

Uma arquitetura de Agentes BDI para auto-regulação de Trocas Sociais em Sistemas Multiagentes Abertos

Luciano V. Gonçalves, Graçaliz P. Dimuro e Antônio Carlos da R. Costa

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

Abstract. *This work presents an hybrid agent architecture for the self-regulation of social exchanges between personality-based agents in a open multiagent system. The architecture proposal follows the concepts of the BDI architecture (Beliefs, Desires, Intentions), adding a module to discover the personality traits, through HMMs (Hidden Markov Models) and a module for specifying new plans for the control of agents, through optimal POMDPs policies (Partially Observable Markov Decision Processes). The joint work of the HMMs and POMDPs creates a new deliberation cycle for the BDI agents, implementing a mechanism for social regulation at the agents's social reasoning level.*

Resumo. *Este trabalho apresenta uma arquitetura de agentes híbrida para a auto-regulação de trocas sociais executadas entre pares de agentes baseados em traços de personalidade que operam em sistemas multiagentes abertos. A arquitetura proposta segue os conceitos da arquitetura BDI (Beliefs, Desires, Intentions), adicionando um módulo de descoberta de traços de personalidade, através dos HMMs (Hidden Markov Models) e um módulo de especificação de novos planos para o controle dos agentes, através de políticas ótimas de POMDPs (Partially Observable Markov Decision Processes). O trabalho conjunto de HMMs e POMDPs cria um novo ciclo de deliberação para os agentes BDI, implementando um mecanismo de regulação social no nível do raciocínio social dos agentes.*

1. Introdução

A modelagem de interações auto-reguladas em sistemas multiagentes baseados em personalidades apresentada nos trabalhos [Dimuro et al. 2007, Pereira et al. 2008], através da teoria das trocas sociais de Jean Piaget [Piaget 1973], representa um caminho para dinamizar o processo de modelagem e controle de interações em Sistemas Multiagentes (SMA). Com base nessa teoria, as interações entre agentes são modeladas como trocas de serviços entre eles, e a avaliação deste serviço dá origem aos chamados valores de trocas sociais, cujos os balanços de trocas contínuas determinam o equilíbrio ou não das interações. Os agentes avaliam o investimento na realização de um serviço, assim como a satisfação associada ao recebimento do serviço. Logo, o sistema está em equilíbrio material se os balanços entre os valores de investimento e satisfação estão equilibrados para cada agente, após uma sucessão de trocas realizadas no tempo.

O primeiro mecanismo de regulação de trocas sociais, baseado no conceito de supervisor de equilíbrio centralizado com um Processo de Decisão de Markov Qualitativo Intervalar associado, foi introduzido em [Dimuro and Costa 2006], em que um

agente analisa e controla as interações estabelecida entre pares de agentes. Já o trabalho [Pereira 2008], deu início a um processo de controle descentralizado, para o mecanismo de regulação de interações, internalizando nos agentes regras de controle interno. Para isso, foi necessário trabalhar com Processos de Decisão de Markov Parcialmente Observáveis (POMDP) [Kaelbling et al. 1998], visto que os agentes não tinham acesso aos estados internos dos outros agentes. Também neste trabalho, foi apresentado um método para extrair planos BDI (Beliefs, Desires, Intentions) a partir de políticas ótimas de POMDPs.

Observa-se, entretanto, que nas propostas anteriores, os traços de personalidade dos agentes eram fixos e conhecidos por todos, não sendo permitido o ingresso na sociedade de agentes com traços de personalidade diferentes dos pré-estabelecidos ou desconhecidos, ou que os agentes pudessem sofrer alterações em seu comportamento, modificando sua personalidade no decorrer das interações.

Assim, este trabalho está focado no desenvolvimento de um mecanismo para o controle e descoberta de traços de personalidade, ampliando o horizonte de aplicação e permitindo a operação em sistema multiagentes abertos. Para viabilizar este processo desenvolveu-se uma arquitetura de agentes híbrida, baseada na arquitetura BDI, nos POMDPs e nos Modelos de estados Ocultos de Markov (HMM - Hidden Markov Model) [MacDonald and Zucchini 1997], onde o agente seja auto-gerenciador de suas trocas sociais (baseadas em personalidades) nas quais executa com outros agentes (sejam estes de personalidade conhecida ou não por ele), negociando dependências e resolvendo conflitos de forma a alcançar e manter o equilíbrio social. Considera-se também que o conjunto de traços de personalidade não é delimitado, e que o agente possa descobrir, construir, atualizar modelos de traços de personalidade desconhecidos, mantendo uma biblioteca de planos BDI extensível que lhe permita lidar com novas personalidades dos agentes.

O artigo está organizado da seguinte forma. Na Seção 2 são discutidos sucintamente os conceitos principais da modelagem das trocas sociais adotadas neste trabalho. A Seção 3 apresenta a nova formalização para os POMDP e também a especificação dos HMM aplicado no processo de descoberta de novos traços de personalidade. Já a Seção 4 descreve a arquitetura interna para agentes BDI auto-reguladores de trocas sociais. 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 Jean Piaget [Piaget 1973], sendo esta abordagem para a regulação de interações em sistemas multiagentes adotada neste trabalho. Com o objetivo de capturar a natureza qualitativa dos conceitos sobre os valores de trocas de Piaget, Dimuro e Costa 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 $[-L \leq x_1 \leq x_2 \leq L]$, para $L \in \mathbb{R}$, sem perder a expressividade dos valores qualitativos na representação de incertezas.

Uma troca social entre dois agentes α e β pode ser representada por dois estágios de trocas. No primeiro estágio $I_{\alpha\beta}$, o agente α realiza um serviço de forma espontânea para o agente β e os valores de trocas envolvidos são os seguintes: $r_{I_{\alpha\beta}}$ é o valor do

investimento feito por α na realização de um serviço para β (um valor de investimento é sempre negativo); $s_{I_{\beta\alpha}}$ é o valor de satisfação de β pelo serviço recebido de α ; $t_{I_{\beta\alpha}}$ é o valor do débito adquirido por β para com α pela satisfação com o serviço recebido; e $v_{I_{\alpha\beta}}$ é o valor de crédito que α adquiriu com β pela realização do serviço.

No segundo estágio $II_{\alpha\beta}$, o agente α solicita o pagamento do serviço realizado anteriormente para β , 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 α e β , 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 α e β 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 o mais provável estado de parceiros através das informações recebidas, que representam as avaliações dos resultados virtuais destes agentes. Assim, considera-se o conjunto de estados materiais internos de um agente parceiro β como $E_\beta \subseteq \{E_\beta^-, E_\beta^0, E_\beta^+\}$.

Já o conjunto das observações dos possíveis resultados virtuais de um agente β foi alterado em relação aos trabalhos anteriores [Dimuro et al. 2005, Pereira 2008], na busca de uma aproximação maior com situações reais. Assim, o agente β avalia seu estado de balanço virtual para definir se aceita ou não uma nova proposta de interação. Dessa forma, o conjunto de observações foi reduzido $\Omega_\beta = \{Ac, Rc\}$, onde Ac representa que o agente β aceitou a proposta de interação, e Rc representa que β recusou a proposta de interação.

2.1. Agentes Baseados em Traços de personalidade

Com a especificação do novo conjunto de observações, viabilizou-se um processo de modelagem diferenciado para os possíveis traços de personalidade presentes em um SMA. Agora, cada agente aceita ou recusa as propostas de interações conforme seus interesses e seu estado interno, marcando de forma clara seus objetivos dentro da sociedade. A seguir, alguns destes traços de personalidade são ilustrados:

- Altruísta: agente com alta probabilidade de aceitar pedidos de serviços (*ask_service*) e baixa probabilidade de aceitar que outros agentes lhe façam serviço (*do_service*);

- Egoísta: agente com alta probabilidade de aceitar que outros agentes executem serviços em seu favor (*do_service*) e baixa probabilidade de aceitar que lhe peçam serviço (*ask_service*);
- Tolerante: agente com igual probabilidade de aceitar pedidos de serviços (*ask_service*), assim como, de aceitar que outros agentes lhe façam serviço (*do_service*).

Os traços de personalidade (Altruísta, Egoísta e Tolerante) formam o conjunto básico de traços de personalidade, e, são a base para a modelagem do processo estabelecido. Com o intuito de operarmos em sistemas multiagentes abertos, no qual existe a constante entrada e saída de agentes, com os mais diversos traços de personalidade, ou ainda, seja possível a estes agentes alterarem seu comportamento durante a execução do processo de interação. Não é viável modelar todos os traços de personalidade, até porque não há um conjunto finito desses traços. Assim, busca uma forma de reconhecer o comportamento dos diferentes traços de personalidade e também gerar regras operatórias para o seu controle.

3. O Processo de Auto-Regulação de Trocas Sociais

Para tornar os agentes independentes e reguladores de processos de trocas com seus parceiros, [Pereira 2008, Pereira et al. 2008] internalizou-se nos agentes regras BDI extraídas de políticas ótimas de POMDPs, de tal forma que um agente (com regras internalizadas) poderia propor as melhores trocas ao seu parceiro de interação, de forma a direcionar ambos para o estado de equilíbrio de seus processos de trocas.

Com o objetivo de ampliar o escopo de aplicação, buscou desenvolver um processo de auto-regulação que visa ter uma dinâmica interna própria para estabelecer novos modelos de traços de personalidade e conseqüentemente gerar novas regras BDI-POMDPs para seu controle interno. Para proporcionar essa dinâmica, busca-se integrar os HMMs com os POMDPs, no qual os HMMs sejam responsáveis pela descoberta de novos traços de personalidade, gerando um modelo aproximado, capaz de representar o comportamento dos novos agentes, apenas baseando-se nas observações recebidas, de aceites e de recusas às propostas de trocas. Os resultados do processo de aprendizagem dos novos traços de personalidade, gerados pelos HMMs, servem de base para modelar os POMDPs correspondentes aos novos traços personalidade encontrados.

3.1. A Descoberta de Traços de personalidade através de HMMs

Agentes baseados em traços de personalidade apresentam ao longo do processo de interação com outros agentes, padrões de comportamento frente às proposta de trocas que lhe são oferecidas. Assim, é possível utilizar HMMs na descoberta de novos traços de personalidade, baseando-se nos aceites e recusas operados pelos agentes.

A adoção de HMMs para executar o processo de descoberta de novos traços de personalidade se torna viável, visto que, no decorrer do processo de interação, os valores de trocas operados pelos agentes participantes, causam transições de estado, através do acúmulo ou perda em seu balanço de trocas material. Este processo de variação do balanço material descreve uma função de transição de estados não observável, em que os valores materiais definidos pelos agentes não são acessíveis aos demais agentes. Entretanto, o processo de tomada de decisão quanto ao aceite ou não das propostas de trocas

é um processo observável, descrito pela distribuição de probabilidade para cada traço de personalidade. Estas transições caracterizam a possibilidade da adoção de um HMM para executar o processo de descoberta de novos traços de personalidade.

Definição 1 *Um modelo de estados ocultos de Markov para as trocas entre agentes baseados em personalidade, no qual o agente α busca descobrir o traço de personalidade do agente β , é uma tupla $HMM_{\alpha\beta} = (E_\beta, \Omega_\beta, \Pi_\beta, F_\beta, G_\beta)$, onde:*

- $E_\beta = \{E_-, E_+, E_0\}$ é o conjunto dos possíveis estados materiais do agente β , sendo que E_0 representa o estado de equilíbrio, E_+ o estado favorável e E_- o estado desfavorável;
- $\Omega_\beta = \{Ac, Rc\}$ é o conjunto das observações disponíveis para β , tal que Ac implica que o agente β aceitou a proposta de troca e Rc representa que o agente recusou a proposta de troca;
- Π_β é a distribuição de probabilidade inicial indexada pelo conjunto de estados E_β ;
- $F_\beta : E_\beta \rightarrow \prod(E_\beta)$ é a função de transição de estados do agente β , que define para cada estado, uma distribuição de probabilidade sobre os estados, de acordo com o traço de personalidade do agente β ;
- $G_\beta : E_\beta \rightarrow \prod(\Omega_\beta)$ é a função de observação, que define para cada estado, uma distribuição de probabilidade sobre o conjunto de observações Ω_β , com base no traço de personalidade do agente β .

O modelo gerado tem capacidade de representar o comportamento de um traço de personalidade aproximado para o agente. Logo, é possível ajustar os parâmetros deste modelo, com base na sequência de observações geradas ao longo de um período de interação, reestimando as funções de transição de estado e de observação, a fim de encontrar um modelo capaz de gerar o mesmo padrão de comportamento que o agente que está apresentando, esta tarefa é alcançada através do algoritmo de Baum Welch [Rabiner and Juang 1989].

3.2. Os POMDPs com base em Aceites e Recusas

Neste trabalho, para viabilizar a especificação de uma arquitetura interna para os agentes auto-reguladores de trocas sociais, com a integração dos modelos HMMs e POMDPs, uma nova abordagem para a modelagem de regras BDI-POMDPs é proposta, com base unicamente nos aceites e recusas dos agentes a cada proposta de troca estabelecida entre pares de agentes. Dessa forma, o conjunto de observações Ω_β é redefinido para considerar as observações (Ac, Rc), simplificando o processo de modelagem de novos POMDP, e possibilitando a descoberta de novos traços de personalidade pelos HMMs.

A função de transição também foi modificada, para refletir o comportamento determinado pela nova definição dos traços de personalidade. Salienta-se agora a ligação forte entre a função de transição e as recusas ou aceites às propostas de trocas. Nesta nova abordagem, as funções de transição são definidas considerando estes índices. Assim, o agente só transita de um estado para outro, se aceitar a proposta de troca que lhe for oferecida, e, caso recusar a interação permanece no seu estado atual.

Os demais parâmetros para especificação dos POMDPs, não foram alterados, mantendo as especificações formalizadas no trabalho de [Pereira 2008], assim como, a

decomposição do cálculo do POMDP em 3 Sub-POMDPs, um para cada estado interno (balanço de trocas desfavorável, equilibrado ou favorável), para reduzir a complexidade na busca de políticas ótimas.

Assim, define-se um POMDP baseado em aceites e recusas para um par de agentes baseados em traços de personalidade:

Definição 2 *Considere um agente α com o processo supervisor internalizado e $E_\alpha^* \in \{E_-, E_+, E_0\}$ o estado corrente de α . Seja β um agente baseado em personalidade que interage com α . O Processo de Decisão Parcialmente Observável Baseado em Aceites e Recusas para o agente α , que realiza trocas com um agente baseado em personalidade β , é definido como uma tupla $POMDP_AR_\alpha^* = (E_\beta; A_\alpha; T_\beta; R_\alpha; \Omega_\beta; O_\beta)$, onde:*

- $E_\beta = \{E_-, E_+, E_0\}$ é o conjunto de estados possíveis para o agente β ;
- $A_\alpha = \{do_service; ask_service\}$ é o conjunto de ações disponíveis para o agente α , de modo que *do_service* significa que α deve oferecer um serviço a β (estágio de troca tipo $I_{\alpha\beta}$), e *ask_service* significa que α deve requisitar um serviço a β (estágio de troca tipo $II_{\alpha\beta}$);
- $T_\beta : E_\beta \times A_\alpha \rightarrow \Pi(E_\beta)$ é a função de transição baseada no traço de personalidade do agente β , ou seja, dada o estado atual do agente β e a ação realizada pelo agente α , a função transição indica qual a probabilidade do agente β mudar de estado, baseado em sua personalidade;
- $\Omega_\beta = \{Ac, Rc\}$ é o conjunto de observações que podem ser realizadas por α no processo de negociação de trocas com β ;
- $O_\beta : E_\beta \times A_\alpha \rightarrow \Pi(\Omega_\beta)$ é a função de observação baseada no traço de personalidade do agente β , ou seja, dado o estado atual do agente β e a ação realizada pelo agente α , a função indica qual a probabilidade do agente β recusar ou aceitar a proposta de α , baseado em sua personalidade.
- $R_\alpha : E_\beta \times A_\alpha \rightarrow \mathbf{R}$ é a função de recompensa para as ações escolhidas pelo agente α em cada estado de β , com maiores recompensas para ações que levem ao estado de equilíbrio do sistema.

A integração das políticas parciais (π_-, π_0, π_+) resultam na política ótima $\pi_{\alpha\beta}$, para um agente α que interagem com um agente β . A política ótima $\pi_{\alpha\beta}$ é específica para o traço de personalidade de β . Utiliza-se o algoritmo *policyToBDIplans* [Pereira 2008] para extração dos planos BDI a partir das políticas ótimas dos POMDP_ARs. Tais planos direcionam o comportamento do agente durante o processo de interação.

3.3. A Integração de POMDP e HMM

Para operação em sistemas multiagentes abertos, existe a necessidade de um mecanismo dinâmico para a descoberta e a geração de novas regras de controle, para tornar os agentes capazes de operar de maneira eficiente suas trocas sociais. Assim, busca-se integração de modelos reconhecedores de padrões (HMM), com modelos geradores de comportamento observáveis em ambientes parcialmente observáveis, onde novas políticas ótimas serão geradas, somente tomando por base o comportamento observável de recusas e aceites.

Na formalização destes modelos, verifica-se a existência de algumas semelhanças relevantes. A primeira delas é que ambas as abordagens possuem a propriedade markoviana em sua definição “O estado atual depende apenas do estado anterior e não de um

conjunto de estados passados [White 2002]”. Em outras palavras, um estado consiste nas informações de que o agente precisa para tornar o futuro independente do passado. Desta forma, a dinâmica do sistema depende da informação do passado apenas através do estado corrente e da última ação escolhida.

A segunda é que ambos atuam em ambientes parcialmente observáveis. O POMDP utiliza as observações (percepções) para inferir o mais provável estado interno do parceiro de interação, tornando viável a tomada de decisão de curso de ações. Já os HMM utilizam as observações para inferir o estado atual do parceiro, e operaram uma aproximação do seu modelo atual a novos modelos, no qual a probabilidade de gerar o comportamento observável, aumenta a cada nova interação, quando possível.

Assim, considerando o POMDP_AR o gestor do processo de controle de propostas de trocas dos agentes α e β , em que o agente α implementa as regras definidas pela política ótima do $POMDP_AR_{\alpha\beta}$, próprias para o traço de personalidade de β (Seção 3.2), e o HMM_{β} aplicado ao modelo baseado em aceites e recusas, como ferramenta capaz descobrir o comportamento de novos traços de personalidade (Seção 3.1). Esta união, proporciona o desenvolvimento de um agente auto-regulador de trocas sociais, que dispõe de meios para interagir com os mais diferentes traços de personalidade presentes em um SMA aberto.

4. Arquitetura interna para um Agente com Processo de Auto-Regulação de Trocas Sociais

No decorrer do processo de interação entre agentes baseados em traços de personalidade, existe a necessidade de conhecer o parceiro de interação, quando o estado interno do agente (balanço material) começa a se distanciar da zona de equilíbrio estabelecida. Os agentes então buscam reconhecer o parceiro de interação, a fim de identificar um conjunto de regras operatórias que os direcionem para o estado de equilíbrio interno. Essas regras podem estar em uma base de regras para os traços de personalidade conhecidos, ou podem ser inferidas através da adoção do processo HMM-POMDP para a geração de novas regras BDI.

Para auxiliar no processo de implementação de agentes auto-reguladores de trocas sociais baseadas em personalidade, uma arquitetura interna é proposta, baseada nos princípios e conceitos da arquitetura BDI, tornando o processo de interação de agentes baseados em personalidades operacional e dinâmico.

O ciclo de raciocínio envolvendo *Objetivos*, *Intenções* e *Planos*, próprios da arquitetura BDI, direcionam a tomada de decisão das ações individuais realizadas pelo agente de forma isolada. Quando o agente precisa formar parcerias para realização de uma ação que está além de suas capacidades individuais, este ciclo se estende para operar planos conjuntos, envolvendo agentes parceiros e a realização de trocas sociais, que podem ou não ocorrer, conforme os aceites e recusas operados.

Conforme [Rodrigues 2003], o processo de raciocínio quanto à escolha de agentes parceiros de interação, com base no sistema de valores de trocas sociais, envolve um conjunto de pressupostos, que vão desde o conhecimento de agentes com quem têm créditos até o conhecimento de agentes com que têm dependências. Na arquitetura proposta (Figura 1), um agente auto-regulador, delibera sobre o conjunto de agentes parceiros

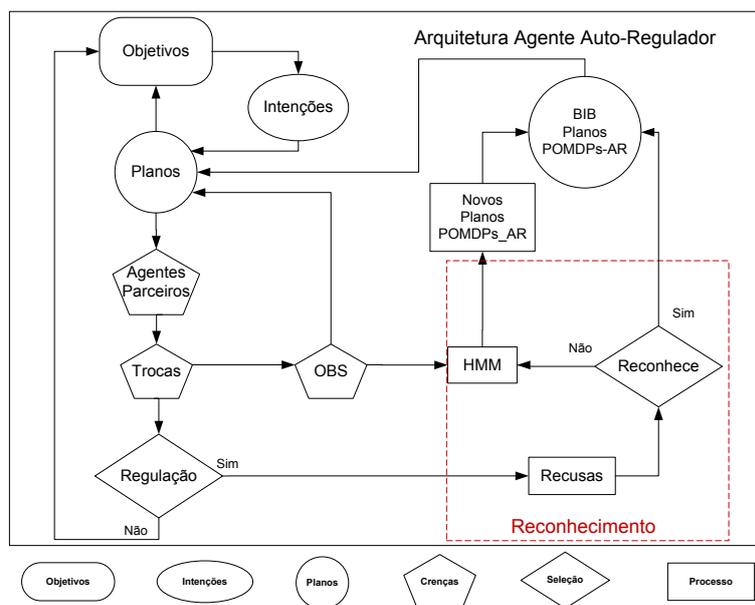


Figura 1. Arquitetura do agente Auto-regulador de trocas sociais

e escolhe um ou mais, segundo um critério de seleção, para estabelecer propostas de trocas. Assim, os *Agentes Parceiro* e as *Trocas* sociais são vistos dentro da arquitetura como crenças do agente.

O agente auto-regulador necessita operar um conjunto de propostas de trocas sociais, para identificar o traço de personalidade do parceiro de interação, mas também pode fazer uso de um histórico de propostas de trocas, caso já tenha um conjunto de trocas estabelecido. Durante o processo de identificação do traço de personalidade o agente auto-regulador releva a segundo plano o seu estado interno (balanço material), operando sem um processo de controle. A função de seleção *Regulação* é responsável por aplicar as diferentes políticas de regulação na arquitetura proposta.

As políticas de regulação descrevem os momentos que o agente auto-regulador opera seguindo seus interesses individuais, definidos pelo seu traço de personalidade, com momentos em que este faz uso de regras de controles BDI-POMDPs, dando início ao processo de auto-regulação de trocas sociais. Este processo compreende a descoberta e geração de planos BDI, ou a inclusão de planos pré-compilados para os traços de personalidade conhecidos.

O processo de auto-regulação é composto pelos seguintes módulos:

- **Reconhecimento** - opera as duas propostas para identificar os traços de personalidade de agentes parceiros:
 - **Recusa** - avalia um conjunto de propostas de trocas executadas pelo agente auto-regulador, definindo os índices de recusas;
 - **Reconhece** - função de seleção de traços já descobertos, selecionando as regras BDI, extraídas de POMDP_AR, adequadas para traço de personalidade em análise;
 - **HMM** - busca descobrir novos traços de personalidade, aproximando modelos conhecidos a novos modelos de traço de personalidade, dando ori-

gem às matrizes de transição e observação do novo traço de personalidade;

- **Novos Planos POMDPs_AR** - é o módulo fundamental na arquitetura do agente auto-regulador de trocas sociais em sistemas multiagentes abertos, em que, de posse das matrizes de transição e observação geradas pelo HMM, proporciona a modelagem de um novo POMDP_AR, dando origem a conjunto de políticas e consequentemente a um novo conjunto de regras BDI-POMDP próprias para o controle do novo agente;
- **Bib de Planos POMDPs_AR** - armazena todos os conjuntos de planos BDI-POMDPs conhecidos, e os definidos pelo módulo *Novos Planos POMDPs_AR*. Ao serem selecionados pela função *Reconhece*, estes são adicionados ao ciclo de raciocínio do agente, para que novos processos de trocas sejam operados na busca do estado de equilíbrio.

O conjunto das observações (recusas e aceites) realizadas durante as propostas de trocas (*OBS*), têm duas funções de grande importância na arquitetura do agente auto-regulador: de um lado, as observações são a base para o processo de descoberta de novos traços de personalidade, sendo o padrão de comportamento a ser aprendido pelo HMM, tornando processo de reestimação das funções de observação e transição operacional; de outro lado, direcionam o processo de seleção das regras BDI-POMDPs, uma vez que o agente adota planos tomando por base as observações que recebe. Assim o ciclo de raciocínio do agente auto-regulador se funde com o ciclo de raciocínio BDI, internalizando planos de ações individuais com planos de ações coletivas.

5. Conclusões e Trabalhos Futuros

Este trabalho segue a modelagem de agentes baseados em personalidade híbridos BDI-POMDP contidos em [Pereira 2008, Pereira et al. 2008] (com a alteração nos modelos dos traços de personalidade de trocas de acordo com os índices de aceites e recusas), ampliando o horizonte de aplicação ao adicionar um módulo de descoberta de traços de personalidade, baseado nos índices de recusas e nos HMMs, proporcionando aos agentes terem um processo interno efetivo de auto-regulação de trocas sociais, internalizando o controle do processo de interação. A especificação de agentes auto-reguladores de trocas sociais viabiliza a modelagem de SMA abertos, direcionados a simulações sociais, onde o fato do agente não conhecer parceiros de interação é um requisito essencial de modelagem, da mesma forma que modelos para o controle e de tomada de decisão se fazem necessários no sistema.

A integração dos HMM, com os POMDP, no processo de regulação de interações dos agente com traço de personalidade proporcionou a modelagem de uma arquitetura interna para agentes auto-reguladores de processo de trocas sociais, mantendo os princípios da arquitetura BDI, para construção de agentes racionais, e ampliando o processo de deliberação dos agentes, com a adição de planos BDI pré-compilados em seu ciclo de raciocínio. Onde os agentes seguem as regras estabelecidas pelas políticas ótimas oriundas dos cálculos de POMDPs_AR, específicos para cada traço de personalidade assumido por parceiros de interação.

A arquitetura proposta aponta um caminho para o desenvolvimento de agentes BDI auto-reguladores de trocas sociais, no qual introduz um novo ciclo de raciocínio ao

agente BDI, proporcionando a operacionalidade de trocas sociais em sistemas multiagentes abertos. A implementação desta proposta está em fase de desenvolvimento, alguns experimentos já foram realizados e apresentaram bons resultados para o processo de controle de interações sociais entre pares de agentes baseados em personalidade.

6. Trabalhos Futuros

A operação do processo de supervisão baseado em aceites e recusas aproxima mais o modelo de controle de interação, a modelos executado em situações reais, no qual indivíduos se mostram mais ou menos disposto auxiliar parceiros conforme seus traços de personalidade. Dessa forma, a especificação e implementação de simulações de situações reais (simulações sociais) e uma linha trabalho que deve ser explorada, tomando por base o modelo de auto-regulação desenvolvido.

Referências

- Dimuro, G. P. and Costa, A. C. R. (2006). Exchange values and self-regulation of exchanges in multi-agent systems: The provisory, centralized model. In Brueckner, S., Serugendo, G. D. M., Hales, D., and Zambonelli, F., editors, *Engineering Self-Organising Systems: Revised Selected Papers of the 3rd Intl. Work., ESOA 2005, Utrecht, 2005*, volume 3910 of *LNCS*, pages 75–89. Springer.
- 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. *JBCS*, 11(1):31–50.
- 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.
- MacDonald, I. L. and Zucchini, W. (1997). *Hidden Markov and Other Models for Discrete-valued Series*. Chapman & Hall/CRC, Boca Raton.
- 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, Brasil.
- Pereira, D. R., Gonçalves, L. V., Dimuro, G. P., and Costa, A. C. R. (2008). Towards the self-regulation of personality-based social exchange processes in multiagent systems. In Zaverucha, G. and da Costa, A. C. P. L., editors, *Advances in Artificial Intelligence, Proc. of the 19th Braz. Symp. on Artificial Intelligence, Salvador, 2008*, volume 5249 of *LNAI*, pages 113–123. Springer, Berlin.
- Piaget, J. (1973). *Estudos Sociológicos*. Forense, Rio de Janeiro.
- Rabiner, L. R. and Juang, B. H. (1989). An introduction to Hidden Markov Models. *IEEE ASSP Magazine*, 3(1):4–16.
- Rodrigues, M. (2003). Um sistema de valores troca para suporte às interações em sociedades artificiais. Dissertação de mestrado, PGCC/UFRGS, Porto Alegre, Brasil.
- White, D. J. (2002). *Markov Decision Processes*. Wiley, New York.