

Juliana Carpes Imperial

**Confiança em Agentes
Inteligentes**

TESE DE DOUTORADO

DEPARTAMENTO DE INFORMÁTICA
Programa de Pós-Graduação em Informática

Rio de Janeiro
Agosto de 2007

Juliana Carpes Imperial

Confiança em Agentes Inteligentes

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Informática do Departamento de Informática da PUC-Rio como requisito parcial para obtenção do título de Doutor em Informática

Orientador: Prof. Edward Hermann Haeusler

Rio de Janeiro
agosto de 2007

Juliana Carpes Imperial

Confiança em Agentes Inteligentes

Tese apresentada ao Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC-Rio como requisito parcial para obtenção do título de Doutor em Informática. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Edward Hermann Haeusler

Orientador

Departamento de Informática — PUC-Rio

Prof. Christiano de Oliveira Braga

Facultad de Informática – UCM – Madri

Prof. Daniel Schwabe

Departamento de Informática – PUC-Rio

Prof. Marco Antonio Casanova

Departamento de Informática – PUC-Rio

Prof. Mario Roberto F. Benevides

Programa de Sistemas – COPPE/UFRJ

Prof. José Eugenio Leal

Coordenador Setorial do Centro Técnico Científico — PUC-Rio

Rio de Janeiro, 10 de agosto de 2007

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Juliana Carpes Imperial

Graduou-se em Engenharia de Computação em julho de 2001 e concluiu o Mestrado em Informática em agosto de 2003 pela PUC-Rio (Pontifícia Universidade Católica do Rio de Janeiro). Sua área de pesquisa é a de Teoria da Computação, tendo participado de vários congressos e seminários, e alguns cursos em sua área de atuação no Brasil e no exterior. Também trabalhou como monitora na graduação e na pós-graduação em cadeiras oferecidas pelo Departamento de Informática. Por fim, participou do projeto de revisão técnica e pós-edição do conteúdo do “Visual Studio 2005 Microsoft MSDN Library”.

Ficha Catalográfica

Imperial, Juliana Carpes

Confiança em Agentes Inteligentes / Juliana Carpes Imperial; orientador: Edward Hermann Haeusler. — Rio de Janeiro : PUC–Rio, Departamento de Informática, 2007.

v., 136 f: il. ; 29,7 cm

1. Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Tese. 2. Agentes Inteligentes. 3. Sistemas Multi-Agentes Abertos. 4. BDI. 5. Lógica Modal. 6. Lógica Epistêmica. 7. Confiança. 8. Inteligência Artificial. I. Haeusler, Edward Hermann. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Agradecimentos

A Deus por ter me permitido concluir esse trabalho e ter me dado a chance de possuir todo o conhecimento que tenho hoje.

Aos meus pais, Célia Carpes Imperial e Ronaldo D'Agostino Imperial por durante muitos anos terem investido em mim para que eu pudesse estudar nas melhores instituições de ensino e, dessa forma, adquirisse as ferramentas necessárias para facilitar meu aprendizado.

À minha irmã, Roberta Carpes Imperial, por abrir mão de usar o computador em muitas oportunidades para que este trabalho fosse concluído.

Ao Hermann, meu orientador, por sempre ter sido atencioso comigo tanto na elaboração deste trabalho quanto durante as aulas nas quais ele foi meu professor, sempre disposto a tirar minhas dúvidas e por ter me ensinado boa parte do que sei em minha área de pesquisa.

A todos os professores do departamento de Informática, por tudo o que me ensinaram e por toda atenção que sempre me dedicaram.

A todos os funcionários do departamento de Informática, por sempre estarem dispostos a ajudar e pela simpatia.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Aos professores que participaram da Comissão Examinadora.

Aos meus colegas e amigos do TECMF que sempre me apoiaram e ajudaram quando eu precisava.

Aos meus gatos Tigresa, Billy e Nicky que sempre me fizeram companhia durante os meus estudos e na elaboração desta tese, especialmente ao último, que muitas vezes ficava me pedindo carinho durante tais tarefas e roubava a minha cadeira ao me levantar.

Resumo

Imperial, Juliana Carpes; Haeusler, Edward Hermann. **Confiança em Agentes Inteligentes**. Rio de Janeiro, 2007. 136p. Tese de Doutorado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Confiança é um aspecto fundamental em sistemas distribuídos abertos de larga-escala. Ela está no núcleo de todas as interações entre as entidades que precisam operar em ambientes com muita incerteza e que se modificam constantemente. Dada a complexidade das interações, esses componentes, e o sistema resultante, são cada vez mais contextualizados, projetados e construídos usando técnicas baseadas em agentes. Portanto, confiança é fundamental em um sistema multi-agentes (MAS) aberto. Logo, este trabalho investiga como se ter um modelo de confiança explicitamente em um agente inteligente, que possui crenças (*Beliefs*), desejos (*Desires*) e intenções (*Intentions*), chamado de agente BDI. Ou seja, o agente passa a ter um quarto componente chamado confiança (*Trust*). Dessa forma, é interessante haver uma lógica para englobar o conceito de confiança em um MAS BDI aberto, para que os diferentes aspectos de um modelo de confiança sejam expressos de forma precisa e formal. Isso é feito usando uma lógica multi-modal indexada, onde os mundos possíveis que modelam um sistema multi-agentes representam quais agentes estão presentes em um dado instante de tempo. E, para cada uma das três componentes originais de um agente BDI, há também uma representação de mundos possíveis, pois as mesmas são tratadas como modalidades. Já a confiança é modelada como um predicado, e não uma modalidade.

Palavras-chave

Agentes Inteligentes. Sistemas Multi-Agentes Abertos. BDI. Lógica Modal. Lógica Epistêmica. Confiança. Inteligência Artificial.

Abstract

Imperial, Juliana Carpes; Haeusler, Edward Hermann. **Trust in Intelligent Agents**. Rio de Janeiro, 2007. 136p. PhD Thesis — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Trust is a fundamental concern in large-scale open distributed systems. It lies at the core of all interactions between the entities that have to operate in such uncertain and constantly changing environments. Given the complexity of the interactions, these components, and the ensuing system, are increasingly being conceptualised, designed, and built using agent-based techniques. Therefore, the presence of trust is imperative in a multi-agent system (MAS). Consequently, this work studies how to have an explicit trust model in an intelligent agent, which has beliefs, desires and intentions (BDI agent). That is, the agent now has a fourth component called Trust. This way, a logic to include the concept of trust in an open BDI MAS is interesting, so that the different aspects of a trust model can be expressed formally and accurately. This is achieved by using an indexed multi-modal logic, where the possible worlds which model a multi-agent system represent which agents are in the system in a given moment. Moreover, for each one of the three original components of a BDI agent, where the components represent beliefs, desires and intentions, there is a representation of possible worlds, because these are treated as modalities. However, trust is modelled as a predicate, not as a modality.

Keywords

Intelligent Agents. Open Multi-Agent Systems. BDI. Modal Logic. Epistemic Logic. Trust. Artificial Intelligence.

Sumário

1	Introdução	10
1.1	Motivação	10
1.2	Objetivos	12
1.3	Trabalhos Relacionados	13
1.4	Organização do Texto	15
2	Agentes Inteligentes	18
2.1	Introdução	18
2.2	O que São Agentes	18
2.3	O que São Agentes Inteligentes	21
2.4	Arquiteturas para Agentes Inteligentes	24
2.5	Aplicações para Agentes Inteligentes	46
3	Implementando Agentes Inteligentes	48
3.1	Introdução	48
3.2	Planos	50
3.3	Estratégias de Comprometimento	52
3.4	Reconsiderando a Intenção	55
4	Confiança no Modelo	59
4.1	Introdução	59
4.2	O que é Confiança	60
4.3	Confiança x Crença	61
4.4	Confiança em um Sistema Multi-Agentes	62
4.5	Confiança em um Agente BDI	65
4.6	Implementação da Confiança	67
5	O Modelo em Lógica Modal	73
5.1	Introdução	73
5.2	Modelagem para o Sistema	74
5.3	A linguagem Lógica para Lidar com o Agente	81
5.4	Modelagem para o Agente	87
5.5	Os Estados Mentais de um Agente	98
5.6	Outras Modalidades	102
5.7	Relacionando Confiança no Modelo	104
5.8	Sucesso de Intenções e Planos	119
5.9	Dois Exemplos Simples	122
6	Conclusões	124
6.1	Decisões de Projeto	125
6.2	Trabalhos Futuros	126
	Referências Bibliográficas	129

Lista de figuras

2.1	Um agente e seu ambiente	20
2.2	Sub-sistemas de ação e percepção	27
2.3	Agentes que contêm estados	28
2.4	Função <i>acao()</i> de um agente lógico	32
2.5	Função <i>acao()</i> de um agente reativo	36
2.6	Diagrama esquemático de uma arquitetura de crenças, desejos e intenções genérica	41
2.7	Função <i>acao()</i> de um agente BDI	43
2.8	Fluxos de informação e controle em três tipos de arquiteturas de agentes em camadas	45
3.1	Função <i>acao()</i> de um agente BDI com planos	51
3.2	<i>Loop</i> de controle de um agente BDI com planos	52
3.3	Introduzindo reatividade: o agente procurará um novo plano quando ele acreditar que um plano falhou por qualquer razão	53
3.4	Descartando intenções quando elas forem impossíveis ou já tiverem sido atingidas	54
3.5	Descartando intenções quando o agente não estiver mais motivado para atingi-las	55
3.6	Um agente cauteloso, que pára para reconsiderar as intenções antes de realizar qualquer ação	56
3.7	Um agente que tenta atingir um balanço entre teimosia e cautela	57
4.1	Diagrama esquemático de uma arquitetura de crenças, desejos e intenções com confiança genérica	65
4.2	<i>Loop</i> de controle de um agente BDI com confiança com comprometimento cego	66
4.3	Agente cabeça-dura com confiança que procurará um novo plano quando ele acreditar que um plano falhou por qualquer razão	67
4.4	Descartando intenções quando elas forem impossíveis ou já tiverem sido atingidas no modelo BDI com confiança	68
4.5	Um agente cabeça-aberta com confiança	69
4.6	Um agente com confiança que sempre reconsidera as intenções	70
4.7	Um agente com confiança que tenta atingir um balanço entre teimosia e cautela	71
5.1	Os elementos da linguagem	85
5.2	Sintaxe das ações	85
5.3	Sintaxe das fórmulas de estado	86
5.4	Sintaxe das fórmulas de caminho	86
5.5	A sintaxe dos predicados de confiança	86
5.6	Exemplo de um modelo temporal ramificado	88
5.7	Regras definindo a semântica de fórmulas de caminho	95
5.8	Regras definindo a semântica de fórmulas de estado	96

Lista de tabelas

3.1 Situações no raciocínio prático

58