

Geração e sequenciamento automáticos de lexemas e sentenças em Sistema Tutor Inteligente para ensino de línguas

Ismael Ávila¹, Ricardo R. Gudwin²

¹Centro de Pesquisas e Desenvolvimento em Telecomunicações (CPqD)
Rodovia SP-340, km 118,5 – 13086-902 – Campinas – SP – Brasil.

²Faculdade de Engenharia Elétrica e Computação – Universidade de Campinas
(UNICAMP) Caixa Postal 6.101 – 13.083-970 – Campinas – SP – Brasil.

avila_an@cpqd.com.br, gudwin@dca.fee.unicamp.br

***Abstract.** The paper briefly describes an ITS architecture for language teaching which is based on three main pillars: (i) the target language (L2) is modeled as a multiagent system where lexeme-agents autonomously attempt to be learned competing for a limited resource, the course interface; (ii) the images used in the course scenes are linked with metadata that describe them semantically and (iii) an ontology of L2 defines the semantic and syntactic relations among its lexemes and then support the inferences made by the multiagent system. We discuss the approach and some of its implications.*

***Resumo.** O artigo descreve resumidamente uma arquitetura STI para ensino de línguas que é baseada em três principais pilares: (i) a língua-alvo (L2) é modelada como um sistema multiagente no qual os agentes-lexemas buscam autonomamente ser aprendidos e competem por um recurso limitado, a interface do curso; (ii) as imagens usadas nas cenas do curso são associadas a metadados que as descrevem semanticamente e (iii) uma ontologia da L2 define as relações semânticas e sintáticas entre seus lexemas e assim embasa as inferências feitas pelo sistema multiagente. Nós discutimos essa abordagem e salientamos algumas de suas implicações.*

1. Introdução

Um dos grandes desafios do ensino de línguas assistido por computador é modelar a complexa natureza de uma língua. Isso se refere, por exemplo, à amplitude do léxico, às nuances semânticas de cada lexema e suas interdependências com outros lexemas. Esses conhecimentos, que nos falantes da língua já se encontram internalizados, devem ser expostos e exemplificados aos alunos em contextos válidos, variados e tantas vezes quantas forem necessárias para permitir sua assimilação como segunda língua (L2). Em situações de ensino presencial o professor é tanto o especialista no domínio, quanto o responsável por avaliar o aprendizado e detectar quando um tópico deve ser reforçado ou contrastado em resposta a dúvidas, hesitações e erros dos alunos. Já nas arquiteturas baseadas em sistemas tutores inteligentes (ITS), sobretudo no ensino a distância, para substituir ou apoiar o professor o sistema tutor deve manter uma representação do domínio a ser ensinado e ser capaz de modelar o processo de aprendizado dos alunos. Em um ITS para o ensino de línguas (ILTS) isso implica possuir um modelo de domínio que capture a riqueza e a sutileza da linguagem humana e um modelo do aluno que correlacione as ações e respostas deste com aspectos da L2 e de seu aprendizado.

O desenvolvimento de uma estratégia abrangente de modelagem do domínio em um ILTS é ainda um desafio. A fim de contribuir com essa área, o objetivo deste estudo em andamento é investigar um novo tipo de modelagem do domínio (da L2), no qual a L2 é modelada como um sistema multiagente cujos agentes representam seus lexemas e perseguem o objetivo individual de serem aprendidos. Para isso, eles devem competir por espaço na interface do curso, um recurso limitado, para que ali possam apresentar-se ao aluno com o apoio de imagens ilustrativas. Nesses termos, a disputa por um lugar na tela do curso pode ser entendida como um processo de seleção de ação.

O agente-lexema também pode aparecer formando sentenças. Isso é importante porque o aprendizado do valor semântico de um lexema dependerá, em muitos casos, de sua articulação com outros lexemas em sentenças e frases. Logo, a despeito de competir por espaço, o agente-lexema deve também cooperar com outros lexemas para favorecer a exibição daqueles com os quais ele normalmente forma sintagmas e sentenças válidas na L2. E assim as sequências de exibição dos lexemas e de formação das sentenças emergirão dessas interações (competições e cooperações). Isso deverá ocorrer de forma autônoma no ambiente, assemelhando-se às dinâmicas de coalizões propostas em Baars (1997) e implementadas, por exemplo, por Dubois (2007) para outro cenário de ensino.

2. Geração e sequenciamento automáticos de lexemas e sentenças

A modelagem multiagente dá ao sequenciamento dos conteúdos do curso uma lógica *bottom-up* baseada em competições e coalizões de lexemas. Isso não significa, contudo, que não possam ser definidas metas didáticas *top-down* para o curso. O gestor do curso pode definir táticas que priorizem certos tópicos e temas, e assim favorecer os lexemas correspondentes na competição por espaço. Ele também pode priorizar palavras mais frequentes na L2, ou as mais parecidas com as correspondentes na L1, por exemplo.

A formação de sentenças resulta de uma coalizão dos lexemas que as formam. Cabe ressaltar que elas só são formadas quando os elementos dos quais são constituídas já tiverem sido trabalhados, ao menos em sua maioria, permitindo que o(s) vocábulo(s) desconhecido(s) seja(m) inferido(s) com base naqueles já conhecidos.

A geração automática de sentenças-exemplo é uma funcionalidade útil a um ILTS, pois ela libera o especialista (tutor humano) da tarefa de elaborar essas sentenças sempre que necessárias, e de forma individualizada para cada aprendiz, ou de ter que elaborar *a priori* um vasto repertório de sentenças para uso no curso. Entretanto, a geração automática não pode prescindir dos conhecimentos que o tutor humano tem da L2, isto é, de sua vivência com a língua e com seus recortes semânticos, suas regras sintáticas e aspectos pragmáticos. Isso significa então que a automatização da geração de sentenças requer formas de sistematizar o conhecimento linguístico do tutor humano de maneira tratável pelo ILTS. Na IA, uma solução sistemática de representação do conhecimento se baseia no uso de ontologias. Assim, o uso de uma ontologia para definir as relações semânticas entre os lexemas da L2 é o segundo pilar da arquitetura. A ontologia define as relações de sinonímia, antonímia, hiponímia e hiperonímia entre os lexemas da L2, além das regras que ligam os diversos conceitos segundo suas relações existenciais e associa a cada substantivo os adjetivos com os quais ele pode se articular, ou os verbos dos quais ele pode ser sujeito ou objeto. Por exemplo, a ontologia pode informar, de modo inteligível a um computador, que o ente *rinoceronte* é um animal mamífero, grande, que tem chifres, quatro patas, rabo e que ele se alimenta de vegetais. Essa ontologia constitui então uma base de conhecimento sobre a L2. Ela pode ser uma

combinação dos conhecimentos que o especialista do domínio possui com outros dados provenientes de gramáticas e dicionários da L2, ou mesmo de registros de uso da L2 (obtidos, por exemplo, por motores de busca na Internet). Um aplicativo específico para criação e uso de ontologias, o Protégé, foi utilizado no ambiente de ensino aqui descrito.

Além do mecanismo multiagente e do uso de ontologias, a arquitetura se baseia em um terceiro pilar, que é o da utilização, no ambiente de ensino, de metadados para seus conteúdos visuais, que são as imagens que o curso usa para ilustrar o significado de lexemas ou de sentenças e dar o contexto ou pano de fundo de cada cena. O mecanismo de competição e cooperação entre agentes leva em conta a informação dos metadados para avaliar a pertinência de um lexema a uma dada cena. Assim, se o lexema é um elemento presente na ou sugerido pela imagem, ou, se a partir da ontologia, o *reasoner* infere que ele é pertinente aos elementos anotados na cena, seu agente recebe *ativação contextual* e passa a levar vantagem na disputa por espaço. A associação de imagens a significados nas cenas é feita por descritores semânticos, inclusive com a segmentação e identificação de elementos ou partes de objetos. Por exemplo, na imagem associada a *rinoceronte* os metadados podem informar quais segmentos da imagem representam objetos ligados ao lexema-foco, como *chifres*, *patas*, *boca*. Esses metadados são criados *a priori* e codificados no padrão MPEG-7 pelo especialista do domínio, que assim pode, considerando cada imagem usada no curso, agregar os descritores e mapas semânticos que identifiquem os principais elementos e as interrelações e papéis desses elementos. As relações mapeadas no MPEG-7 (*agentOf*, *beneficiaryOf*, *componentOf*, *dependsOn*, por exemplo) são codificadas em XML, podendo assim ser interpretadas pelos agentes.

As informações do contexto da cena somam-se àquelas contidas na ontologia, e permitem ao ILTS formar sentenças semanticamente válidas e que reflitam o contexto ilustrado. Os exemplos gerados tratam primeiro de elementos retratados na imagem, na forma de palavras isoladas ou sintagmas curtos, e em seguida revisitam esse vocabulário na forma de sentenças mais longas, baseadas em dados da ontologia sobre o objeto ou conceito retratado. No exemplo do *rinoceronte*, na imagem este está parado, com seu dorso sujo de lama, e não aparece comendo. Mas considerando-se, pela ontologia, que animais dessa espécie comem vegetais, são rápidos, irritadiços e habitam savanas africanas, pode-se concluir que sentenças adequadas à cena serão: *O rinoceronte está parado* ou *o rinoceronte está sujo de lama*. Além dessas, sentenças aplicáveis à espécie como um todo serão: *Rinocerontes vivem na África*, *Rinocerontes comem vegetais* ou *Rinocerontes são nervosos e velozes*. Por outro lado, sentenças como *Rinocerontes são carnívoros* ou *O rinoceronte está voando* não serão formadas porque são anômalas ao violarem informações da base de conhecimento (ontologia) ou da descrição contextual (metadados), respectivamente. As sentenças anômalas dificultariam a compreensão do significado dos lexemas, pois o sentido desses enunciados se choca com conhecimentos do aluno sobre o objeto em questão, ou com as informações mostradas na imagem.

3. Configuração esquemática da arquitetura proposta

A arquitetura proposta é vista na Figura 1. Nela há um aplicativo para a ontologia, um para associar metadados às imagens (descritores semânticos de segmentos da imagem e um mapa semântico da imagem), e um para organizar os dados internos dos agentes (lexema, tipo, tema, etc.). Essas informações alimentam um simulador multiagentes, um *reasoner* da ontologia, e um aplicativo que interpreta os metadados. Os aplicativos são integrados na retaguarda da interface do curso, com a qual os alunos interagem.

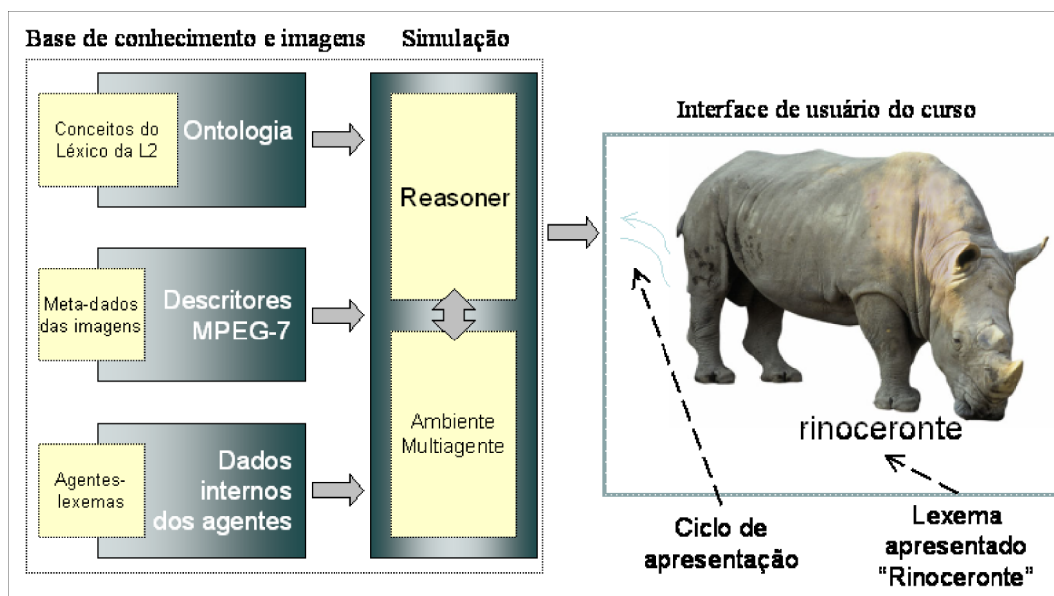


Figura 1. Arquitetura proposta para o ILTS

4. Conclusão

A arquitetura ILTS aqui proposta tem características interessantes para o ensino de uma língua. Em primeiro lugar, ela combina a capacidade de gestão dos conteúdos lexicais em seu nível elementar a recursos de alto nível que permitem direcionar o processo de ensino para atender necessidades e temas específicos. Em segundo lugar, ela facilita o acompanhamento do aprendizado ao garantir aos tutores humanos acesso direto ao ciclo de vida instrucional de cada lexema. Por fim, a flexibilidade da abordagem *bottom-up* possibilita à aplicação lidar com a complexidade da área de ensino de línguas.

A estratégia automática de sequenciamento de lexemas e formação de sentenças significa, em última análise, que o ensino se baseará na exposição gradativa do aluno a elementos da L2, em contextos e sequências que façam sentido na lógica interna da L2 e que estejam fortemente ancorados em informações do ambiente. Isso busca reproduzir o primeiro aprendizado linguístico, a aquisição da língua materna que, segundo Bloom (2000), é um processo gradual, contextualizado e em grande medida inconsciente.

Os três pilares da arquitetura, agentes, metadados e ontologias, são também os conceitos mais fundamentais da Web Semântica. Isso significa que, num certo sentido, a arquitetura aqui proposta incorpora, em um âmbito mais restrito, muitos dos elementos que norteiam a evolução da Internet, sendo assim uma aplicação que testa em pequena escala ideias que podem vir a ser expandidas para a Web no futuro.

Referências

- Baars, B.J. (1997), *In the theater of consciousness: The Workspace of the Mind*. Oxford University Press, 1st edition.
- Bloom, P. (2000), *How Children Learn the Meanings of Words*, MIT Press, 1st edition.
- Dubois, D. (2007), *Constructing an Agent Equipped with an Artificial Consciousness: Application to an ITS*. PhD Thesis at the University of Quebec, Montreal.
- Protégé: <http://protege.stanford.edu/>