

Uma Proposta de Arquitetura Baseada em Ontologias para Chatterbots com Intenção e Personalidade

Renê N.S. Gadelha, Hilário Tomaz
UNIPÊ / WRCO-UFPB
João Pessoa - PB - Brasil
{renegadelha, hilariotomaz}@gmail.com

Ryan Ribeiro de Azevedo
SWORD-CIn-UFPE / SI-UFPI / WRCO-UFPB
Recife - PE - Brasil
rra2@cin.ufpe.br

João Bosco Delfino Júnior, Guilherme Ataíde Dias
IFPB / DCI-UFPB / WRCO-UFPB
João Pessoa - PB - Brasil
jbdj2006@gmail.com, guilherme@dci.ccsa.ufpb.br

RESUMO

Chatterbots são uma realidade, sendo utilizados em diversos contextos, do entretenimento ao educacional. Neste artigo, relatamos o desenvolvimento de uma arquitetura baseada em ontologias para chatterbots com intenção e personalidade. Este projeto possui como principal objetivo ser um instrumento didático na complementação do aprendizado de diversas áreas do conhecimento por estudantes de cursos presenciais e à distância.

Categories and Subject Descriptors

H.4 [Information Systems Applications and Semantic Web applications]: Miscellaneous

Keywords

Chatterbots, Ontologias, Informática e Educação.

1. INTRODUÇÃO

Chatterbots são robôs destinados a conversação com seres humanos e contemplam três princípios básicos de um diálogo real: abertura, desenvolvimento e fechamento e possuem como principal objetivo dialogar ou simular conversas com usuários em linguagem natural e em tempo real [1]. Com o avanço das TDICs-Tecnologias Digitais da Informação e Comunicação e a popularização dos computadores pessoais bem como da Internet, os *chatterbots* são inseridos em um cenário promissor e são utilizados em interações com seres humanos, possuindo assim, um papel educativo já que muitos possuem diversos conhecimentos sobre conteúdos acadêmicos distintos.

O foco deste trabalho é apresentar e utilizar a arquitetura de um *chatterbot* baseado em ontologias, com conceitos genéricos e abstratos de determinado domínio de conhecimento, no qual possui como principal objetivo ser um instrumento didático na complementação do aprendizado de estudantes de cursos presenciais e à distância, apoiando-os na fixação, compreensão e re(construção) de conhecimento e habilidades.

As demais seções deste artigo estão estruturadas da seguinte forma: Na Seção 2 são apresentadas as tecnologias utilizadas no desenvolvimento da proposta e a metodologia de pesquisa utilizada. A proposta (arquitetura e funcionalidades) é tratada na Seção 3, na Seção 4 descrevemos os experimentos e resultados e por fim na Seção 5 são delineadas as conclusões e trabalhos futuros.

2. METODOLOGIA E TECNOLOGIAS UTILIZADAS

Para que o trabalho aqui proposto pudesse ser desenvolvido, foram executadas as seguintes atividades a seguir elencadas: (i) Revisão da literatura; (ii) Definição dos componentes e modelagem da arquitetura proposta; (iii) Implementação da arquitetura e seus componentes, por fim; (iv) Testes e validação da proposta.

As tecnologias envolvidas na realização deste trabalho foram: AIML (*Artificial Intelligence Markup Language*), uma linguagem totalmente baseada em XML (*Extensible Markup Language*) [7]. Sua estrutura é representada por *tags* (marcações) nomeadas categorias, formando bases de conhecimento. A tecnologia iAIML (*Intention Artificial Intelligence Markup Language*) na qual consiste na incorporação do AIML padrão, adicionado mecanismos de tratamento de intenções. O tratamento de intenções é um dos principais problemas dos *chatterbots* e constitui um fator de suma importância na interpretação de diálogos naturais sendo capaz de solucionar alguns dos principais problemas relacionados à conversação com *chatterbots*.

Usamos o Persona-AIML [4] no qual possui características marcantes para criação de *chatterbots*, tais como: modularização, flexibilidade, reutilização e extensibilidade. *Chatterbots* que apresentam traços de personalidade ganham mais desempenho na sua execução. Persona-AIML é uma arquitetura baseada em AIML que introduz ao *chatterbot* aspectos de personalidade, podendo ser baseados em modelos de personalidade. Utilizamos o modelo composto de cinco elementos distintos, proposto por [6]: traços, atitudes, humor, emoções e estados físicos.

Foram desenvolvidas e utilizadas ontologias como componentes de representação e bases de conhecimento. A utilização de ontologias vem das vantagens citadas por [3], são elas: - Oportunidade de os desenvolvedores reusarem ontologias e bases de conhecimento, mesmo com adaptações e extensões. Portanto, reusar ontologias promove um ganho significativo em termos de esforços e de investimentos; - A grande disponibilidade de “ontologias de prateleira”, prontas para uso, reuso e comunicação entre agentes, podendo estas serem estendidas e complementadas com conceitos de domínios específicos; - Possibilidade de tradução entre diversas linguagens e formalismos de representação do conhecimento tais como: *CLIPS*, *Jess*, *Prolog*, *XML*, *RDF*, *OWL*, *OIL*, *DAML-OIL* e *FLogic*;

Na seção seguinte, é apresentada a proposta, sua arquitetura e suas funcionalidades.

3. PROPOSTA

Apresentamos neste trabalho a arquitetura de um *chatterbot* inteligente baseado em ontologias cujo principal objetivo é auxiliar alunos e professores no processo de ensino/aprendizagem, provendo conhecimento específico por meio de diálogos mais próximos de interlocutores humanos. A arquitetura, componentes e funcionalidades do sistema proposto serão apresentadas na subseção a seguir.

3.1 Arquitetura

A proposta aqui apresentada, desenvolvida sobre a plataforma web por meio da tecnologia *Java Enterprise Edition (JEE)*¹, é composta por quatro módulos, de acordo com a Figura 1. Os módulos são descritos a seguir:

Módulo AIML – Responsável pelo tratamento das perguntas emitidas pelos usuários, sendo dividido em seis partes, facilitando a manutenção e o reuso da arquitetura.

- **Arquivos de Abertura:** responsáveis pelo tratamento de perguntas referentes à fase de abertura do diálogo com o usuário (iAIML).
- **Arquivos de Desenvolvimento:** possuem o tratamento de perguntas referentes ao domínio da Ontologia.
- **Arquivos de Fechamento:** contém o tratamento de perguntas com intenção de fechamento do diálogo com usuário (iAIML).

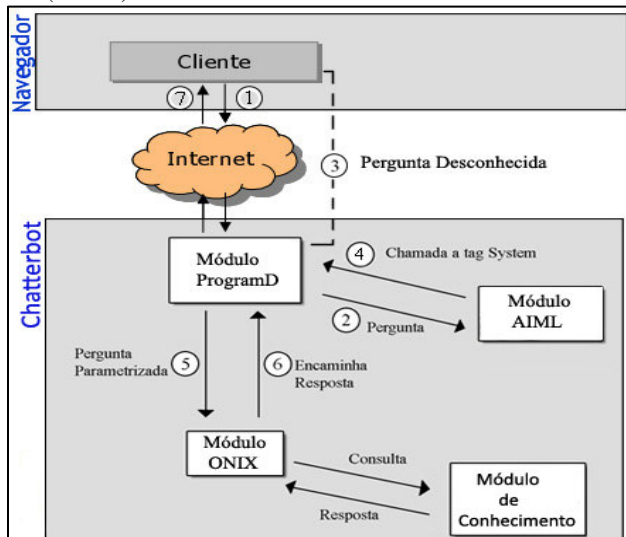


Figura 1: Interações entre os módulos da Arquitetura.

- **Componente de Personalidade:** define as crenças, os elementos de personalidade e as regras que determinam o comportamento do *chatterbot* (Persona-AIML).

¹ <http://java.sun.com/javaea>

² API responsável por acessar informações a partir de uma ontologia.

- **Componente de Raciocínio:** após identificar a sentença do usuário, baseada nas perguntas, verifica as regras do componente de personalidade, de forma a determinar a resposta ao usuário (Persona-AIML).

- **Histórico de Conversação:** Possui um registro de todas as perguntas e respostas possibilitando ao componente de raciocínio associar perguntas dos usuários a futuras respostas do *chatterbot* (Persona-AIML).

Módulo ProgramD – Interpretador da base AIML, gerencia o fluxo de perguntas e respostas de forma a fornecer a conexão entre o usuário e os demais componentes do sistema. Além de prover suporte a linguagem AIML, este módulo permite ao sistema integrar diversas bases de perguntas, que aliado a capacidade da API ONIX² em suportar múltiplas ontologias, possibilita ao *chatterbot* ser aplicado a mais de uma área de conhecimento.

Módulo Onix: responsável por acessar informações das ontologias de forma simples, criando uma interface geral para acesso da mesma pelo módulo ProgramD.

Módulo de Conhecimento: base de conhecimento composta por ontologias que provêm respostas para perguntas referentes a um determinado domínio, sendo acessada por meio da ONIX.

3.2 Funcionalidades

O funcionamento desta arquitetura se inicia com o usuário enviando uma mensagem ao *chatterbot*, passo (1), sendo o módulo ProgramD responsável por encaminhá-la à base AIML, passo (2). De acordo com o conteúdo desta mensagem, o fluxo da aplicação ocorre de três maneiras distintas como segue:

- O padrão da pergunta não é encontrado em nenhum dos arquivos existentes na base AIML, sendo assim o ProgramD responde ao usuário por meio de uma resposta padrão informando-o da inexistência da informação desejada, passo (3).
- O padrão da pergunta é encontrado nos arquivos de abertura ou fechamento. Neste caso o ProgramD retornará ao usuário a respectiva mensagem contida na *tag template*, passo (7).
- O padrão da pergunta é encontrado no arquivo de desenvolvimento. Neste arquivo, cada padrão está associado a uma consulta parametrizada por meio da *tag System*, passo (4). O módulo ProgramD encaminha esta consulta ao ONIX, que processará a requisição de forma a identificar qual ontologia será consultada, passo (5) no Módulo de Conhecimento. Após a obtenção da informação contida na respectiva ontologia, o ONIX encaminha ao ProgramD a resposta, passo (6), sendo esta repassada ao usuário por meio da interface *web*, passo (7).

4. EXPERIMENTOS E RESULTADOS

Com o objetivo de validar a arquitetura proposta, apresentamos nesta seção os testes e experimentos realizados. Foram contempladas características para fins de validação, entre estas se destaca a possibilidade de utilizar uma ou mais ontologias como base de conhecimento, permitindo ao *chatterbot* responder perguntas de domínios distintos. Desta forma, desenvolvemos

duas ontologias de aplicação segundo classificação adotada por [5], uma específica para o domínio da Psiquiatria e outra para o domínio de Segurança da Informação, testando assim, o *chatterbot* com questões referentes aos dois domínios de conhecimento.

Para o desenvolvimento das ontologias utilizamos a metodologia Methontology [2] baseada no padrão IEEE para desenvolvimento de *software* e desenvolvimento de sistemas de informação baseados em conhecimento. É apresentado na Figura 2 uma das ontologias desenvolvida.

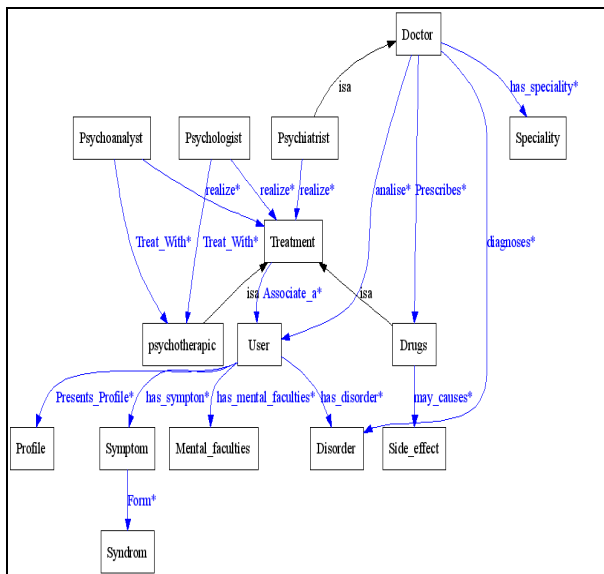


Figura 2: Classes e relacionamentos que formam a ontologia de Psiquiatria.

Foram realizados experimentos com interlocutores reais, somando um total de 21 usuários que mantiveram diálogo aberto com o *chatterbot*, realizando perguntas referentes ao domínio das supracitadas ontologias, no qual 11 foram do curso de Ciência da Computação do Centro de Informática da Universidade Federal do Pernambuco, 6 do curso de Psicologia do Centro Universitário de João Pessoa e 5 do curso de Medicina da Universidade Federal da Paraíba. Foram realizadas perguntas acerca dos diferentes domínios aqui abordados, sendo estas previamente conhecidas pelos participantes do experimento e posteriormente divididas em três categorias: referentes à Segurança da Informação, à Psiquiatria e a perguntas contidas nos arquivos de abertura e encerramento do diálogo.

Após a interação com o *chatterbot*, 90% dos usuários indicaram êxito na obtenção das respostas, no entanto percebeu-se que os restantes não informaram ao *chatterbot* o padrão correto da pergunta, sendo assim, excluídos dos resultados do experimento. Outro quesito atestado foi a qualidade do diálogo provido pelos módulos IAIML e Persona-AIML, em que constatou-se o funcionamento adequado de cada módulo de acordo com sua proposta. Na Figura 3, verifica-se um trecho de diálogo no qual o interlocutor realiza perguntas referentes aos dois domínios, sendo estas respondidas através da respectiva ontologia. Também é possível perceber o tratamento inteligente de sentenças ausentes da base de conhecimento.

Histórico de Conversação	
Bot	Olá <Nome do Usuário>, em que posso lhe ajudar?
Usuário	Qual o tratamento mais indicado ao usuário para depressão?
Bot	Antidepressivos, algo mais ?
Usuário	Qual a contramedida para ataques UPD Packet Storm ?
Bot	Analisar tráfego de rede, utilizar IDS, algo mais ?
Usuário	O que é Engenharia de Software ?
Bot	Não tenho como responder sua pergunta. Refaça sua pergunta.
Usuário	Até logo.
Bot	Até logo, espero ter lhe ajudado.

Figura 3: Histórico de um diálogo com o usuário.

5. CONCLUSÕES E TRABALHOS FUTUROS

A arquitetura proposta cumpre o papel ao qual se propõe, constituindo-se de uma ferramenta computacional inovadora que pode ser utilizada em contextos educacionais, permitindo ao usuário obter conhecimento específico a respeito de determinado domínio. Os objetivos do projeto foram contemplados possibilitando a seus usuários abordarem diferentes áreas de conhecimento por meio de diálogos consistentes. Os resultados obtidos indicam que a arquitetura proposta realmente permite suporte a múltiplas ontologias, além de prover ao *chatterbot* um diálogo mais coerente com os usuários. Como pesquisa futura, propomos e estamos trabalhando no desenvolvimento de uma interface por comando de voz e 3D a fim de melhorar comunicação e a interação entre usuários no processo de ensino/aprendizagem.

6. REFERÊNCIAS

- [1] Dias, G., A; Neves, D., A., B.; Silva, J., W., M.; Neto, G., H., N.; Almeida, M., J., S., C. Representando o Conhecimento Através de Ontologias: o Caso do Chatterbot Lunmi. VIII Encontro Nacional de Pesquisa em Ciência da Informação, 2007.
- [2] Fernández, M. A.; Gómez-Pérez, A.; Juristo, N. 1997. Methontology: From ontological art towards ontological engineering. In Proceedings of the AAI Spring Symposium Series, p. 33-40;
- [3] Freitas, F. Ontologias e a web semântica. In: Renata Vieira; Fernando Osório. (Org.). *Anais do XXIII Congresso da Sociedade Brasileira de Computação*. Campinas: SBC, 2003. v. 8, p. 1-52.
- [4] Galvão, A. "Persona-AIML: Uma Arquitetura para Desenvolver Chatterbots com Personalidade". Dissertação de Mestrado, Centro de Informática, UFPE, Recife, Brasil, 2003.
- [5] Guarino, N. 1998. Formal Ontologies and Information Systems. In Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Itália. 6-8 June. P3-15. IOS Press;
- [6] McCrae, R; John, O. *An Introduction to the Five-Factor Model and its Applications*. Journal of Personality, Vol 60, pg. 175-213, 1992
- [7] Neves, André M. M., Barros, Flávia de A. iAIML: Um Mecanismo para Tratamento de Intenção em *Chatterbots*. XXV Congresso da Sociedade Brasileira de Computação. 2005.