

TrustE - An Emotional Trust Model for Agents

Guilherme K. Bitencourt
Departamento de Informática
e Estatística
Universidade Federal
de Santa Catarina
Email: bitencourt@inf.ufsc.br

Ricardo A. Silveira
Departamento de Informática
e Estatística
Universidade Federal
de Santa Catarina
Email: silveira@inf.ufsc.br

Jerusa Marchi
Departamento de Informática
e Estatística
Universidade Federal
de Santa Catarina
Email: jerusa@inf.ufsc.br

Abstract—Trust and Reputation has been proposed in the Multiagent System area as a way for assisting agents to select good partners in order to increase the well succeeded interactions between agents. As well as trust, agent emotions has been studied with the intention to turn actions and reactions of the agents more human like. In this paper, we present a trust emotional based model. This proposal is an hybrid model that congregates a mathematical and a symbolical models for capturing the complexity of the reasoning. Quantitative and qualitative evaluations are mixed through the incorporation of some emotional flavors in the trust evaluation.

I. INTRODUÇÃO

O estudo e modelagem da confiança tem atraído o interesse de pesquisadores em diversas áreas, tais como psicologia, sociologia, filosofia e economia, e possui grande importância nas relações sociais e comerciais [1]. Em Ciências da Computação esse interesse ocorre principalmente nos Sistemas Multiagente (SMA).

Sistemas Multiagente Abertos podem receber novos agentes a qualquer momento. Esta característica torna difícil, para um agente, verificar se outro recém ingressado no sistema é confiável, pois não existem informações suficientes referentes ao novo agente. Para amenizar essa dificuldade, vários modelos de confiança e reputação vem sendo desenvolvidos na área [2]–[4]. Considerando que os sistemas computacionais estão se tornando sistemas de larga escala, abertos, dinâmicos e distribuídos, contendo uma grande quantidade de agentes que agem por interesse próprio [3], a utilização da confiança e da reputação nesses sistemas torna-se fundamental para que haja uma efetiva interação entre os agentes [5].

Contudo, a maioria dos modelos de confiança e reputação existentes possui um enfoque essencialmente matemático [6], não levando em consideração a complexidade presente na maneira como nós, humanos, confiamos ou não uns nos outros. Para nós, as emoções influenciam diretamente o ato de confiar. Isto, de certa forma, impõe um caráter subjetivo à avaliação da confiança. Além disso, os modelos propostos na literatura, determinam, através de pesos arbitrários e/ou dependentes de funções, a relevância das informações, oriundas de outros agentes, utilizadas no cálculo de confiança pelo agente avaliador. Esse fato torna parte da ação de confiar independente do comportamento do agente e portanto dissociada do contexto no qual o agente está inserido.

Neste trabalho, é proposto o modelo **TrustE** de confiança que utiliza informações de natureza simbólica, que estão relacionadas diretamente com o contexto no qual o agente está inserido. Tais informações compõem sensações ou emoções que possibilitam ao agente associar às avaliações quantitativas ou racionais, avaliações qualitativas ou subjetivas, que são fruto da introspecção das situações vividas pelo agente.

O artigo está organizado da seguinte forma: na seção II, são apresentados os conceitos de confiança e reputação entre agentes. Na seção III será introduzido o conceito de emoções em agentes, apresentando o modelo OCC. O modelo proposto é apresentado na seção IV. Por fim, a seção V apresenta as considerações finais.

II. MODELOS DE CONFIANÇA

Os modelos de confiança existentes podem ser divididos em dois grupos: modelos baseados em confiança e modelos baseados em reputação. Embora os termos confiança e reputação possam se confundir, a principal diferença entre eles está na origem da informação. A confiança está relacionada a uma relação direta entre dois agentes a e b , e é gerada através de experiências e interações ocorridas entre eles. Portanto a ação de um agente a confiar ou não em um agente b depende somente da análise dessas interações pelo agente a (a origem da informação é o próprio agente a), não sendo considerada a reputação de b . A reputação, por sua vez, é uma confiança socializada (a origem da informação são os outros agentes), que são transmitidas entre os agentes, possibilitando que a confie em b , sem a necessidade de ambos terem tido alguma interação direta no passado. Assim, a reputação de um agente se constrói a partir de informações provenientes de vários agentes, e através dela, o agente é capaz de decidir se confia ou não em um outro agente [6].

São vários os modelos de confiança e reputação encontrados na literatura, dentre os quais destacam-se: o modelo de Marsh [7], que considera apenas a confiança de cada agente; o modelo SPORA [8], que considera somente a reputação dos agentes; os modelos REGRET [2], Referral Network [9] e TRAVOS [3] que levam em consideração tanto a confiança quanto a reputação, combinando esses valores para chegar a um resultado; e finalmente o modelo FIRE [5], que introduz dois conceitos adicionais, a confiança baseada em papéis e

a reputação certificada. A seguir, apresenta-se uma descrição sucinta destes modelos.

A. Modelo de Marsh

O modelo proposto por [7] foi um dos primeiros modelos desenvolvidos sobre confiança local (considera apenas a interação direta entre agentes para medir a confiança). Sua arquitetura é distribuída, uma vez que, cada agente é o responsável pelo cálculo da sua confiança perante os outros agentes.

Ele diferencia três tipos de confiança: *confiança básica* - representa a disposição de um agente confiar em outro; *confiança geral* - a confiança que um agente exerce sobre outro sem levar em conta qualquer situação específica; e *confiança situacional* - a confiança que um agente tem em relação a outro, levando em consideração uma situação específica.

Esse modelo, por levar em consideração apenas a confiança local (direta) entre agentes, é limitado em relação a capacidade de calcular a confiança quando os agentes nunca interagiram entre si. No entanto, sua citação é importante devido ao seu pioneirismo na área e a sua definição de confiança direta entre agentes.

B. TRAVOS

O modelo de confiança e reputação TRAVOS [3] foi desenvolvido para ser utilizado, principalmente, em SMA abertos e assume que o comportamento dos agentes não mudam com o tempo, no entanto, essa hipótese nem sempre é verdadeira. A sua principal característica - que também está presente nos modelos REGRET [2] e FIRE [5] - é a possibilidade de um agente avaliar a confiança em outro agente de forma direta, através de experiências passadas, ou através da reputação, quando tais experiências não existirem ou forem insuficientes. Ao estabelecer a confiança em outros agentes e escolher aquele que é mais confiável, o agente tem a capacidade de maximizar a probabilidade de que sua interação seja bem sucedida. Outra característica presente é a possibilidade de filtrar as opiniões imprecisas de outros agentes, permitindo que a reputação seja utilizada para aumentar significativamente o desempenho do sistema, já que as informações indesejáveis serão descartadas.

Neste modelo, um agente a_{tr} (*truster*) possui dois métodos para calcular a confiança em outro agente a_{te} (*trustee*) em um contexto específico. Primeiramente o agente a_{tr} faz a avaliação baseado nas interações diretas com o agente a_{te} , depois o agente a_{tr} avalia a confiabilidade de a_{te} através da reputação de a_{te} .

A reputação é necessária quando um agente a_{tr} quer fazer uma avaliação sobre o comportamento de um agente a_{te} , e possui poucas informações sobre ele ou a confiança entre ambos possui uma baixa confiabilidade. Assim, a reputação pode aumentar a precisão da confiabilidade do valor de confiança que a_{tr} tem em a_{te} , auxiliando a_{tr} na sua tomada de decisão.

A confiança é modelada através de uma abordagem probabilística, baseada nas experiências passadas de um agente

sendo avaliado. Se um agente avaliador (a_{tr}), tem todas as informações sobre o agente sendo avaliado (a_{te}), então, de acordo com a_{tr} , a probabilidade de a_{te} cumprir com suas obrigações é expressa por $B_{a_{tr},a_{te}}$. Contudo, normalmente não se tem toda a informação necessária sobre a_{te} , logo, o melhor caminho a seguir é utilizar o *valor esperado* de $B_{a_{tr},a_{te}}$ dado o conhecimento (conjunto de todos os resultados das interações observadas) de a_{tr} .

C. FIRE

O modelo FIRE, proposto por [5], é um modelo de confiança e reputação integrado, com uma arquitetura de tomada de decisão distribuída entre os agentes, semelhante ao modelo REGRET [2]. Este modelo incorpora quatro fontes de informação: confiança por interação, reputação baseada em testemunho, confiança baseada em papéis e reputação certificada, as quais são combinadas para fornecer uma métrica de confiança em praticamente todas as circunstâncias. Essa variedade de fontes torna-se importante, visto que, em várias situações nem todas estarão prontamente disponíveis, além de permitir aos agentes combiná-las para lidar com as incertezas do ambiente.

D. REGRET

Neste modelo, a reputação é vista como uma opinião ou visão de um agente sobre algo, sendo formada e atualizada ao longo do tempo através das interações com os outros agentes do sistema. As interações fornecem como resultado *impressões* que são registradas pelos agentes e refletem como eles avaliam suas experiências com outros agentes, de acordo com o resultado de um diálogo (contrato inicial que estabelece os termos e condições de uma transação) firmado entre os agentes. Como cada agente possui uma opinião diferente dos demais, pode-se dizer que a reputação assume um caráter mais subjetivo.

Um resultado, o_b , de um diálogo relacionado a uma transação comercial, entre dois agentes a e b , do ponto de vista do agente comprador b , poderia ser:

$$o_b = (DataEntrega =_c 10/02 \wedge Preco =_c 2000 \wedge Qualidade =_c A \wedge DataEntrega = 15/02 \wedge Preco = 2000 \wedge Qualidade = C)$$

Nesse exemplo, as variáveis com o subscrito c representam o acordo inicial entre ambos os agentes. Assim o agente b esperava um produto com qualidade A (boa) porém recebeu um produto de qualidade C (ruim), além de ter recebido o produto com 5 dias de atraso.

O modelo utiliza o termo Reputação Individual (RI) para representar a confiança direta entre dois agentes, e Reputação Social (RS) para representar a reputação propriamente dita. A abordagem descentralizada implementada por este modelo permite a cada agente calcular a RI e a RS de outro agente, podendo utilizar ambas ou apenas uma delas para se chegar a um resultado final.

A RS leva em consideração três fontes de informação: a interação do agente a (avaliador) com os membros do grupo que o agente b (avaliado) pertence, expressa por $R_{a \rightarrow B}(\varphi)$, equação 1. O que os membros do grupo A (o grupo do agente a) pensam sobre o agente b , expressa por $R_{A \rightarrow b}(\varphi)$, equação 2. E o que os membros do grupo A pensam sobre o outro grupo B , expressa por $R_{A \rightarrow B}(\varphi)$, equação 3.

$$R_{a \rightarrow B}(\varphi) = \sum_{b_i \in B} \omega^{ab_i} \cdot R_{a \rightarrow b_i}(\varphi) \quad (1)$$

$$R_{A \rightarrow b}(\varphi) = \sum_{a_i \in A} \omega^{a_i b} \cdot R_{a_i \rightarrow b}(\varphi) \quad (2)$$

$$R_{A \rightarrow B}(\varphi) = \sum_{a_i \in A} \omega^{a_i B} \cdot R_{a_i \rightarrow B}(\varphi) \quad (3)$$

Esses 3 valores provenientes dessas fontes de informação são combinados com a RI do agente, denotada por $R_{a \rightarrow b}(\varphi)$, para se chegar ao valor final de confiança, representado por $SR_{a \rightarrow b}(\varphi)$, equação 4.

$$SR_{a \rightarrow b}(\varphi) = \xi_{ab} \cdot R_{a \rightarrow b}(\varphi) + \xi_{aB} \cdot R_{a \rightarrow B}(\varphi) + \xi_{Ab} \cdot R_{A \rightarrow b}(\varphi) + \xi_{AB} \cdot R_{A \rightarrow B}(\varphi) \quad (4)$$

onde $\xi_{ab} + \xi_{aB} + \xi_{Ab} + \xi_{AB} = 1$, e representam a importância de cada uma das fontes de informação para o agente avaliador, sendo essa decisão dependente de aplicação.

As características deste modelo permitem que à ele sejam incorporadas avaliações simbólicas, na forma de emoções. Na seção seguinte é apresentado um modelo de emoções para agentes, que conjuntamente ao modelo REGRET, forma a base para a proposta do modelo *TrustE*.

III. EMOÇÕES EM AGENTES

O estudo das emoções está presente em várias disciplinas como psicologia, economia, neurociência cognitiva, e nos últimos anos esse estudo também está presente nas pesquisas em IA e Ciência da Computação. Tal estudo visa a criação de sistemas de interação emocional, como por exemplo, robôs com comportamento emocional e agentes virtuais para entretenimento [10].

O modelo psicológico de emoções, conhecido como *OCC*, proposto por [11], tem ganhado popularidade entre pesquisadores que desenvolvem sistemas de raciocínio sobre emoções ou que incorporam emoções em agentes artificiais. O modelo classifica 22 tipos de emoções, sendo metade destas positivas (ex.: alegria e esperança) e metade negativas (ex.: tristeza e medo). Duas das emoções descritas no modelo *OCC* são:

- *ALEGRIA* - *satisfeito* em relação a um evento desejável;
- *MEDO* - *insatisfeito* em relação a um evento indesejável.

Há uma relação de temporalidade nas emoções, enquanto a *ALEGRIA* está relacionada a algo que está acontecendo, o *MEDO* refere-se a algo que poderá acontecer.

Quanto aos aspectos quantitativos das emoções, estes são descritos no modelo *OCC* em termos de potencialidades, limiares e intensidades. Para cada uma das 22 emoções, é fornecida uma lista de variáveis que afetam a intensidade da emoção e quais as condições necessárias para que a emoção ocorra.

Assim, intensidade de uma emoção é definida subtraindo-se o *limiar* de seu *potencial*. O modelo *OCC* não especifica como são calculados os *limiares* das emoções, porém acredita-se que eles dependam de variáveis globais que indicam o humor do agente [12]. Por exemplo, se um agente está de bom humor, os limiares das emoções negativas aumentam, causando uma diminuição na intensidade dessas emoções. Quando uma condição necessária para disparar uma emoção ocorre, mas o seu *potencial* está abaixo do seu *limiar*, um agente pode reconhecer que essa emoção foi desencadeada porém ela não o afetará. Por exemplo, "o humor de um agente estava tão bom que mesmo ele tendo praticado uma ação ruim, ele não foi afetado pela vergonha".

Para cada emoção, são necessárias 3 funções para o cálculo da *intensidade*. São elas: *função de potencialidade*, *função do limiar* e *função de intensidade*. A maneira como essas 3 funções são calculados é dependente de aplicação, porém, de um modo geral, a *função de intensidade* de uma emoção, denotada por $I(P(E), L(E), t)$, que por simplificação será representada por $I(E)$, pode ser declarada como:

$$I(P(E), L(E), t) \rightarrow \mathbb{R}^+ \quad (5)$$

onde o parâmetro $P(E)$ representa a *função de potencialidade* da emoção E , $L(E)$ a *função do limiar* de E , e t o tempo corrente. Como resultado, a função retorna um real positivo, incluindo 0 (zero), que representa, de forma quantitativa, a intensidade da emoção. Segundo [12], o valor de $I(E)$ persiste ao longo do tempo, e tende a diminuir com o passar dele, diferentemente dos valores de $P(E)$ e $L(E)$ que são recalculados a cada vez que E é disparada novamente, portanto esses valores não persistem com o passar do tempo. Para suportar a temporalidade da intensidade cada agente deve possuir uma memória (MEA) que armazene os valores de cada uma das emoções sentidas por ele.

A. Nova hierarquia de emoções

Contudo, para que as emoções possam ser utilizadas em agentes inteligentes é necessário formalizá-las, possibilitando a sua implementação. Steunebrink et al. [13] propuseram uma revisão do modelo *OCC*, visando adequar o modelo a implementações computacionais. Neste trabalho são identificadas as ambiguidades existentes na estrutura lógica do *OCC* e são propostas alterações, conduzindo o modelo a uma estrutura baseada em herança (Figura 1), suportada por uma nova estrutura lógica, e novas especificações dos tipos de emoções. A partir dessas alterações, as 22 emoções do modelo *OCC* foram formalizadas, criando-se um modelo qualitativo de emoções, descrevendo precisamente quando as emoções são desencadeadas [14].

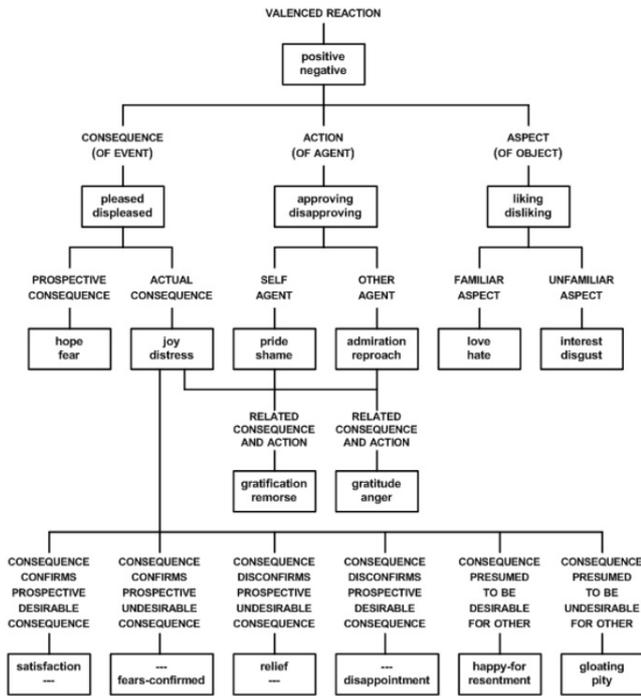


Figure 1. Hierarquia baseada em herança de emoções do modelo OCC modificado [13]

As principais alterações na estrutura do modelo foram: a introdução de herança explícita na composição hierárquica das emoções; o uso de rótulos em cada ponto da hierarquia e definição de nós filhos como superconjunto dos nós pais. A partir dessa nova estrutura foram criadas novas especificações dos tipos de emoções, conforme a figura 2. Tais modificações possibilitam uma melhor avaliação das emoções pois torna o modelo mais formal e com uma maior linearidade na composição das emoções.

Como pode ser visto nas figuras 1 e 2, as emoções SATISFEITO (“PLEASSED”), INSATISFEITO (“DISPLEASED”), APROVAÇÃO (“APPROVING”), DESAPROVAÇÃO (“DISAPPROVING”), GOSTO (“LIKING”) e DESGOSTO (“DISLIKING”) que no modelo original OCC fazem parte do processo de avaliação e não representam uma emoção propriamente dita, na alteração proposta por Steunebrink et al. [13], esses sentimentos são tratados como emoções e são a base para todas as outras emoções. A diferença entre a ALEGRIA e ESPERANÇA, por exemplo, é que a ALEGRIA está relacionada a uma consequência que está ou já aconteceu, enquanto a ESPERANÇA está relacionada a uma consequência que poderá acontecer. O mesmo raciocínio vale para TRISTEZA e MEDO. Por exemplo, “O homem estava alegre porque seu time ganhou e tinha esperança de ver seu time ser campeão ao final do campeonato”.

positive and negative are valenced reactions (to “something”)
 pleased is being positive about a consequence (of an event)
 displeased is being negative about a consequence (of an event)
 hope is being pleased about a prospective consequence (of an event)
 fear is being displeased about a prospective consequence (of an event)
 joy is being pleased about an actual consequence (of an event)
 distress is being displeased about an actual consequence (of an event)
 satisfaction is joy about the confirmation of a prospective desirable consequence
 fears-confirmed is distress about the confirmation of a prospective undesirable consequence
 relief is joy about the disconfirmation of a prospective undesirable consequence
 disappointment is distress about the disconfirmation of a prospective desirable consequence
 happy-for is joy about a consequence (of an event) presumed to be desirable for someone else
 resentment is distress about a consequence (of an event) presumed to be desirable for someone else
 gloating is joy about a consequence (of an event) presumed to be undesirable for someone else
 pity is distress about a consequence (of an event) presumed to be undesirable for someone else
 approving is being positive about an action (of an agent)
 disapproving is being negative about an action (of an agent)
 pride is approving of one’s own action
 shame is disapproving of one’s own action
 admiration is approving of someone else’s action
 reproach is disapproving of someone else’s action
 gratification is pride about an action and joy about a related consequence
 remorse is shame about an action and distress about a related consequence
 gratitude is admiration about an action and joy about a related consequence
 anger is reproach about an action and distress about a related consequence
 liking is being positive about an aspect (of an object)
 disliking is being negative about an aspect (of an object)
 love is liking a familiar aspect (of an object)
 hate is disliking a familiar aspect (of an object)
 interest is liking an unfamiliar aspect (of an object)
 disgust is disliking an unfamiliar aspect (of an object)

Figure 2. Especificações dos tipos de emoções do modelo OCC modificado [13]

IV. TRUSTE: UM MODELO DE CONFIANÇA BASEADO EM EMOÇÕES

O modelo TrustE, proposta deste trabalho, agrega emoções ao cálculo da confiança e da reputação em agentes. Todos os modelos encontrados na literatura fazem uso de análise algébrica para determinar o ato de confiar, o que torna tais modelos essencialmente matemáticos e desconectados do histórico de agente. A proposta do modelo TrustE é permear o modelo de confiança com avaliações oriundas de raciocínio simbólico, tornando o ato de confiar mais dinâmico e dependente do histórico do agente.

A incorporação de emoções pode ser feita de várias formas, contudo optou-se por incorporar ao modelo REGRET, fatores emocionais, baseados nas *intensidades das emoções* do agente. Assim, os pesos definidos pelo programador, ξ e ω das equações do modelo REGRET (Subseção II-D) são substituídos por símbolos emocionais.

A. Formalização das emoções no TrustE

Embora o modelo OCC defina 22 tipos de emoções, para implementação do modelo TrustE foram utilizadas 4 tipos de emoções, ALEGRIA, TRISTEZA, ADMIRAÇÃO e DESCONSIDERAÇÃO. Forem escolhidas essas emoções pois as duas primeiras representam um tipo de emoção que não está ligada diretamente a ação de outro agente (ex. O agente estava feliz por ter atingido seus objetivos), enquanto as duas últimas são emoções ligadas a atitude de um determinado agente em relação a outro (ex. O agente comprador admirou a honestidade do agente vendedor). A seguir serão dadas algumas definições necessárias para a formalização dessas 4 emoções.

- Sejam a e b agentes, X um evento, Y uma ação.
- Se a confia em b então $CONF_a(b)$ senão $\neg CONF_a(b)$.
- $POS_a(Y)$ - uma ação Y positiva realizada por a .

- $NEG_a(Y)$ - uma ação Y negativa realizada por a .
- $EVENT(X)$ - um evento (fato) X já realizado.
- $SAT_a(X)$ - a $SATISFACAO$ de a em relação a consequência de $EVENT(X)$.
- $\neg SAT_a(X)$ - a $INSATISFACAO$ de a em relação a consequência de $EVENT(X)$.
- $AP_a(Y, b)$ - a $APROVACAO$ de a em relação a uma ação Y de b . O agente a irá ter uma *reação positiva* em relação a Y , quando b realizar uma *ação positiva* do ponto de vista do agente a .

$$AP_a(Y, b) \leftrightarrow POS_b(Y) \wedge \neg NEG_b(Y) \quad (6)$$

- $DES_a(Y, b)$ - a $DESAPROVACAO$ de a em relação a uma ação Y de b . O agente a irá ter uma *reação negativa* em relação a Y , quando b realizar uma *ação negativa* do ponto de vista do agente a .

$$DES_a(Y, b) \leftrightarrow NEG_b(Y) \wedge \neg POS_b(Y) \quad (7)$$

Logo será admitido que $AP_a(Y, b) \leftrightarrow \neg DES_a(Y, b)$. Assim será utilizado a notação $AP_a(Y, b)$ para $APROVACAO$ e $\neg AP_a(Y, b)$ para $DESAPROVACAO$.

A formalização e a condição de ocorrência de cada uma das 4 emoções são descritas a seguir:

Sendo $ADM_a(Y, b)$ a $ADMIRACAO$ de a em relação a uma ação Y realizada por b . O agente a irá $ADMIRAR$ Y , quando a $APROVAR$ Y .

$$ADM_a(Y, b) \leftrightarrow AP_a(Y, b) \quad (8)$$

Sendo $DES_a(Y, b)$ a $DESCONSIDERACAO$ de a em relação a uma ação Y realizada por b . O agente a irá $DESCONSIDERAR$ Y , quando a $REPROVAR$ Y .

$$DES_a(Y, b) \leftrightarrow \neg AP_a(Y, b) \quad (9)$$

Sendo $AL_a(X)$ o sentimento de $ALEGRIA$ do agente a em relação a consequência atual de um evento X . O agente a irá sentir-se $ALEGRE$ quando ficar $SATISFEITO$ em relação a X .

$$AL_a(X) \leftrightarrow SAT_a(X) \wedge EVENT(X) \quad (10)$$

Sendo $TR_a(X)$ o sentimento de $TRISTEZA$ do agente a em relação a consequência atual de um evento X . O agente a irá sentir-se $TRISTE$ quando ficar $INSATISFEITO$ em relação a X .

$$TR_a(X) \leftrightarrow \neg SAT_a(X) \wedge EVENT(X) \quad (11)$$

No escopo utilizado para a explicação do modelo TrustE, as ações e eventos importantes no processo do cálculo de confiança são:

- 1) a ação Y (positiva) de um agente em cumprir os termos firmados em um diálogo;
- 2) a ação Y (negativa) de um agente em não cumprir os termos firmados em um diálogo;
- 3) a ação Y de um agente confiar em outro agente;
- 4) o fato X de um agente receber a confiança de outro agente;

- 5) o fato X de um agente ter cumprido os termos firmados em um diálogo;
- 6) o fato X de um agente não ter cumprido os termos firmados em um diálogo;

De certa forma, os eventos (5 e 6) se confundem com as ações (1 e 2), no entanto essas ações, que são realizadas pelo agente avaliado, são as responsáveis por dispararem os eventos que produzem consequências sentidas pelo agente avaliador.

Utilizando-se o modelo em um escopo mais abrangente, como por exemplo, numa simulação de bolsa de valores, existiriam outras ações e/ou eventos que não estariam relacionados diretamente com o cálculo da confiança. Por exemplo, $POS_b(Y)$ poderia representar a ação de b indicar ao agente a , uma venda de ações com baixo preço, enquanto $EVENT(X)$ poderia representar o lucro de a por ter comprado essas ações. Como consequência, a sentiria-se $ALEGRE$ e teria uma $ADMIRACAO$ por b .

B. Detalhamento do Modelo TrustE

A figura 3 mostra uma visão conceitual do TrustE que demonstra o fluxo de ações dos agentes. Para explicar esse fluxo será utilizado um cenário no qual um agente quer comprar um produto, de boa qualidade, pela Internet. Seja a um agente comprador que irá escolher um agente vendedor b entre todos os vendedores existentes (etapa 1). Para saber se a deve ou não comprar o produto de b , ele irá calcular a confiança em b (etapa 2). Caso o valor da confiança seja baixo, a irá procurar outro agente vendedor, caso ela seja alta, a irá comprar o produto de b (etapa 3). Ao receber a confiança de a o agente b irá aumentar suas emoções positivas (etapa 4). Passado algum tempo, b irá entregar o produto ao agente a (etapa 5), esse por sua vez irá avaliar a qualidade do produto (etapa 6). Se ele considerar a qualidade boa, suas emoções positivas irão aumentar (etapa 7) e sua avaliação a respeito de b será positiva (etapa 8), caso contrário suas emoções negativas irão aumentar e sua avaliação a respeito de b será negativa.

A figura 4 mostra os dois módulos existentes no TrustE, o módulo de confiança (MC) e o módulo de emoções (ME). O primeiro é o responsável pelo cálculo da Reputação Individual (RI) e Reputação Social (RS), que juntas irão resultar no valor final da confiança (VFC). O segundo contém as funções de intensidades das emoções $I(E)$ e a memória das emoções dos agentes (MEA), sendo que cada agente possui sua própria MEA.

Conforme explicado anteriormente, o modelo REGRET possui 4 valores para formar a confiança final, representados na figura por $R_{a \rightarrow b}$, $R_{a \rightarrow B}$, $R_{A \rightarrow b}$ e $R_{A \rightarrow B}$. O primeiro peso utilizado no cálculo da RI $R_{a \rightarrow b}$ é o mesmo utilizado no REGRET, que atribui valores maiores às avaliações dadas recentemente. Já os primeiros pesos utilizados nas Reputações Sociais $R_{a \rightarrow B}$, $R_{A \rightarrow b}$ e $R_{A \rightarrow B}$, são calculados a partir das IEA. A escolha por esse grupo de emoções (*emoções-de-ações*) para esses pesos, se deve ao fato desse tipo de emoção estar relacionado a algum agente em específico, assim é possível que cada agente consiga avaliar os outros levando em consideração as emoções relacionadas a cada um deles.

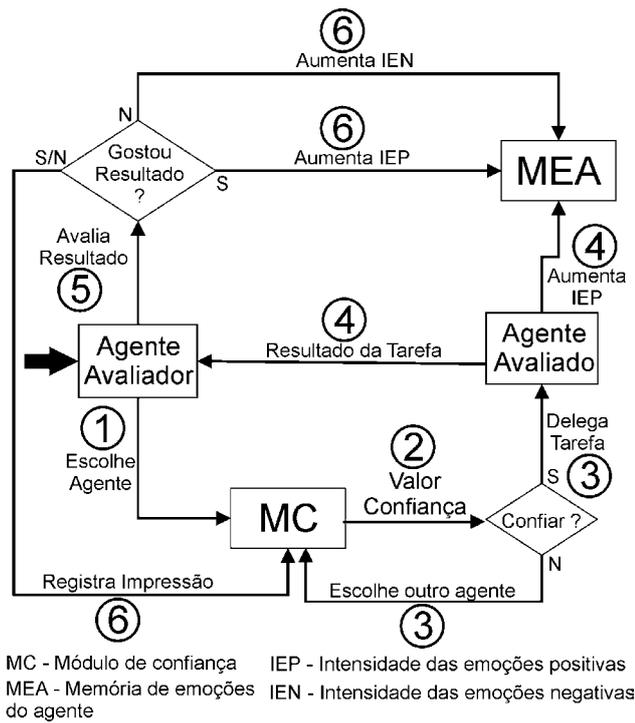


Figure 3. Visão Conceitual do Modelo TrustE

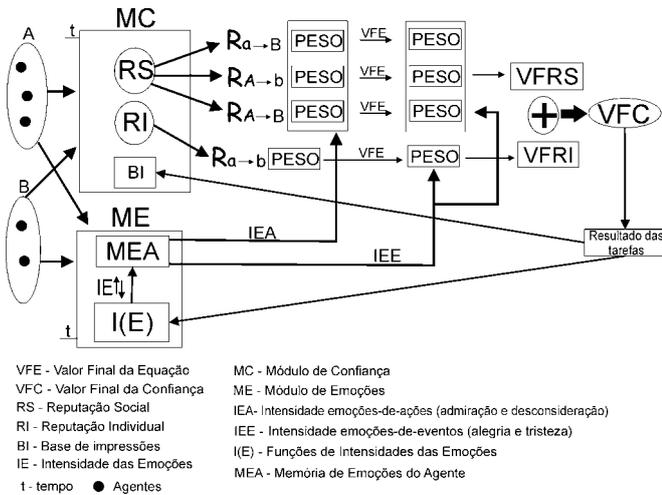


Figure 4. Modelo TrustE

Depois de calculados esses 4 valores, eles são combinados para formar o VCF. Os pesos utilizados nessa parte do cálculo levam em consideração as IEE, já que esse grupo de emoções (emoções-de-eventos) não está relacionado diretamente a nenhum agente em específico e sim a eventos ocorridos que modificam o estado emocional do agente. Devido ao fato de acreditar-se que quando estamos alegres temos uma maior tendência em confiar nos outros, quando um agente sentir alegria ele irá dar um peso maior para a RS, e quando sentir tristeza será dado um peso maior a RI.

V. CONSIDERAÇÕES FINAIS

Esse artigo apresentou o modelo híbrido TrustE, que é um modelo de confiança, baseado no REGRET, que utiliza as emoções do agente no mecanismo de cálculo da confiança, visando capturar a complexidade do raciocínio humano e flexibilizar as avaliações quantitativas, introduzindo elementos de natureza qualitativa ao modelo.

Acredita-se que a inserção de emoções e a utilização das suas intensidades possam propiciar um maior realismo ao modelo, pois a tomada de decisão do agente estará diretamente ligada ao seu estado emocional. Como próxima etapa, o modelo deverá ser implementado e validado em cenários de negociação entre agentes. O modelo originalmente proposto pode ainda ser expandido para comportar um maior número de emoções.

REFERENCES

- [1] E. Ostrom, "A behavioral approach to the rational choice theory of collective action: Presidential address, american political science association, 1997," *American Political Science Review*, pp. 1-22, 1998.
- [2] J. Sabater and C. Sierra, "Regret: reputation in gregarious societies," in *Proceedings of the fifth international conference on Autonomous agents*. ACM, 2001, pp. 194-195.
- [3] W. Teacy, J. Patel, N. Jennings, and M. Luck, "Travos: Trust and reputation in the context of inaccurate information sources," *Autonomous Agents and Multi-Agent Systems*, vol. 12, no. 2, pp. 183-198, 2006.
- [4] J. Sabater and C. Sierra, "Review on computational trust and reputation models," *Artificial Intelligence Review*, vol. 24, no. 1, pp. 33-60, 2005.
- [5] T. Dong-Huynha, N. Jennings, and N. Shadbolt, "Fire: An integrated trust and reputation model for open multi-agent systems," in *ECAI 2004: 16th European Conference on Artificial Intelligence, August 22-27, 2004, Valencia, Spain*, vol. 110. Ios Pr Inc, 2004, p. 18.
- [6] G. Lu, J. Lu, S. Yao, and Y. Yip, "A review on computational trust models for multi-agent systems," *The Open Information Science Journal*, vol. 2, pp. 18-25, 2009.
- [7] S. Marsh, "Formalising trust as a computational concept," Ph.D. dissertation, Department of Mathematics and Computer Science, University of Stirling, 1994.
- [8] G. Zacharia, A. Moukas, and P. Maes, "Collaborative reputation mechanisms for electronic marketplaces," *Decision Support Systems*, vol. 29, no. 4, pp. 371-388, 2000.
- [9] B. Yu and M. Singh, "Searching social networks," in *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*. ACM, 2003, pp. 65-72.
- [10] E. Lorini, "Agents with emotions: a logical perspective," *ALP Newsletter*, vol. 12, no. 2-3, 2008.
- [11] A. Ortony, G. Clore, and A. Collins, *The cognitive structure of emotions*. Cambridge university press, 1990.
- [12] B. R. Steunebrink, M. Dastani, and J.-J. C. Meyer, "Towards a quantitative model of emotions for intelligent agents," in *Proceedings of the 2nd Workshop on Emotion and Computing-Current Research and Future Impact, Osnabrück, Germany, 2007*.
- [13] B. Steunebrink, M. Dastani, and J. Meyer, "The occ model revisited," in *Proceedings of the 4th Workshop on Emotion and Computing*, vol. 65, 2009, pp. 2047-2056.
- [14] B. R. Steunebrink, M. Dastani, and J. Meyer, "A logic of emotions for intelligent agents," in *Proceedings of the National Conference on Artificial Intelligence*, vol. 22, no. 1. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2007, p. 142.