

# A NLG-based Application for Walking Directions



Michael Roth, Anette Frank  
Department of Computational Linguistics  
Heidelberg University, Germany

**Introduction** This work describes a web application that makes use of third party resources for computing routes and landmarks, as a knowledge base for the generation of walking directions. The generation model is statistically trained on a corpus of walking directions annotated with POS and syntactic and frame semantic information.

## Route Directions

- Goal: Informing a person of how to get to his goal
- Navigation systems:  
Usage of street names, distance measures and directions
- Cognitive science:  
*Landmarks* allow for more efficient navigation

## Data Sources

- Routes: Google Maps API
- Landmarks:
  - Wikipedia WikiProject Geographical Coordinates
  - Google AJAX Search API
  - Wikimapia
  - OpenStreetMaps



## Corpus-based Generation

- Corpus of route directions (Marciniak & Strube, 2005)
- Annotation:
  - PoS
  - Syntactic (verb and phrase types, arguments, connectives)
  - Frame semantics (frames and semantic roles)
  - Temporal relations (between discourse units)

## Current Work

Collecting additional data for pre-selected routes in order to:

- Map route segments to linguistic realizations
- Extension to multilingual setting (cf. Schuldes et al., UCNLG'09; Talk: August 6, 5:00pm)

### Example: CHIJMES in wikimapia

```
<description>
<![CDATA[ CHIJMES (Old CHIJ) ... ]]>
</description>
<Region>...</Region>
<MultiGeometry><LineString>
<coordinates>
103.8515364,1.2956241,0 103.8515364,1.2947401,0
103.8527644,1.2947401,0 103.8527644,1.2956241,0
103.8527644,1.2956241,0 103.8515364,1.2956241,0
</coordinates>
...
```



### Raffles Hotel

From Wikipedia, the free encyclopedia



Categories: 1899 architecture | Downtown Core | Hotels in Singapore | Visitor attractions in Singapore | National Monuments of Singapore | Museums in Singapore

## Integration of Generation Model

- Current model is only able to generate step-by-step directions
- Hardcoded set of rules that map route segments to frames:

Starting point

→ SELF\_MOTION + *path*  
+ ?landmark as *source*

Landmarks along the route

→ SELF\_MOTION  
+ landmark as *path*

Turning point

→ SELF\_MOTION + *direction*  
+ ?landmark as *source*

Goal with/without landmark

→ SELF\_MOTION  
+ landmark/address as *goal*

2-step generation:  
parameters estimated  
from annotated corpus

1) SELF\_MOTION + *path*

→ walk (17%),  
follow (13%), ...

2) walk+ [*street*,*path*]

→ walk + down + np  
= walk down Raffles Blvd

SELF\_MOTION + [*building*,*path*]

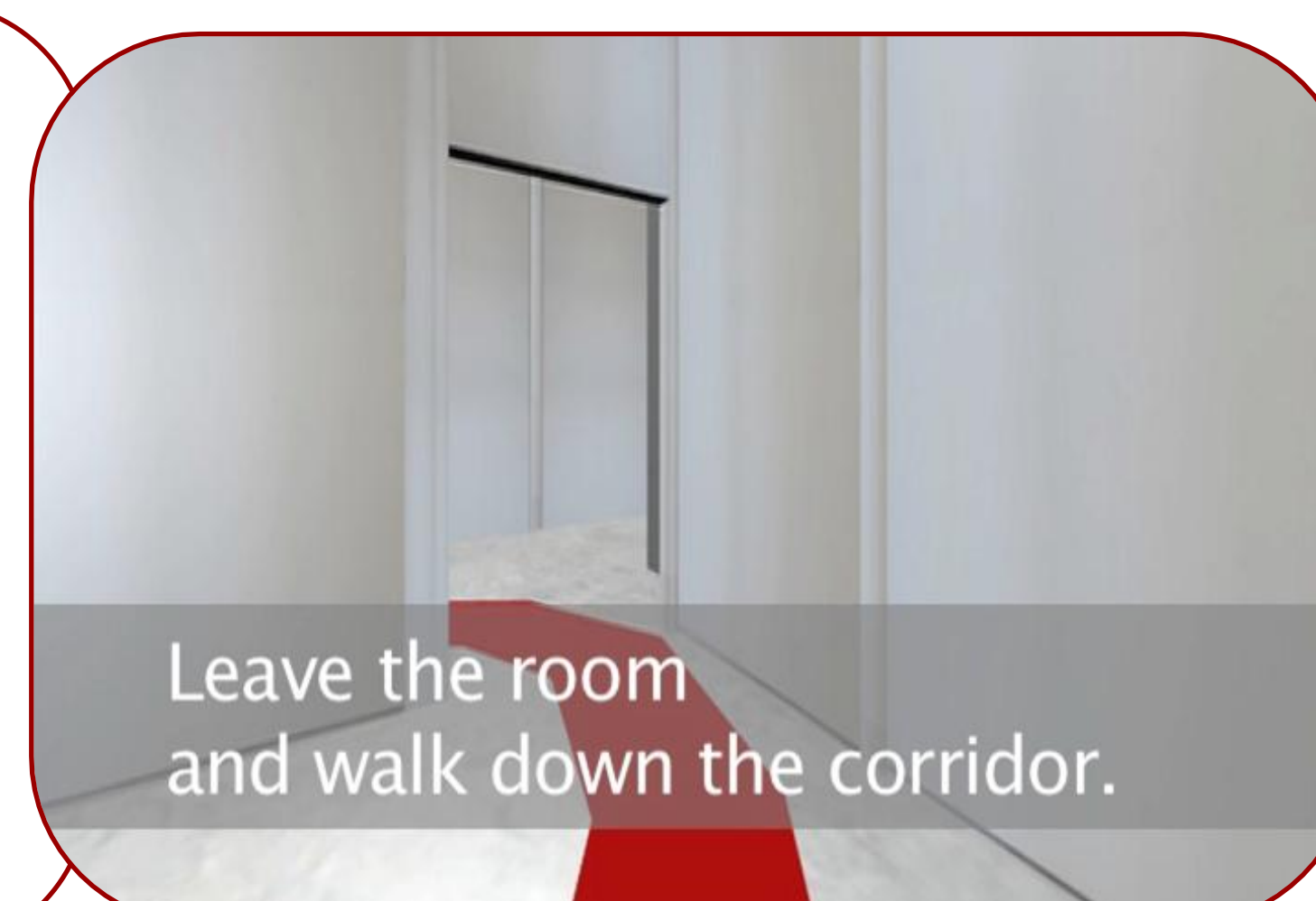
→ pass the New Asia Bar

SELF\_MOTION + [*right*,*direction*]

→ make a right

## Conclusions & Outlook

- Research prototype for web-based data acquisition and evaluation
- Future work:
  - Induction of a mapping from route segments to frames
  - Data-driven integrated generation system
  - Generation in authentic 3D navigation settings



Leave the room  
and walk down the corridor.

## References

- Marciniak, T., & Strube, M. (2005). Using an annotated corpus as a knowledge source for language generation. Proceedings of the Workshop on Using Corpora for Natural Language Generation. Birmingham, UK.
- Schuldes, S., Roth, M., Frank, A., and Strube, M. (2009). Creating an Annotated Corpus for Generating Walking Directions. Proceedings of the ACL-IJCNLP'09 Workshop on Language Generation and Summarisation. Singapore.