

Uma proposta para extração de perguntas e respostas de textos

Everton Moschen Bada

UFES
Brasil
ebada@inf.ufes.br

Crediné Silva de Menezes

Universidade Federal do Espírito Santo
Brasil
credine@gmail.com.br

ABSTRACT

This paper presents a proposal for construction of Personal Assistants (PA) to support the clarification of doubts. Clarifying questions will be based on knowledge produced by a user, which will be extracted using techniques from natural language processing and will feed the knowledge bases of the PA. Based on this knowledge the PA can perform clarify questions about the matters described in the user's productions. Clarifying the doubts is driven by a chatterbot like Alicebot, and the questions that chatterbot can not answer are stored in a database of questions, which could be resolved with the support of the ontology or knowledge base.

RESUMO

Este trabalho apresenta uma proposta para construção de Assistentes Pessoais (AP) para dar apoio ao esclarecimento de dúvidas. O esclarecimento de dúvidas será baseado no conhecimento produzido por um usuário, que será extraído utilizando técnicas de processamento de linguagem natural e alimentará as bases de conhecimento do AP. Com base neste conhecimento o AP poderá realizar o esclarecimento de dúvidas sobre os assuntos descritos nas produções do usuário. O esclarecimento das dúvidas é conduzido por um chatterbot do tipo Alicebot, sendo que as perguntas que o chatterbot não souber responder são armazenadas em uma base de dúvidas, que poderão ser solucionadas com o apoio da ontologia ou da base de conhecimento..

PALAVRAS-CHAVE

Agentes Inteligentes, Chatterbots, de Texto para AIML, Prolog, Ontologia.

INTRODUÇÃO

Com a crescente utilização de ambientes virtuais para apoiar a aprendizagem, as interações se tornaram constantes e cada vez mais intensas. A utilização das ferramentas de autoria providas pelos ambientes, além de incentivar a produção de textos, também auxilia no registro da aprendizagem. Desta forma, muitos documentos ficam disponíveis nestes ambientes, gerados pelos indivíduos e de forma cooperativa, resultante das interações. Com isso, um acervo valioso de conhecimento fica registrado o qual poderia ser usado para ajudar outros usuários que estejam em busca de respostas, sem que seja necessária a

leitura exaustiva de cada documento do acervo ou mesmo utilizar as convencionais máquinas de busca.

Vale ressaltar que a opção pelo uso de máquina de busca para esta finalidade, implica no uso de palavras chave e apresenta diversos inconvenientes. Nesta alternativa, a interação homem-máquina é dispendiosa, pois não é natural para os usuários essa forma de comunicação. Além disso, as máquinas de busca não produzem a resposta a uma pergunta, ao invés disso, retornam uma seleção de textos que podem possuir a resposta. Assim, ainda é necessário realizar a leitura dos documentos selecionados, para encontrar a resposta à dúvida.

Segundo Teixeira [14] muitos estudantes, da educação a distância, desestimulam-se pela falta de interações síncronas (presenciais ou não), devido ao grande volume de informações a serem estudadas. Como também muitos estudantes tem a sensação de isolamento por esta falta de interação [7].

O Projeto MOrFEu busca a produção de ambientes inteligentes para mediação na web de atividades cooperativas, com ênfase no apoio à aprendizagem e à gestão do conhecimento. No presente momento estamos desenvolvendo um ambiente voltado para a descrição e a realização de espaços virtuais interativos, com ênfase na gerência das produções individuais e na organização flexível desses espaços. Esta concepção está pautada pelos seguintes elementos: plasticidade, ergonomia e redução da sobrecarga cognitiva, onde as noções clássicas de veículos de comunicação possam ser definidas pelos seus próprios usuários.

Segundo esta nova concepção, qualquer ferramenta para apoio à produção cooperativa de artefatos digitais na web, seja ela síncrona ou assíncrona, será modelada através de veículos de comunicação [8], [10], que comportam as Unidades de Produção Intelectual (UPI). Logo, qualquer produção que o usuário fizer será uma UPI e esta será armazenada na sua biblioteca de UPIs. Estas podem ser versionadas e reutilizadas em diferentes VComs [6],[11].

O projeto MOrFEu também visa prover ferramentas inteligentes que forneçam serviços que auxiliem o usuário em atividades dentro do ambiente, com o intuito de diminuir a sobrecarga cognitiva.

Este artigo inicia apresentando alguns sistemas de esclarecimento de dúvidas, destacando o modo como eles trabalham e suas fragilidades. Posteriormente apresentamos trabalhos correlatos, que tratam da questão por nós endereçada, usando a tecnologia

de chatterbot, que é a mesma escolhida por nós. A seguir apresentamos a arquitetura e o funcionamento do nosso trabalho que está pautado na construção semiautomática da base de conhecimento para chatterbots. Na sequência trazemos a seção que apresenta os padrões morfossintáticos, que é utilizada em nosso trabalho para encontrar as estruturas do conhecimento. Prosseguindo apresentamos a visão geral da proposta, trazendo a metodologia, a arquitetura do sistema e as tecnologias utilizadas para criação de um protótipo. Por fim apresentamos uma avaliação preliminar dos resultados obtidos e as conclusões.

SISTEMA DE ESCLARECIMENTO DE DÚVIDAS

O sistema de esclarecimento de dúvidas mais conhecido é denominado como Perguntas mais Frequentes, ou como é mais citado, Frequently Asked Question (FAQ). Neste mecanismo são cadastradas as dúvidas mais frequentes em um contexto e suas respectivas respostas. Neste tipo de sistema de esclarecimento, as perguntas geralmente são apresentadas em forma de lista. A pessoa que tiver com dúvida, tenta encaixá-la em uma das cadastradas. Ao selecionar uma, clicando sobre ela, o usuário tem acesso à uma resposta, cabendo-lhe a tarefa de verificar se a resposta resolve sua dúvida.

A Internet por se tratar de um local onde a distância não é o limite, tornou-se um local que possui diversos usuários, dispostos a ajudar, com diferentes níveis de conhecimento. Por esse motivo, os fóruns de esclarecimento de dúvidas, tornaram-se populares e bastante utilizados, pois, geralmente, atingem seus objetivos, ou seja, o esclarecimento das dúvidas dos usuários. Entretanto, dependendo do ambiente, este tipo de mecanismo não se mostra muito eficiente, pois muitos potenciais usuários tem vergonha de postar uma dúvida que seja muito banal, temendo a crítica e a chacota dos outros usuários. Essa situação ocorre bastante em fóruns presentes em Ambientes Virtuais de Aprendizagem (AVA), acarretando ao estudante o prejuízo de continuar com a dúvida, e que, com isso, não consiga avançar mais em seu aprendizado.

Em [2] é abordado o problema da falta de tempo de membros de comunidades de esclarecimento de duvidas, que embora desejem ajudar, encontram dificuldade em identificar as perguntas que eles tem conhecimento para responder. Com isso, muitas perguntas ficam sem resposta e o sistema tende a cair na reputação de seus potenciais usuários. Como solução a esta deficiência, foi apresentado em [2] o sistema QSabe para dar suporte a criação e gerenciamento de uma comunidade de esclarecimento de dúvidas. O ambiente visa dar suporte a comunidade de esclarecimento de dúvidas, roteando perguntas para pessoas que tem a maior probabilidade de saber respondê-la, baseado nos perfis dos usuários e dos a seus históricos de participação no ambiente. O sistema possui várias ferramentas para auxiliar na organização das perguntas no sistema.

Uma classe de sistemas que se tornou bastante popular, para tarefas relacionadas ao esclarecimento de dúvidas, foi a dos chatterbots. Os chatterbots são programas de computador orientados à conversação com usuários. Estes programas surgiram inicialmente para tentar responder ao questionamento levantado

por Alan Turing, *As máquinas podem pensar?*[15] A partir da tentativa de gerar resposta para questionamento esta categoria de softwares foi evoluindo cada vez mais, chegando ao modelo mais utilizado atualmente, baseado na tecnologia ALICEbot e o padrão AIML, para representação do conhecimento.

Tanto nos chatterbots pioneiros quanto nos atuais, baseados em AIML, a base de conhecimento é criada e mantida pelo gestor do chatterbot (*Botmaster*), uma pessoa que, manualmente, insere o conhecimento na base. Como no modelo AIML o conhecimento é representado da forma (<pattern><template>), ou seja, o criador do chatterbot tem que pensar nas interações que o chatterbot vai ser capaz de responder, e desta maneira inserir na base. Este é um dos motivos pelos quais os chatterbots são limitados a responder sobre área de conhecimento bastante restritas.

ALICEBOT E AIML

A base de conhecimento de um chatterbot, criado com a tecnologia ALICE, é escrita no padrão AIML que possui algumas características. O padrão AIML utiliza-se do XML características. O padrão AIML utiliza-se do XML (*eXtensible Markup Language*), e especifica as tags que são interpretadas pelo chatterbot.

O padrão AIML funciona basicamente informando ao agente os padrões que o agente deve encontrar nas sentenças de entrada, e a respectiva resposta para este padrão. Os padrões e as respostas são representadas pelas tags <pattern> e <template>, respectivamente. Partindo disto teremos várias *tags* que visam fornecer facilidades ao *botmaster* (pessoa responsável por alimentar a base de conhecimento do chatterbot). Estas tags podem ser melhor estudadas em [17].

Este tipo de base de conhecimento para os chatterbots grande parte das vezes é alimentada pelo *botmaster* que insere o conhecimento na base, pensando nas perguntas e nas possíveis respostas.

Com base na forma como o conhecimento é organizado na base de conhecimento AIML é fácil ver como funciona uma conversa com um chatterbot da tecnologia ALICE. Toda sentença de entrada é analisada pelo chatterbot procurando na base AIML um *pattern* correspondente à sentença de entrada, ao encontrar ele retorna a respectiva resposta (<template>).

TRABALHOS CORRELATOS

A utilização de chatterbots para responder perguntas sobre um determinado assunto tornou-se bastante popular em sistemas que tentam prover uma interação mais natural com os usuários. Por exemplo a Conpet em parceria com a Petrobras criou o robô Ed que tem o objetivo de esclarecer dúvidas sobre meio ambiente e energias renováveis, além de incentivar o uso destas [3].

Com o mesmo propósito de esclarecer dúvidas, a utilização de chatterbots em ambientes educacionais também está se tornando frequente, podemos encontrar alguns chatterbots em português do Brasil tais como: Elektra [7], Tuxbot [14], Maria [12], dentre outros.



Em geral, neste tipo de sistema não ocorre o aprendizado a partir de perguntas para as quais o sistema não possui uma resposta apropriada. A base de perguntas e respostas é incrementada de forma manual, muitas das vezes não acompanhando as dúvidas e as expectativas dos usuários.

Teixeira [14], em seu trabalho, apresenta uma proposta para construção de bases de conhecimento para chatterbots, que utilizam a tecnologia Alicebot e o Padrão AIML. Este autor propôs uma metodologia que visa recolher as unidades de conhecimento de fóruns de esclarecimento de dúvidas, afim de alimentar a base de conhecimento do chatterbot.

A estratégia, segundo Teixeira, “[...]está apoiada na modelagem do conhecimento através da recuperação de padrões de conversação baseado na linguística e Corpus, desses padrões com a utilização do modelo vetorial ou do método Latent Semantic Indexing (LSI) e na conversão supervisionada desses padrões em conhecimento de um chatterbot.”

Como estudo de caso, Teixeira criou o tuxBot, chatterbot para esclarecimento de dúvidas sobre o sistema operacional GNU/Linux [14]. Um aspecto interessante deste trabalho, foi a semi-automatização do processo de construção da base de conhecimento do chatterbot, utilizando uma abordagem eficaz neste processo.

Entretanto, a aprendizagem deste chatterbot proposto ocorre apenas no início do processo, quando a base de conhecimento é produzida. O que limita, em partes, o crescimento da base de conhecimento do chatterbot.

Em [5] Gasperis, apresenta uma metodologia para criação das bases de conhecimento de chatterbots a partir de FAQs e Glossários. A metodologia visa encontrar palavras-chave das perguntas para poder organizá-las em classes e ramos, para melhor realizar a procura e assim obter uma resposta mais eficiente. Já nos glossários, a metodologia visa encontrar os termos e suas respectivas definições, com isso criando um corpora dos termos específicos do domínio ao qual o chatterbot tem conhecimento para responder questões.

Em [13], Schumaker fez a análise de métodos de aquisição de conhecimento em massa. Em sua análise ele apresenta algumas vantagens e desvantagens das abordagens, e destaca que cada abordagem pode ser mais apropriada para uma situação específica. Não é possível classificar a melhor ou pior abordagem em geral, somente podemos analisar se pensarmos no problema que ela deseja atacar.

Amorim [1] apresenta um Sistema Inteligente baseado em Ontologia para apoio ao Esclarecimento de Dúvidas. Este trabalho visa fornecer esclarecimento de dúvidas em determinado domínio de conhecimento apoiado por uma ontologia. Ao receber uma pergunta o sistema através do chatterbot AIML procura uma resposta na base AIML, caso não encontre uma resposta, ele procura resposta na Ontologia ou na Web, transformando o conhecimento recolhido da ontologia ou da Web, em uma resposta,

respondendo ao questionamento do usuário, e registrando também o conhecimento na base AIML. Como estudo de caso Amorim propõe a construção de um helpdesk inteligente, esclarecendo dúvidas dos usuários, referente à sistemas operacionais.

Construção Semiautomática de Bases de Conhecimento para Chatterbots

Partimos do pressuposto que as dúvidas são expressas através de perguntas, com o intuito de receber uma resposta que promova o seu esclarecimento. Neste sentido, consideramos que a estrutura de armazenamento do conhecimento na tecnologia AIML é organizada em tuplas que representam:

- o padrão que o chatterbot vai procurar nas interações do usuário associado (pattern) e,
- a resposta que o sistema deverá oferecer a este padrão (template).

Nossa proposta está pautada em resolver o problema de extrair conhecimento, a partir de textos em Língua Portuguesa do Brasil, necessário para popular a base de conhecimento de um chatterbot, que terá a função de auxiliar no esclarecimento de dúvidas.

Deste modo precisamos descobrir o conhecimento que conseguimos extrair destes textos, e a seguir, identificar as possíveis perguntas que podem ser formuladas a este corpo de conhecimento.

Pesquisa por Padrões Morfossintáticos

Com o intuito de responder ao questionamento levantado anteriormente, propomos a pesquisa por padrões morfossintáticos, para identificar as estruturas do conhecimento do texto. E deste modo conseguir extrair as informações necessárias para preenchimento das bases de conhecimento do chatterbot.

Os textos, escritos em Língua Portuguesa do Brasil, possuem um conjunto de estruturas válidas, denominada estrutura sintática. Estas estruturas ajudam a definir a semântica dos textos, entre os elementos que esta estrutura determina, temos: sujeito, predicado, aposto, objetos (diretos e indiretos), etc.

Aliado à estrutura sintática, também temos a morfologia das palavras que compõem o texto, e esta define a classe gramatical das palavras, tais como: substantivo, verbo, adjetivo, artigo, etc.

Propomos então utilizar-se da pesquisa por padrões que integrem estes dois elementos, no intuito de encontrar as perguntas que um texto pode responder, e quais as respectivas respostas.

VISÃO GERAL DA ABORDAGEM

Visamos prover um Assistente Pessoal para cada usuário do MORFEu, com o intuito de auxiliar outros usuários no esclarecimento de dúvidas suas dúvidas, em relação ao conhecimento produzido por ele no ambiente. Para isto visamos a construção automática de conhecimento, necessário ao esclarecimento de dúvidas, a partir de textos em Língua Portuguesa do Brasil. Nossa abordagem busca resguardar a fidelidade ao conhecimento produzido pelo usuário. Cada Assistente Pessoal poderá responder, em nome de

seu proprietário, às dúvidas formuladas a ele por outros usuários.

Metodologia da Pesquisa

Inicialmente realizamos um levantamento das propostas para sistemas de esclarecimento de dúvidas mais promissoras, identificando suas vantagens e desvantagens. Percebeu-se então que a tecnologia de chatterbot era a que mais se aproxima de nossos objetivos.

Em um passo posterior, estabelecemos os requisitos funcionais do sistema que desejávamos construir, focando principalmente no diferencial que deseja vamos oferecer.

Este diferencial está associado à forma em que a construção do conhecimento destes sistemas é normalmente realizada, pois na maioria das vezes é realizada de maneira manual, ou automatizado apenas inicialmente, mas entretanto não havia métodos de aprendizado contínuo.

Partimos então para elaboração de um modelo de aprendizagem para os Assistentes Pessoais, com a intenção de fornecer a cada usuário do MOrFEu um assistente pessoal para auxiliá-lo em tarefas de esclarecimento de dúvidas para outros usuários.

Para atender aos requisitos definidos no projeto, o Assistente Pessoal deve aprender apenas com o que o que é produzido por seu proprietário, visto que o assistente só deve responder sobre o que usuário-dono manifestou em suas produções.

Proposta

Como proposta de solução ao desafio apresentado anteriormente, apresentamos a seguinte arquitetura composta por agentes inteligentes, que utilizam técnicas de processamento de Linguagem Natural para extrair conhecimento dos textos.

A proposta de solução foi organizada em etapas, de forma que cada etapa é realizada por agentes inteligentes, com exceção da etapa da conversação que é realizada por um chatterbot, utilizando a tecnologia AliceBot. A utilização de agentes inteligentes justifica-se pela complexidade da proposta, necessitando uma modularização do sistema.

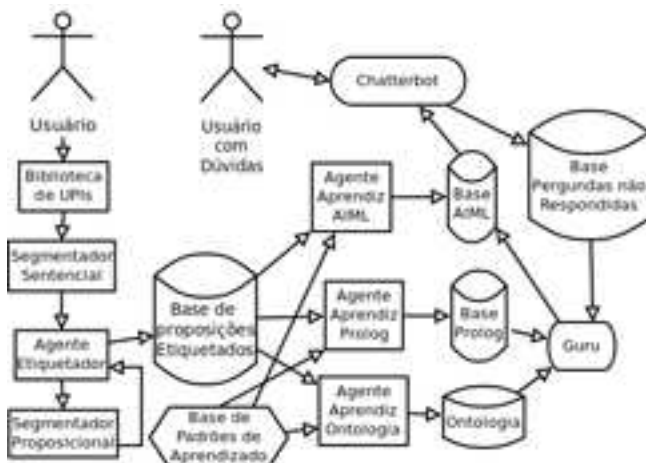


Figura 1: Arquitetura do Assistente Pessoal para Esclarecimento de Dúvidas

A modularização pode ser melhor compreendida analisando a

Figura 1 que sintetiza o nosso conceito de Assistente Pessoal (AP), o qual, na verdade, é um sistema multiagentes composto por um chatterbot e por sete agentes. Um Agente Segmentador Sentencial, um Agente Etiquetador, um Agente Segmentador Proposicional, um Agente Aprendiz AIML, um agente Aprendiz Prolog, um Agente Aprendiz Ontologia e um Agente Guru. Cada um destes componentes possui uma função, como:

- **Agente Segmentador Sentencial:** Este agente é responsável por varrer a biblioteca de UPIs de seu dono e segmentar as UPIs em sentenças, utilizando o SENTER. As frases recolhidas da biblioteca de UPIs são então repassadas ao Agente Etiquetador.

- **Agente Etiquetador:** O Agente Etiquetador determina para cada palavra as etiquetas morfológicas e sintáticas correspondentes, utilizando o VISL [16]. A etiqueta morfológica identifica um: substantivo, verbo, pronome, preposição, advérbio, conjunção ou artigo. Já as etiquetas sintáticas determinam a função sintática das palavras na frase, tais como: sujeito, objeto direto, etc. Então após esta análise é gerada uma base de arquivos com as sentenças etiquetadas.

- **Agente Segmentador Proposicional:** Este agente é responsável por dividir as sentenças etiquetadas, em proposições simples. Com o intuito de prover maior eficiência no processo de extração do conhecimento.

- **Agente Aprendiz AIML:** É responsável por transformar as Proposições Etiquetadas, em conhecimento no formato AIML para alimentar o chatterbot dando a este a possibilidade realizar conversas.

- **Agente Aprendiz Prolog:** É responsável por recolher o conhecimento presente nas Proposições Etiquetadas, e alimentar uma base Prolog.

- **Agente Aprendiz Ontologia:** É responsável por recolher o conhecimento presente nas Proposições Etiquetadas e representar em forma de ontologia, utilizando o RDF.

- **Agente Guru:** É responsável por tentar sanar as dúvidas sem respostas, geradas pelas conversas entre o Chatterbot e os usuários, quando o usuário faz uma pergunta que o chatterbot não tem resposta na base AIML. Este agente analisa estas perguntas e tenta encontrar uma resposta na Base Prolog e na Ontologia. Para pesquisar na base Prolog utilizamos o PCI_Prolog [9], que nos possibilita realizar consultas na base Prolog a partir de um programa em Java. Para acessar a ontologia utilizamos o JENA. Caso encontre uma resposta, esta é inserida na base AIML, para que numa próxima conversa o chatterbot consiga responder. Caso contrário, essa dúvida é marcada, para que seja sanada com o dono do AP.

- **Chatterbot:** Este é responsável por responder às perguntas do seu dono (usuário) e também de outros usuários do sistema que tenham interesse no conhecimento do dono do AP. Como também, verificar que perguntas da base de dúvidas sem respostas estão marcadas para serem sanadas com o dono.

A estratégia de extração de conhecimento através dos padrões

de aprendizado, utiliza-se das estruturas sintáticas e das análises morfológicas. Para melhor entender vamos considerar a seguinte proposição: “Paulo vai ao parque com Ana.” Esta proposição tem informação suficiente para responder, dentre outras, pelo menos as seguintes perguntas:

- Questão 1: Paulo vai ao parque com Ana?
 - Resposta: Sim
- Questão 2: Quem vai ao parque com Ana?
 - Resposta: Paulo

Para encontrar a proposição apresentada, é necessário o padrão, utilizando a nomenclatura descrita em [16],

```
<hum> N @SUBJ>; V [ir];$*,
```

sendo que o ponto e vírgula separa as etiquetas de cada palavra, e o \$* indica qualquer coisa. Com isso é possível encontrar aquela sentença.

Sintaxe para pesquisa de padrões morfossintáticos

A sintaxe utilizada para pesquisa por padrões morfossintáticos foi criada neste projeto com a finalidade de atender as necessidades, entretanto a mesma não está completamente desenvolvida, podendo ser evoluir ao longo do tempo para atender novas demandas.

A criação da linguagem para a realização da pesquisa por padrões foi baseada na sintaxe do sistema etiquetador morfossintático utilizado no protótipo, denominado VISL [16].

Criamos o sistema desta maneira tentando não vincular a abordagem a um sistema específico de análise morfossintática, possibilitando a utilização do analisador que melhor atenda as necessidades de utilização.

A linguagem foi pensada para possibilitar uma definição bem específica do padrão a ser pesquisado, podendo utilizar diversas etiquetas que definam as características das palavras, como por exemplo: Classe gramatical, gênero, morfologia, etc. As descrições entre as palavras são separadas pelo sinal de “ponto e vírgula”.

Toda linha do padrão é composta por três partes necessárias a extração do conhecimento do texto, sendo elas: Padrão morfossintático, ação, parâmetros. Deste modo, o agente inteligente saberá que ao encontrar o padrão morfossintático, deverá executar uma ação, seguindo os parâmetros apresentados no fim da regra.

Extraindo conhecimento baseado em padrões morfossintáticos

Possuindo tais informações, podemos criar as perguntas e as respostas, por exemplo na Questão 1 anteriormente apresentada, bastou pegar a proposição e acrescentar um sinal de interrogação no fim, e a resposta é “sim”. Para a segunda questão basta substituir o sujeito da frase pelo pronome relativo “Quem”, e acrescentar a interrogação no fim da frase, sendo a resposta o sujeito da frase.

Estas perguntas são bem utilizadas, a primeira sempre exigirá resposta afirmativa, cujas respostas serão sim/não. Na segunda

a pergunta tem o objetivo de descobrir quem praticou a ação, sendo o sujeito da proposição a resposta.

No mesmo sentido, ocorre a extração de conhecimento, utilizando padrões. A utilização da base Prolog para responder à perguntas que necessitam de informações presentes em diferentes bases proposições, e que foram sendo adquiridas

gradativamente ao logo do tempo. Por exemplo, imaginemos que tenhamos as seguintes proposições: 1) Computador possui memória; 2) Computador possui processador; 3)Computador possui HD. Cada uma delas será tratada individualmente, por se tratarem de proposições distintas.

No caso da base AIML, será criada um par <pergunta, resposta> por proposição. Desta forma, caso o AP receba uma pergunta do tipo “Que partes tem o computador?”, a base AIML não conseguiria responder a esta questão, com isso seria então recorrido à base Prolog para resolução da pergunta, que por exemplo, criaria uma resposta, com todos os componentes listados acima.

Neste sentido, as técnicas de extração também são válidas para a construção da Ontologia. Propomos também a utilização de ontologias para representação do conhecimento do AP, com o objetivo de utilizá-la para inserir etiquetas semânticas nas palavras que compõem as proposições, de modo a aumentar a expressividade da análise morfossintática. Como pretendemos inserir etiquetas semânticas nos textos, identificamos que a maioria das etiquetas semânticas estão relacionadas a hiperonímia e hiponímia. Neste sentido propomos utilizar as estratégias descritas no trabalho de [4], para identificação de hiperônimos e hipônimos.

AVALIAÇÃO DOS RESULTADOS

Nosso protótipo foi desenvolvido utilizando várias tecnologias, como apresentado nas seções anteriores. A linguagem de programação utilizada foi Java, utilizando o Framework JADE, para prover toda a camada necessária para programarmos orientado a agentes. Identificamos que essa abordagem de extração do conhecimento mostrou-se interessante, pois para aumentar a capacidade de aprendizado do sistema basta inserir mais Padrões de Aprendizado para que o sistema consiga tratar novos tipos de proposições.

Por questão de espaço não é possível apresentar o modo como realizamos a extração para todos os tipos de base, mas apresentamos a seguir alguns exemplos que conseguimos extrair de alguns textos. Na figura 2, apresenta uma unidade de

```
<category>  
  <pattern>0 que são softwares autônomos que tem a capacidade aprender!</pattern>  
  <template>Agentes Inteligentes</template>  
</category>
```

Figura 2: Unidade de Conhecimento AIML, extraído de texto

conhecimento em AIML, apresentando uma possibilidade de extração, não se limitando a esta.

