

Maga Vitta – um assistente inteligente para jogo educacional

Taiuã Pires¹, Otto Lopes Braitback de Oliveira¹, Lisardo Sallaberry Kist², Alexsandro Rosa dos Anjos³, Leonivan Moraes³, Paloma Dias Silveira⁴, Carla Inez Lima Freitas⁴ e Daniel Nehme Müller³

Curso de Jogos Digitais, Universidade do Vale do Rio dos Sinos (UNISINOS)¹

Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS)²

Conexum Ltda., Centro de Empreendimentos em Informática, UFRGS³

Laboratório de Estudos em Linguagem, Interação e Cognição (LELIC), Faculdade de Educação (FACED), Pós-Graduação em Educação (PPGEDU), UFRGS⁴

taiuapires@gmail.com, ottolb@gmail.com, lisardokist@gmail.com,
alex@conexum.com.br, leonivanmoraes@gmail.com, paloma.dias@gmail.com,
carlainlima@uol.com.br, daniel@conexum.com.br

Abstract. *The ASSISTIVA – Vocal Intelligent Assistant to Maga Vitta Project aims to accomplish the free software development, by CONEXUM company, of an intelligent agent to assist children and users of educational game cities simulator Città, developed by LELIC (Laboratory of Studies in Language, Interaction and Cognition at the Faculty Education from UFRGS). Maga Vitta is an assistant that performing a dialogue with the user, reporting issues on the virtual city model in development. With the technology developed through this project, it is intended for the future establishment of other virtual assistants and other mobile gaming or monitoring systems.*

Resumo. *O Projeto ASSISTIVA – Assistente Inteligente Vocal para Maga Vitta visa o desenvolvimento em software livre pela empresa CONEXUM de um agente inteligente para assistência às crianças usuárias do jogo educacional e simulador de cidades Città, desenvolvido pelo LELIC (Laboratório de Estudos em Linguagem, Interação e Cognição da Faculdade de Educação da UFRGS). A Maga Vitta é uma assistente que dialoga com o usuário, relatando questões sobre a maquete virtual em desenvolvimento. Com a tecnologia desenvolvida através deste projeto pretende-se a futura criação de outros assistentes virtuais para dispositivos móveis e outros sistemas de jogo ou monitoramento.*

1. Introdução

O projeto ASSISTIVA – Assistente Inteligente Vocal para Maga Vitta propõe a integração de diferentes áreas de pesquisa para o desenvolvimento de um assistente virtual de diálogo falado. O ASSISTIVA insere-se no Projeto CIVITAS – Cidades Virtuais com Tecnologias para Aprendizagem e Simulação, em andamento na Faculdade de Educação da UFRGS, o qual possui um jogo educacional chamado Città, de onde parte a concepção da agente inteligente virtual Maga Vitta. O projeto CIVITAS é uma abordagem de ensino e pesquisa que gira em torno do desenvolvimento de maquete de cidades projetadas por alunos das séries iniciais [Axt e Martins, 2008].

Dentro deste contexto, o jogo Città insere-se como uma maquete digital que permite um diferente nível de interação entre os alunos. Città é um jogo educacional colaborativo, que visa à construção coletiva, voltada para a ação cooperativa e negociada. No jogo, a assistente virtual Maga Vitta foi projetada para ser um agente inteligente para ser seu regulador ecológico. A agente interage com o usuário na forma de alertas, de forma a promover a problematização de situações críticas que poderiam trazer prejuízos à população

da cidade virtual.

2. A agente inteligente Maga Vitta

Diz-se que um agente inteligente é um agente de conversação (conversational agent) quando possui a capacidade de manter uma conversa “humana”, face-a-face com o usuário humano, através do reconhecimento da fala e dos gestos envolvidos durante uma conversação. Portanto, um *Embodied Conversational Agent* (ECA) é um agente autônomo, inteligente, que tem uma representação gráfica (embodied) em um mundo virtual e é capaz de manter alguma forma de conversação com o usuário. Cassel (2001) propõe um modelo conversacional, conhecido como modelo FMBT (leia-se fembot - *Functions, Modalities, Behaviors, Time*), que incorpora as principais características presentes numa conversação. O modelo FMBT é caracterizado por: (1) funções proposicionais e funções interacionais que correspondem, respectivamente, ao conteúdo da conversa (inclui fala, entonação da voz e gestos corporais) e as *deixas* que regulam o processo de conversação, as quais compreendem tanto comportamentos sonoros, como *hã?*, *óóó!*, *hein?*, *iih*, quanto não-sonoros, como movimento de ombros, mãos e face para expressar dúvida, alegria, etc.; (2) modalidades de comunicação, responsáveis por construir as funções proposicionais e interacionais, podem ser verbais (palavra, canto, sons, etc.) e não-verbais (gestos manuais, faciais e corporais, piscar os olhos, olhar fixo, etc.); (3) comportamentos na conversação, onde a mesma função de comunicação pode assumir vários significados, por exemplo, baixar a cabeça pode denotar concordância, saudação ou ênfase em uma palavra; (4) tempo e sincronização da conversação, na medida em que a duração excessiva de emissão da fala ou de um sinal pode, eventualmente, indicar diferentes sentimentos como ansiedade, nervosismo e tédio, e atrasos na resposta podem ser entendidos como não cooperação, desatenção, desacordo, dúvida, desistência, etc.

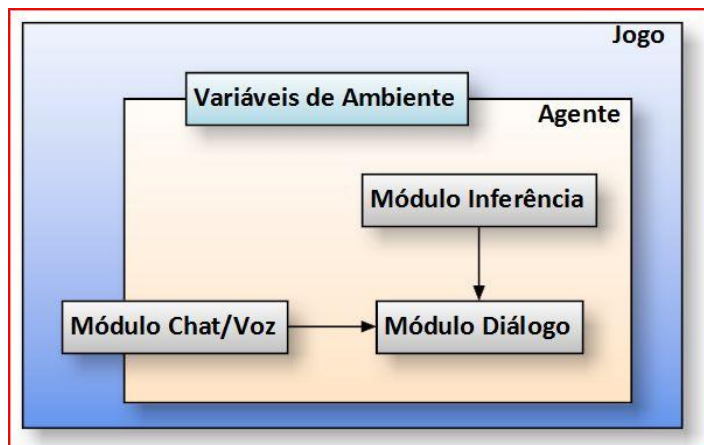


Figura 1 - Esquema de dependências dos módulos

Durante a construção da cidade virtual, o aluno conta com auxílio de um agente que observa suas ações no ambiente virtual. A ECA Maga Vitta não intervém na construção da cidade, mas está sempre atenta para indagar, aconselhar, informar e provocar reflexões sobre os processos de construção da cidade. Além disso, Maga Vitta é uma agente que possui como padrão interno estar eternamente preocupado com a ecologia: ao memorizar ações contra o equilíbrio ecológico, adverte sobre possíveis tragédias ambientais. Em sua interação com o jogo Città, Maga Vitta tem um motor de inferência que trabalha em conjunto com um módulo de conversação, que pode interagir também com voz dentro do ambiente de construção da cidade virtual. Veja a figura 1.

A agente Maga Vitta incorpora a arquitetura padrão de agente autônomo, inteligente e

as propriedades de um ECA. Tais propriedades são essenciais para que a agente Maga Vitta expresse o que *deseja* falar, bem como para que construa o plano de interação com o usuário, de forma a se revestir de credibilidade.

3. O jogo Città

Città é um jogo educacional colaborativo, que visa à construção coletiva, voltada para a ação cooperativa e negociada. Ele tem por objetivo apoiar pedagogicamente na construção de conhecimento dos alunos do 4º ano do ensino fundamental, no que se refere ao desenvolvimento de capacidades que lhes permitam pensar - construir - organizar uma cidade virtual a partir da realidade infantil e seus conhecimentos incorporados através dos conteúdos oferecidos no aprendizado dos sistemas que regulam uma cidade e suas relações [Müller et al, 2009].

O objetivo principal do Città é que o aluno aprenda através da construção e da organização de uma cidade. Para que o mesmo passe por diversas situações que proporcionem este aprendizado ele terá que completar tarefas fornecidas pelos educadores ou pelo próprio jogo. Nessas tarefas os alunos terão que satisfazer a população e ter uma sustentabilidade que respeite o meio ambiente.

3.1. Banco de Dados

Foi criado um banco de dados SQLite¹ para armazenar informações de todos os eventos do jogo sobre cada mapa de cidade diferente, permitindo armazenamento de textos de extensão indeterminada para cada ação feita sobre o mapa do jogo que está sendo manipulado durante a execução.

Cada mapa é composto de 100x100 células, totalizando dez mil tabelas no banco de dados. Há a necessidade de tantas tabelas porque as construções no jogo têm tamanhos diferentes, ocupando não somente posições, mas áreas arbitrárias. Isso ocorre porque o espaço de uma grande construção pode ser ocupado por várias construções menores. Por exemplo, o que antes era uma fazenda passou a ser várias casas. Todo este histórico de construções fica armazenado nas tabelas das células envolvidas.

Para o histórico individual do jogador é usada uma tabela principal para os eventos feitos pelo jogador que são adicionados a medida que o jogador atua no mapa, ficando consequentemente em ordem cronológica, e tabelas de texto para os textos escritos pelo jogador. Esse histórico será a base para a inteligência artificial da Maga criar uma personalidade para lidar com o jogador. O histórico individual do jogador será usado em todos os mapas em que este jogar, tornando a personalidade da Maga persistente ao usuário.

3.2. Módulo de Diálogo

O módulo de diálogo é uma parte importante tanto para o jogo quanto para a integração das áreas envolvidas no projeto, pois para a construção das frases é importante ressaltar que ocorre um planejamento conjunto entre a equipe do ASSISTIVA e do CIVITAS para que as respostas da Maga não sejam dadas diretamente, mas sim, que estimulem a mente das crianças.

Em uma simplificação para explicar o módulo de diálogo, a figura 2 mostra as etapas para construção do diálogo onde a frase entra no sistema através da interface do chat e o controlador que dispara processos para a construção do diálogo.

¹ <http://www.sqlite.org/>

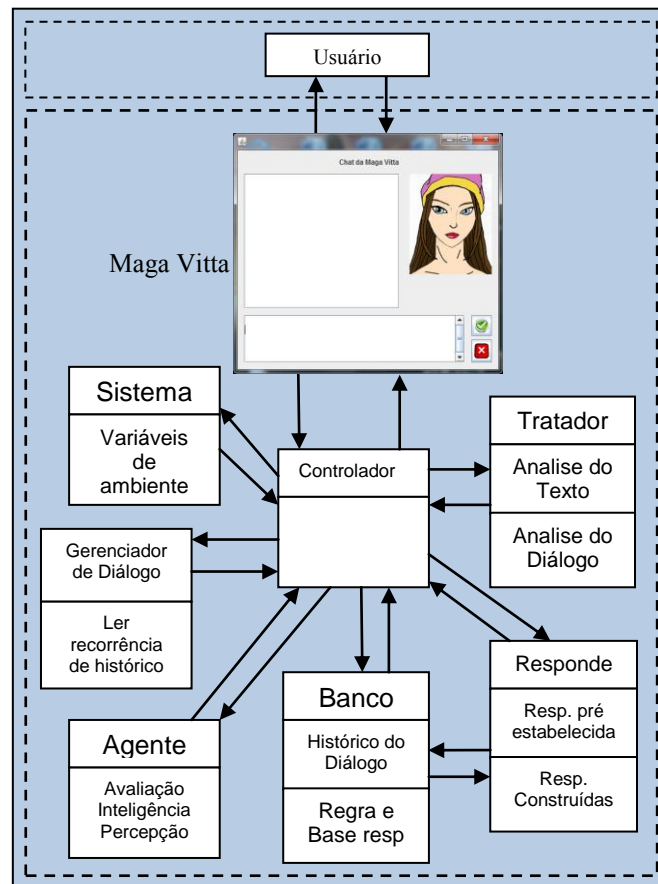


Figura 2 – Diagrama de construção da frase (Maga)

O processo começa com o controlador enviando o texto recebido através da fala ou da escrita para o tratador para que possamos identificar sobre que assunto estará sendo conversado e qual a necessidade solicitada. Em paralelo a esse processo, o controlador dispara outro processo para gerenciador de diálogo, que fará o controle informando se o diálogo está aberto, se é o mesmo assunto durante o diálogo, e se já houve ocorrência do mesmo problema. Dessa forma chegamos à formulação da resposta. O controlador utiliza a consulta aos históricos sobre o assunto, os últimos casos similares ocorridos e as regras para formulação da resposta, assim resultando a resposta completa a ser dada através do chat para o aluno.

Para que todo esse processo de formulação da resposta obtenha resultados, foram necessários estudos em Processamento de Linguagem Natural (PLN) para entender o processo de desenvolvimentos das respostas para as conversações [Jurafsky e Martin, 2000], as quais tratam de multifuncionalidade, multidimensionalidade, anotação multidimensional, especificações do ato de diálogo, lingüística computacional e interação de voz em jogo [Petukhova e Bunt, 2009].

Dentro do desenvolvimento da Maga, a equipe do CIVITAS está trabalhando em diferentes personalidades para influenciar na emoção com que serão geradas as respostas para os alunos.

3.2.1. Da construção dos Diálogos

Ao criar diálogos tendo como base o estudo do Método Clínico Experimental de Jean Piaget, nossa ideia é possibilitar que a interação da agente inteligente virtual Maga Vitta e as

crianças construtoras da cidade virtual através do Città ocorram por meio de interferências criativas da agente, durante o jogo, seja pela solicitação da criança ou pela necessidade que surgir a partir das ações da criança sobre o jogo.

As crianças usuárias frequentam o 4º ano do Ensino Fundamental (antiga 3ª série), com nove anos de idade aproximadamente, das escolas participantes do Projeto CIVITAS. Segundo Piaget (1993), no período dos nove aos 11 anos aproximadamente, pode-se observar que a criança já possui uma organização assimilativa rica e funcionando em equilíbrio com um mecanismo assimilativo. A criança encontra-se na fase das operações concretas, que é uma etapa marcada por grandes aquisições intelectuais. As formas de explicação para os fenômenos que ocorrem a sua volta deixam de ser egocêntricas e passam a ser por identificação, é um estágio intermediário entre a heteronomia moral e a autonomia plena.

Considerando, portanto, a proposta de que o diálogo da agente Maga Vitta com as crianças possibilite uma interação criativa, nosso estudo toma como base o Método Clínico Piagetiano e os estudos deste epistemólogo sobre a formação dos mecanismos mentais da criança.

3.3. Módulo de Inferência

A agente Maga Vitta é a consciência do jogo Città, promovendo a regulação do sistema ecológico. A sua proposta não é realizar a orientação direta, mas provocar ações internas no aluno de forma a fazê-lo se preocupar com a situação indicada no momento do jogo. Dependendo da gravidade da situação, indicada pelo nível de equilíbrio da variável de ambiente, e ainda da prioridade do elemento que está sendo considerado, é indicado ao módulo de diálogo a importância e nível de expressividade que o agente terá que executar.

3.4. Módulo de Voz

O módulo consiste em um transcritor de voz para texto, que será enviado para o assistente do jogo. Para realizar a transcrição é utilizado o software livre Julius². Por outro lado, a decisão do projeto foi por usar a API brasileira Coruja, baseada no Julius. Trata-se do projeto Falabrasil³, do Laboratório de Processamento de Sinais (LAPS), que permitiria o trabalhar com o reconhecimento na forma de API. Houve muitas dificuldades para conseguir compilar o projeto e obter os primeiros resultados. Foi escrito um pequeno tutorial de como realizar a compilação, que foi enviado para a lista de e-mail do Coruja para os usuários que estavam tendo problemas semelhantes, mas por decidiu-se utilizar diretamente o Julius para o reconhecimento.

Apesar do reconhecimento de voz estar bastante desenvolvido para o inglês, houve dificuldade em obterem-se modelos acústicos para português. Os primeiros testes foram realizados com o modelo acústico e linguístico do próprio LAPS (versão 1.3, na época), mas o reconhecimento era bastante ruim. Assim, está sendo treinado um modelo acústico próprio para português a partir das gravações que a Conexum tem disponível.

Como mostra a figura 3, o módulo de voz funciona recebendo um stream de áudio (arquivo ou sinal do computador) e retornando a transcrição na forma de stream de texto. O programa presume um modelo acústico e um modelo linguístico sobre qual o reconhecimento será realizado. Modelo acústico trata-se de um modelo estocástico de cadeias ocultas de markov sobre qual o reconhecimento é feito. O modelo linguístico é uma gramática formal especificada para aplicação. Texto reconhecido poderá ser acessado como

² <http://julius.sourceforge.jp/>

³ <http://www.laps.ufpa.br/falabrasil/>

variável do programa e utilizado na assistência ao aluno.

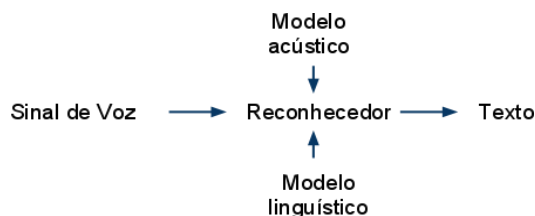


Figura 3 - Esquema funcional do módulo de reconhecimento

As respostas processadas também devem ser sintetizadas de volta para o estudante. No nível de síntese de voz, tem-se realizados testes com software livre Espeak⁴ e obtidos resultados satisfatórios

4. Conclusões

Como o Projeto ASSISTIVA está utilizando o jogo Città, o qual tem seu código no SourceForge⁵, na medida em que os testes dos protótipos forem consolidados, os códigos desenvolvidos para a assistente serão disponibilizados juntamente com o jogo. Maiores informações podem ser obtidos em seu sítio na internet do ASSISTIVA⁶.

O modelo de agente proposto no presente artigo é a tentativa de criação do assistente inteligente Maga Vitta para o jogo educacional Città, ambos concebidos dentro do Projeto CIVITAS, do LELIC/UFRGS. A Conexum Ltda., através do Projeto ASSISTIVA, entra como parceira deste software livre na produção do conhecimento de uma nova tecnologia de assistentes virtuais, com o intuito de promover o debate acadêmico e comercial deste novo campo que se abre na interação humano-computador. Espera-se que, com esta iniciativa de tecnologia aberta, sejam desenvolvidos trabalhos derivados, permitindo assim a evolução da concepção de um assistente pessoal conversacional para uso nos mais diversos dispositivos eletrônicos.

Agradecimentos. O projeto Assistiva é financiado com o apoio do CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico. Agradecemos a toda a equipe do Projeto CIVITAS, por sua assessoria a este projeto.

Referências

- Axt, M. e Martins, M. A. R. (2008) Coexistir na diferença: De quando a formação em serviço pensa modos de habitar a sala de aula. In: TRINDADE, Iole Maria Faviero. (Ed), *Múltiplas Alfabetizações e Alfabetismos*, pp. 133-158, Editora da UFRGS, Porto Alegre, 1 ed., v. 1.
- Cassel, J. (2001), *Embodied Conversational Agents: Representation and Intelligence in User Interface*. In *AI Magazine*, 22(3), pages 67-83, AAAI Press.
- Jurafsky, D. e Martin, J. H. *Speech and Language Processing*. University of Colorado. Boulder, 2000
- Petukhova, V. e Bunt, H. (2009). The independence of dimensions in multidimensional dialogue act annotation Proceedings of NAACL HLT 2009: Short Papers, pages 197–200,. Boulder, Colorado, June 2009.
- Piaget, J. *Seis estudos de Psicologia*. Rio de Janeiro, Forense Universitária, 1993.

⁴ <http://espeak.sourceforge.net>

⁵ <http://civitas.sourceforge.net>

⁶ <http://assistiva.conexum.com.br>