

IV Workshop de Dissertações de Mestrado do PPgSI (2017)

FICHA DA PESQUISAⁱ

DADOS GERAIS					
Título do projeto de pesquisa	Geração de Expressões de Referência em situações de comunicação com restrição de tempo				
Orientando	Andre Costa Mariotti				
Orientador(es)	Ivandr� Paraboni				
Semestre no curso, na data do workshop	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input type="checkbox"/> 4º semestre	<input checked="" type="checkbox"/> 5º semestre	<input type="checkbox"/> 6º semestre
Qualificação	<input checked="" type="checkbox"/> Qualificação já realizada em: 20/09/2016 <input type="checkbox"/> Realização da qualificação planejada para:				
Defesa	Prazo máximo para depósito: 26/02/2018		Realização da defesa planejada para: 13/09/2017		
Linha e Área de pesquisa	Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input type="checkbox"/> Eng. de Software		Inteligência de Sistemas: <input checked="" type="checkbox"/> IA <input type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico		
Área de aplicação	<input type="checkbox"/> Ambientes Corporativos	<input type="checkbox"/> Educação	<input checked="" type="checkbox"/> Linguagem Natural	<input type="checkbox"/> Redes Sociais	
	<input type="checkbox"/> Bioinformática	<input type="checkbox"/> Educação a Distância	<input type="checkbox"/> Linguística	<input type="checkbox"/> Robótica	
	<input type="checkbox"/> Biometria	<input type="checkbox"/> Internet	<input type="checkbox"/> Processos de Negócio	<input type="checkbox"/> Saúde	
	<input type="checkbox"/> Economia	<input type="checkbox"/> Jogos	<input type="checkbox"/> Química	<input type="checkbox"/> [outro – escrever]	
Publicações associadas ao projeto de mestrado	No momento foi submetida uma publicação sob o título de “Geração de Expressões de Referência em situações de comunicação com restrição de tempo” ao congresso STIL. Além disso, há outra publicação em fase de elaboração que será submetida à conferência LREC.				

DESCRIÇÃO DO PROJETO DE PESQUISA	
Contextualização / motivação	A motivação consiste de estudar os aspectos da linguagem que podem se manifestar em situações de comunicação diferentes do habitual, podendo contribuir para algoritmos de Processamento de Linguagem Natural mais realistas.
Problema de pesquisa	Os algoritmos tradicionais utilizados na tarefa de Geração de Expressões de Referência (um dos componentes fundamentais da arquitetura de sistemas geradores de linguagem), tratam da tarefa de descrever objetos de forma irrestrita, sem considerar a situação de comunicação na qual a tarefa ocorre.
Objetivo geral da pesquisa	O objetivo da pesquisa é o estudo e desenvolvimento de modelos computacionais de GER em situações de comunicação com restrição de tempo, de modo a produzir descrições mais próximas do comportamento humano do que as que seriam produzidas por algoritmos de GER tradicionais.
Trabalhos relacionados	A situação de comunicação com restrição de tempo tratada nesse trabalho guarda certa afinidade com as condições de comunicação críticas exploradas em Arts et al. (2011). Além disso, estudos apresentados em Krahmer (2014) também mostram que foi possível notar uma influência da emoção sobre a tarefa de GER. Entretanto, apesar de observada a influência do contexto de comunicação sobre a produção de linguagem, algoritmos de GER bem difundidos, como o Incremental (DALE; REITER, 1995) não realizam a tarefa considerando a situação de comunicação.
Justificativa e relevância	O desenvolvimento de algoritmos para Geração de Expressões de Referência, tal como o proposto nesta pesquisa, proporciona a criação de sistemas de Geração de Linguagem Natural mais próximos ao comportamento humano.
Proposta para Solução	(i) Realização de um estudo exploratório sobre como sujeitos humanos produzem descrições em condições de comunicação com restrição de tempo. (ii) Proposta de um algoritmo de GER com parâmetros adicionais que permitam exercer maior controle sobre a saída produzida. (iii) Avaliação da proposta como alternativa a algoritmos tradicionais da área.
Dados	Durante o estudo exploratório foram coletadas descrições realizadas por sujeitos humanos em um experimento controlado que visou observar como a tarefa de gerar expressões é realizada em situações de comunicação com restrição de tempo. Posteriormente, esses mesmos dados também foram utilizados para testes dos modelos.
Validação	O modelo computacional proposto foi executado de forma a comparar seus resultados com um algoritmo tradicional da área de GER definido como baseline. As saídas de ambos os algoritmos foram então comparadas com descrições reais realizadas por sujeitos humanos, sendo que estes dados passaram por testes estatísticos usando Dice e Wilcoxon.
Limitações	O projeto tratou especificamente da Geração de Expressões de Referência, que é um componente específico do Geração de Linguagem Natural. Portanto, fica em aberto uma possível generalização das ideias aqui apresentadas para outros subproblemas da área de Processamento de Linguagem.
Resultados esperados	Contribuições científicas: Avanços na produção de modelos computacionais de GER mais próximos do comportamento humano. Contribuições tecnológicas: Um algoritmo de GER para gerar expressões sob restrição de tempo, denominado IAL. E também, um corpus dessas expressões coletadas a partir de um experimento com humanos.

! Esta ficha é uma adaptação da usada no “VIII Workshop de Teses e Dissertações em Sistemas de Informação (WTDSI 2015)” realizado como parte do “XI Simpósio Brasileiro de Sistemas de Informação (SBSI 2015)”