

LING/C SC/PSYC 438/538

Lecture 14

Prof. Sandiway Fong

Today's Topics

- Ungraded Homework
- Predicate-Argument Structure *contd.*
 - Framenet: *verbs and their semantic roles*
 - Stanford CoreNLP: *dependency parses*
 - (seems no longer in operation)
<http://nlp.stanford.edu:8080/corenlp/>
 - *For Python-savvy users, can run it locally as:*
 - `pip3 install stanza`
 - `python3`
 - `>>> import stanza`
 - `>>> stanza.install_corenlp()`
 - `https://stanfordnlp.github.io/stanza/client\_usage.html`

Ungraded regex Homework

Exercises (from the textbook)

If you'd like
a bit more
practice

...

2.1 Write regular expressions for the following languages. You may use either Perl/Python notation or the minimal “algebraic” notation of Section 2.3, but make sure to say which one you are using. By “word”, we mean an alphabetic string separated from other words by whitespace, any relevant punctuation, line breaks, and so forth.

1. the set of all alphabetic strings;
2. the set of all lower case alphabetic strings ending in a *b*;
3. the set of all strings with two consecutive repeated words (e.g., “Humbert Humbert” and “the the” but not “the bug” or “the big bug”);
4. the set of all strings from the alphabet *a, b* such that each *a* is immediately preceded by and immediately followed by a *b*;
5. all strings that start at the beginning of the line with an integer and that end at the end of the line with a word;
6. all strings that have both the word *grotto* and the word *raven* in them (but not, e.g., words like *grottos* that merely *contain* the word *grotto*);
7. write a pattern that places the first word of an English sentence in a register. Deal with punctuation.

Background: Framenet: *drive*

Operate_vehicle

[Lexical Unit 1](#)

Definition:

The words in this frame describe motion involving a **Vehicle** and someone who controls it, the **Driver**. Some words normally allow the **Vehicle** to be expressed as a separate constituent.

Tim **DROVE** **his car** **all the way across North America**.

Tom **PADDLED** **my canoe** **across the Canadian border**.

Other words in this domain are based on the names of vehicles, and do not normally allow the **Vehicle** to be expressed as a separate constituent.

The group **BIKED** **all the way across the country**.

However, a separate **Vehicle** constituent can occur if it adds information not included in the verb.

Tim **BIKED** **across the country** **on an old 10-speed**.

Background: Framenet: *drive*

Core:

Area [Area]

This frame element is used for expressions which describe a general area in which motion takes place when the motion is understood to be irregular and not to consist of a single linear path. Locative setting adjuncts of motion expressions may also be assigned this frame element.

The bikers **RODE** *all over the place*.

Driver [Drive]

Semantic Type: Sentient

This is the being, typically human, that controls the **Vehicle** as it moves.

Kim **DROVE** my old car cross-country.

Goal [Goal]

Semantic Type: Goal

Excludes: Area

Any expression which tells where the moving object(s) ends up as a result of the motion expresses the frame element **Goal**. Some particles imply the existence of a **Goal** which is understood in the context of utterance.

Kim **DROVE** *into the parking lot*.

Kim **DROVE** *in*.

Background: Framenet: *drive*

Path [Path]

Excludes: Area

Any description of a trajectory of motion which is neither a **Source** nor a **Goal** expresses the frame element **Path**. This includes directional expressions and "middle of path" expressions, e.g.:
The bikers **RODE** **through the desert**.

Source [Src]

Semantic Type: Source

Excludes: Area

Any expression which implies a definite starting-point of motion expresses the frame element **Source**. In prepositional phrases, the prepositional object expresses the starting point of motion. With particles, the starting point of motion is understood from context.

Kim **DROVE** **out of town**.

Kim **DROVE** **away**.

Vehicle [Veh]

This is the means of conveyance controlled by the **Driver**. It may move in any way or medium. With verbs, **Vehicle** is usually expressed as Object:

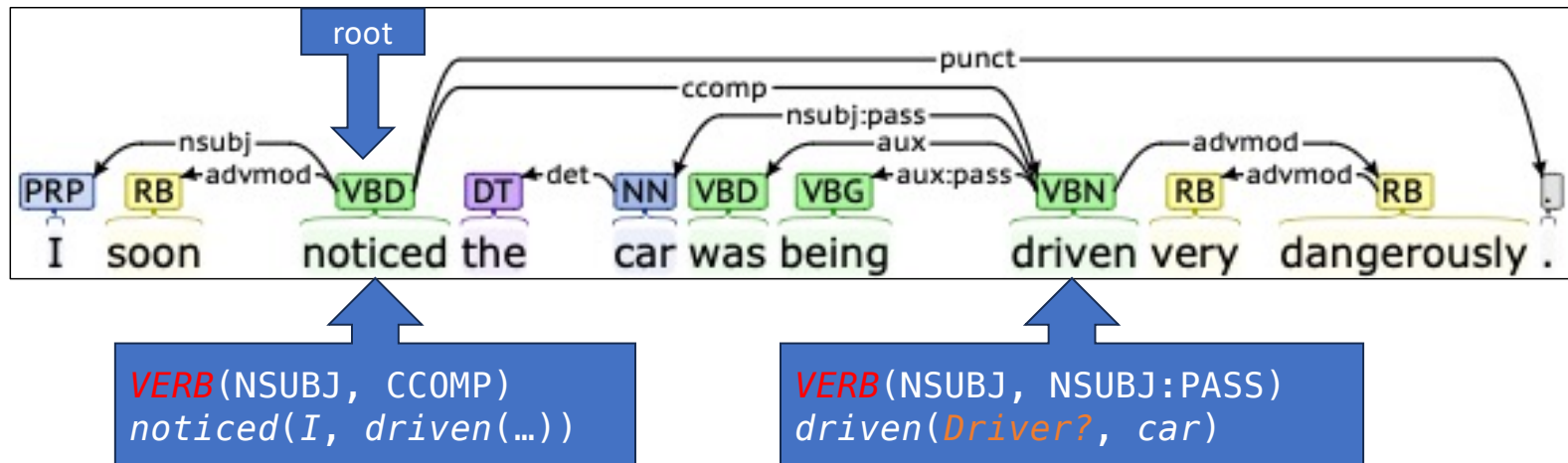
I **DROVE** **my car** to Canada.

Background: CoreNLP

- Examples (from Framenet):

1. [Cognizer]I soon **NOTICED** [Phenomenon the car was being driven very dangerously] .

- Verb (VBD) = noticed, Cognizer = nsubj, Phenomenon = ccomp
- Verb(VBN) = driven, Vehicle = nsubj:pass



ChatGPT



I soon noticed the car was being driven very dangerously. Who drove the car?



I'm sorry, but I don't have enough information to answer your question. Could you please provide more context or details about the situation, such as who was present in the car or any other relevant information?



ChatGPT

I soon noticed the car was being driven very dangerously.
Who drove the car?

The sentence **doesn't say who drove the car.**

Where am I?

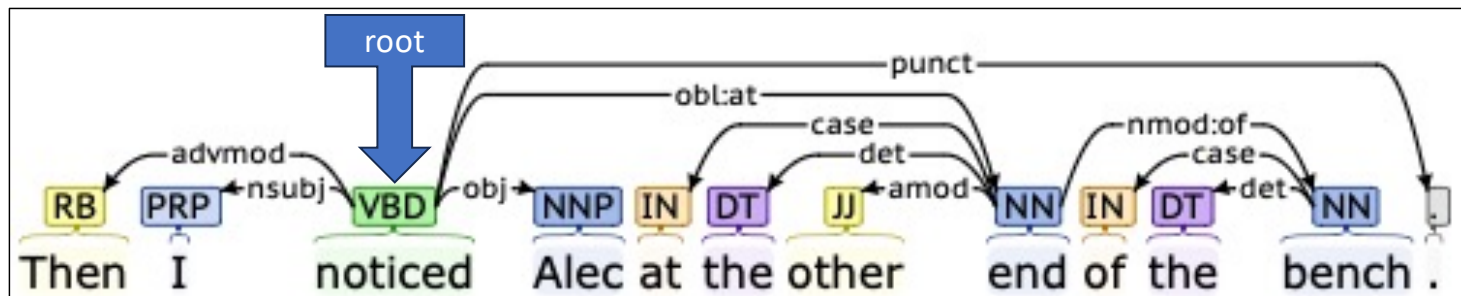
From that sentence alone, **we can't determine where you are.**

You could be **inside the car, another vehicle, on the street, or observing the car from somewhere else.** The sentence only tells us that you noticed the car being driven dangerously.

Background: CoreNLP

- Examples (from Framenet):

2. Then [Cognizer] **NOTICED** [Phenomenon] Alec [Ground] at the other end of the bench .



VERB(NSUBJ, OBJ)
noticed(I, Alec)

*event(e, noticed) &
cognizer(e, I) &
phenomenon(e, Alec) &
ground(e, other end of the bench)*

ChatGPT

 SA

Then I noticed Alec at the other end of the bench. Where was I?



Based on the information provided, you were likely at the same bench where you noticed Alec but at a different end of it.



You're **sitting on a bench, at the opposite end from Alec.** GPT-5.6 Luna.

Background: Stanford Dependencies

- Some definitions you may find useful
https://nlp.stanford.edu/software/dependencies_manual.pdf
 - **ccomp**: clausal complement
A clausal complement of a verb or adjective is a dependent clause
 - **dobj**: direct object
The direct object of a VP is the noun phrase which is the (accusative) object of the verb.
 - **iobj**: indirect object
The indirect object of a VP is the noun phrase which is the (dative) object of the verb.
 - **nsubj**: nominal subject
A nominal subject is a noun phrase which is the syntactic subject of a clause.
 - **rcmod**: relative clause modifier
A relative clause modifier of an NP is a relative clause modifying the NP. The relation points from the head noun of the NP to the head of the relative clause, normally a verb.
 - **vmod**: reduced non-finite verbal modifier
A reduced non-finite verbal modifier is a participial or infinitive form of a verb heading a phrase (which may have some arguments, roughly like a VP).

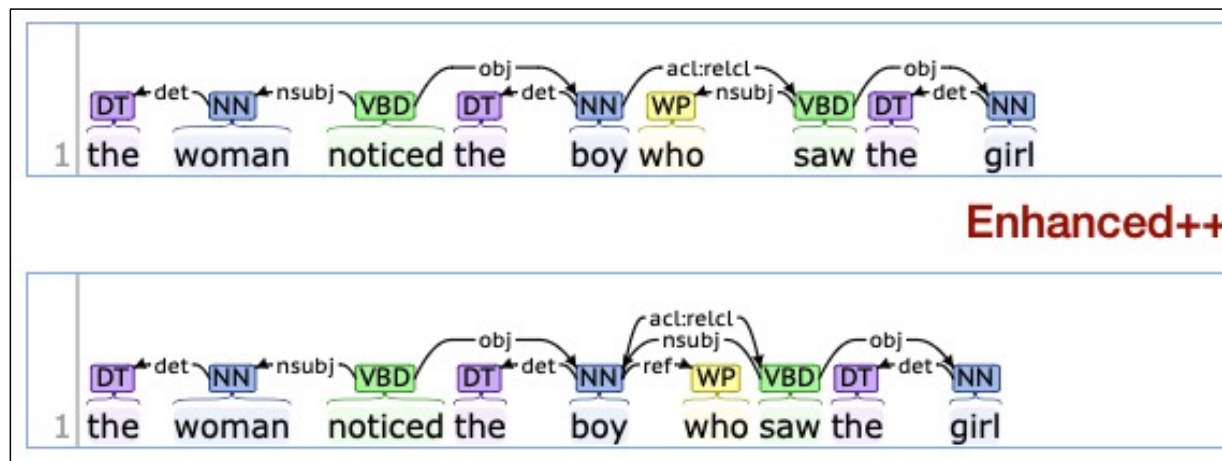
Background: Universal Dependencies

<https://universaldependencies.org/u/dep/index.html>

	Nominals	Clauses	Modifier words	Function Words
Core arguments	nsubj obj iobj	csubj ccomp xcomp		
Non-core dependents	obl vocative expl dislocated	advcl	advmod* discourse	aux cop mark
Nominal dependents	nmod appos nummod	acl	amod	det clf case
Coordination	MWE	Loose	Special	Other
conj cc	fixed flat compound	list parataxis	orphan goeswith reparandum	punct root dep

Background: CoreNLP

- Root: noticed(woman, boy)
- ACL:RELCL points back to NOUN boy
- ACL:RELCL/**VERB**(NSUBJ/PRON, OBJ)
- We infer saw(boy, girl)



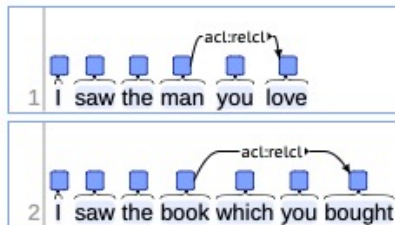
Background: Universal Dependencies

- `acl` = adnominal clause (basically, a sentence that modifies a noun)

`acl:relcl`: relative clause modifier

A relative clause modifier of a nominal is a clause that modifies the nominal, whereas the nominal is coreferential with a constituent inside the relative clause (here the constituent may be realized as a relative pronoun, another relative word, or it may not be overtly realized at all). The `acl:relcl` relation points from the head of the modified nominal to the head of the relative clause.

Depending on language, it may be required that relative clauses are finite. For example, English non-finite clauses are traditionally not termed relative; therefore, *the girl **that was born today*** is a relative clause because it is finite, while *the girl **born today*** is non-finite (the participle is not accompanied by a finite auxiliary) and it uses the plain `acl` relation. In other languages however, the distinction between finite and non-finite clauses may not exist or may not be used as a criterion for relative clauses.



Subject relative clauses and Perl

- Examples:

- 1. *the man noticed the girl*

- noticed(man, girl)

- 2. *the man noticed the girl who noticed the boy*

- noticed(man, girl)

- noticed(girl, boy)

- Perl regex patterns:

- /the (\w+) (\w+) the (\w+)/ ➡ print "\$2(\$1, \$3)"

- /the (\w+) who (\w+) the (\w+)/ ➡ print "\$2(\$1, \$3)"

Subject relative clauses and Perl

- Note the regex overlap:

the man noticed the girl who noticed the boy who found the person
the (\w+)(\w+) the (\w+)
the (\w+) who (\w+) the (\w+)
the (\w+) who (\w+) the (\w+)



- Recall regex matching goes from left to right (*keeping track of a pointer*)
- We want to iterate this matching using the g (global) flag
- One solution: make the overlapping part a **lookahead** (so the pointer is not advanced):
 - i.e. `(?=the (\w+))`

Subject relative clauses and Perl

- Began with two regex patterns:

- `/the (\w+) (\w+) the (\w+)/` ➡ `print "$2($1, $3)"`
- `/the (\w+) who (\w+) the (\w+)/` ➡ `print "$2($1, $3)"`

- One regex pattern:

- `/the (\w+) (who )?(\w+) the (\w+)/`

You can also use a non- capturing group
(?:*regexp*)

- With lookahead:

- `/the (\w+) (who )?(\w+) (?=the (\w+))/`

- Global match using a while loop:

- `while (/the (\w+) (who )?(\w+) (?=the (\w+))/g) {`
- `print "$3($2, $4)"`
- `}`

Subject relative clauses and Perl

- Examples:

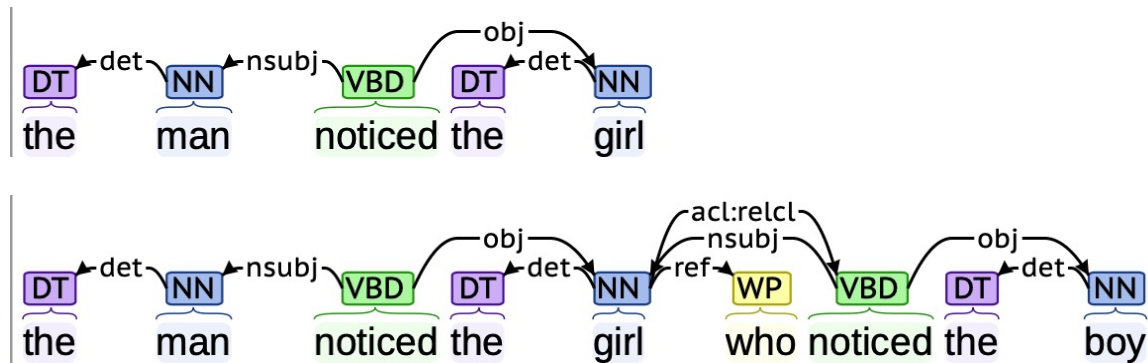
```
$ perl -le '$_ = qq/@ARGV/; while (/the (\w+) (who )?(\w+) (?:the (\w+))/g){  
print "$3($1, $4)"}' the man noticed the girl  
noticed(man, girl)
```

```
$ perl -le '$_ = qq/@ARGV/; while (/the (\w+) (who )?(\w+) (?:the (\w+))/g){  
print "$3($1, $4)"}' the man noticed the girl who noticed the boy  
noticed(man, girl)  
noticed(girl, boy)
```

```
$ perl -le '$_ = qq/@ARGV/; while (/the (\w+) (who )?(\w+) (?:the (\w+))/g){  
print "$3($1, $4)"}' the man noticed the girl who noticed the boy who found the  
person  
noticed(man, girl)  
noticed(girl, boy)  
found(boy, person)
```

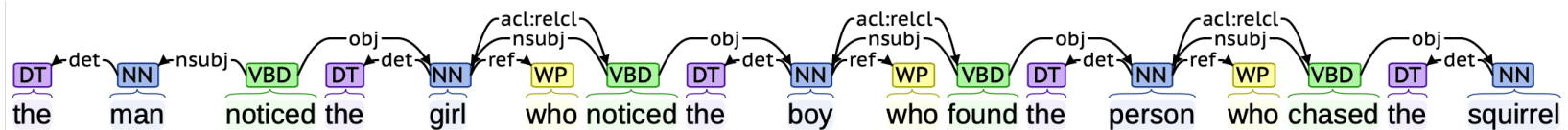
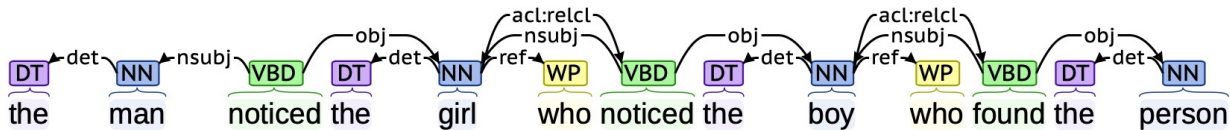
CoreNLP

1. the man noticed the girl
2. the man noticed the girl who noticed the boy



CoreNLP

3. the man noticed the girl who noticed the boy who found the person
4. the man noticed the girl who noticed the boy who found the person who chased the squirrel



Object relative clauses and Perl

Examples:

5. *the man discovered the girl*

- discovered(man, girl)

6. *the man discovered the girl the boy noticed*

- discovered(man, girl)
- noticed(boy, girl)

7. *the man discovered the girl the boy the man found noticed*

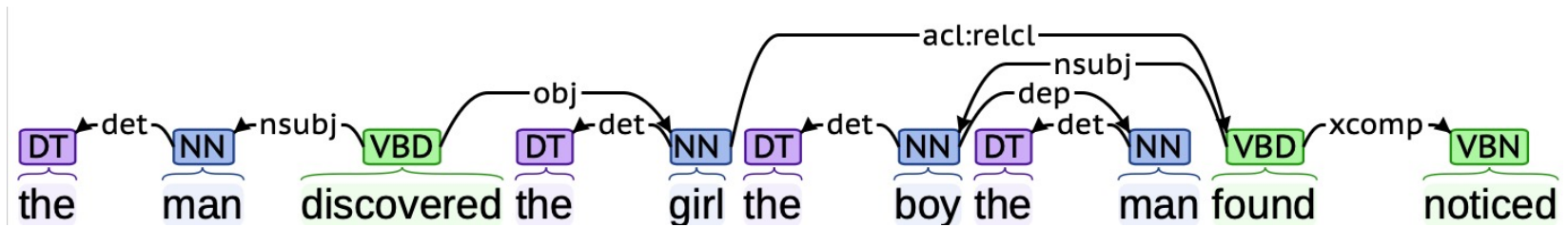
- discovered(man, girl)
- noticed(boy, girl)
- found(man, boy)

8. *the man discovered the girl the boy the man the dog chased found noticed*

- ...
- chased(dog, man)

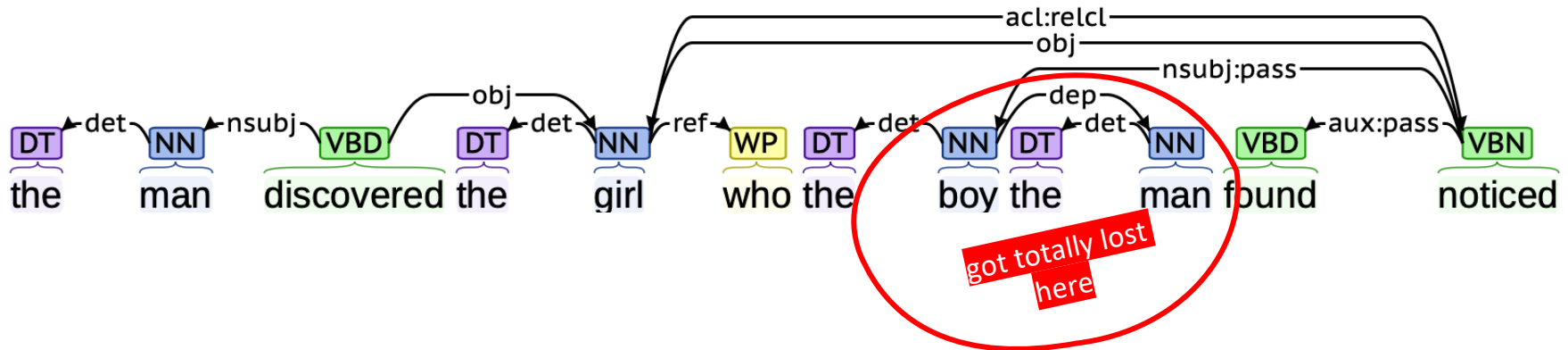
CoreNLP

7. *the man discovered the girl the boy the man found noticed*



CoreNLP

7. (7') *the man discovered the girl who the boy the man found noticed*

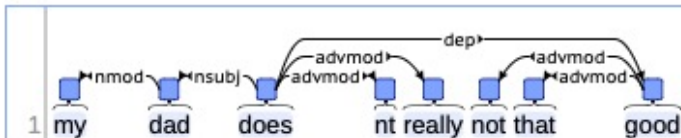


Background: Universal Dependencies

- *i.e. when we have no clue what's going on!*

dep: unspecified dependency

A dependency can be labeled as **dep** when it is impossible to determine a more precise relation. This may be because of a weird grammatical construction, or a limitation in conversion or parsing software. The use of **dep** should be avoided as much as possible.



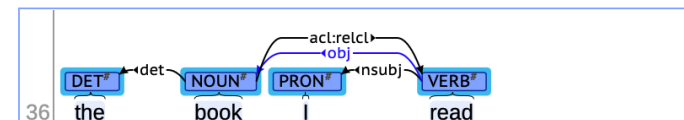
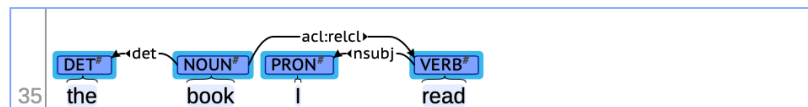
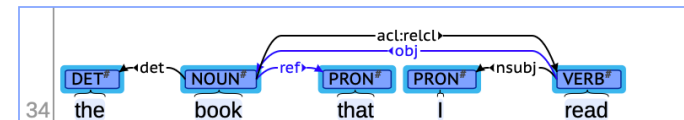
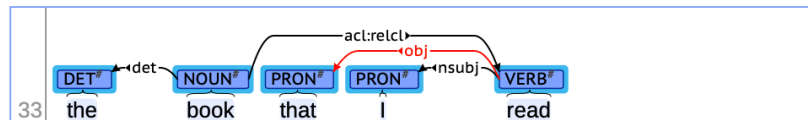
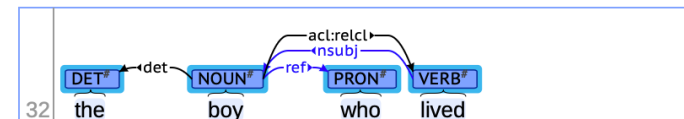
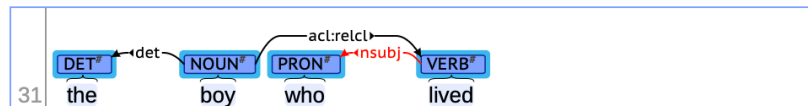
Background: Universal Dependencies: *ref*

Relative clauses

In *basic* trees, relative pronouns are attached to the main predicate of the relative clause (typically with a *nsubj* or *obj* relation). In the corresponding *enhanced* graphs, the relative pronoun is attached to its antecedent with the special *ref* relation and the antecedent is attached as a dependent of the node that is the parent of the relative pronoun in the basic tree. Typically this parent is the main predicate of the relative clause, but it is not always so (see examples below).

In the case where there is no explicit relative pronoun, there is no *ref* relation in the enhanced graph but the antecedent is still annotated as a dependent of a node in the relative clause, depending on the role it plays in the relative clause.

Note that such graphs contain a cycle.



CoreNLP

8. *the man discovered the girl the boy the man the dog chased found noticed*

