

# Auto-Regulação de Trocas Sociais Baseadas em Personalidades em Sistemas Multiagentes

Luciano V. Gonçalves, Diego R. Pereira e Graçaliz P. Dimuro

<sup>1</sup>Programa de Pós-Graduação em Informática – Universidade Católica de Pelotas  
Rua Felix da Cunha 412 – Pelotas 96010-000 – RS – Brazil

**Abstract.** *This work extends the personality-based social exchange simulator, by introducing a mechanism of self-regulation of social exchanges in the agents themselves. This system is based on a hybrid approach BDI-POMDP, where BDI agent plans are extracted from POMDP optimal policies, for the agents' decision making process. Results of simulations between pairs of agents with different personality traits are also presented. The implementation was done in AgentSpeak using Jason.*

**Resumo.** *Este trabalho estende o simulador de trocas sociais baseadas em personalidades, com a introdução de um mecanismo de auto-regulação de trocas sociais internalizado em cada agente. Este sistema utiliza uma abordagem híbrida BDI-POMDP, em que planos dos agentes BDI são extraídos de políticas ótimas de POMDPs, no processo de tomada de decisão dos agentes BDI. Apresentam-se também resultados de simulações de trocas sociais entre pares de agentes com diferentes traços de personalidades. A implementação foi realizada em AgentSpeak usando o Jason.*

## 1. Introdução

A regulação das interações de agentes baseada na teoria social de *Piaget* [Piaget 1995] foi proposta primeiramente por [Dimuro et al. 2005]. Segundo esta teoria social, as interações são vistas como trocas de serviços entre pares de agentes, e as avaliações qualitativas destas trocas pelos agentes geram os chamados valores de trocas sociais: valor de investimento pela ação realizada e o valor de satisfação pelo serviço recebido. As trocas também geram valores de débitos e créditos que ajudam na garantia da continuidade de interações futuras. Um sistema multiagente (MAS) é dito estar em equilíbrio se o balanço dos valores de trocas de investimento e satisfação estiverem em equilíbrio nas trocas ocorridas ao longo do tempo. Um mecanismo (centralizado) para a regulação de trocas sociais em MAS, baseado no conceito de supervisor de equilíbrio com um Processo de Decisão de Markov Qualitativo Intervalar associado (QI-MDP), foi introduzido em [Dimuro et al. 2005]. Esta abordagem foi estendida para considerar agentes baseados em traços de personalidades [Dimuro et al. 2006, Dimuro et al. 2007].

O objetivo deste trabalho é ir em direção a auto-regulação das trocas sociais (distribuída), descentralizando o supervisor de equilíbrio através da internalização do processo de decisão das interações dos agentes. Para reduzir a complexidade do modelo, consideraremos que as trocas entre cada par de agentes não influem nas trocas ocorridas por outros pares de agentes, o que nos permite simplificar a distribuição do processo de decisão. Além disso, os agentes não têm necessariamente acesso ao estado interno (balanços de

valores trocas) de outros agentes, e, assim, o processo de decisão deve operar em um modelo parcialmente observável.

Dando continuidade aos trabalhos anteriores, escolheu-se a arquitetura BDI para os agentes, de forma a estender o simulador de trocas sociais já desenvolvido [Dimuro et al. 2006, Dimuro et al. 2007]. Na arquitetura BDI, os agentes têm um conjunto de crenças sobre o estado do mundo e um conjunto de desejos, que identificam os estados que os agentes têm como objetivos. Para um processo de deliberação, o agente gera uma ou mais intenções (consideradas aqui como os estados que o agente quer alcançar). Logo, o agente constrói um plano para atingir estas intenções (p.ex através de um raciocínio meio-e-fim), e o executa. Assim, sendo esta uma aproximação heurística, os agentes BDI apresentam freqüentemente um desempenho pior que os agentes que utilizam Processos de Decisão de Markov Parcialmente Observáveis (POMDP) [Kaelbling et al. 1998], quando uma solução POMDP é tratável. Entretanto, o modelo BDI pode abranger um conjunto de problemas bem maior e mais complexo que as soluções POMDPs, e, além disso, podem ter um desempenho aproximado, para alguns problemas relativamente pequenos [Simari and Parsons 2006].

Modelos híbridos tem sido propostos para unirem as vantagens do modelo POMDP aos da arquitetura BDI [Schut et al. 2002, Simari and Parsons 2006, Pereira and Dimuro 2007, Pereira 2008]. Em [Simari and Parsons 2006], o relacionamento formal existente entre certos componentes do modelo BDI e os do POMDP foram discutidos, mostrando como planos BDI e políticas POMDPs podem ser relacionados. Assim, foi apresentado que os planos derivados de uma política são também adotados pelo agente BDI, que seleciona planos com a maior utilidade, e que possui uma estratégia ótima de reconsideração. Em [Pereira 2008], foi introduzido um algoritmo que constrói planos BDI em AgentSpeak a partir de grafos de políticas de POMDP.

Neste trabalho, a decisão sobre a melhor troca que um agente deve sugerir a seu parceiro de interação, a fim de alcançar o equilíbrio social, ou para promover novas interações, é modelado através de um POMDP para cada traço de personalidade, que seu parceiro de interação pode assumir. Adotando-se um conjunto de seis traços de personalidades (egoísta, egoísta forte, altruísta, altruísta forte, tolerante e fanático), cada um destes POMDP foi dividido em 3 sub-POMDPs, de acordo com o balanço do estado corrente (favorável, equilibrado, desfavorável), visando reduzir a complexidade dos cálculos.

O artigo está organizado da seguinte forma. Na Seção 2 são discutidos sucintamente os conceitos principais da modelagem das trocas sociais adotadas nesse trabalho. A Seção 3 apresenta os traços de personalidades que os agentes podem assumir durante os processos de trocas. Já a Seção 4 mostra os resultados e simulações realizadas no simulador de trocas. Por fim, a Seção 5 relata as principais conclusões.

## **2. A Modelagem das Trocas Sociais**

O sistema de valores de trocas apresentado em [Dimuro et al. 2005] está centrado na teoria dos valores de trocas sociais de Piaget [Piaget 1995], sendo esta a abordagem para a regulação de interações em sistemas multiagentes adotada neste trabalho.

Piaget modelou as interações entre os indivíduos como trocas de serviços, envolvendo não somente a realização de um serviço, mas também a avaliação de tal serviço

pelos envolvidos. Esta avaliação do serviço é realizada com base em escalas de valores de trocas. Na teoria de Piaget, os valores são de natureza qualitativa (bom, ruim, melhor que, pior que, etc.).

Com o objetivo de capturar a natureza qualitativa dos conceitos dos valores de trocas de Piaget, Dimuro e Costa [Dimuro et al. 2005] buscaram nas técnicas da Matemática Intervalar uma forma de representá-los em sistemas quantitativos. Assim, desenvolveram uma álgebra, no qual todo valor de troca pode ser representado por um intervalo real  $x = [x_1; x_2]$ , em que  $x_1, x_2 \in \mathbb{R}$  e  $[-L \leq x_1 \leq x \leq x_2 \leq L]$ , para  $L \in \mathbb{R}$ , sem perder a expressividade dos valores qualitativos na representação de incertezas.

## 2.1. Estágios de Trocas

Uma troca social entre dois agentes  $\alpha$  e  $\beta$  pode ser representada por dois estágios de trocas. No primeiro estágio  $I_{\alpha\beta}$ , o agente  $\alpha$  realiza um serviço de forma espontânea para o agente  $\beta$  e os valores de trocas envolvidos são os seguintes:  $r_{I_{\alpha\beta}}$  é o valor do investimento feito por  $\alpha$  na realização de um serviço para  $\beta$  (um valor de investimento é sempre negativo);  $s_{I_{\beta\alpha}}$  é o valor de satisfação de  $\beta$  pelo serviço recebido de  $\alpha$ ;  $t_{I_{\beta\alpha}}$  é o valor do débito adquirido por  $\beta$  para com  $\alpha$  pela satisfação com o serviço recebido; e  $v_{I_{\alpha\beta}}$  é o valor de crédito que  $\alpha$  adquiriu com  $\beta$  pela realização do serviço.

No segundo estágio  $II_{\alpha\beta}$ , o agente  $\alpha$  solicita o pagamento do serviço realizado anteriormente para  $\beta$ , e os valores gerados nesta troca são:  $v_{II_{\alpha\beta}}$ ,  $t_{II_{\beta\alpha}}$ ,  $r_{II_{\beta\alpha}}$  e  $s_{II_{\alpha\beta}}$ , que possuem significado semelhante aos valores do estágio  $I_{\alpha\beta}$ . Os valores de troca  $r_{I_{\alpha\beta}}$ ,  $s_{I_{\beta\alpha}}$ ,  $r_{II_{\beta\alpha}}$ , e  $s_{II_{\alpha\beta}}$  são chamados de valores materiais, negociados durante a realização do serviço, já os valores  $t_{I_{\alpha\beta}}$ ,  $v_{I_{\beta\alpha}}$ ,  $t_{II_{\beta\alpha}}$ , e  $v_{II_{\alpha\beta}}$  são chamados de valores virtuais, que podem ser negociados em futuro próximo. Não há uma ordem na ocorrência dos estágios de trocas  $I - II$ .

Durante um intervalo de tempo  $T$  do processo de troca entre os agentes  $\alpha$  e  $\beta$ , os balanços materiais e virtuais, sob o ponto de vista de cada agente, são acumulados e denotados por  $bm_{\alpha\beta}^T$  e  $bm_{\beta\alpha}^T$ , e  $bv_{\alpha\beta}^T$  e  $bv_{\beta\alpha}^T$ , onde:

$$bm_{\alpha\beta}^T = \sum_{t \in T} r_{(\alpha\beta)}^t + s_{(\alpha\beta)}^t \quad \text{e} \quad bv_{(\alpha\beta)}^T = \sum_{t \in T} t_{(\alpha\beta)}^t + v_{(\alpha\beta)}^t.$$

O processo de troca social entre um par de agentes  $\alpha$  e  $\beta$  é dito estar em equilíbrio material se  $bm_{\alpha\beta}^T \cong 0$  e  $bm_{\beta\alpha}^T \cong 0$ .

Neste trabalho, consideram-se três estados para descrever os resultados de trocas sociais. O estado de equilíbrio ( $bm \cong 0$ ) é denotado por  $E^0$ . O estado favorável ( $bm > 0$ ) é denotado por  $E^+$ , e o desfavorável ( $bm < 0$ ) é denotado por  $E^-$ .

O resultado material constitui o estado interno de um agente. Cada agente tem acesso somente ao seu próprio estado interno, mas é capaz de inferir sobre as informações que os outros agentes informam de seus resultados virtuais. Assim, considera-se o conjunto de estados internos de um agente  $\beta$  como  $E_\beta = \{E_\beta^-, E_\beta^0, E_\beta^+\}$ . O conjunto de observações dos possíveis resultados virtuais de um agente  $\beta$  é dado por  $\Omega_\beta = \{D_\beta, N_\beta, C_\beta\}$ , onde  $D_\beta$  e  $C_\beta$  representam que o agente  $\beta$  tem débitos ou créditos, respectivamente, assim como  $N_\beta$  representa que  $\beta$  não tem nem débitos e nem créditos.

### 3. Trocas entre Agentes Baseados em Traços de Personalidades

Os agentes podem ter diferentes traços de personalidades que os levam a terem diferentes funções de transição de estado, as quais especificam, dado o estado atual do agente e a troca proposta, uma distribuição de probabilidade sobre o conjunto de estados que os agentes podem atingir após a interação. A seguir, alguns destes traços de personalidade são ilustrados:

- **Egoísta:** o agente na maioria das vezes busca o seu benefício próprio, com alta probabilidade de aceitar trocas que representem transições para resultados favoráveis (trocas nas quais o outro agente realize um serviço para ele); se o agente também tiver alta probabilidade de super-avaliar seus créditos e sub-avaliar seus débitos, este é denominado de **egoísta forte**.
- **Altruísta:** o agente na maioria das vezes busca o benefício de seu parceiro, com alta probabilidade de aceitar trocas que representem transições para resultados favoráveis de seu parceiro (trocas nas quais ele possa realizar um serviço para seu parceiro); se o agente também tiver alta probabilidade de sub-avaliar seus créditos e super-avaliar seus débitos, este é denominado de **altruísta forte**.
- **Fanático pelo Equilíbrio:** o agente tem alta probabilidade de aceitar trocas que o leve para o estado de equilíbrio, rejeitando outras trocas que o tirem desse estado.
- **Tolerante:** o agente tem alta probabilidade de aceitar as trocas propostas.

**Tabela 1. Função de transição de estados para um agente baseado em personalidade  $\beta$**

(a) Estágios de troca tipo - $I_{\alpha\beta}$						
$\Pi(E_\beta)$	Agente Egoísta			Agente Altruísta		
	$E^0$	$E^+$	$E^-$	$E^0$	$E^+$	$E^-$
$E^0$	muito-baixa	muito-alta	0.0	muito-alta	muito-baixa	0.0
$E^+$	0.0	1.0	0.0	0.0	1.0	0.0
$E^-$	muito-baixa	alta	muito-baixa	muito-baixa	muito-baixa	alta
$\Pi(E_\beta)$	Agente Fanático equilíbrio			Agente Tolerante		
	$E^0$	$E^+$	$E^-$	$E^0$	$E^+$	$E^-$
$E^0$	muito-alto	muito-baixa	0.0	muito-baixa	muito-alta	0.0
$E^+$	0.0	1.0	0.0	0.0	1.0	0.0
$E^-$	alta	muito-baixa	muito-baixa	alta	alta	muito-baixa
(b) Estágios de troca tipo - $II_{\alpha\beta}$						
$\Pi(E_\beta)$	Agente Egoísta			Agente Altruísta		
	$E^0$	$E^+$	$E^-$	$E^0$	$E^+$	$E^-$
$E^0$	muito-alta	0.0	muito-baixa	muito-baixa	0.0	muito-alta
$E^+$	baixa	alta	muito-baixa	muito-baixa	muito-baixa	alta
$E^-$	0.0	0.0	1.0	0.0	0.0	1.0
$\Pi(E_\beta)$	Agente Fanático equilíbrio			Agente Tolerante		
	$E^0$	$E^+$	$E^-$	$E^0$	$E^+$	$E^-$
$E^0$	muito-alta	0.0	muito-alta	muito-baixa	0.0	muito-baixa
$E^+$	muito-baixa	alta	muito-baixa	alta	muito-baixa	alta
$E^-$	0.0	0.0	1.0	0.0	0.0	1.0

A Tabela 1(a) apresenta a distribuição de probabilidade  $\Pi(E_\beta)$  sobre o conjunto de estados  $E_\beta$ , determinado pela função de transição de estado que caracteriza um agente baseado em personalidade  $\beta$ , quando outro agente  $\alpha$  oferecer a realização de um serviço para  $\beta$ , (estágio de troca  $I_{\alpha\beta}$ ). Observa-se que, para um agente egoísta  $\beta$ , as transições que terminam em resultados favoráveis ( $E_+$ ) são as mais prováveis, logo significa que

existe uma alta probabilidade de que  $\beta$  aceite a realização de um serviço proposto por  $\alpha$ . Se  $\beta$  é um agente altruísta então há uma alta probabilidade que este recuse tal proposta, permanecendo no mesmo estado. Veja a Tabela 1(b) para comparar o comportamento de um agente  $\beta$  egoísta/altruísta, quando  $\alpha$  requisita um serviço para  $\beta$  (estágio de troca  $II_{\alpha\beta}$ ). Agora, se  $\beta$  é fanático pelo equilíbrio, então existe uma alta probabilidade que recuse transições que o levem a deixar o estado de equilíbrio ( $E^0$ ). Já o agente tolerante apresenta uma alta probabilidade de aceitar qualquer tipo de troca que lhe seja proposta.

A Tabela 2 mostra as funções de observação de valores virtuais, que geram uma distribuição de probabilidade  $\Pi(O_\beta)$  sobre o conjunto de observações  $\Omega_\beta$  que um agente  $\alpha$  pode fazer sobre os valores virtuais de um agente  $\beta$  egoísta/altruísta (forte) em qualquer um dos estados (não observáveis) de  $\beta$ .

**Tabela 2. Função de observação para um agente baseado em personalidade  $\beta$**

$\Pi(\Omega_\beta)$	Agente egoísta/altruísta			Agente egoísta forte			Agente altruísta forte		
	$N$	$D$	$C$	$N$	$D$	$C$	$N$	$D$	$C$
$E^0$	mto-alta	mto-baixa	mto-baixa	mto-baixa	mto-baixa	mto-alta	mto-baixa	mto-alta	mto-baixa
$E^+$	mto-baixa	mto-alta	mto-baixa	mto-baixa	mto-baixa	mto-alta	0.0	1.0	0.0
$E^-$	mto-baixa	mto-baixa	mto-alta	0.0	0.0	1.0	mto-baixa	mto-alta	mto-baixa

#### 4. Auto-Regulação das Trocas Sociais

O mecanismo de controle social introduzido em [Dimuro et al. 2007] foi realizado por um supervisor de equilíbrio social que, a cada instante, decide quais ações devem ser recomendadas aos agentes baseados em traços de personalidades para serem executadas de forma a conduzir o sistema para o estado de equilíbrio, reparando os balanços dos valores de trocas envolvidos em suas interações. O supervisor toma suas decisões baseadas em um (QI-MDP) [Dimuro et al. 2005] totalmente observável, que teria um completo acesso aos estados internos dos agentes.

Na busca da auto-regulação das trocas sociais é necessário descentralizar o supervisor de equilíbrio, para distribuir o processo de decisão, de forma a internalizar o mecanismo de controle social nos agentes BDI, sendo esta a arquitetura utilizada no simulador de trocas sociais [Dimuro et al. 2007]. A distribuição do processo de decisão é pura e simples, uma vez que se considera que as trocas entre cada par de agentes não têm influência sobre as trocas ocorridas entre outros pares de agentes. Entretanto, assumiu-se que para internalizar o processo de decisão, os agentes atuam em um contexto parcialmente observável, porque um não deve ter necessariamente acesso ao estado interno de outro. No entanto, assume-se que os agentes são capazes de tratar as observações como as manifestações que outros agentes fazem sobre os seus balanços virtuais, negociados em cada interação, partindo do princípio que os agentes necessariamente fazem essas manifestações, embora não necessariamente de uma forma fiel.

O processo de decisão sobre melhores trocas que um agente  $\alpha$  deve propor a um agente  $\beta$ , com intuito de alcançar o equilíbrio ou promover novas interações, é então modelado como um POMDP, denotado por  $POMDP_{\alpha\beta}$ , conforme proposto em [Pereira 2008]. Para qualquer par de traços de personalidades, a representação dos estados internos dos agentes é dada por  $(E_\alpha, E_\beta)$  em que  $E_\alpha$  é de conhecimento de  $\alpha$ , mas  $E_\beta$  não é observável por  $\alpha$ . Então, para cada traço de personalidade que  $\beta$  pode assumir, decompõe-se o  $POMDP_{\alpha\beta}$  em três sub-POMDPs, um para cada estado interno do

agente  $\alpha$ , representado por  $POMDP_{\alpha\beta}^-$ ,  $POMDP_{\alpha\beta}^0$  e  $POMDP_{\alpha\beta}^+$ , se  $\alpha$  está no estado favorável, equilibrado e desfavorável respectivamente. Assim, para cada traço de personalidade discutido na Seção 3, define-se:

**Definição 1** *Seja  $\alpha$  um agente baseado em personalidade, com um supervisor de equilíbrio internalizado, e  $E_\alpha^* \in \{E_\alpha^-, E_\alpha^0, E_\alpha^+\}$  o estado interno corrente de  $\alpha$ . Seja  $\beta$  o agente baseado em personalidade que interage com  $\alpha$ . O POMDP para o supervisor de equilíbrio internalizado em  $\alpha$  é definido como uma estrutura  $POMDP^* = (E_\beta, A_\alpha, T_\beta, \Omega_\beta, O_\beta, R_\beta)$  onde:*

- $E_\beta = \{E_\beta^-, E_\beta^0, E_\beta^+\}$  é o conjunto de estados internos de  $\beta$ ;
- $A_\alpha = \{\text{faz-serviço}, \text{pede-serviço}\}$  é o conjunto de ações disponíveis para  $\alpha$ , em que *faz-serviço* significa que  $\alpha$  faz uma oferta de serviço para  $\beta$  (troca do estágio  $I_{\alpha\beta}$ ), e *pede-serviço* significa que  $\alpha$  fez uma requisição de um serviço para  $\beta$  (estágio de troca  $II_{\alpha\beta}$ );
- $T_\beta = E_\beta \times A_\alpha \rightarrow \Pi(E_\beta)$  é a função de transição de estado baseada no modelo de personalidade de  $\beta$ , onde dado o estado corrente de  $\beta$  e a ação realizada por  $\alpha$ , resulta em uma distribuição de probabilidade sobre  $E_\beta$  (a probabilidade que  $\beta$  aceita ou recusa a troca proposta por  $\alpha$ );
- $\Omega_\beta = \{N_\beta, C_\beta, D_\beta\}$  é o conjunto de observações que podem ser inferidas por  $\alpha$  sobre a informação dada por  $\beta$  sobre os seus valores virtuais;
- $O_\beta = E_\beta \times A_\alpha \rightarrow \Pi(\Omega)$  é a função de observações baseada no modelo de personalidade do agente  $\beta$ , que dada ação realizada por  $\alpha$  e o estado resultante de  $\beta$ , indica a distribuição de probabilidade sobre o conjunto de observações possíveis  $\Omega_\beta$  (a probabilidade que  $\beta$  informará que têm débitos, créditos ou nenhum deles);
- $R_\alpha = E_\beta \times A_\alpha \rightarrow \mathbb{R}$  é a função de recompensa para o agente  $\alpha$ , dado a recompensa imediata adquirida por  $\alpha$  para cada ação recebida.

A solução de um  $POMDP_{\alpha\beta}^*$  é uma política ótima, que irá auxiliar o agente  $\alpha$  a elaborar propostas de trocas que possam levar cada agente na direção do estado de equilíbrio interno. Foi utilizado o algoritmo *Witness* [Kaelbling et al. 1998] para definir o grafo de política para dezoito modelos. Desses grafos de políticas, foram construídos planos BDI na linguagem *AgentSpeak*, utilizando o algoritmo *PolicyToBDIPlan*, introduzido em [Pereira 2008]. Esses planos foram adicionados aos agentes do Simulador de Trocas Sociais, de forma a serem analisadas durante o processo de deliberação para a tomada de decisão sobre que ações que devem ser tomadas para auto-regulação das trocas.

## 5. Simulações de Trocas Sociais Baseadas em Personalidade

A implementação do simulador de trocas sociais baseadas em personalidades foi desenvolvida em *AgentSpeak*, usado a plataforma de agentes Jason. Para cada traço de personalidade foram realizadas 10.000 simulações não supervisionadas e 10.000 simulações com 100% de supervisão, em cada simulação calculou-se o balanço médio (AV - Average balance) ocorrido durante a interação, assim como analisou o número de trocas que foram aceitas ou recusadas pelos agentes durante o processo de troca. Ao final de cada simulação geraram-se gráficos da evolução do balanço valores trocas (BV - Balance values) dos resultados materiais do agente, mostrando as primeiras 500 interações realizadas na simulação.

Um gráfico é a relação entre o balanço dos valores de trocas materiais do agente (BV) e as devidas interações (INT) que os geraram. Para sua representação gráfica, utilizou-se o ponto médio do balanço material de cada agente e dividiu-se o gráfico em três zonas: zona de equilíbrio  $[-2,2]$ , zona favorável  $]2,+\infty[$  e zona desfavorável  $]-\infty,-2[$ .

### 5.1. Simulação das trocas sociais não supervisionadas

Para gerar simulações não supervisionadas, criou-se um agente Teste  $\alpha$  que propõem trocas sociais de forma randômica (estágios  $I_{\alpha\beta}$  e  $II_{\alpha\beta}$ ) para o agente com traço de personalidade  $\beta$ , com o intuito de analisar os comportamentos dos modelos de personalidades apresentados na Tabela 1. Nestas simulações os valores trocas negociados em módulo foram os mesmos por ambos os agentes: o valor de investimento  $r_{\alpha\beta}$  foi representado pelo intervalo  $[-4,-2]$ , com ponto médio em  $-3$ , o valor de satisfação  $s_{\beta\alpha}$  foi representado pelo intervalo  $[2,4]$ , com ponto médio em  $3$ , e similarmente para os valores de débito  $t_{\beta\alpha}$  e crédito  $v_{\beta\alpha}$ .

Na simulação do agente Teste  $\alpha$  com o agente Tolerante  $\beta$ , o balanço médio encontrado após as 10.000 interações foi de aproximadamente  $AV = -148.27$  e o índice de recusa (P-R1) nas interações proposta por  $\alpha$  do estágio  $I_{\alpha\beta}$  é de somente 19.50% e as do estágio  $II_{\alpha\beta}$  (P-R2) foi de 19.80%. Por outro lado, na interação do agente Teste, com o agente Fanático o índice de recusa chegou a 82.68% das interações sugeridas pelo agente Teste (para P-R1 e P-R2), de forma que o agente se manteve em grande parte do tempo na zona de equilíbrio, com um balanço médio  $AV = -0.03$ . Já o balanço material de um agente com o modelo Altruísta alcançou o balanço médio entorno de  $AV = -6045.69$ , mostrando a sua preferência em realizar serviços para seu parceiro (trocas do estágio  $II_{\alpha\beta}$ ), ao invés de aceitar que ele faça serviço em seu favor (trocas do estágio  $I_{\alpha\beta}$ ). De maneira contrária, o agente com modelo Egoísta apresentou balanço médio dos valores trocas materiais de  $AV = 4318.93$ , mostrando a preferência por receber serviços do agente Teste (rejeitando apenas 19.74% das trocas do estágio  $I_{\alpha\beta}$ , contra 79.71% das trocas do estágio  $II_{\alpha\beta}$ ). A Figura 1 mostra os gráficos das simulações de trocas não supervisionadas do agente Teste  $\alpha$ , com os respectivos agentes  $\beta$  altruísta (a), egoísta (b), tolerante (c) e fanático (d).

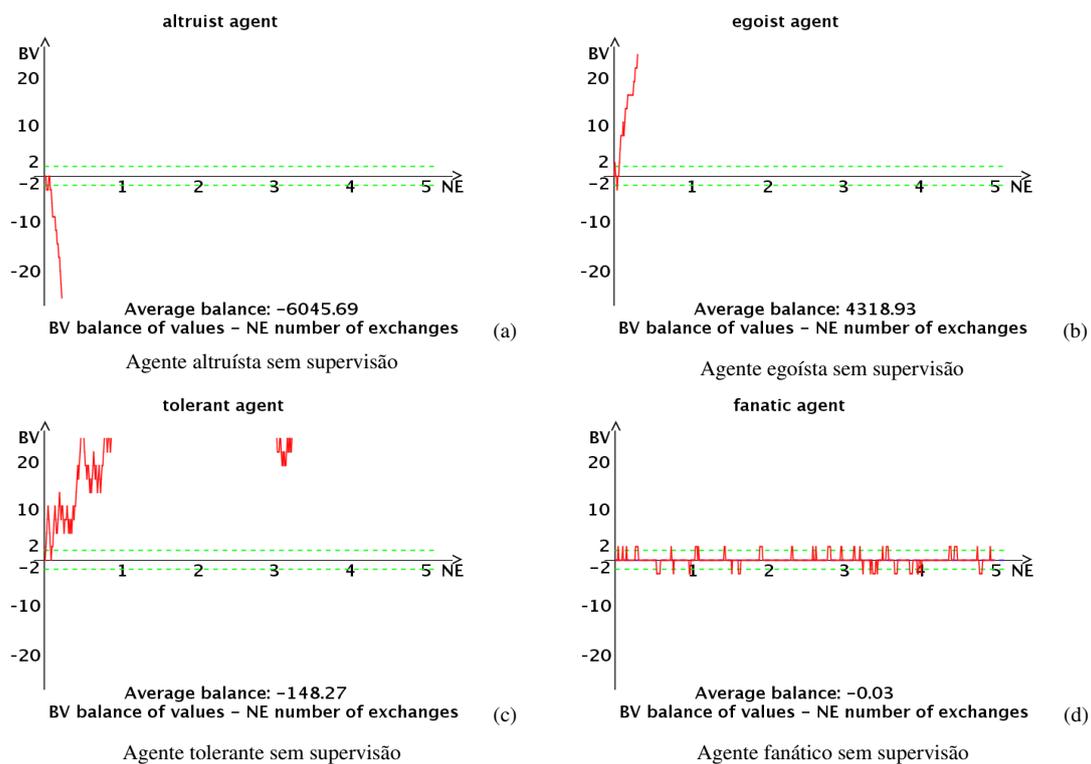
**Tabela 3. Valores médios após 10.000 interações**

Agente	AV	INT-I	INT-II	R1	R2	P-R1	P-R2
Altruísta	-6045.69	4957	5053	4465	500	90.10%	9.91%
Egoísta	4318.93	4902	5098	968	4059	19.74%	79.71%
Tolerante	-148.27	4951	5049	965	1000	19.50%	19.80%
Fanático	-0.03	5000	5000	4134	4134	82.68%	82.68%

A Tabela 3 resume os resultados médios obtidos nas 10.000 interações dos agentes baseados em traços de personalidades. Esta tabela apresenta o balanço médio obtido (AV), o número de interações do estágio  $I_{\alpha\beta}$  (INT-I) e as do estágio  $II_{\alpha\beta}$  (INT-II) propostas pelo agente  $\alpha$ , bem como as respectivas recusas dos agentes  $\beta$  a estas interações: recusas às interações do tipo  $I_{\alpha\beta}$  (R1) e recusas às interações do tipo  $II_{\alpha\beta}$  (R2). Por fim, tem-se o cálculo da porcentagem de recusas às interações propostas (P-R1, P-R2).

### 5.2. Simulação das Trocas Sociais Supervisionadas

Em trocas supervisionadas utilizou-se um grau de supervisão de 100%. Dessa forma, todas as interações proposta pelo agente  $\alpha$  (com o supervisor de equilíbrio internalizado)



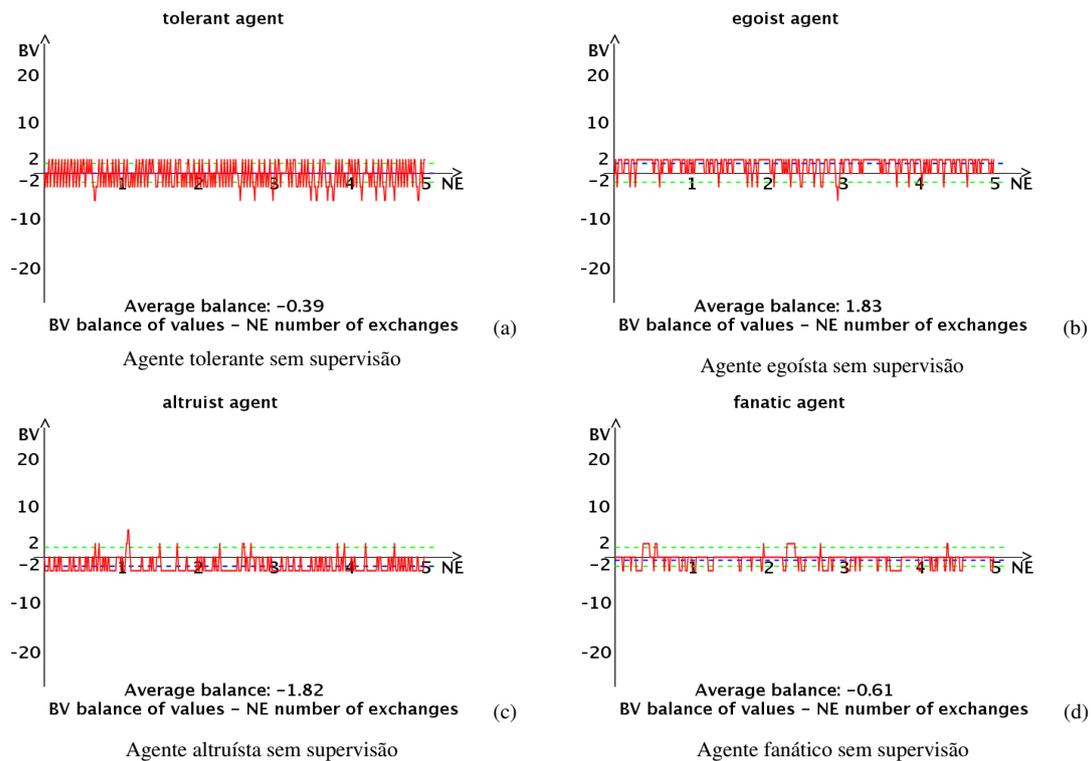
**Figura 1. Simulações dos comportamentos de agentes baseados em personalidades  $\beta$  nas primeiras 500 interações**

foram definidas pela política para o traço de personalidade de  $\beta$ . A política utilizada tem o objetivo de controlar o agente  $\alpha$  mantendo-o no equilíbrio, e, da mesma forma, propor trocas que levem o seu parceiro  $\beta$  ao equilíbrio.

Da mesma forma que nas trocas não supervisionadas geraram-se simulações com 10.000 interações, analisando-se o desempenho da política sobre os agentes com traços de personalidades. A Figura 2(a) apresenta a simulação do agente  $\alpha$  (com a política internalizada para o agente com traço de personalidade tolerante) para a interação com um agente Tolerante  $\beta$ . Nesta simulação a média de aceitação as trocas sugeridas por  $\alpha$  foi superior a 80% e o balanço médio dos valores de trocas materiais de  $\beta$  ( $AV = -0.36$ ) ficou dentro da faixa de equilíbrio esperada  $[-2,2]$ .

Na simulação com o agente  $\beta$  com traço de personalidade Altruísta, a política gerada para este agente conseguiu controlar o agente. Foram geradas 7963 interações do tipo  $I_{\alpha\beta}$  das quais o agente recusou 6349; já para as interações do tipo  $II_{\alpha\beta}$ , foram geradas 2037 interações, das quais apenas 422 foram recusadas. Dessa forma o balanço médio do agente Altruísta ficou dentro da faixa de equilíbrio ( $AV = -1.82$ ). Na interação com o agente Egoísta, geraram-se 7956 interações do tipo  $II_{\alpha\beta}$ , das quais o agente recusou 6324. Para as interações do tipo  $I_{\alpha\beta}$ , foram geradas 2044 interações,  $\beta$  recusou somente 412 interações, e o balanço médio ( $AV = 1.83$ ) também ficou compreendido na zona de equilíbrio.

Assim como na simulação não supervisionada, o agente com traço de personalidade fanático apresentou bons resultados e baixa dificuldade para alcançar o estado de equilíbrio. A Tabela 4 apresenta um resumo dos resultados obtidos nas simulações.



**Figura 2. Simulações dos comportamentos de agentes baseados em personalidades  $\beta$  nas primeiras 500 interações supervisionadas**

**Tabela 4. Valores médios após 10.000 interações**

Agente	AV	INT-I	INT-II	R1	R2	P-R1	P-R2
Altruísta	-1.82	7963	2037	6349	422	79.73%	20.71%
Egoísta	1.83	2044	7956	412	6324	20.15%	79.48%
Tolerante	-0.39	5009	4991	1005	987	20.06%	19.77%
Fanático	-0.61	2976	7024	2223	6271	74.69%	89.27%

## 6. Conclusões e Trabalhos Futuros

Este trabalho apresentou uma extensão do Simulador de Trocas Sociais [Dimuro et al. 2007] para possibilitar a auto-regulação das trocas sociais baseadas em personalidades, modelando o problema de decisão em contextos parcialmente observáveis, através de uma abordagem híbrida, em que agentes BDI, com um supervisor de equilíbrio internalizado, têm planos construídos de acordo com as políticas ótimas de POMDPs definidos para resolver o problema do equilíbrio das trocas sociais baseadas em traços de personalidades. Para isto, considerou-se um conjunto de 18 modelos POMDPs para internalizar o supervisor de equilíbrio [Pereira 2008], de forma que um agente seja capaz de se relacionar com todos traços de personalidades modelados.

As simulações mostraram que abordagem é viável e pode ser uma boa solução em aplicações contextualizadas baseadas na teoria trocas sociais, tais como as apresentadas em [Dimuro et al. 2007] (cenário candidato/eleitor), [Rodrigues and Luck 2007] (seleção de parceiros, interações cooperativas no domínio da bioinformática) e [Grimaldo et al. 2007] (sociabilidade em um cenário de bar).

Neste trabalho, partiu-se da suposição de que os agentes não têm acesso ao estado interno dos parceiros de interação, mas eles conhecem o traço de personalidade de seu par-

ceiro. Em trabalhos futuros deverão ser considerados agentes não-transparentes (agentes que restringem totalmente o acesso aos seus estados interno a outros agentes). Neste caso, o supervisor terá que ser capaz de reconhecer e manipular um modelo adequado para o traço de personalidade do seu parceiro, baseando-se somente nas observações recebidas.

**Agradecimentos.** Este trabalho é financiado pela CAPES, Petrobrás (Projeto COPPETEC) e CNPq (Proc. 473201/2007-0).

## Referências

- Dimuro, G. P., Costa, A. C. R., Gonçalves, L. V., and Hübner, A. (2006). Regulating social exchanges between personality-based non-transparent agents. In Gelbukh, A. F. and Garça, C. A. R., editors, *MICAI 2006: Advances in Artificial Intelligence*, volume 4293 of *LNCS*, pages 1105–1115. Springer, Berlin.
- Dimuro, G. P., Costa, A. C. R., Gonçalves, L. V., and Hübner, A. (2007). Centralized regulation of social exchanges between personality-based agents. In *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *LNCS*, pages 338–355. Springer, Berlin.
- Dimuro, G. P., Costa, A. C. R., and Palazzo, L. A. M. (2005). Systems of exchange values as tools for multi-agent organizations. *Journal of the Brazilian Computer Society*, 11(1):31–50.
- Grimaldo, F., Lozano, M., and Barber, F. (2007). Coordination and sociability for intelligent virtual agents. In Sichman, J., Noriega, P., Padget, J., and Ossowski, S., editors, *Coord., Organizations, Institutions, and Norms in Agent Systems III*, number 4870 in *LNAI*, pages 58–70, Berlin.
- Kaelbling, L. P., Littman, M. L., and Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2):99–134.
- Pereira, D. R. (2008). Construção de planos BDI a partir de políticas ótimas de POMDPs, com aplicação na auto-regulação de trocas sociais em sistemas multiagentes. Dissertação de mestrado, PPGINF/UCPel, Pelotas, RS.
- Pereira, D. R. and Dimuro, G. P. (2007). Um algoritmo para extração de um plano BDI que obedece uma política MDP Ótima. In *Anais do Workshop-Escola de Sistemas de Agentes para Ambientes Colaborativos*, Pelotas. PPGINF/UCPel.
- Piaget, J. (1995). *Sociological Studies*. Routledge, London.
- Rodrigues, M. and Luck, M. (2007). Cooperative interactions: an exchange values model. In *Coordination, Organizations, Institutions and Norms in MAS II*, volume 4386 of *LNCS*, pages 16–23, Berlin. Springer.
- Schut, M. C., Wooldridge, M., and Parsons, S. (2002). On partially observable MDPs and BDI models. In d’Inverno, M., Luck, M., Fisher, M., and Preist, C., editors, *Foundations and Applications of Multi-Agent Systems, Selected Papers of UKMAS Workshop 1996-2000*, volume 2403 of *LNCS*, pages 243–260. Springer, Berlin.
- Simari, G. I. and Parsons, S. (2006). On the relationship between MDPs and the BDI architecture. In Nakashima, H., Wellman, M. P., Weiss, G., and Stone, P., editors, *Proc. of the 5th Intl. Joint Conf. on Autonomous Agents and Multiagent Systems, AAMAS 2006, Hakodate, Japan, May 8-12, 2006*, pages 1041–1048. ACM.