

Uma Revisão Teórica sobre Classificação Textual de Atos de Fala para Atores Virtuais Utilizando Aprendizagem de Máquina

Dino Magri¹, Rogério Eduardo da Silva¹, Claudio Cesar de Sa¹, Milton Heinen¹

¹Departamento de Ciência da Computação
Universidade do Estado de Santa Catarina (UDESC)
Caixa Postal 631 – 89217-301 – Joinville – SC – Brasil

dinomagri@gmail.com, {rsilva,claudio,miltonh}@joinville.udesc.br

Abstract. *This research aims at contributing to the creation of an autonomous digital actor. An important step is the automatic interpretation of script via text classification. Therefore, this paper presents a study for text classification of speech acts using natural language processing and machine learning methods. Each text can be classified into one of five possible classes of illocutionary acts using reinforcement learning that enables learning from criticism. The proposed system defines the use of lexical, syntactic and semantic analysis in the natural language process. As the scope and test set we opt to use the utterances of an animated film. We expect with this approach, that the automatic classification would present a better accuracy than purely random methods when compared to manual classification.*

Resumo. *Esta pesquisa visa contribuir na criação de um ator digital autônomo. Para tal, um importante passo é a interpretação automática de scripts através de classificação de textos. Sendo assim, este trabalho apresenta um estudo para a classificação textual de atos de fala utilizando processamento de linguagem natural e métodos de aprendizagem de máquina. Cada texto pode ser classificado em uma de cinco possíveis classes dos atos ilocutórios utilizando aprendizagem por reforço que possibilita aprender com as críticas. O sistema proposto define a utilização da análise léxica, sintática e semântica no processamento da linguagem natural. Para o escopo e o conjunto de testes escolheu-se utilizar as falas de um filme de animação. Espera-se que com esta abordagem, a classificação automática tenha uma melhor precisão do que o puramente aleatório, quando comparados à classificação manual.*

1. Introdução

Humanos Virtuais são modelos computacionais dedicados a reproduzir certos aspectos do comportamento humano, como por exemplo, expressão de emoções, tomada autônoma de decisão, entre outros. As possíveis aplicações são robótica, entretenimento digital, simulação, etc.

Este trabalho faz parte de um projeto que tem por objetivo a criação de um ator digital autônomo (ADA), no qual um roteiro é submetido a um conjunto de atores que produzem sugestões individuais para atuação de seus respectivos personagens, as quais serão posteriormente combinados em uma *timeline* não-linear a ser submetida à uma ferramenta

de animação, na qual o usuário pode então visualizar e alterar o resultado em um processo iterativo. Uma das habilidades fundamentais de um ADA é a interpretação autônoma de roteiros. Além do uso em ADA's, é possível utilizá-lo em aplicações de Processamento de Linguagem Natural, como por exemplo, tradução automática, interpretação de linguagem natural, correção ortográfica, dataming entre outros.

Em geral, os atos de fala são atos de comunicação. Para comunicar, é preciso expressar uma certa atitude, e o tipo de ato de fala que está sendo executado corresponde ao tipo da atitude expressada [Tsohatzidis 1994]. Existem técnicas computacionais que possibilitam diversos tipos de análise contextual, extraíndo informações que podem auxiliar na classificação automática de sentenças em atos de fala. Além disso, pode-se utilizar métodos que permitem à ferramenta aprender. A hipótese deste trabalho é que a abordagem de classificação automática tenha um melhor desempenho do que o puramente aleatório em relação à classificação manual.

1.1. Trabalhos Correlacionados

Pesquisas de reconhecimento de atos de fala em textos focam em descobrir e utilizar palavras-chaves. Esta é a forma mais intuitiva de reconhecê-los em um texto e vários métodos já propõem a utilização de palavras-chave. Em [Tao and Tan 2004], dividiu-se as palavras em palavras de conteúdo e palavras funcionais de emoção, ao invés de utilizarem as palavras-chave emocionais. Em [Seol et al. 2008] definiu-se um sistema híbrido, onde utilizam a classificação por palavras-chave e um método de aprendizagem de máquina, o KBANN (*Knowledge-based Artificial Neural Network*). E em [Cohen et al. 2004] apresentou-se o uso do método de classificação de texto para detectar atos de fala em email.

1.2. Organização do Trabalho

A Seção 1 apresenta o problema, a proposta de solução, trabalhos correlacionados e a estrutura deste trabalho. A fundamentação teórica é apresentada na Seção 2. Uma possível solução para o problema descrito e os critérios de avaliação estão na Seção 3. E por fim, as conclusões são descritas na Seção 4.

2. Fundamentação Teórica

Esta seção define os atos de fala, os níveis de processamento de linguagem natural utilizados e o modelo de classificação.

2.1. Atos de fala

Um ato de fala (*Speech acts*) é o ato que o locutor emite quando pronuncia uma expressão, que inclui o modo de proferir uma palavra, fazendo uma referência e/ou uma previsão, com uma intenção para produzir o enunciado. Como um ato de comunicação, um ato de fala é bem sucedido se o ouvinte identifica, através da intenção do locutor, a atitude que está sendo expressada [Tsohatzidis 1994].

Searle [Searle 1969] define que atos de fala dividem-se em dois segmentos interdependentes: a força ilocutória e conteúdo proposicional. Na frase “*I assert that John was here yesterday*”, a parte “*I assert that*” refere-se a força ilocutória F e “*John was here yesterday*” ao conteúdo proposicional p . Assim, o ato ilocutório é simbolizado por:

$$F(p)$$

Searle [Searle 1976] define doze dimensões significativas de variação em que os atos ilocutórios diferem uns dos outros e constrói sua taxonomia em torno de três dimensões, *ponto ilocutório* que difere no ponto do tipo de ato de acordo com a condição essencial de cada ato de fala, a *direção de ajuste* é definida pelas diferenças na direção de ajuste entre palavras e o mundo. Alguns atos de fala tentam obter palavras (conteúdo proposicional) que coincidam com o mundo (como no caso das descrições) enquanto outros tentam fazer o mundo coincidir com as palavras (como no caso de promessas e pedidos) e por fim, a *condição de sinceridade* que define as diferenças nos estados psicológicos expressados. O estado psicológico expressado é simbolizado com as letras iniciais dos verbos correspondentes, B para *Belief*, I para *Intention*, D para *Desire*, etc.

Com base na teoria de [Austin and Urmson 1978], [Searle 1976] procedeu a divisão e a classificação dos atos ilocutórios em cinco classes: *assertivo*, *diretivo*, *compromissivo*, *expressivo* e *declarativo*.

A classe *assertivo* determina a dimensão da avaliação, no qual inclui verdadeiro e falso. Este conceito é formalizado por:

$$\vdash \downarrow B(p),$$

leia-se: afirmação, direção de ajuste (\downarrow) palavras-ao-mundo, o estado psicológico expressado é *Belief* e o conteúdo proposicional (p). Já a classe *diretivo* consiste no fato do locutor persuadir o ouvinte a fazer algo. Os questionamentos são uma sub-classe dos diretivos. Este conceito é formalizado por

$$! \uparrow W(H \text{ does } A),$$

leia-se: diretivo, a direção de ajuste é mundo-a-palavras (\uparrow), o estado psicológico expressado é *Will* e conteúdo proposicional atribui um futuro ato A ao ouvinte H . *Compromissivo* é a classe em que o locutor denota a intenção de realizar uma determinada ação no futuro e é formalizado por:

$$C \uparrow I(S \text{ does } A),$$

leia-se: compromissivo, a direção de ajuste é mundo-a-palavras (\uparrow), o estado psicológico expressado é *Intention* e o conteúdo proposicional atribui uma ação futura A ao locutor S . A classe *expressivo* apresenta o estado psicológico expressado na *condição de sinceridade* sobre um estado (sentimentos ou emoções) especificado no conteúdo proposicional. Este conceito é formalizado por:

$$E \phi (P) (S/H + \text{property}),$$

leia-se: expressivo, a direção de ajuste é nula ϕ , o P indica os diversos estados psicológicos expressados e o conteúdo proposicional atribui alguma propriedade (não necessariamente uma ação) para S ou H . A propriedade especificada no conteúdo proposicional de um ato expressivo deve estar relacionado com S ou H . E por fim a classe *declarativo*, que muda a realidade de acordo com o conteúdo proposicional, por exemplo, se eu executar com êxito o ato de nomear você para presidente, então você é presidente. Este conceito é formalizado por:

$$D \updownarrow \phi (P),$$

leia-se: declarativo, a direção de ajuste é para ambos os lados (\updownarrow), a condição de sinceridade é nula (ϕ) e o conteúdo proposicional é expressado por P .

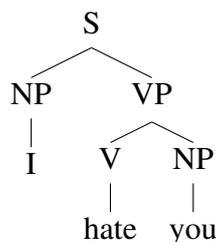
2.2. Análise Léxica e Sintática

Na gramática, as partes do discurso (*parts of speech* ou *POS*) são categorias linguísticas das palavras, as quais, geralmente, são definidas pelo comportamento sintático ou morfológico do item léxico em questão. Nesta etapa as palavras são analisadas em termos de seus componentes e os sinais são separados das palavras.

POS tagging é o processo de marcação das palavras em um texto (*corpus*) correspondendo a uma determinada parte do discurso, baseada na sua definição, bem como em seu contexto. Utilizando o conjunto de ferramentas NLTK¹, a frase *I hate you!!*, pode ser classificada com as seguintes tags:

‘I’ → ‘PRO’, ‘hate’ → ‘VBP’, ‘you’ → ‘PRO’, ‘!’ → ‘.’, ‘!’ → ‘.’

Através da gramática da linguagem a ser analisada e das informações do analisador morfológico, constrói-se árvores de derivação para cada sentença, mostrando como as palavras estão relacionadas entre si. Para a frase *I hate you*, a árvore sintática construída através de uma gramática livre de contexto, apresentada em (1), é mostrada abaixo.

$$\begin{aligned}
 S &\rightarrow \langle NP \rangle \langle VP \rangle \\
 VP &\rightarrow \langle V \rangle \langle NP \rangle \mid \langle V \rangle \langle NP \rangle \langle PP \rangle \\
 PP &\rightarrow \langle P \rangle \langle NP \rangle \\
 V &\rightarrow \text{saw} \mid \text{hate} \mid \text{ate} \mid \text{walked} \\
 NP &\rightarrow \text{John} \mid \text{Mary} \mid \text{Bob} \mid \text{I} \mid \text{you} \mid \langle Det \rangle \langle N \rangle \mid \langle Det \rangle \langle N \rangle \langle PP \rangle \\
 Det &\rightarrow \text{a} \mid \text{an} \mid \text{the} \mid \text{my} \\
 N &\rightarrow \text{man} \mid \text{dog} \mid \text{cat} \mid \text{park} \mid \text{telescope} \\
 P &\rightarrow \text{in} \mid \text{on} \mid \text{by} \mid \text{with}
 \end{aligned}
 \tag{1}$$


A produção da análise sintática de uma frase não conclui o processo de compreensão. Ainda é preciso produzir uma representação do significado da frase.

2.3. Análise Semântica

De acordo com a estrutura sintática de uma sentença, é possível estabelecer uma representação lógica correspondente, onde o verbo indica uma relação entre os argumentos expressados por sujeito e complemento verbal (objeto direto ou indireto). Os atos de fala incidirão sobre essa representação [Vieira and Vera 2001].

Como linguagem de representação é utilizado Lógica de Primeira Ordem, portanto a interpretação semântica se torna um processo de associar uma expressão lógica

¹Natural Language Toolkit - www.nltk.org

de primeira ordem a um sintagma, onde determina-se quais representações semânticas queremos associar a cada um dos sintagmas [Russell and Norvig 2004].

Considere a frase “*John loves Mary*”. O NP (*Noun Phrase*) *John* tem como interpretação semântica o termo lógico *John* e a sentença como um todo deve ser interpretada como uma sentença lógica *Loves(John,Mary)*. A parte VP (*Verbal Phrase*) *loves Mary* da frase é uma descrição que poderia ou não se aplicar a uma pessoa em particular (se aplica a *John*). Isso significa que *loves Mary* é um **predicado** que, quando combinado com um termo que representa uma pessoa, produz uma sentença lógica completa. Utilizando-se a notação λ (fórmula com falta de argumento) o predicado *loves Mary* é representado por:

$$\lambda x \text{ Loves}(x, \text{Mary}).$$

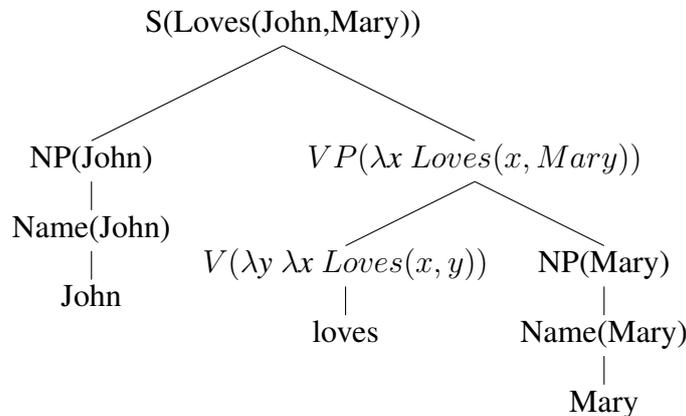
A regra $S(\text{rel}(\text{obj})) \rightarrow NP(\text{obj}) VP(\text{rel})$ informa que “um NP com semântica *obj* seguido por um VP com semântica *rel* gera uma sentença cuja semântica é o resultado de se aplicar *rel* a *obj*”. A interpretação semântica de “*John loves Mary*” é $(\lambda x \text{ Loves}(x, \text{Mary}))(\text{John})$, que se equivale a *Loves(John, Mary)*.

O verbo *loves* é representado por $\lambda y \lambda x \text{ Loves}(x, y)$, o predicado que, ao receber o argumento *Mary*, retorna o predicado $\lambda x \text{ Loves}(x, \text{Mary})$. Logo, a regra $VP \rightarrow Verb NP$ aplica o predicado que é a interpretação semântica do verbo ao objeto, que por sua vez, é a interpretação semântica de NP para se obter a interpretação semântica do VP de forma completa, resultando na gramática (2).

$$\begin{aligned} S(\text{rel}(\text{obj})) &\rightarrow \langle NP(\text{obj}) \rangle \langle VP(\text{rel}) \rangle \\ VP(\text{rel}(\text{obj})) &\rightarrow \langle Verb(\text{rel}) \rangle \langle NP(\text{obj}) \rangle \\ NP(\text{obj}) &\rightarrow \langle Name(\text{obj}) \rangle \end{aligned} \tag{2}$$

$$\begin{aligned} Name(\text{John}) &\rightarrow \text{John} \\ Name(\text{Mary}) &\rightarrow \text{Mary} \\ Verb(\lambda x \lambda y \text{ Loves}(x, y)) &\rightarrow \text{loves} \end{aligned}$$

Uma árvore de análise com interpretações semânticas para a cadeia “*John loves Mary*” é apresentada abaixo:



Outros entraves na análise semântica de uma frase são os tempos verbais e a eliminação da ambiguidade que serão explorados na seção 2.3.1.

2.3.1. Tempo Verbal e Ambiguidade

A língua inglesa² utiliza **tempos verbais** (passado, presente e futuro) para indicar a hora relativa de um evento (e). Para a representação de eventos é utilizado a notação de cálculo de eventos [Russell and Norvig 2004]. Considere duas frases: “*John loves Mary*” e “*John loved Mary*”, no cálculo de eventos temos, respectivamente (3) e (4):

$$e \in Loves(John, Mary) \wedge During(Now, e) \quad (3)$$

$$e \in Loves(John, Mary) \wedge After(Now, e) \quad (4)$$

Logo, as duas regras léxicas³ para as palavras *loves* (5) e *loved* (6) são:

$$Verb(\lambda x \lambda y e \in Loves(John, Mary) \wedge During(Now, e)) \rightarrow loves \quad (5)$$

$$Verb(\lambda x \lambda y e \in Loves(x, y) \wedge After(Now, e)) \rightarrow loved \quad (6)$$

Estudos mais avançados dessa questão incluem representação semântica para os quantificadores da linguagem natural. E esses devem ser traduzidos para os quantificadores lógicos. A frase “*Every agent smells a wumpus*” (retirada de [Russell and Norvig 2004]) é ambígua e o significado preferencial é o de que *os agentes poderiam estar sentindo o cheiro de diferentes wumpus* (7), ou um significado alternativo é o de que *existe um único wumpus do qual todo mundo sente cheiro* (8). Podem ser representados por (Considere a como agente, w como wumpus):

$$\forall a a \in Agents \implies \exists w w \in Wumpuses \wedge \exists e e \in Smells(a, w) \wedge During(Now, e) \quad (7)$$

$$\exists w w \in Wumpuses \forall a a \in Agents \implies \exists e e \in Smells(a, w) \wedge During(Now, e) \quad (8)$$

A primeira interpretação é analisada de modo composicional (NP e VP), logo, “*Every agent*” e “*smells a wumpus*” é representada, respectivamente por (9) e (10).

$$NP(\forall a a \in Agents \implies P) \quad (9)$$

$$VP(\exists w w \in Wumpuses \wedge \exists e (e \in Smells(a, w) \wedge During(Now, e))) \quad (10)$$

Segundo [Russell and Norvig 2004], existem duas dificuldades:

1. A semântica de toda a sentença parece ser semântica do NP, e a semântica do VP preenchendo a parte P, ou seja, não é possível formar a semântica da sentença com $rel(obj)$.
2. E é preciso obter a variável a como um argumento para a relação *Smells*, ou seja, a semântica da sentença é formada inserindo-se a semântica do VP no encaixe do argumento correto do NP, enquanto também insere-se a variável a do NP no encaixe do argumento correto da semântica do VP.

²A língua inglesa foi adotada para o desenvolvimento deste projeto

³Regras léxicas são usadas para expressar as relações entre as descrições [Bredenkamp et al. 1996]

Para resolver essa ambiguidade, utiliza-se a forma **quase-lógica** que inclui toda a lógica de primeira ordem e é aumentada por expressões lambda e por uma nova construção, chamado **termo quantificado**. Por exemplo, o termo quantificado de *every agent* é escrito $[\forall a a \in Agents]$ e a forma quase-lógica para “*Every agent smells a wumpus*” é visualizada em (11).

$$\exists e ((e \in Smells[\forall a a \in Agents], [\exists w w \in Wumpsuses]) \wedge During(Now, e)) \quad (11)$$

A Tabela 1 (retirada de [Russell and Norvig 2004]) mostra os tipos semânticos e exemplos de formas correspondentes a cada categoria sintática sobre a abordagem de forma quase-lógica.

Categoria	Tipo Semântico	Example	Forma quase-lógica
S	sentence	the I sleep	$\exists e e \in Sleep(Speaker) \wedge During(now, e)$
NP	object	a dog	$[\exists d Dog(d)]$
PP	$object^2 \rightarrow sentence$	in [2, 2]	$\lambda x In(x, [2, 2])$
RelClause	object \rightarrow sentence	that sees me	$\lambda x \exists e e \in See(x, Speaker) \wedge During(Now, e)$
VP	$object^n \rightarrow sentence$	sees me	$\lambda x \exists e e \in See(x, Speaker) \wedge During(Now, e)$
Adjective	object \rightarrow sentence	smelly	$\lambda x Smelly(x)$
Adverb	event \rightarrow sentence	today	$\lambda e During(e, Today)$
Article	quantifier	the	$\exists!$
Conjunction	$sentence^2 \rightarrow sentence$	and	$\lambda p q (p \wedge q)$
Digit	object	7	7
Noun	object \rightarrow sentence	wumpus	$\lambda x x \in Wumpsuses$
Preposition	$object^2 \rightarrow sentence$	in	$\lambda x \lambda y In(x, y)$
Pronoun	object	I	Speaker
Verb	$object^n \rightarrow sentence$	eats	$\lambda y \lambda x \exists e e \in Eat(x, y) \wedge During(Now, e)$

Tabela 1. Categorias Sintática, Tipo Semântico e Forma quase-lógica

Na Tabela 1 a notação $t \rightarrow r$ denota uma função que recebe um argumento do tipo t e retorna um resultado do tipo r . Por exemplo, o tipo semântico de Preposição é $objeto^2 \rightarrow sentence$, o que significa que a semântica da preposição é uma função que quando for aplicada a dois objetos lógicos, produzirá uma sentença lógica [Russell and Norvig 2004]. A gramática completa com semântica em forma quase-lógica é apresentada em 12.

$$\begin{aligned}
S(\text{rel}(\text{obj})) &\rightarrow \langle NP(\text{obj}) \rangle \langle VP(\text{rel}) \rangle \\
S(\text{conj}(\text{sem}_1, \text{sem}_2)) &\rightarrow \langle S(\text{sem}_1) \rangle \langle \text{Conjunction}(\text{conj}) \rangle \langle S(\text{sem}_2) \rangle \\
NP(\text{sem}) &\rightarrow \langle \text{Pronoun}(\text{sem}) \rangle \\
NP(\text{sem}) &\rightarrow \langle \text{Name}(\text{sem}) \rangle \\
NP([q \ x \ \text{sem}(x)]) &\rightarrow \langle \text{Article}(q) \rangle \langle \text{Noun}(\text{sem}) \rangle \\
NP([q \ \text{xobj} \wedge \text{rel}(x)]) &\rightarrow \langle NP([q \ \text{xobj}]) \rangle \langle PP(\text{rel}) \rangle \\
NP([q \ \text{xobj} \wedge \text{rel}(x)]) &\rightarrow \langle NP([q \ \text{xobj}]) \rangle \langle \text{RelClause}(\text{rel}) \rangle \\
NP([\text{sem}_1, \text{sem}_2]) &\rightarrow \langle \text{Digit}(\text{sem}_1) \rangle \langle \text{Digit}(\text{Sem}_2) \rangle \\
VP(\text{sem}) &\rightarrow \langle \text{Verb}(\text{sem}) \rangle \\
VP(\text{rel}(\text{obj})) &\rightarrow \langle VP(\text{rel}) \rangle \langle NP(\text{obj}) \rangle \\
VP(\text{sem}_1(\text{sem}_2)) &\rightarrow \langle VP(\text{sem}_1) \rangle \langle \text{Adjective}(\text{sem}_2) \rangle \\
VP(\text{sem}_1(\text{sem}_2)) &\rightarrow \langle VP(\text{sem}_1) \rangle \langle PP(\text{sem}_2) \rangle \\
\text{RelClause}(\text{sem}) &\rightarrow \mathbf{that} \langle VP(\text{sem}) \rangle \\
PP(\lambda x \ \text{rel}(x, \text{obj})) &\rightarrow \langle \text{Preposition}(\text{rel}) \rangle \langle NP(\text{obj}) \rangle
\end{aligned} \tag{12}$$

E por fim é necessário converter a forma quase-normal em lógica de primeira ordem. Isto é feito com uma regra simples:

para cada termo quantificado $[q \ x \ P(x)]$ dentro de uma forma quase-lógica (FQL), substitua o termo quantificado por x e substitua FQL por $q \ x \ P(x)$ op FQL, onde op é \implies quando q é \forall e \wedge quando q é \exists ou $\exists!$.

Por exemplo, a sentença *Every dog has a day* tem a forma quase-lógica em (13). Onde e , d , a , são respectivamente, evento, cão e dia.

$$\exists e (e \in \text{Has}([\forall d \ d \in \text{Dogs}], [\exists a \ a \in \text{Days}], \text{Now})) \tag{13}$$

Logo, existem duas interpretações lógicas possíveis, pois não foi especificado qual dos dois termos quantificados seria extraído primeiro:

$$\forall d \ d \in \text{Dogs} \implies \exists a \ a \in \text{Days} \wedge \exists e \ e \in \text{Has}(d, a, \text{Now}) \tag{14}$$

$$\exists a \ a \in \text{Days} \wedge \forall d \ d \in \text{Dogs} \implies \exists e \ e \in \text{Has}(d, a, \text{Now}) \tag{15}$$

Portanto, a interpretação (14), diz que cada cão tem seu próprio dia, enquanto que a (15) diz que existe um dia especial que todos os cães compartilham. A escolha entre elas é o trabalho de eliminação da ambiguidade.

2.3.2. Classificação

Para realizar a classificação do texto nas cinco classes dos atos ilocutórios, utiliza-se a abordagem de classificação por templates, apresentados na Seção 2. Cada template definido é comparado com a classe semântica encontrada, gerando então uma anotação.

Para a frase “*John loves Mary*”, uma possível abordagem é comparar o quão próximo o template $E \ \phi \ (P) \ (S/H \ + \ \text{property})$, se relaciona com a interpretação

semântica $S(\text{Loves}(\text{John}, \text{Mary}))$. Pode-se entender que a estrutura semântica criada na árvore identifica o verbo *Loves* como um membro pertencente da classe **Expressivo** dos atos ilocutórios.

Suponha que existam dois *templates* (A, B) e que comparado com a interpretação semântica (S), têm resultados idênticos, ou seja, classificar S em A ou S em B é uma questão de escolha aleatória. Nesses casos, uma possível abordagem, seria utilizar algoritmos de aprendizagem por reforço, pelo fato de que os atores virtuais podem ser criticados por um diretor e/ou animador sugerindo um modo diferente de interpretação.

3. Sistema Proposto

A Figura 1, define a arquitetura do sistema proposto, que é dividida em duas partes: processamento de linguagem natural e classificação.

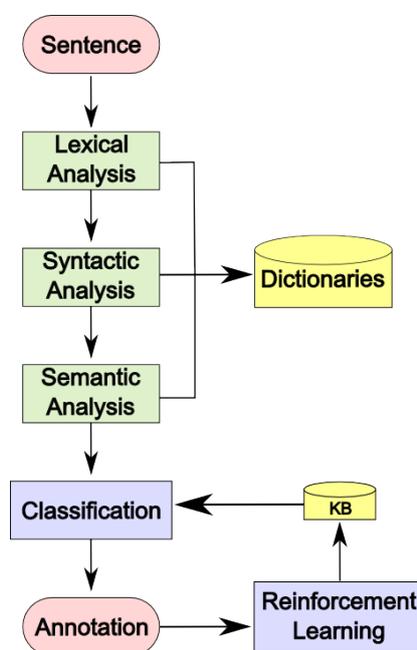


Figura 1. Arquitetura do módulo de anotação

Para a etapa de processamento de linguagem natural é proposto a utilização da análise léxica e sintática, abordados na subseção 2.2 e análise semântica, na subseção 2.3. Na fase de classificação, utiliza-se um algoritmo de aprendizagem por reforço para classificar sentenças nos atos ilocutórios, abordados na subseção 2.3.2.

Para a validação do sistema proposto, um conjunto de falas de um filme de animação é utilizado. Para avaliar o sistema utiliza-se duas abordagens: o puramente aleatório e a ferramenta a ser desenvolvida. A primeira, já em desenvolvimento, consiste em uma ferramenta que classifique aleatoriamente as falas em 5 classes dos atos ilocutórios. Para a segunda etapa será utilizadas algumas técnicas computacionais, ainda em estudo, para a implementação da ferramenta proposta. Ambas terão os resultados comparados com o método de anotação manual.

A hipótese deste trabalho é que o método de classificação automática apresente resultados mais próximos da classificação manual do que a anotação puramente aleatória.

4. Conclusões

Neste trabalho apresentou-se um estudo sobre os atos ilocutórios que possibilitou compreender como identificar as condições comunicativas e também reconhecer as intenções em atos de fala.

Também foi feito um estudo sobre o processamento de linguagem natural, compreendendo as etapas envolvidas no processo para conseguir extrair informações de texto. Estas etapas são descritas como Análise Morfológica, que separa cada palavra individualmente e a classifica utilizando o método de *POS tagging*, a Análise Sintática que cria uma árvore através de uma Gramática Livre de Contexto e por fim a Análise Semântica, que tenta extrair relações semântica das árvores criadas na fase anterior.

Outro tópico abordado neste trabalho foi a utilização de aprendizagem de máquina, que se inserida na fase de classificação da arquitetura do módulo de Anotação possibilita analisar a classificação entre a interpretação semântica e os *templates* dos atos de fala e então aprender com uma crítica, por exemplo.

O próximo passo será comparar as ferramentas em desenvolvimento com os resultados da classificação manual, o qual será disponibilizado na web para a captação dos mesmo durante um período estipulado. Após o desenvolvimento será possível analisar e comparar qual ferramenta obteve melhores resultados.

Referências

- Austin, J. and Urmson, J. (1978). *How to do things with words*. Harvard University Press.
- Bredenkamp, A., Markantonatou, S., and Sadler, L. (1996). Lexical rules: What are they? pages 163–168.
- Cohen, W., Carvalho, V., and Mitchell, T. (2004). Learning to classify email into "speech acts".
- Russell, S. and Norvig, P. (2004). *Inteligência artificial*. Elsevier.
- Searle, J. (1969). *Speech acts: an essay in the philosophy of language*. Cambridge University Press.
- Searle, J. R. (1976). A Classification of Illocutionary Acts. *Language in Society*.
- Seol, Y. S., Kim, D. J., and Kim, H. W. (2008). Emotion Recognition from Text Using Knowledge-based ANN. In *Proceedings of 23rd International Technical Conference on Circuits/Systems Computers and Communications*, pages 1569–1572. ITC-CSCC 2008.
- Tao, J. and Tan, T. (2004). Emotional chinese talking head system. In *Proceedings of the 6th international conference on Multimodal interfaces, ICMI '04*, pages 273–280, New York, NY, USA. ACM.
- Tsohatzidis, S. (1994). *Foundations of speech act theory: philosophical and linguistic perspectives*. Routledge.
- Vieira, R. and Vera (2001). Linguística computacional: princípios e aplicações. In *As Tecnologias da informação e a questão social*.