Natural Language Processing 1

Live Q & A: Morphology and syntax

Katia Shutova

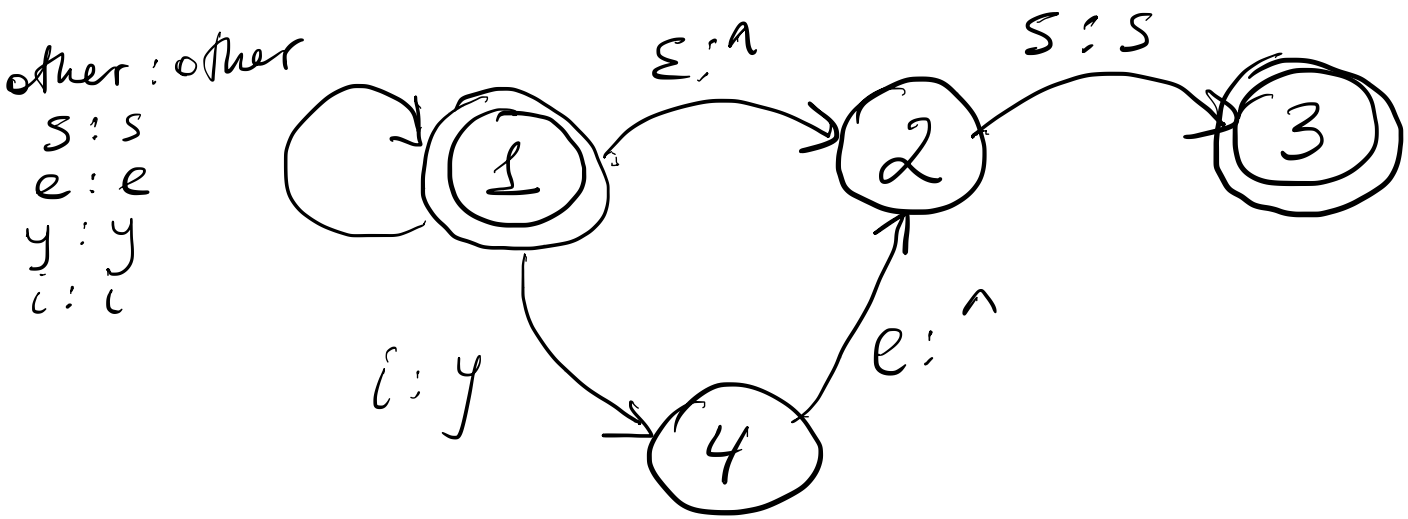
ILLC
University of Amsterdam

11 November 2020

Outline.

Another example of FST

party –> party
parties –> party^s
parts –> part^s



Question 1: Morphology

Split the following words into morphological units, labelling each as stem, suffix or prefix. If there is any ambiguity, give all possible splits.

1. dries
2. cartwheel
3. uncaring
4. intruders
5. reattaches
6. anticipated
7. feed

Question 2: FSTs for morphology

Circumfixes in German: *ge-* + *-t* for past participle

stem	surface	underlying
kauf	gekauft	kauf^P
arbeit	gearbeitet	arbeit^P

1. How would you design an FST to handle this?
2. What does this example illustrate about limitations of the FST approach?

Question(s) 3: Discussion

1. What sources of information can you use to PoS-tag unknown words?
2. How are n-gram language modelling, PoS tagging and probabilistic syntactic parsing methods affected by the differences in genre? Which of the three tasks is more sensitive to this difference and why?
3. For the challenges and applications listed in lecture 1, which problems can n-gram language modelling, PoS tagging and syntactic parsing solve?

Question 4: Context-free grammars

rules

```
S -> NP VP
VP -> VP PP
VP -> V
VP -> V NP
VP -> V VP
NP -> NP PP
PP -> P NP
```

lexicon

```
V -> can
V -> fish
NP -> fish
NP -> rivers
NP -> pools
NP -> December
NP -> Scotland
NP -> it
NP -> they
P -> in
```

Question 4: Context-free grammars

How would you modify this CFG to handle the following examples?

They can fish early.

Unexpectedly, they can fish.

Question 5: FSA and syntax

Why can't we use FSA to model syntax?

Provide two reasons.