha racional – compartilhada nesse aspecto pela teoria dos jogos.

O requisito de consistência das preferências não é compatível com o conceito de razão de David Hume, uma vez que não diz respeito a relações entre idéias ou fatos, mas apenas a relação entre impressões, nos termos de Hume. No entanto, sem esgotar essa controvérsia, mesmo com o requisito de consistência das preferências essas permanecem aracionais no sentido anteriormente definido: não são passíveis de

justificativa nem de explicação posterior. É possível pelo critério da consistência afirmar que é irracional para um agente preferir X a Y, Y a Z e Z a X; esse critério de racionalidade no entanto é apenas formal: se é possível pelo critério formal justificar porque o agente prefere X a Z dadas as suas outras preferências, não é possível justificar ou explicar seja a ordem completa de suas preferências, seja cada uma delas tomadas par a par – nesse último caso, sejam elas consistentes ou não. Nesse sentido, independente da interpretação que se dê ao papel da racionalidade na teoria da ação de David Hume, as preferências permanecem aracionais da mesma forma que os desejos.

Mas as preferências não são apenas um estado mental e aracionais como os desejos em David Hume; da mesma forma que os desejos as preferências geram a disposição para agir. Segundo Aumann (1985, p.11),

Briefly put, game and economic theory are concerned with the interactive behavior of *Homo rationalis*—rational man. *Homo rationalis* is the species that always acts both purposefully and logically, has well-defined goals, is motivated solely by the desire to approach these goals as closely as possible [grifo nosso], and has the calculating ability required to do so.

A afirmação de Aumann sintetiza a relação entre disposição para agir e preferências, embora não se refira explicitamente a essas, na teoria dos jogos: os agentes são motivados, - i.e, estão dispostos a agir, - pelo desejo de obter os seus objetivos; esses objetivos são identificados com a maximização da utilidade, o melhor resultado de acordo com suas preferências. Assim, na teoria dos jogos as preferências fornecem a motivação da mesma forma que em Hume os desejos despertam a vontade, a disposição para agir.

Mas na teoria dos jogos a ação não depende apenas das preferências mas também das crenças dos agentes. Cabe então saber como as ações dependem das crenças, o sentido dado a esse conceito na teoria dos jogos e se essa relação de dependência entre as ações e as crenças também se encontra na teoria da ação de David Hume.

Nos jogos não-cooperativos com informação incompleta, as crenças são abordadas como um cálculo probabilístico e estão sujeitas a restrição das leis de coerência formal da probabilidade – aditividade finita, não-negatividade – e condicionadas às informações possuídas pelos jogadores. A partir disso é possível entender os elementos presentes nesse conceito de crença contrastando-os com os elementos presentes no conceito de preferência.

Como as preferências, as crenças são um estado mental; mas se diferenciam daquelas por estarem voltadas para o mundo: são suposições sobre o mundo. Embora a teoria dos jogos não-cooperativos com informação incompleta diferencie o conhecimento – entendido como informações – das crenças – entendidas como probabilidades bayesianas -, ambos tem em comum o fato de dizerem respeito a “matérias de fato”, no sentido de David Hume. Assim, ao contrário das preferências as crenças não são “inerentemente pessoais”: dois jogadores com o mesmo conjunto de informações terão as mesmas crenças sobre os estados do mundo.

Enquanto as preferências são aracionais, não sendo passíveis de julgamento, o mesmo não ocorre com as crenças. É possível afirmar que uma crença é falsa ou verdadeira, seja demonstrativamente pela verificação do uso das regras da probabilidade, seja por meio da observação das relações entre eventos e a verificação de sua correspondência com a probabilidade atribuída aos jogadores para ocorrência dos eventos.

As crenças são entendidas como o componente racional da ação: para satisfazer suas preferências, os jogadores devem fazer suposições sobre o mundo de forma a escolher a melhor ação que leve ao seu objetivo. As crenças são relevantes para a escolha dos meios, tem um caráter instrumental e, portanto, não são capazes de exercer o papel de motivador da ação ao contrário das preferências.

Crenças e preferências são estritamente separadas nos jogos não cooperativos: cada jogador parte de suas preferências – que são conhecidas por eles – e escolhe sua estratégia considerando também as suposições sobre o mundo, incluindo as preferências dos outros jogadores. As crenças de um agente não dizem respeito as suas preferências e não influenciam na formação destas; da mesma forma a crença de outros jogadores a respeito de suas preferências não influenciam um jogador, bem como a sua crença sobre a preferência dos outros não influenciam as preferências destes. As crenças não só não exercem um papel de motivador ao lado das preferências como também não são um motivador de segunda ordem, capaz de influenciar de alguma forma as preferências do próprio jogador ou dos outros jogadores.

De que modo a tematização das crenças na teoria dos jogos pode ser compatível com a teoria da ação de David Hume? Para Hume (2000) a crença é “uma maneira particular de formar uma idéia”(p.126); essa maneira não está relacionada a regras lógicas, probabilísticas ou não, mas a “uma maneira diferente de sentir”; de acordo com Hume (2000) “a crença não acrescenta nada à idéia, mas apenas transforma nossa maneira de conceber, tornando-a mais forte e vívida”(p.131). Assim, para Hume a

crença está fora do campo do entendimento e da razão, sendo um sentimento em relação a uma idéia.

O conceito de crença de David Hume não é equivalente ao conceito de crença presente na teoria dos jogos. No entanto, a teoria dos jogos permanece compatível com a teoria da ação de David Hume na medida em que a formação bayesiana de crença se enquadra no campo do entendimento, dos juízos a respeito de matérias de fato. Nesse sentido as crenças exercem na teoria dos jogos não-cooperativos com informação incompleta o papel auxiliar da ação que Hume atribui à razão.

2.2 – Causalidade humeana e ação na teoria dos jogos

Nessa seção é argumentado que a relação de dependência entre as preferências e a ação na teoria dos jogos é uma relação de dependência causal, na qual a causalidade é entendida no mesmo sentido que em David Hume. Início então com a exposição da noção de causalidade em David Hume para posteriormente argumentar que a teoria da ação presente na teoria dos jogos atende aos requisitos da causalidade humeana.

No livro I de seu “Tratado sobre a Natureza Humana” “ Do entendimento”, Hume define os conteúdos mentais como percepções que se originam dos sentidos ou da reflexão. As percepções dos sentidos são denominadas de impressões; as percepções que se originam da reflexão tem por origem primeira uma impressão e são chamadas de idéias. As percepções segundo Hume se relacionam pela associação, pela faculdade da mente de estabelecer relações entre percepções.

Para Hume a causalidade consiste no estabelecimento de três relações entre percepções. A primeira delas é a de contigüidade, a proximidade física entre causa e efeito; a segunda é a de sucessão no tempo: as causas precedem os efeitos. A terceira e mais importante relação presente na causalidade segundo Hume é a idéia de conjunção constante. Segundo Hume a experiência nos fornece exemplos frequentes de “uma ordem regular de contigüidade e sucessão” entre dois objetos. A recordação dessa conjunção constante nos faz passar da impressão sensível de um objeto para a idéia de outro que não está presente aos sentidos e essa relação é o que Hume denomina de relação de causa e efeito (Hume, 2000, p.116-117). Segundo Hume: “(...) a transição que fazemos de uma impressão presente à memória ao aos sentidos, para a idéia de um objeto que denominamos de causa ou efeito está fundada na *experiência* passada e em nossa lembrança de sua *conjunção constante*.”(p.117)

É importante assinalar que a causalidade para Hume é uma relação estabelecida na mente: “Nenhum objeto implica a existência de outro se considerarmos esses objetos em si mesmos (...)” (Hume, 2000, p.115) e ainda “Considerados em si mesmos, todos os objetos que chamamos de causas e efeitos são tão distintos e separados uns dos outros quanto de qualquer outra coisa na natureza.” (Hume, 2000, p.441). Pode se inferir daí que para Hume, fora da associação feita pela mente, os eventos se relacionam de forma atomística, sem nenhuma conexão necessária.

Como então a mente pode estabelecer uma relação necessária entre causa e efeito se essa relação não se apresenta na natureza? Para Hume a experiência da conjunção não é o suficiente pois, por mais experiências que tenhamos tido de uma conjunção constante nada nos autoriza que ela seja venha a ser observada novamente. Somente o costume, a repetição da experiência de uma conjunção constante permite a mente fazer inferências causais (Hume, D. 2000, p.188-189).

Dos parágrafos acima podemos chegar a uma definição de causalidade de David Hume. Para Hume a causalidade pode ser entendida como uma relação de conjunção constante estabelecida na mente e formada pelo hábito. A partir da definição de causalidade Hume estabelece regras para identificar uma causa: a contigüidade entre causa e efeito, a anterioridade da causa, a conjunção constante e a regra de que “A mesma causa sempre produz o mesmo efeito” (Hume, D. 2000, p.207) são as mais importantes.

Segundo Hume, as mesmas relações que nos levam a atribuir a causalidade entre corpos físicos também podem ser estabelecidas entre as ações e os motivos: “Nenhuma união pode ser mais constante e certa que a de algumas ações com determinados motivos e caracteres” (p.440) e “(...) frequentemente existe uma conjunção constante das ações da vontade com seus motivos(...)” (p.698). Como os motivos para Hume são as paixões e as paixões indiretas derivam do desejo, Hume oferece uma explicação da ação em que o desejo figura como causa.

Da mesma forma que o conceito de preferências na teoria dos jogos é equivalente ao conceito de desejo em David Hume, a relação entre preferências e ação na teoria dos jogos é a mesma relação causal entre desejos e ações em David Hume. Nesse sentido será examinado a seguir se é possível identificar causas da ação na teoria dos jogos a partir das regras definidas por David Hume.

A relação de contigüidade presente na noção humeana de causalidade não se coloca no caso das ações uma vez que esta não se refere a relação entre objetos físicos. A relação de precedência das causas com relação aos efeitos está presente na relação

entre preferências e ação na teoria dos jogos. As preferências são um dado no início de um jogo e os jogadores escolhem suas estratégias a partir dessas preferências; em outras palavras a formação de preferências precede a escolha da melhor ação. Além disso a escolha de uma estratégia não modifica as preferências; novas preferências significam um novo payoff e portanto um novo jogo.

A regra da conjunção constante na teoria dos jogos não pode ser examinada independente da regra da “mesma causa para o mesmo efeito”. Pelo menos a primeira vista esta última regra é violada na teoria dos jogos: nas estratégias mistas, dadas as suas preferências, não há um único curso de ação para os jogadores; estes mudam seu curso de ação com determinada probabilidade ao longo do jogo. No caso de equilíbrios múltiplos podem haver diferentes estratégias que levam o agente a satisfazer suas preferências de forma ótima. A conjunção constante entre a ação e as preferências só pode ser observada caso seja considerado o elemento adicional das crenças; mas a primeira vista isso violaria a regra da mesma causa para o mesmo efeito.

A violação da regra da mesma causa para o mesmo efeito pode ser interpretada como sendo resultado da adoção pela teoria dos jogos de uma variante da causalidade humeana que admite a existência de uma pluralidade de causas (Mill, J. S.; 1979). Nos jogos com informação completa a ação depende das preferências uma vez que as crenças são supostas como dadas e verdadeiras; nos jogos com informação incompleta e com as preferências dadas a ação depende das crenças.

Embora a variante da pluralidade de causas admita a ausência da regularidade na ação e portanto que essa seja incerta, essa incerteza é atribuída a descrição incompleta das condições que determinam a ação. Na interpretação dominante,

A player's mixed strategy is thought of as a plan of action which is dependent on private information which is not specified in the model. Although the player's behavior appears to be random, it is actually deterministic. If we add this information structure to the model, the mixed strategy becomes a pure strategy in which the action depends on the extraneous information. (Rubinstein, 1991, p.914).

Assim a idéia de que as ações dependem das preferências e das crenças não viola o requisito de regularidade de eventos da causalidade humeana; as ações estão sempre em conjunção constante com dadas preferências e crenças.

Mas é possível também interpretar a relação entre preferências, crenças e ação na teoria dos jogos de forma mais fiel a teoria original de Hume. Segundo Hume para que seja estabelecida a relação de conjunção constante entre causa e efeito é preciso que haja uma “repetição de objetos similares em condições similares”(Hume, D.; 2000, p.

197).Ao explicar o que entende por conjunção constante no famoso exemplo da bola de bilhar, Hume afirma que:

A primeira bola está em movimento; toca a segunda; imediatamente a segunda se movimenta; e quando repito o experimento com a mesma bola, ou com bolas semelhantes, **na mesma situação ou em circunstâncias semelhantes** [grifo nosso], constato que, quando uma bola se move e toca a outra, segue-se sempre um movimento da segunda bola(Hume, D. 2000, p.687)

Na interpretação aqui proposta, as crenças se referem as circunstâncias em que ocorre a conjunção entre preferências e ações e não a causa das ações; dessa forma, as mesmas preferências com crenças diferentes implicam em ações diferentes não porque haja mais de uma causa para o mesmo efeito, mas porque as circunstâncias não são as mesmas ou semelhantes. Defendo essa interpretação a partir de dois argumentos apresentados a seguir.

O primeiro argumento se refere ao conceito de crenças na teoria dos jogos. Como visto anteriormente as crenças são suposições sobre o mundo formadas a partir das informações dos jogadores; sendo assim, as crenças se referem às circunstâncias em que os jogadores decidem agir. Diferentes informações significam diferentes circunstâncias em que os jogadores decidem suas estratégias dadas suas preferências, e portanto em condições dessemelhantes em que ocorre a conjunção entre preferências e crenças.

O segundo argumento se refere ao caráter auxiliar, instrumental, das crenças com relação a ação. Embora com diferentes informações, e portanto diferentes crenças, as mesmas preferências possam causar diferentes ações, as crenças sozinhas não são capazes de causar qualquer ação. Da forma com as preferências são conceituadas é difícil imaginar um agente que não as tenha; a situação mais próxima disso seria aquela em que um jogador - um monge budista, por exemplo – fosse totalmente indiferente com relação ao resultado do jogo. Não se pode dizer que o monge budista não tenha crenças, que não tenha suposições sobre o mundo com base em suas informações; no entanto, sendo indiferente a qualquer resultado dos jogos, nosso hipotético monge não se decidiria com relação a qualquer estratégia.

A partir dos argumentos apresentados nos parágrafos anteriores me parece ser mais razoável interpretar que a teoria dos jogos assume as preferências como causa da ação, no sentido de David Hume, sendo as crenças um elemento derivado das circunstâncias em que ocorre a conjunção entre as preferências e as ações.

Um último argumento em favor da adoção da noção humeana de causalidade na teoria dos jogos é a adesão explícita de alguns autores à visão de mundo implícita na referida noção de causalidade. Conforme afirmado anteriormente para Hume os eventos se relacionam de forma atomística, sem nenhuma conexão necessária, sendo a causalidade estabelecida por uma associação mental. Em uma entrevista, Robert Aumann deixa explícita sua adesão aos supostos ontológicos implícitos na noção humeana de causalidade:

What we do in science is that we organize things, we relate them to each other. (...) These relationships exist in the mind of human beings, in the mind of observer, in the mind of scientist. The world without the observer is a chaos, it is just a bunch of particles flying around, it is “Tohu Vavohu” - the biblical description of situation before creation (Aumann, R. apud Van Damme, 1998, p.182).

Rubinstein (1991), após citar o filósofo positivista Rudolf Carnap, afirma: “Game Theory should be employed as a descriptive science only after the social scientist observe or expect some regularities in family or “similar” events.” (Rubinstein, 1991, p.921).

3. Teoria dos jogos e explicação causal da ação : uma avaliação preliminar.

Nessa seção procura-se realizar uma avaliação preliminar da explicação para a ação humana adotada pela teoria do jogos. É apresentado o argumento de que as preferências não fornecem um elemento suficiente para explicar a ação e que a adoção dos critérios humeanos de causalidade omitem o elemento da alternativa como característico da ação.

Conforme argumentado na seção anterior as preferências, concebidas como aracionais e capazes de dispor a ação, figuram na teoria dos jogos como causa da ação. Nessa explicação, as preferências são constitutivas dos motivos e da finalidade dos agentes: o que leva alguém a agir é a finalidade de satisfazer suas preferências. O problema é que as preferências por si só não são capazes de comprometer o agente com a execução de uma determinada ação.

Um agente pode preferir bebida alcoólica à refrigerantes, mas não há nada nessa caracterização das preferências que leve o agente a executar a ação de ir a loja de bebidas para comprar vinho. Tal como caracterizadas, as preferências podem dispor o agente a ir a loja para comprar vinhos mas não implicam no comprometimento do agente em executar essa ação; esse comprometimento requer uma intencionalidade, uma

posição de finalidade por parte do agente. A formação da intencionalidade pressupõe a decisão consciente de executar determinada ação e a partir dessa decisão o comprometimento em realizar o esforço necessário para alcançar a finalidade intencionalmente posta. Dessa forma a intencionalidade não pode ser reduzida ao ordenamento das preferências uma vez que essas são aracionais e portanto, não são objeto de deliberação e justificativa.

Se a formação da intencionalidade requer a deliberação consciente quanto aos fins, ao decidir por uma ação o agente escolhe entre fins alternativos, o que leva a escolha dos meios e ao comprometimento com a ação. Assim, nosso agente que prefere vinhos a refrigerantes, a partir de uma decisão consciente de beber uma garrafa de vinho irá analisar os meios para atingir sua finalidade – verificando suas possibilidades orçamentárias, o acesso a uma loja de bebidas, etc. - e se comprometerá com a ação de ir até a loja de bebidas e adquirir uma garrafa de vinho.

Se a ação é entendida como resultado da posição de finalidade, uma série de considerações não apreendidas pela simples caracterização das preferências surge como elemento determinante da ação da decisão entre fins alternativos. A decisão entre fins alternativos pode ser influenciada pela disponibilidade dos meios – por exemplo a ausência de recursos orçamentários suficientes para adquirir a garrafa de vinho mas não para um refrigerante pode levar o agente a deliberar pela finalidade de adquirir o refrigerante em lugar do vinho. Outro elemento que pode ser considerado pelo agente na deliberação da finalidade da ação é a fraqueza de vontade – embora o agente queira satisfazer a preferência ou desejo por uma garrafa de vinho não está disposto a realizar o esforço necessário para alcançar tal finalidade. Finalmente, a incompatibilidade entre fins alternativos também aparece como elemento determinante da posição de finalidade; por exemplo, se nosso agente é um alcoólatra em recuperação, embora tenha uma preferência intrínseca por vinhos delibera conscientemente por adquirir uma garrafa de refrigerante². A intencionalidade implica necessariamente no caráter de alternativa das finalidades da ação.

A argumentação anterior não significa a substituição das preferências pela intencionalidade como causa da ação; ao contrário o reconhecimento da intencionalidade como elemento determinante da ação implica na rejeição da explicação causal humeana. A incompatibilidade entre a explicação causal da ação e a intencionalidade decorre de que o elemento da alternativa está ausente na teoria causal

da ação. Segundo a explicação causal da ação adotada pela teoria dos jogos o agente escolhe a ação dadas suas preferências; mas não há deliberação consciente, e portanto alternativa, quanto as preferências: estas são aracionais e funcionam como um dado antecedente a decisão de agir. Além disto só há uma ação possível para o agente – considerando-se as informações e conseqüentemente as crenças dos agentes - dadas as suas preferências. Em outras palavras o requisito de regularidade ou conjunção constante entre motivos e ação presente na explicação causal elimina a alternativa e a posição de finalidade da ação humana.

O argumento apresentado até aqui leva a consideração de que uma explicação da ação que seja capaz de incluir o comprometimento e a alternativa, e portanto com maior poder explicativo do que a explicação causal humana, deve se fundamentar na posição de finalidades, ou seja na explicação da ação como teleologicamente guiada. Essa explicação para a ação está presente na obra de Marx e foi desenvolvida e sistematizada pelo filósofo húngaro Gyorg Lukács em sua obra *Ontologia do Ser Social*. Não cabe aqui apresentar todas as dimensões e conseqüências dessa explicação para a ação, deve-se primeiramente ressaltar que tal explicação não elimina a dimensão causal presente na ação e em segundo lugar indicar algumas conseqüências da ausência do compromisso e da posição de finalidades na explicação causal da ação para a teoria dos jogos.

Partindo do trabalho como modelo de toda práxis social, Lukács aponta a posição de finalidades como elemento diferenciador da ação humana em relação a simples atividade reprodutiva dos animais. Na origem da posição de finalidade se encontra uma necessidade como no caso dos animais; no caso da ação humana, no entanto, duas diferenças se colocam em relação a simples atividade reprodutiva dos animais. Primeiramente a resposta à se apresenta como produto do comportamento consciente e não da mera espontaneidade; os homens não respondem a suas necessidade da forma programada em seus genes mas através da deliberação e posição de alternativas: alternativas com relação aos meios para atender determinados fins, alternativas em relação aos fins (Lukács, G. p.18). Em segundo lugar a ação humana responde a necessidades para além da reprodução biológica, necessidades de natureza social.

A posição de uma finalidade inicia um processo que é guiado pela obtenção do fim, um processo teleológico; pode se falar então de uma ação teleologicamente guiada.

² Esse exemplo poderia ser absorvido na explicação causal por meio da teoria das preferências reveladas. No entanto, como aponta Hausman, D. (2000) a teoria dos jogos é incompatível com a teoria das

Isso explica o comprometimento do sujeito com a ação; para que se obtenha o fim é necessário que o agente monitore sua ação, não ponha novas finalidades no curso de sua ação ou ceda a afetos que impeçam a realização do fim posto inicialmente, verifique os resultados parciais de sua ação e se preciso corrija o curso de sua ação, etc. (Lukács, G. p.23).

A ação teleologicamente guiada, no entanto, não está livre do enfrentamento com causalidades³, nem em sua origem nem em sua efetivação. Como afirmado anteriormente, na origem da posição de finalidade se encontra uma necessidade – de caráter biológico ou social – que não foi posta pelo sujeito que age. A efetivação da finalidade por sua vez requer o conhecimento das causalidades presentes no meio sobre o qual se age e o seu direcionamento para o fim a ser alcançado. Duas implicações podem ser retiradas dessa relação entre causalidade e teleologia: por um lado, a causalidade passa a ser uma causalidade posta, capaz de gerar efeitos que não estariam presentes se não houvesse a intervenção da ação teleologicamente guiada; e por outro lado, a finalidade posta, para que seja efetiva, permanece submetida ao conhecimento das causalidades existentes: “(...) o homem só pode pôr aqueles fins cujos meios adequados a sua efetivação realmente domina” (Lukács, G. p.11).

Assim, na ação teleologia e causalidade interagem formando um todo complexo:

“Sem dúvida, estas permanecem contrapostas, mas apenas no interior de um processo real unitário, cuja mobilidade é fundada na interação destes opostos e que, para tornar real essa interação, age de tal modo que a causalidade, sem ver atingida a sua essência, também ela se torna posta” (Lukács, G. p.7-8).

A interação entre causalidade e teleologia delineia um quadro explicativo bastante diferente do explicação das preferências como causa. Na explicação causal está ausente qualquer causalidade posta; a ação é explicada por causalidades do mesmo tipo que a natural sem nenhuma modificação. Assim, não há qualquer tipo de diferenciação entre a ação humana e a atividade dos animais; basta lembrar que Hume argumenta em favor de sua explicação da ação humana a presença dos mesmos elementos na atividade animal (Hume, 2000, livro I ,parte III, seção XV; livro II, parte I, seção XII e parte II seção XII). Por outro lado, sendo as preferências algo inerentemente pessoal e aracional não apresenta qualquer tipo de constrangimento, seja na sua origem a partir das necessidades seja por uma consideração racional dos meios para efetivação dos fins.

preferências reveladas.

³ Não cabe aqui contrapor o conceito de causalidade presente na explicação da ação como teleologicamente guiada ao conceito humeano de causalidade. Seguindo Lukács, compreende-se a causalidade como “um princípio de automovimento que repousa sobre si mesmo”(p.5) em contraposição à teleologia.

Se no trabalho o meio sobre o qual se age e o caráter da finalidade posta se referem primordialmente à natureza, à transformação de objetos materiais para a satisfação das necessidades, no caso da ação puramente social, o meio e a natureza da finalidade se modificam. Aqui o fim é obter determinados comportamentos e ações de outros sujeitos, fazer com que estes efetivem determinadas posições teleológicas (Lukács, G. p.26). Aqui os elementos apontados anteriormente também estão presentes, de forma modificada: a ação responde a uma necessidade - social ou mesmo biológica – e sua efetivação requer o conhecimento e direcionamento de causalidades de natureza social (Lukács, G. p.65-66).

Esse último ponto nos remete a deficiências explanatórias específicas da teoria dos jogos, em especial dos jogos não cooperativos estáticos. A teoria dos jogos pretende ir além da teoria neoclássica e da escolha racional ao proclamar a tematização da interação social como seu objetivo fundamental. Ao tematizar a interação omitindo a posição de finalidade, a teoria dos jogos deixa de abordar um aspecto fundamental da interação social/humana, a ação dos homens uns sobre os outros. A teoria dos jogos compreende a interação como a situação em que o resultado da ação de um depende da ação dos outros; esse caráter da interação aparece na influência que as crenças em relação aos outros tem na determinação da estratégia dos agentes. As preferências dos agentes no entanto permanecem como um dado e não são alteradas pela interação; em outras palavras a interação social não é capaz de influenciar as motivações dos agentes.

4. Considerações finais

Nesse artigo foi argumentado que a teoria dos jogos adota uma explicação insuficiente para a ação humana. A teoria dos jogos surge a princípio como uma possível alternativa a economia Robinson Crusoe do modelo neoclássico padrão, ao colocar os agentes em uma situação de interação. Ao se basear na concepção humeana da ação no entanto, a teoria dos jogos é incapaz de explicar o comprometimento e a alternativa como características da ação humana. As preferências, apreendidas como causa da ação, não são suficientes para comprometer um agente com uma determinada ação e o requisito da conjunção constante, presente no conceito humeano de causalidade, é incompatível com a decisão entre alternativas.

A adoção da explicação causal para a ação é especialmente contraproducente para a teoria dos jogos que se propõe tematizar a interação humana; em última análise ao eliminar as características específicas da ação humana de seu quadro explicativo a

teoria dos jogos termina por oferecer uma explicação meramente formal da interação, aplicável a interação entre animais em sua atividade de reprodução biológica, como nos jogos evolucionários da biologia.

O argumento acima remete a um último ponto, não explorado em todas suas implicações nesse artigo. A adoção da explicação causal humeana para a ação parte de uma ontologia implícita que apreende o mundo como um caos de eventos não-estruturados; nesse caso o que resta ao pesquisador é construir definições formais que sejam úteis. Uma alternativa crítica se funda no conhecimento ontologicamente fundado, que apreenda as categorias realmente presentes na realidade. Tal alternativa é oferecida pela explicação da ação como teleologicamente guiada

5. Bibliografia

- AUMANN, R. "Game Theory" In: Eatwell, J.; Milgate, M. e Newmann, P.(ed); **The New Palgrave: A Dictionary of Economics**. Vol. 2, Macmillan, London, 1987, pp.460-482.
- AUMANN, R. "Correlated Equilibrium as a Expression of Bayesian Rationality". **Econometrica**, vol.55, nº1, Jan.1987, p.1-18
- AUMANN, R. "What is Game Theory Trying to Accomplish?" Arrow, K. e Honkapohja, S. **Frontiers of Economics**, Basil Blackwell, Oxford, 1985
- COLANDER, D. *ET ALII*. "The Changing Face of Mainstream Economics". **Review of Political Economy**, vol.16, nº4, October, 2004, pp.485-499.
- DAMME, E. "On the State of Art in Game Theory: An Interview with Robert Aumann" **Games And Economic Behavior**, nº24, 1998 , pp.181-210
- HARSANYI, J. "Games With Incomplete Information" **Nobel Lecture**, December, 1994.
- HAUSMAN, D. "Revealed Preference, Belief, and Game Theory". **Economics and Philosophy**, vol.16, 2000, p.99-115.
- HUME, D. **Tratado Sobre da Natureza Humana**. Unesp, São Paulo, 2000.
- LEONARD, R. "From Parlor Games to Social Science: Von Newmann, Morgenstein, and the Creation of Game Theory 1928-1944." **Journal of Economic Literature**, Vol. 33, nº2, June 1995, pp.730-761.
- LUKÁCS, G. **Para Ontologia do Ser Social. O Trabalho**. Tradução de Ivo Tonet, mimeo, s.d.

- MILL, J. ST. **Sistema de Lógica Dedutiva e Indutiva**. Os Pensadores. Abril Cultural, São Paulo, 1979.
- RUBINSTEIN, A. "Comments on the Interpretation of Game Theory." **Econometrica**, vol.59, n°4, July, 1991, pp. 909-924.
- RUBINSTEIN, A. **Lectures Notes in Microeconomic Theory**. Princenton University Press, Princenton, 2006.
- SUGDEN, R. "Hume's Non-instrumental and Non-propositional Decision Theory." **Economics and Philosophy**, vol.22, 2006, pp.365-391.
- VON NEWMANN, J. AND MORGENSTEIN, O. **Theory of Games and Economic Behavior**. Science Editions, New York, 1964.