Towards Machine-based Matching of Addresses Expressed in Natural Languages

Farid Karimipour   fkarimipr@ut.ac.ir

Assistant Professor
Department of Surveying and Geomatics Engineering, University of Tehran, Iran

Ali Javidaneh   ajavidaneh@ut.ac.ir

PhD Student
Department of Surveying and Geomatics Engineering, University of Tehran, Iran

Andrew U. Frank   frank@geoinfo.tuwien.ac.at

Professor
Department of Geodesy and Geoinformation, Technical University of Vienna, Austria

Outline

- **Problem definition**   What is an address expressed in natural language?

- **Approach**   Defining a grammar for addresses expressed in natural language

- **Results and discussion**

- **Conclusion and future work**

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

1

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shariati Ave., Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shariati Ave., After **Zafar Ave.**, Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shariati Ave., Before **Yakhchal Ave.**, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shahrzad Blvd., Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Daroos, Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

2

- Specifies the destination **and** tells how to reach there

- Helps in building up the cognitive map through exposing the user to the environment and its spatial elements

Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shariati Ave., Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2
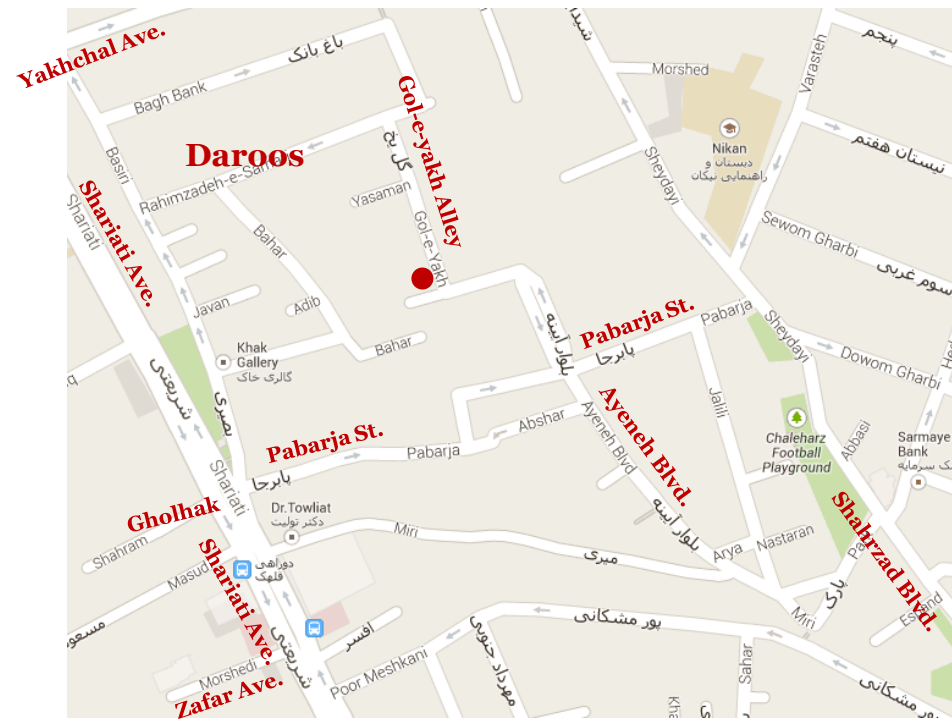
Shariati Ave., After **Zafar Ave.**, Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shariati Ave., Before **Yakhchal Ave.**, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shahrzad Blvd., Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Daroos, Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

3

Address → Map

Map → Address

Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2
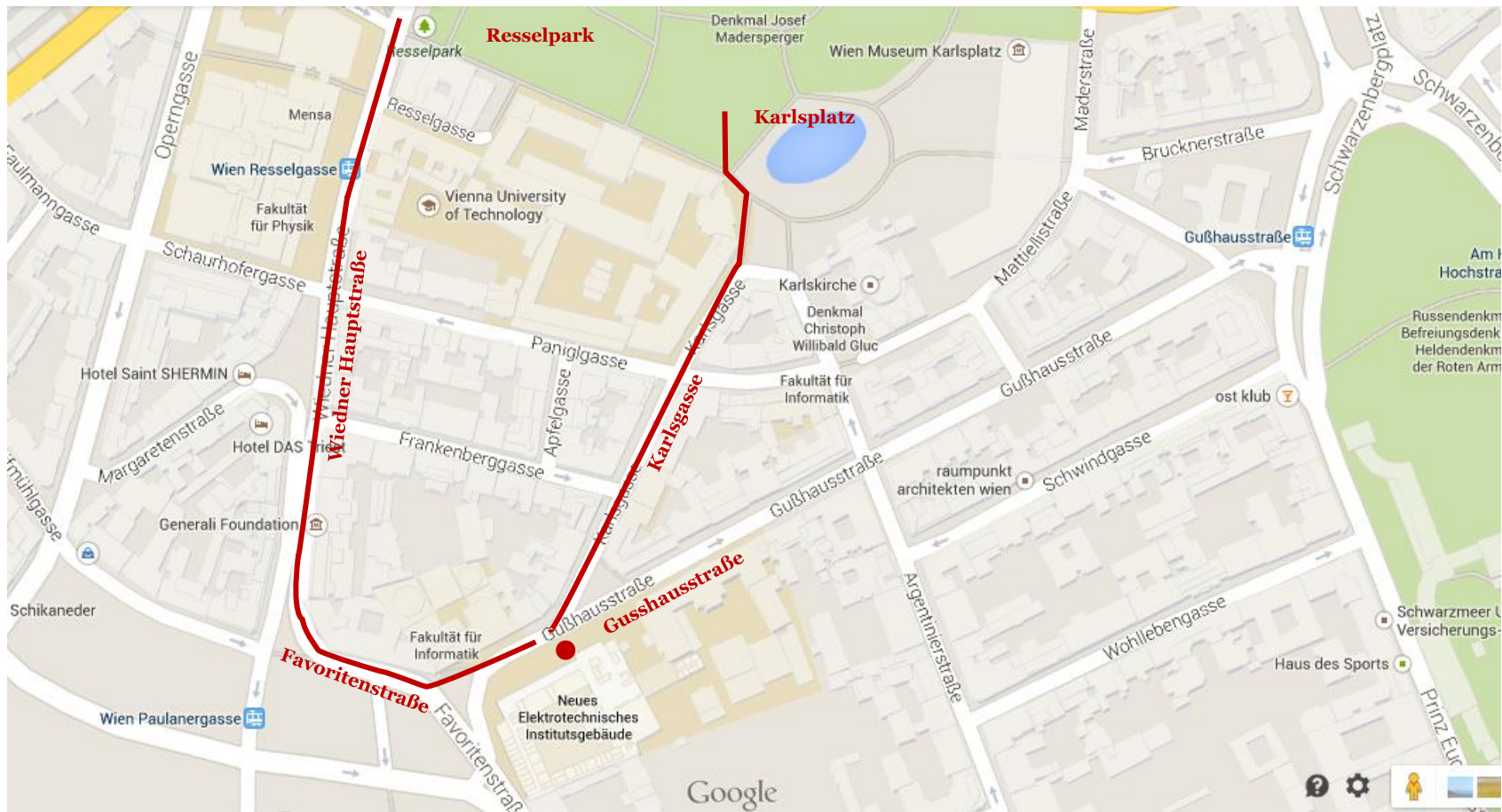
Shariati Ave., Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shariati Ave., After **Zafar Ave.**, Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shariati Ave., Before **Yakhchal Ave.**, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Shahrzad Blvd., Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Daroos, Gholhak, Pabarja St., Ayeneh Blvd., West corner of Gol-e-yakh Alley, No. 2

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

4

Karlsplatz, Karlsgasse, Gusshausstraße, No. 27-29

Resselpark, Wiedner Hauptstraße, Favoritenstraße, Gusshausstraße, No. 27-29

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank
5

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

Non-standard descriptive addresses

Expressed in natural language

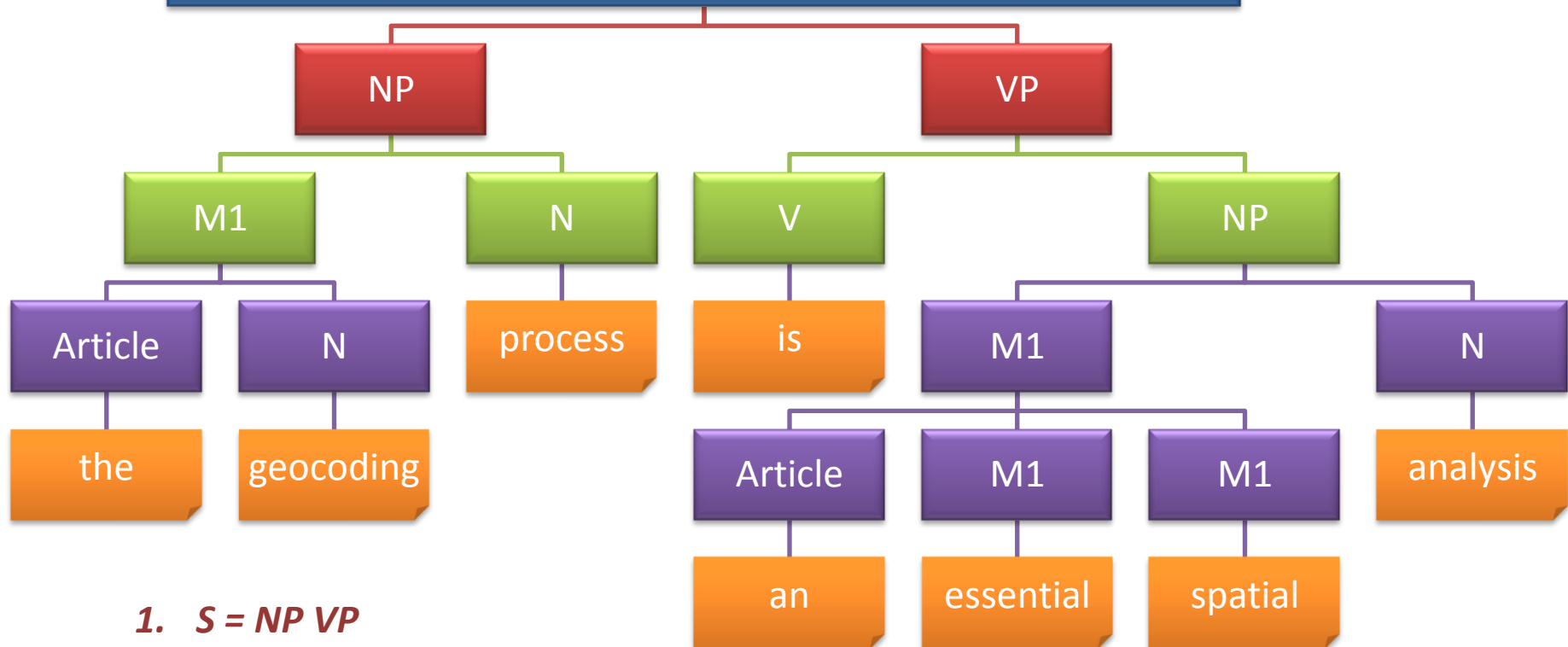Need a formal language to be parsed

Formal language

Symbols

Combination rules

BNF

- Backus-Naur Form expresses formal languages
- A formal meta-language to define other languages

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

6

The geocoding process is an essential spatial analysis

1. *S = NP VP*
2. *NP = M1 N M2*
3. *VP = V NP*
4. *M1 = article | quantifier | N | demonstration | separation*
5. *M2 = appositive | relative clause*

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

7

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

A Sample Grammar for Descriptive Addresses

Descriptive address

GeoNames (GN)

Spatial relations (SR)

Constant GeoNames (cGN)

Variable GeoNames (vGN)

avenue, street, alley, etc.

Names of the constant GeoNames
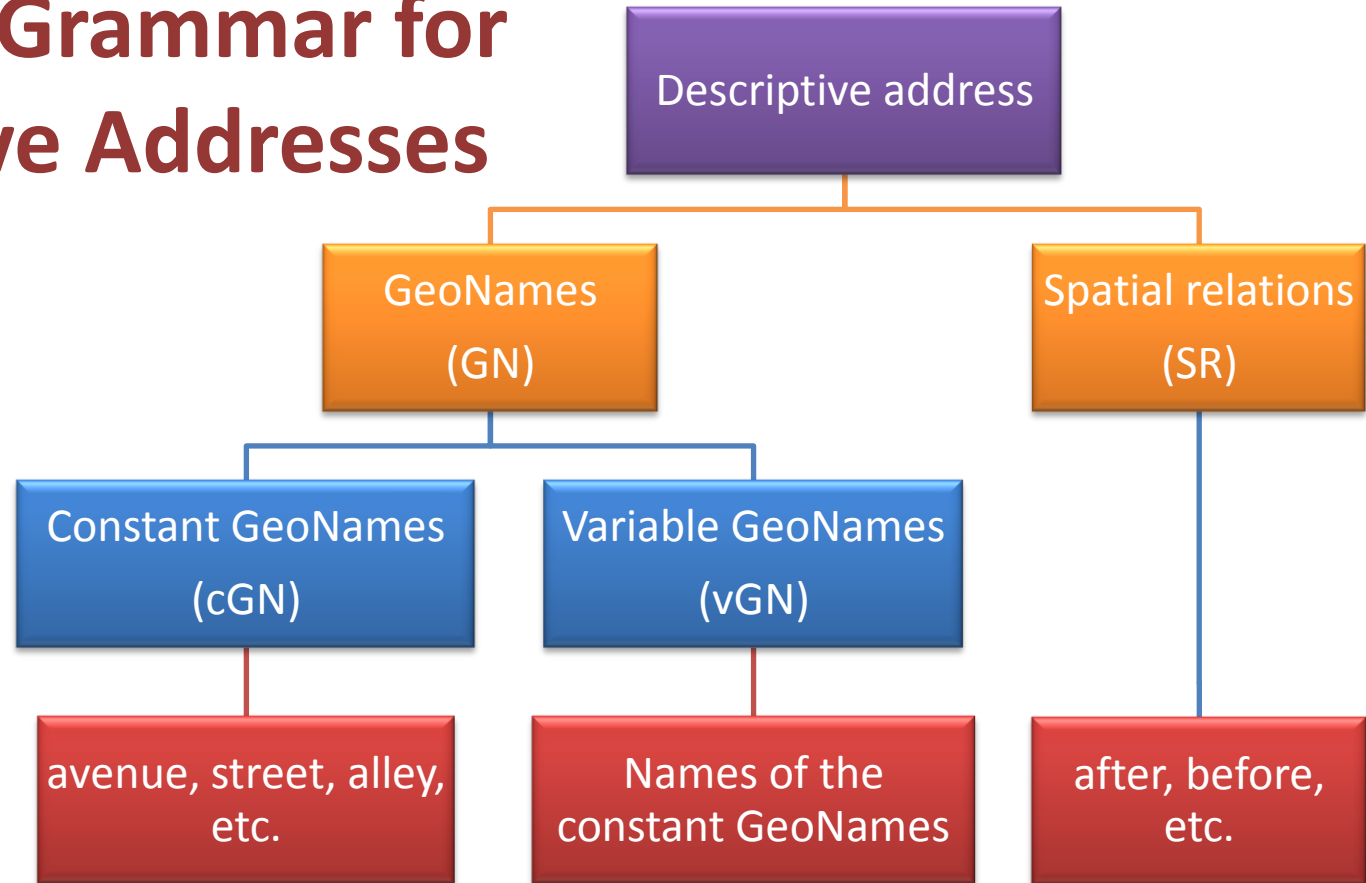
after, before, etc.

1. *S := {SG}*
2. *SG := [SR] GN*
3. *GN := cGN vGN | vGN cGN*
4. *cGN := "avenue" | "ave." | "street" | "st." | "alley" | "number" | "unit"*
5. *SR := "after" | "before" | "in front of" | "left of" | "right of"*

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

8

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

Piroozi avenue, First Street, After Khansari Mosque, Alley 34, Number 22, Unit 12

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

9

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

Descriptive address entered by the user

Constant GeoNames

Spatial Relations

Parsing of the textual address

address parsing

Address Specifications

Constatnt GeoNames

Add Constatnt GeoNames

avenue
ave.
street
st.
alley
number
unit

Delete Selected Item

Spatial Realations

Add Spatial Realations

after
before
in front of
left of
right of

Delete Selected Item

Address Parsing

Address

Piroozi avenue, First street, after Khansari mosque, alley 34, number 22, unit 12

parse the address

clear the address

Spatial Groups
(Address components)

	index	GeoName	Type	Value
	8	avenue	constant geo-name	Piroozi
	22	street	constant geo-name	First
	30	after	spatial relation	Khansari mosque
	53	alley	constant geo-name	34
	63	number	constant geo-name	22
	74	unit	constant geo-name	12
▶*				

clear the parsed address

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

10

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

Conclusion

- Initial results of an ongoing research on developing machine-based matching for addresses expressed in natural languages

- Definition of a formal language - including symbols and combination rules - to parse such descriptive addresses

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

11

LBS 2014
11ᵗʰ Symposium on Location-Based Services
Vienna, 26–28 November 2014

Future Work

- Role of context (user, environment, etc.)

- Natural languages do not completely obey the rules provided by formal languages

- Defining a corpus for the addressing language to capture as many irregularities as possible

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

12

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

Thank you for your attention!

Towards Machine-based Matching of Addresses Expressed in Natural Languages
Farid Karimipour. Ali Javidaneh and Andrew U. Frank

LBS 2014
11th Symposium on Location-Based Services
Vienna, 26–28 November 2014

Towards Machine-based Matching of Addresses Expressed in Natural Languages

Farid Karimipour

fkarimipr@ut.ac.ir

Assistant Professor
Department of Surveying and Geomatics Engineering, University of Tehran, Iran

Ali Javidaneh

ajavidaneh@ut.ac.ir

PhD Student
Department of Surveying and Geomatics Engineering, University of Tehran, Iran

Andrew U. Frank

frank@geoinfo.tuwien.ac.at

Professor
Department of Geodesy and Geoinformation, Technical University of Vienna, Austria