

Geração de instruções em mundos virtuais: primeiros passos

Diego Cardozo Sandrim, Felipe Freire, Ivandré Paraboni

¹School of Arts, Sciences and Humanities (EACH)
University of São Paulo (USP)
Av. Arlindo Bettio, 1000 - São Paulo, Brasil

(diego.sandrim, felipe.freire, ivandre)@usp.br

Abstract. *The generation of navigational instructions in natural language concerns the computational task of producing sentences to guide a user to a target point in virtual or real worlds. This paper describes the development of a system of this kind, which will participate in a shared task, and it is intended to motivate further research in the field.*

Resumo. *A geração de instruções de navegação em língua natural trata da tarefa computacional de produção de sentenças para guiar um usuário até um ponto de destino em ambientes reais ou virtuais. Este artigo descreve o desenvolvimento de um sistema deste tipo, com o objetivo de fomentar esta linha de pesquisa e para participação em uma competição da área.*

1. Introdução

Com o avanço em áreas da ciência da computação como processamento gráfico e interação humano-computador, problemas de geração de língua natural (GLN) puramente textuais encontraram domínios mais realistas - e efetivamente novas aplicações - em diversas áreas que utilizam sistemas baseados em mundos virtuais interativos. Dentre estas aplicações, destacamos a geração de instruções de navegação em língua natural em ambientes como *Gruve* (*Giving Route instructions in Uncertain Virtual Environments*) [Janarthanam et al. 2012] ilustrado na Figura 1.

A geração de instruções em ambientes virtuais consiste na tarefa computacional de produzir sentenças na forma verbal ou escrita para guiar um usuário de um ponto de origem até um ponto de destino. A tarefa apresenta uma ampla gama de desafios computacionais ainda pouco explorados, como a incerteza inerente aos dados de entrada¹, a questão da granularidade da instrução gerada, a natureza dinâmica da aplicação (que requer que instruções sejam reformuladas em tempo de execução), o tratamento do histórico de navegação e conhecimento prévio do domínio etc.

Com o objetivo de divulgar esta linha de pesquisa e, se possível, motivar novos pesquisadores a colaborar na investigação destes e outros desafios, este artigo descreve o estudo e desenvolvimento prático de um sistema de geração de instruções utilizando a plataforma *Gruve* [Janarthanam et al. 2012], e que foi projetado especificamente para participar do desafio *Gruve Challenge* conforme detalhado nas seções seguintes.

¹Em sistemas baseados em *smartphones*, por exemplo, as coordenadas fornecidas pelo dispositivo de GPS naturalmente apresentam um elevado grau de incerteza [Zandbergen and Barbeau 2011]



Figura 1. Interface do ambiente Gruve [Janarthanam et al. 2012]

2. Trabalho relacionado

Gruve [Janarthanam et al. 2012] é um ambiente de desenvolvimento e teste de vários tipos de sistemas de geração de instruções em mundos virtuais, incluindo sistemas ditos ‘a priori’ (i.e., que geram, em um único passo, uma descrição completa de como atingir o objetivo), incrementais (onde cada instrução é fornecida passo a passo, e.g., [Dale et al. 2003]) ou interativos (que, além de fornecer instruções a cada passo, permitem a interação na forma de perguntas e respostas, e.g., [Malaka and Zipf 2000]).

O sistema inclui um modelo de jogo onde são representados os personagens, regras e objetivos da uma aplicação-exemplo, um mediador cliente-servidor que instancia a aplicação para múltiplos usuários *on-line*, um sistema de navegação a ser desenvolvido pelo pesquisador de *GLN*, um modelo do mundo virtual, um sintetizador de voz e um cliente web que utiliza imagens fotográficas reais fornecidas pela API *Google Maps*.²

O modelo de mundo do sistema é representado na forma de um grafo cujos nós são pontos de decisão (e.g., esquinas etc.) e arestas de ligação, o qual é enriquecido com informações passíveis de referência (e.g., nomes e tipos de logradouros, distâncias, áreas ocupadas etc.) fornecidas por um modelo geográfico aberto derivado do projeto *OpenStreetMap*.³ Dada uma estrutura deste tipo e uma rota (ou sequência de nós) para guiar um usuário de um ponto inicial *a* até um ponto final ou objetivo *b*, a geração de instruções consiste em transformar esta rota em instruções em língua natural.

O sistema *Gruve* forma a infraestrutura básica do desafio computacional *Gruve Challenge*, que é parcialmente inspirado na série de desafios *GIVE (Generating Instructions in Virtual Environments)* [Byron et al. 2007]. Diferentemente do seu antecessor, no entanto, o presente desafio enfoca o problema de descrição de rotas de navegação, e não de um plano de ações de manipulação de objetos (e.g., destravar alarmes, abrir portas etc.)

²<https://maps.google.com/>

³<http://www.openstreetmap.org/>

3. O sistema proposto

O ambiente *Gruve* foi utilizado para implementação de um sistema de geração de instruções passo a passo que aceita um grau limitado de interação com o usuário na forma de botões de perguntas.⁴ O estudo em andamento possui duas frentes de ação: (1) o tratamento computacional da incerteza e (2) o aprimoramento das instruções de navegação. Estas etapas são discutidas individualmente a seguir.

3.1. Tratamento da incerteza

Sistemas de GPS existentes em *smartphones* da atualidade fornecem coordenadas com uma margem de erro considerável [Zandbergen and Barbeau 2011]. Esta característica é simulada na plataforma *Gruve*, que fornece coordenadas com variação de até 50 metros em relação à posição real do usuário, o que significa que a posição deste no grafo representando o ambiente virtual é apenas estimado a partir de um grupo de nós candidatos.

Uma primeira alternativa para o tratamento computacional da incerteza seria levar em conta todas as posições candidatas para então produzir uma série de instruções condicionais cobrindo todas as possibilidades. Como resultado, seria produzida uma descrição completa porém bastante extensa, como em *'You must go to Windmill st. If you are already in Chapel st. with Nicolson st., turn around to your right and walk onto Windmill st. But if you are on Chapel st., turn right. If you are on Windmill st., continue walking'*.

Considerando-se que as sentenças candidatas são na forma (*SE-Condição-Instrução*), e que várias delas podem repetir a mesma instrução, uma alternativa a considerar seria a identificação da instrução mais frequente, que seria então eleita como a saída a ser fornecida ao usuário. Em testes realizados, entretanto, o número de repetições foi relativamente baixo, tornando esta alternativa pouco efetiva.

Em virtude destas dificuldades, optamos por não produzir descrições alternativas. Ao invés disso, a estratégia adotada consiste em instruir o usuário a encontrar o ponto de partida (esteja ele certo ou errado) com base no qual a instrução foi gerada, e a partir deste ponto seguir normalmente a instrução produzida (que a partir deste ponto estará necessariamente correta). Por exemplo, supondo-se que as coordenadas fornecidas pelo sistema localizam o usuário em *Marshall st.*, o sistema poderia gerar a instrução *'You should start from Marshall st. Then, continue walking onto Potterrow'*. A desvantagem desta abordagem é o maior ônus ao usuário, que precisa encontrar o início da rota por conta própria e sem maior apoio do sistema.

Além disso, procuramos também ajustar situações em que o usuário possa estar em um ponto mais avançado do percurso do que o indicado pelo sistema.⁵ O sistema neste caso informa não apenas a instrução atual como também antecipará a próxima, como em *'Continue walking onto Potterrow. After that, go to Nicolson Square'*. A desvantagem desta estratégia é a geração de instruções ligeiramente mais longas do que o necessário nos casos em que a coordenada fornecida estava realmente correta.

⁴Mas não oferecendo suporte para entrada de dados em língua natural, o que exigiria funcionalidades adicionais como resolução anafórica [Paraboni 1997, Cuevas and Paraboni 2008] e outras.

⁵Casos em que o usuário possa ser atrasado em relação às coordenadas fornecidas já são tratados pela implementação nativa, que impede a geração de instruções duplicadas.

3.2. Aprimoramento das instruções de navegação

Diversas outras funcionalidades para melhoria da experiência de navegação do usuário foram acrescentadas ao sistema original. Dentre estas, destacamos a detecção de erros de orientação, o uso de pontos de referência nas instruções geradas e o tratamento de situações em que o sistema não conseguiu produzir uma rota adequada.

A detecção de erros cometidos pelo usuário verifica eventuais desvios de rota - medidos pelo ângulo do usuário em relação ao local de destino - e produz instruções para sua correção. Na implementação atual, quando este ângulo ultrapassa 90 graus, o usuário é informado de que está no caminho errado e a instrução original é repetida, como em *'You are going the wrong way. Turn around and walk onto Potterrow'*.

O uso de pontos de referência nas instruções é uma forma de tentar enriquecer a informação fornecida (particularmente importante quando lidando com situações de incerteza, cf. seção anterior) e assim minimizar o número de erros cometidos pelo usuário. Para este fim, é computada uma lista de todas entidades salientes que se encontram nas proximidades, e que possuem um nome próprio passível de referência, tais como bancos, lojas e restaurantes. Uma vez obtida esta lista, o sistema escolhe de forma aleatória dois⁶ pontos de referência e os acrescenta à instrução corrente, como em *'Turn left onto Potterrow. Things you will see near your objective: KFC and Barclays bank'*.⁷

Finalmente, o planejamento de rota entre dois pontos pode ser complexo, especialmente quando o usuário está posicionado em um ponto muito afastado de uma rua principal. Em casos como estes, o planejador do sistema *Gruve* pode retornar uma rota vazia, não havendo assim nenhuma instrução a fornecer ao usuário. Como forma de contornar esta dificuldade, o sistema proposto detecta situações de ausência de informação e solicita ao usuário que se repositone em um local mais acessível para que a instrução possa ser gerada.

4. Considerações

Este artigo descreveu o estudo e desenvolvimento de um sistema de geração de instruções de navegação em mundos virtuais, com o objetivo de fomentar o estudo na área e para participação em uma competição internacional de aplicações deste tipo. O sistema proposto encontra-se em fase de teste, e será submetido para avaliação junto ao *Gruve Challenge*. Segundo [Janarthnam et al. 2012], a avaliação será realizada *on-line* com uma grande massa de usuários (possivelmente recrutados via *crowd-sourcing*) levando em conta métricas como tempo de navegação, número de erros cometidos etc., e critérios subjetivos como a opinião dos usuários.

Como trabalho futuro, esperamos agregar ao sistema um módulo completo de geração de expressões de referência [Paraboni et al. 2006] e um módulo de realização textual de maior cobertura [Pereira and Paraboni 2007, Pereira and Paraboni 2008, de Novais and Paraboni 2012]. Outra possibilidade contemplada é a adaptação do ambiente para aplicações com maior interatividade, incluindo a interpretação de requisições do usuário e outros.

⁶Este limite foi estabelecido como forma de evitar a geração de instruções excessivamente longas.

⁷Na implementação atual, entretanto, ainda não há geração de expressões de referência propriamente dita [Paraboni 2003, de Lucena et al. 2010, Paraboni and van Deemter 2013].

Referências

- Byron, D., Koller, A., Oberlander, J., Stoia, L., and Striegnitz, K. (2007). Generating instructions in virtual environments (GIVE): A challenge and evaluation testbed for NLG. In *Proceedings of the Workshop on Shared Tasks and Comparative Evaluation in Natural Language Generation*.
- Cuevas, R. and Paraboni, I. (2008). A machine learning approach to portuguese pronoun resolution. *Advances in Artificial Intelligence–IBERAMIA 2008*, LNAI 5290:262–271.
- Dale, R., Geldof, S., and Prost, J.-P. (2003). CORAL : Using natural language generation for navigational assistance. In *Proceedings of the Twenty-Sixth Australasian Computer Science Conference (ACSC2003)*.
- de Lucena, D. J., Pereira, D. B., and Paraboni, I. (2010). From semantic properties to surface text: The generation of domain object descriptions. *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, (45):48–58.
- de Novais, E. M. and Paraboni, I. (2012). Portuguese text generation using factored language models. *Journal of the Brazilian Computer Society*, 19(2):135–146.
- Janarthanam, S., Lemon, O., and Liu, X. (2012). A web-based evaluation framework for spatial instruction-giving systems. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pages 49–54.
- Malaka, R. and Zipf, E. (2000). Deep Map - challenging IT research in the framework of a tourist information system. *Information and Communication Technologies in Tourism*, pages 15–27.
- Paraboni, I. (1997). Uma arquitetura para a resolução de referências pronominais possessivas no processamento de textos em língua portuguesa. Master’s thesis, PUCRS, Porto Alegre.
- Paraboni, I. (2003). *Generating references in hierarchical domains: the case of Document Deixis*. PhD thesis, University of Brighton.
- Paraboni, I., Masthoff, J., and van Deemter, K. (2006). Overspecified reference in hierarchical domains: measuring the benefits for readers. In *INLG-2006*, pages 55–62.
- Paraboni, I. and van Deemter, K. (2013). Reference and the facilitation of search in spatial domains. *Language and Cognitive Processes*, online.
- Pereira, D. B. and Paraboni, I. (2007). A language modelling tool for statistical NLP. In *5th Workshop on Information and Human Language Technology (TIL-2007)*. *Anais do XXVII Congresso da SBC*, pages 1679–1688.
- Pereira, D. B. and Paraboni, I. (2008). Statistical surface realisation of portuguese referring expressions. LNAI 5221:383–392.
- Zandbergen, P. A. and Barbeau, S. J. (2011). Positional accuracy of assisted gps data from high-sensitivity gps-enabled mobile phones. *Journal of Navigation*, 64(3):381–399.