A Context-aware Natural Language Generation Dataset for Dialogue Systems

Ondřej Dušek and Filip Jurčíček

Institute of Formal and Applied Linguistics
Charles University in Prague

May 28, 2016
LREC RE-WOCHAT workshop

Introduction

- A new NLG dataset for dialogue systems
    - English public transport domain

Introduction

- A new NLG dataset for dialogue systems
  - English public transport domain
- “Ordinary” NLG dataset (in our setting):
  - input DA (meaning) + natural language sentence(s)

inform(from_stop="Fulton Street", vehicle=bus, direction="Rector Street", departure_time=9:13pm, line=M21)

Go by the 9:13pm bus on the M21 line from Fulton Street directly to Rector Street

Introduction

- A new NLG dataset for dialogue systems
  - English public transport domain
- “Ordinary” NLG dataset (in our setting):
  - input DA (meaning) + natural language sentence(s)
- Our set:
  - input DA + natural language sentences + preceding context

NEW→*I'm headed to Rector Street*

inform(from_stop="Fulton Street", vehicle=bus, direction="Rector Street", departure_time=9:13pm, line=M21)

Go by the 9:13pm bus on the M21 line from Fulton Street directly to Rector Street

Introduction

- A new NLG dataset for dialogue systems
  - English public transport domain
- “Ordinary” NLG dataset (in our setting):
  - input DA (meaning) + natural language sentence(s)
- Our set:
  - input DA + natural language sentences + preceding context

I'm headed to Rector Street
inform(from_stop="Fulton Street", vehicle=bus, direction="Rector Street", departure_time=9:13pm, line=M21)
Go by the 9:13pm bus on the M21 line from Fulton Street directly to Rector Street

- If the generator knows how the user asked, it should be able to produce a more natural response

Introduction

- A new NLG dataset for dialogue systems
  - English public transport domain
- “Ordinary” NLG dataset (in our setting):
  - input DA (meaning) + natural language sentence(s)
- Our set:
  - input DA + natural language sentences + preceding context

I'm headed to Rector Street
inform(from_stop="Fulton Street", vehicle=bus, direction="Rector Street", departure_time=9:13pm, line=M21)
Heading to Rector Street from Fulton Street, take a bus line M21 at 9:13pm.

- If the generator knows how the user asked, it should be able to produce a more natural response

Outline of this talk

1. Why should we look at preceding context: *entrainment*

Outline of this talk

1. Why should we look at preceding context: *entrainment*
2. How to obtain naturally looking contextual data
   - collecting our set

Outline of this talk

1. Why should we look at preceding context: *entrainment*
2. How to obtain naturally looking contextual data
   - collecting our set
3. A summary of the collected set

Why context? Entrainment

Entrainment/alignment/adaptation in dialogue

Why context? Entrainment

Entrainment/alignment/adaptation in dialogue

- “Mutual linguistic convergence”
  - speakers *primed* (influenced) by previously said

Why context? Entrainment

Entrainment/alignment/adaptation in dialogue

- "Mutual linguistic convergence"
  - speakers *primed* (influenced) by previously said
- Reusing words and syntax

Why context? Entrainment

Entrainment/alignment/adaptation in dialogue

- "Mutual linguistic convergence"
  - speakers *primed* (influenced) by previously said
- Reusing words and syntax
- Occurs naturally, subconscious

how bout the next ride

Sorry, I did not find a later option.

I'm sorry, the next ride was not found.

Why context? Entrainment

Entrainment/alignment/adaptation in dialogue

- "Mutual linguistic convergence"
  - speakers *primed* (influenced) by previously said
- Reusing words and syntax
- Occurs naturally, subconscious

how bout the next ride

Sorry, I did not find a later option.

I'm sorry, the next ride was not found.

what is the distance of this trip

The trip covers a distance of 10.4 miles.

It is around 10.4 miles.

The distance is 10.4 miles.

Why context? Entrainment

Entrainment/alignment/adaptation in dialogue

- "Mutual linguistic convergence"
  - speakers *primed* (influenced) by previously said
- Reusing words and syntax
- Occurs naturally, subconscious
- Found to help dialogue success (Friedberg et al. '12)

Why context? Entrainment

Entrainment/alignment/adaptation in dialogue

- "Mutual linguistic convergence"
  - speakers *primed* (influenced) by previously said
- Reusing words and syntax
- Occurs naturally, subconscious
- Found to help dialogue success (Friedberg et al. '12)

Entrainment in dialogue systems

- Several experiments, successful
  (Lopes et al. '13, '15; He et al. '14)
- Limited, partially or completely rule-based

What we aim for

- Fully trainable NLG that allows entrainment
  - let the system adapt to users' words *and* syntax
  - let the data handle the rules

What we aim for

- Fully trainable NLG that allows entrainment
  - let the system adapt to users’ words *and* syntax
  - let the data handle the rules
- We hope for:
  - more natural system responses
  - possibly higher task success
  - applicability to other domains, chat-based systems

What we aim for

- Fully trainable NLG that allows entrainment
  - let the system adapt to users' words *and* syntax
  - let the data handle the rules
- We hope for:
  - more natural system responses
  - possibly higher task success
  - applicability to other domains, chat-based systems
- We need training data

What we aim for

- Fully trainable NLG that allows entrainment
  - let the system adapt to users' words *and* syntax
  - let the data handle the rules
- We hope for:
  - more natural system responses
  - possibly higher task success
  - applicability to other domains, chat-based systems
- We need training data
- …that is why we collected this dataset!

Our domain

- Alex English SDS for NYC public transport
  - https://github.com/UFAL-DSG/alex

Our domain

- Alex English SDS for NYC public transport
  - https://github.com/UFAL-DSG/alex
- Bus/subway services on Manhattan
  - Alex can do more, limited just for this set

Our domain

- Alex English SDS for NYC public transport
  - https://github.com/UFAL-DSG/alex
- Bus/subway services on Manhattan
  - Alex can do more, limited just for this set
- Users ask for a schedule, may request details/modify search

Our domain

- Alex English SDS for NYC public transport
  - https://github.com/UFAL-DSG/alex
- Bus/subway services on Manhattan
  - Alex can do more, limited just for this set
- Users ask for a schedule, may request details/modify search
- 13 slots
  - from_stop, to_stop
  - departure_time
  - vehicle
  - duration
  - …

Collecting the set

Getting natural utterances cheap and fast

- Using crowdsourcing (CrowdFlower)

Collecting the set

Getting natural utterances cheap and fast

- Using crowdsourcing (CrowdFlower)

Addressing data sparsity

- Delexicalization (places, times etc. $\rightarrow$ “X”)
  - both context and response

Collecting the set

Getting natural utterances cheap and fast

- Using crowdsourcing (CrowdFlower)

Addressing data sparsity

- Delexicalization (places, times etc. $\rightarrow$ “X”)
  - both context and response
- Limiting context to previous sentence
  - likely to have the strongest entrainment impact

Collecting the set

Getting natural utterances cheap and fast

- Using crowdsourcing (CrowdFlower)

Addressing data sparsity

- Delexicalization (places, times etc. $\rightarrow$ “X”)
  - both context and response
- Limiting context to previous sentence
  - likely to have the strongest entrainment impact

Collection progress

Collecting the set

Getting natural utterances cheap and fast

- Using crowdsourcing (CrowdFlower)

Addressing data sparsity

- Delexicalization (places, times etc. $\rightarrow$ "X")
  - both context and response
- Limiting context to previous sentence
  - likely to have the strongest entrainment impact

Collection progress

1. Get natural user utterances in calls to a live dialogue system

Collecting the set

Getting natural utterances cheap and fast

- Using crowdsourcing (CrowdFlower)

Addressing data sparsity

- Delexicalization (places, times etc. $\rightarrow$ “X”)
  - both context and response
- Limiting context to previous sentence
  - likely to have the strongest entrainment impact

Collection progress

1. Get natural user utterances in calls to a live dialogue system
2. Generate response DA

Collecting the set

Getting natural utterances cheap and fast

- Using crowdsourcing (CrowdFlower)

Addressing data sparsity

- Delexicalization (places, times etc. $\rightarrow$ “X”)
  - both context and response
- Limiting context to previous sentence
  - likely to have the strongest entrainment impact

Collection progress

1. Get natural user utterances in calls to a live dialogue system
2. Generate response DA
3. Collect natural language paraphrases

Getting natural user requests

1. Record calls to live Alex SDS
   - assign tasks to people on CrowdFlower

You want a connection – your departure stop is *Marble Hill*, and you want to go to *Roosevelt Island*. Ask how long the journey will take. Ask about a schedule afterwards. Then modify your query: Ask for a ride at six o'clock in the evening. Ask for a connection by bus. Do as if you changed your mind: Say that your destination stop is *City Hall*.

You are searching for transit options leaving from *Houston Street* with the destination of *Marble Hill*. When you are offered a schedule, ask about the time of arrival at your destination. Then ask for a connection after that. Modify your query: Request information about an alternative at six p.m. and state that you prefer to go by bus.

Tell the system that you want to travel from *Park Place* to *Inwood*. When you are offered a trip, ask about the time needed. Then ask for another alternative. Change your search: Ask about a ride at 6 o'clock p.m. and tell the system that you would rather use the bus.

Getting natural user requests

1. Record calls to live Alex SDS
   - assign tasks to people on CrowdFlower
   - varying synonyms in task description
   - people unaware that wording is important

Getting natural user requests

1. Record calls to live Alex SDS
   - assign tasks to people on CrowdFlower
   - varying synonyms in task description
   - people unaware that wording is important
2. Manually transcribe on CrowdFlower

Getting natural user requests

1. Record calls to live Alex SDS
    - assign tasks to people on CrowdFlower
    - varying synonyms in task description
    - people unaware that wording is important
2. Manually transcribe on CrowdFlower
3. Parse using Alex handcrafted SLU
    - parsing transcriptions gives better results than ASR n-best lists

Generating response DA

- Handcrafted simple rule-based bigram policy

Generating response DA

- Handcrafted simple rule-based bigram policy
- All possible replies for a single context utterance

what about a connection by bus

iconfirm(vehicle=bus)

inform(from_stop="Dyckman Street", direction="Park Place",
vehicle=bus, line=M103, departure_time=7:05pm)

inform_no_match(vehicle=bus)

request(to_stop)

Generating response DA

- Handcrafted simple rule-based bigram policy
- All possible replies for a single context utterance
  - confirmation
  - answer
  - apology
  - request for additional information

what about a connection by bus

iconfirm(vehicle=bus)

inform(from_stop="Dyckman Street", direction="Park Place",
 vehicle=bus, line=M103, departure_time=7:05pm)

inform_no_match(vehicle=bus)

request(to_stop)

Generating response DA

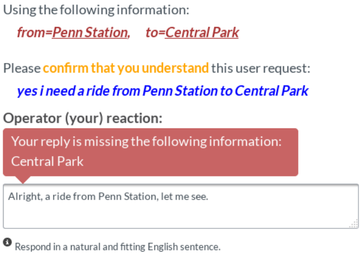
- Handcrafted simple rule-based bigram policy
- All possible replies for a single context utterance
  - confirmation
  - answer
  - apology
  - request for additional information
- In a real dialogue, the correct reply would depend on longer history, but here we try them all

Collecting natural language responses

- CrowdFlower interface

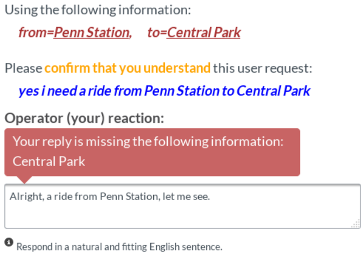
Collecting natural language responses

- CrowdFlower interface



Collecting natural language responses

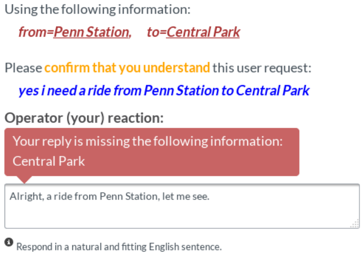
- CrowdFlower interface



- Context displayed at hand

Collecting natural language responses

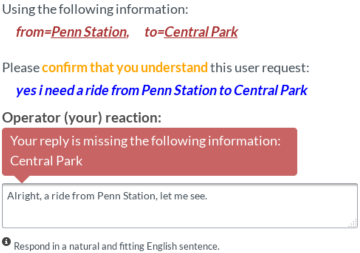
- CrowdFlower interface



- Context displayed at hand
- Minimal slot name description

Collecting natural language responses

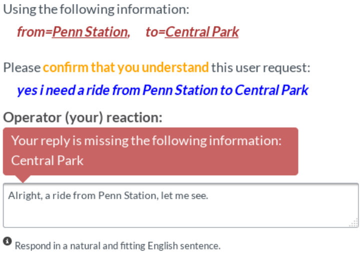
- CrowdFlower interface



- Context displayed at hand
- Minimal slot name description
- Short instructions

Collecting natural language responses

- CrowdFlower interface



- Context displayed at hand
- Minimal slot name description
- Short instructions
- Checks: contents, spelling; automatic + manual
  - ca. 20% overhead (repeated submissions)

Dataset summary

Size

total response paraphrases	5,577
unique (delex.) context + response DA	1,859
unique (delex.) context	552
unique (delex.) context with min. 2 occurrences	119
unique response DA	83
unique response DA types	6
unique slots	13

Entrainment

Syntactic	$\sim$59%
Lexical	$\sim$31%
Both	$\sim$19%

- subjective, based on word & phrase reuse, word order, pronouns

Thank you for your attention

Dataset available for download

- JSON + CSV
- CC BY-SA 4.0
- GitHub: `bit.ly/nlgdata` (link given in the paper)

Contact us

Ondřej Dušek & Filip Jurčíček
Charles University in Prague
`odusek@ufal.mff.cuni.cz`

References

Friedberg et al. (2012). Lexical entrainment and success in student engineering groups. *SLT*, pp. 404–409.

Hu et al. (2014). Entrainment in pedestrian direction giving: How many kinds of entrainment. *IWSDS*, pp. 90–101.

Lopes et al. (2013). Automated two-way entrainment to improve spoken dialog system performance. *ICASSP*, pp. 8372–8376.

Lopes et al. (2015). From rule-based to data-driven lexical entrainment models in spoken dialog systems. *Computer Speech & Language*, 31(1):87–112.