Intro to SRL

Semantic Role Labeling

Wz -10/04/21

In 2002, Donald Trump, in partnership with NBC television, took control of the Miss Universe Organization, which produces the Miss Universe, Miss USA and Miss Teen USA contests.

What would be the main idea?

Who did what to whom (when and where)

Who

In 2002, Donald Trump, in partnership

Did what

with NBC television, took control of the

Whom

Miss Universe Organization, which produces the Miss Universe, Miss USA and Miss Teen USA contests.

Donald Trump took control of the MUO

Who
Argument/Agent

Did what
Predicate

Whom
Argument/Patient

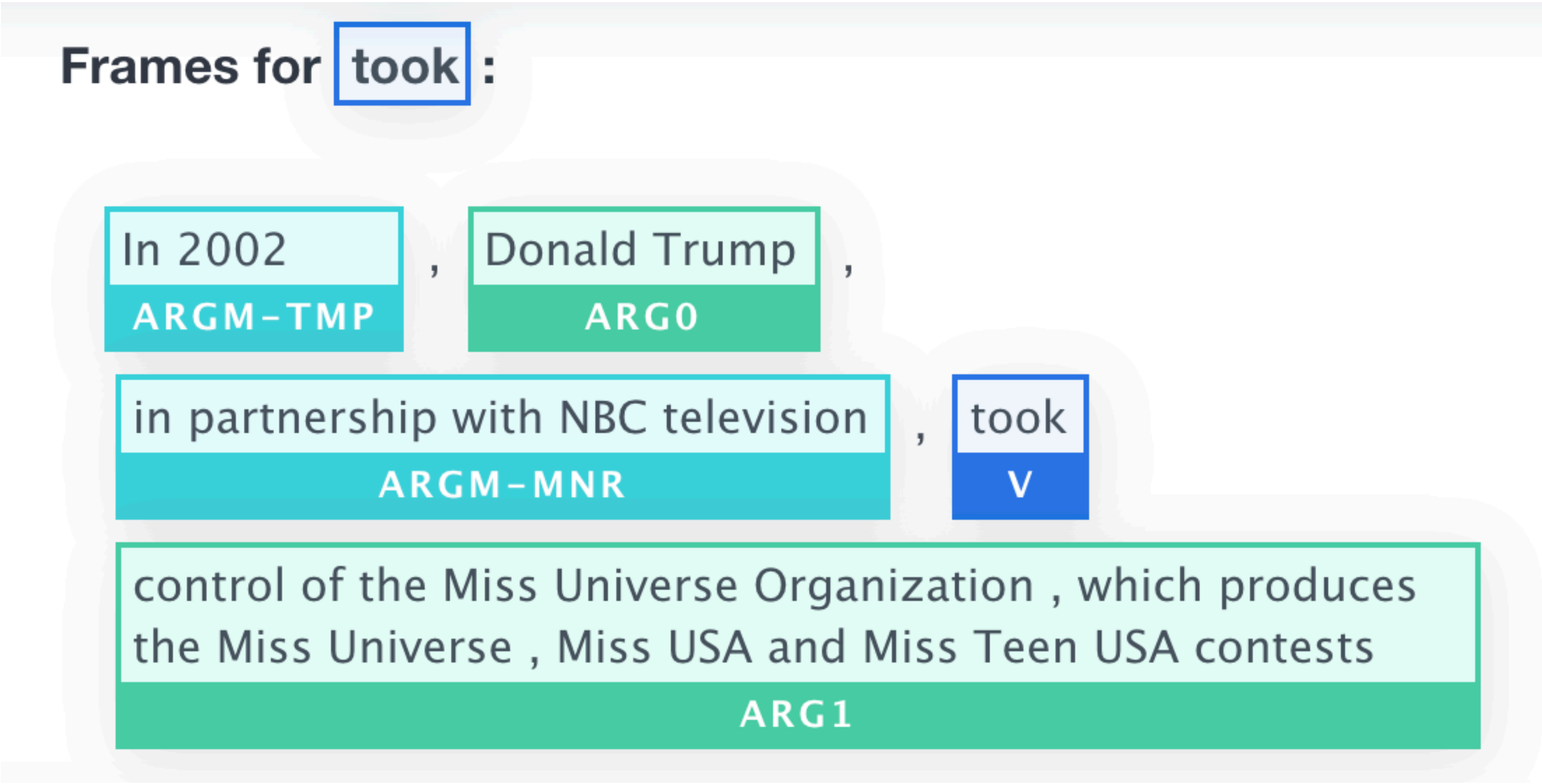
Semantic role (abstract)

(Argument/Agent Predicate Argument/Patient)

In practice (AllenNLP)

https://demo.allennlp.org/semantic-role-labeling

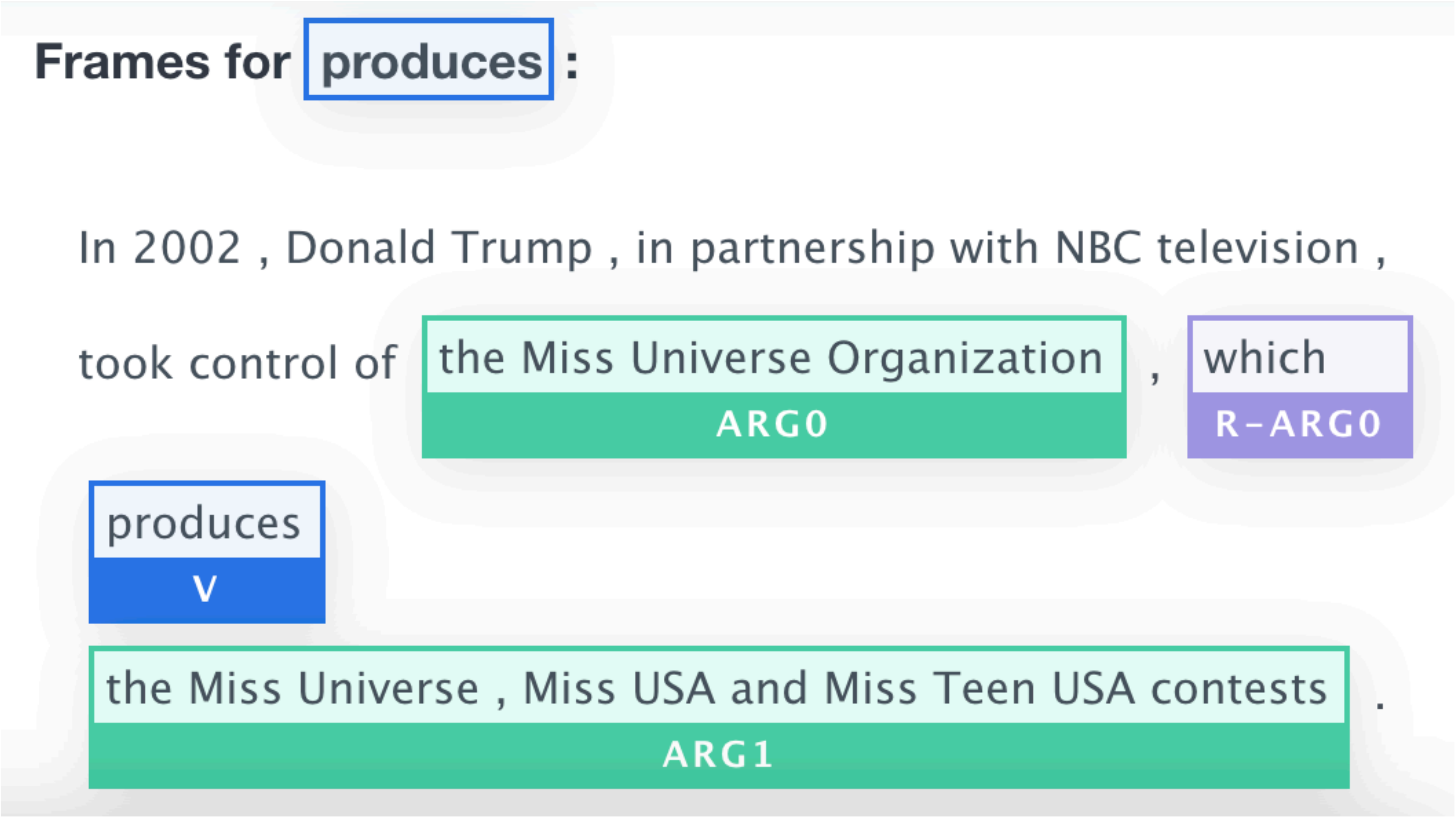
In 2002, Donald Trump, in partnership with NBC television, took control of the Miss Universe Organization, which produces the Miss Universe, Miss USA and Miss Teen USA contests.



(Argument/Agent Predicate Argument/Patient)

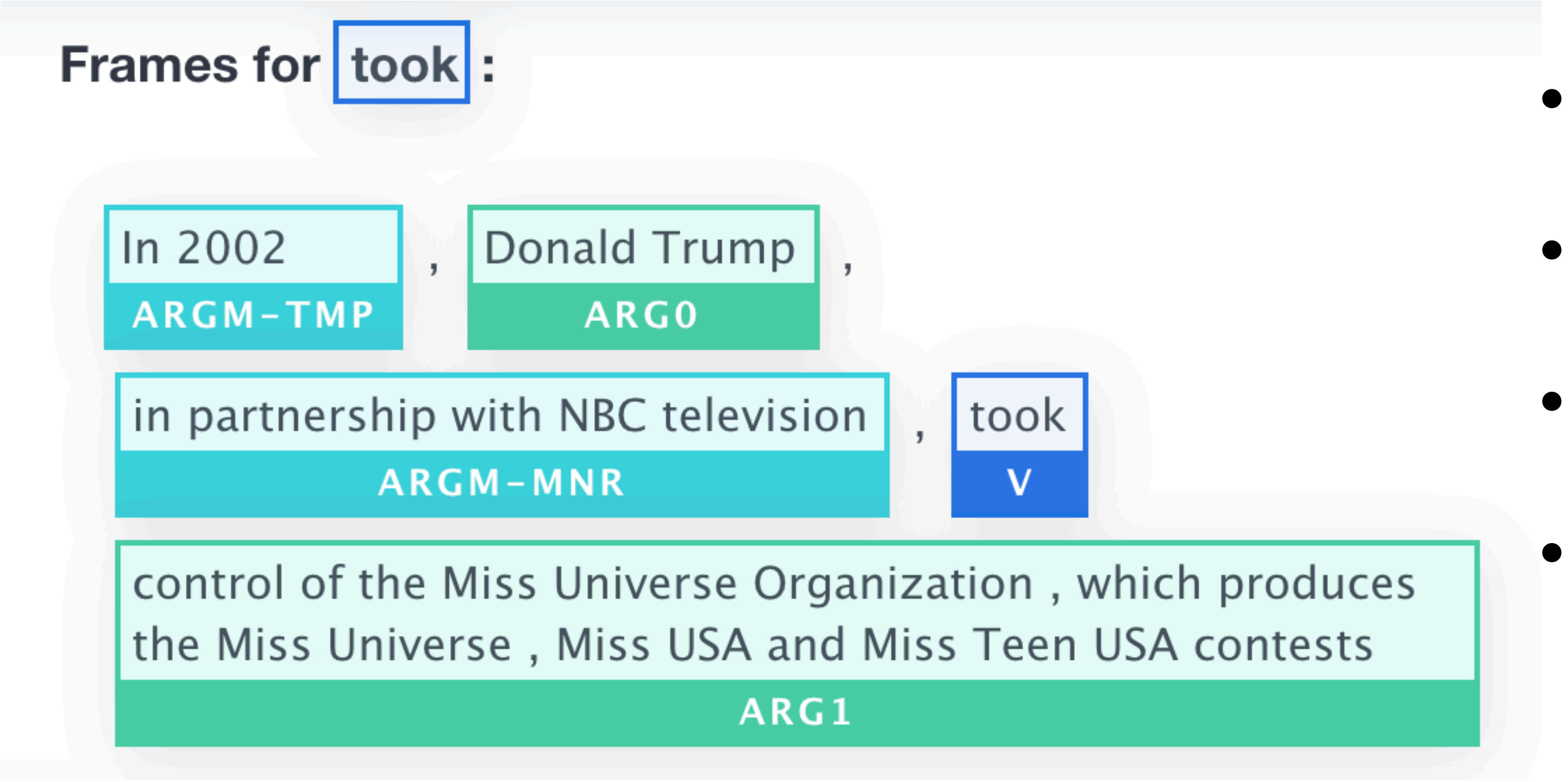
In practice (AllenNLP)

In 2002, Donald Trump, in partnership with NBC television, took control of the **Miss Universe Organization**, which **produces** the **Miss Universe, Miss USA and Miss Teen USA contests.**



(Argument/Agent Predicate Argument/Patient)

In practice (AllenNLP)

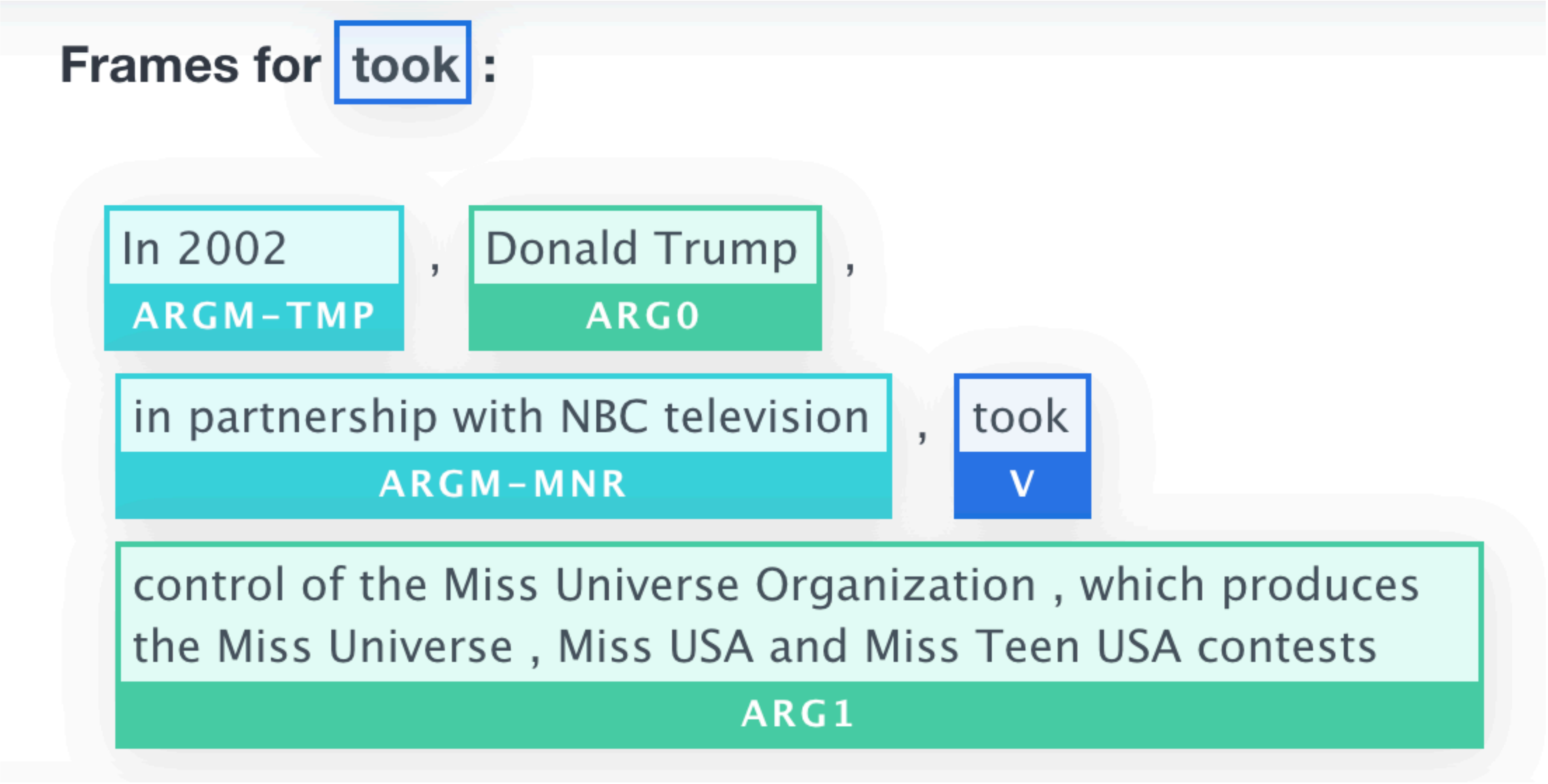


Ex1: [$_{\text{Arg1}}$ Sales] *fell* [$_{\text{Arg4}}$ to \$25 million] [$_{\text{Arg3}}$ from \$27 million].
Ex2: [$_{\text{Arg1}}$ The average junk bond] *fell* [$_{\text{Arg2}}$ by 4.2%].

- Adopted from **PropBank**
- Arg0 : Agent
- Arg1 : Patient
- Arg2 : Benefactive, instrument, attribute or end state
- Arg3 : start point, benefactive, instrument, or attribute
- Arg4 : the end point

(Argument/Agent Predicate Argument/Patient)

In practice (AllenNLP)



TMP	when?	yesterday evening, now
LOC	where?	at the museum, in San Francisco
DIR	where to/from?	down, to Bangkok
MNR	how?	clearly, with much enthusiasm
PRP/CAU	why?	because ... , in response to the ruling
REC		themselves, each other
ADV	miscellaneous	
PRD	secondary predication	...ate the meat raw

Ex1: $[_{\text{Arg1}}$ Sales] *fell* $[_{\text{Arg4}}$ to \$25 million] $[_{\text{Arg3}}$ from \$27 million].

Ex2: $[_{\text{Arg1}}$ The average junk bond] *fell* $[_{\text{Arg2}}$ by 4.2%].

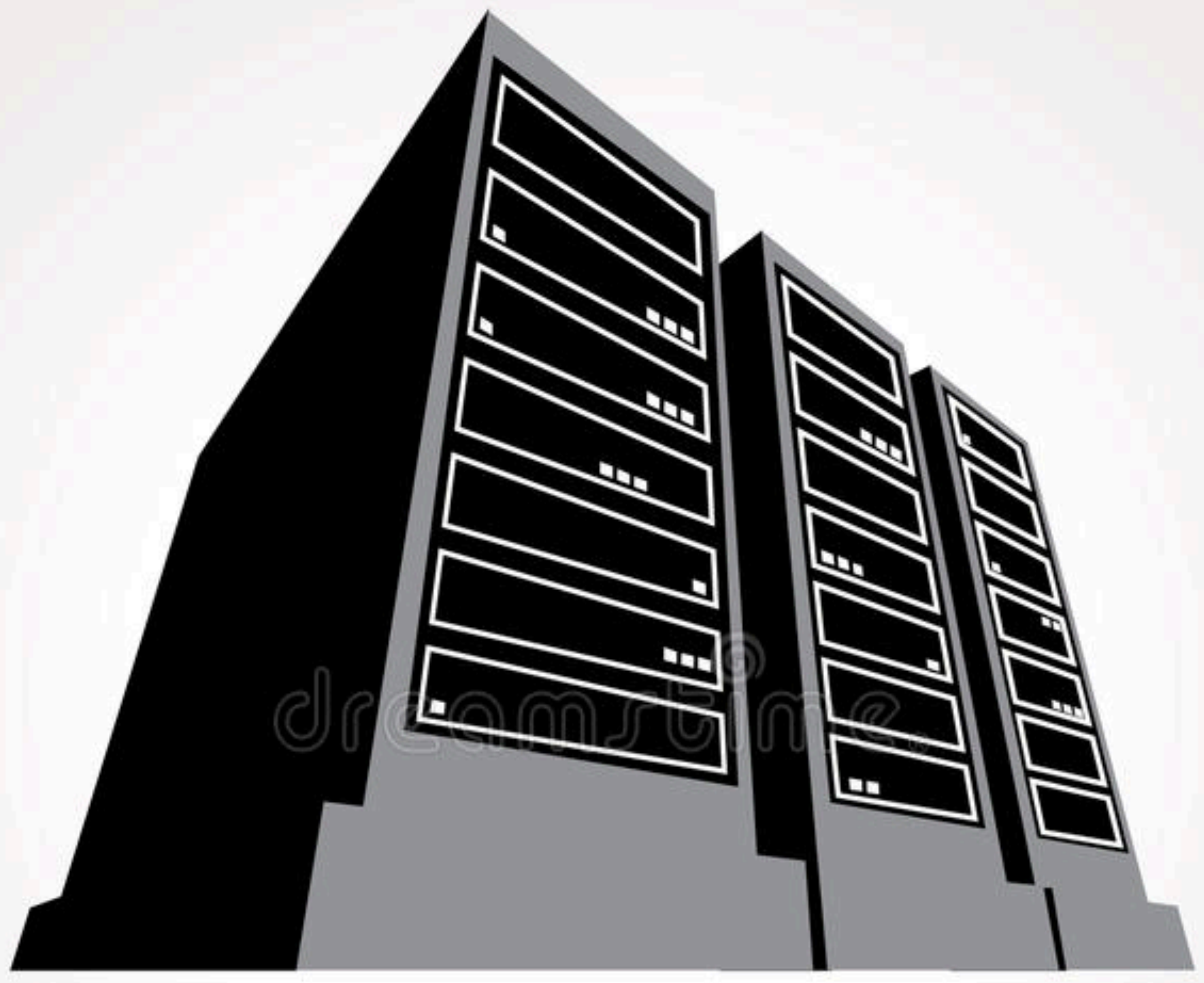
Sentence1
Sentence2
Sentence3
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences
Sentences

Sentence

Input →

Labeled datasets,
including PropBank, FrameNet,
or CoNLL05/09/11/12

Supervised Learning



Input →

(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)
(Arg, v, Arg)

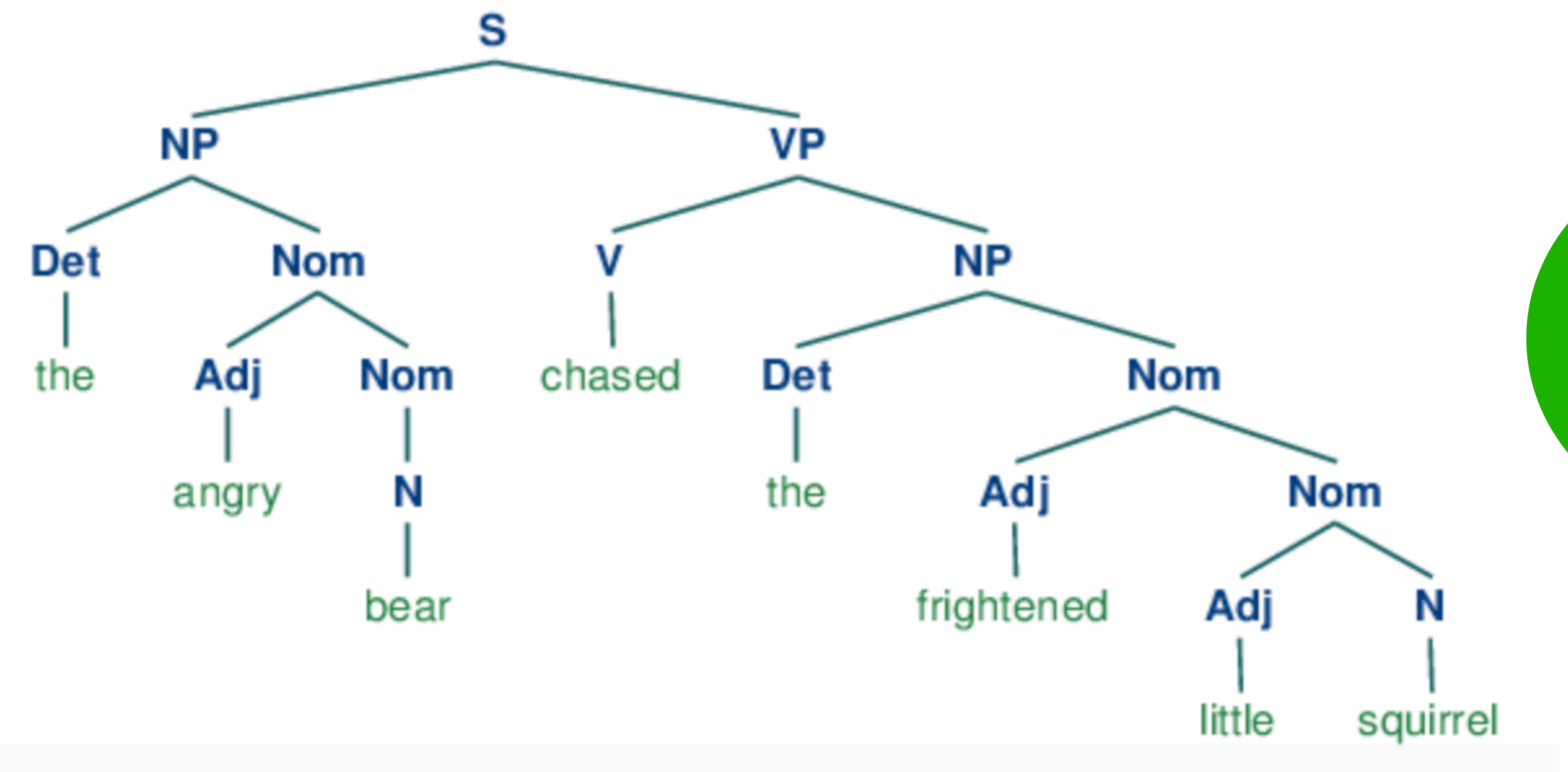
(Arg, v, Arg)

Applications

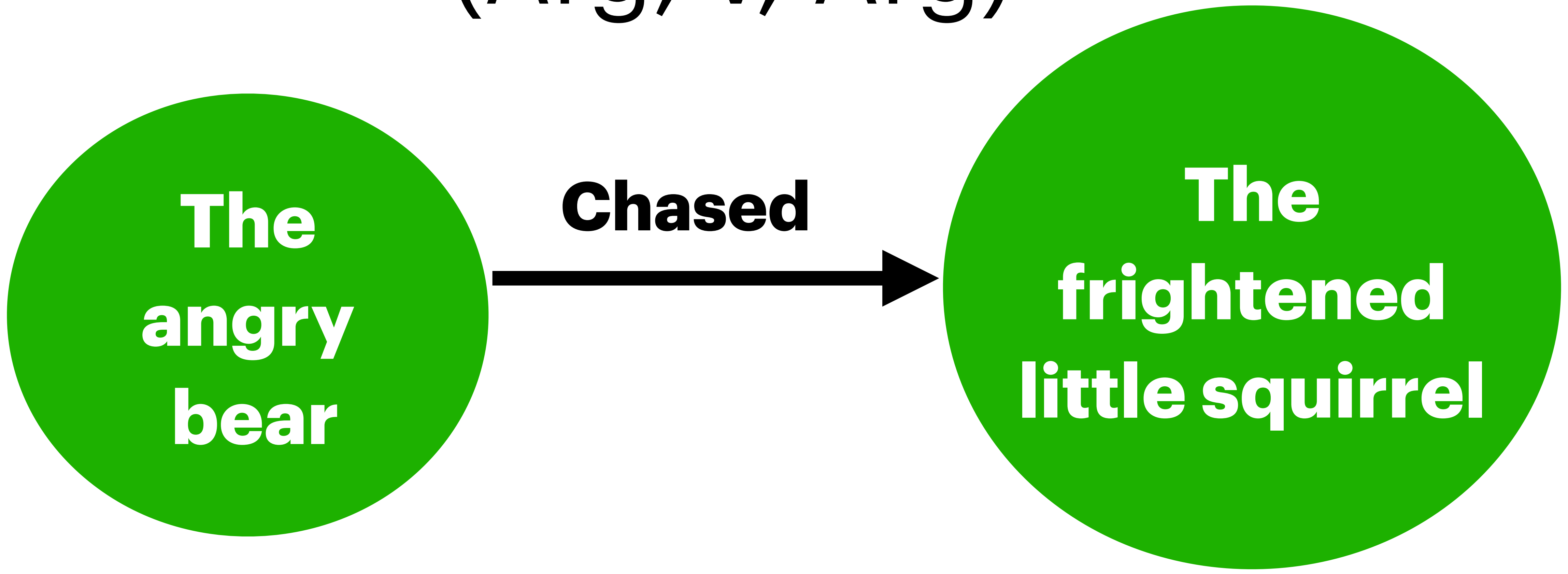
Question-answer system,
Information retrieval,
Machine translation,
Computational reasoning
....

Semantic role labeling

Parsing vs SRL



(Arg, v, Arg)



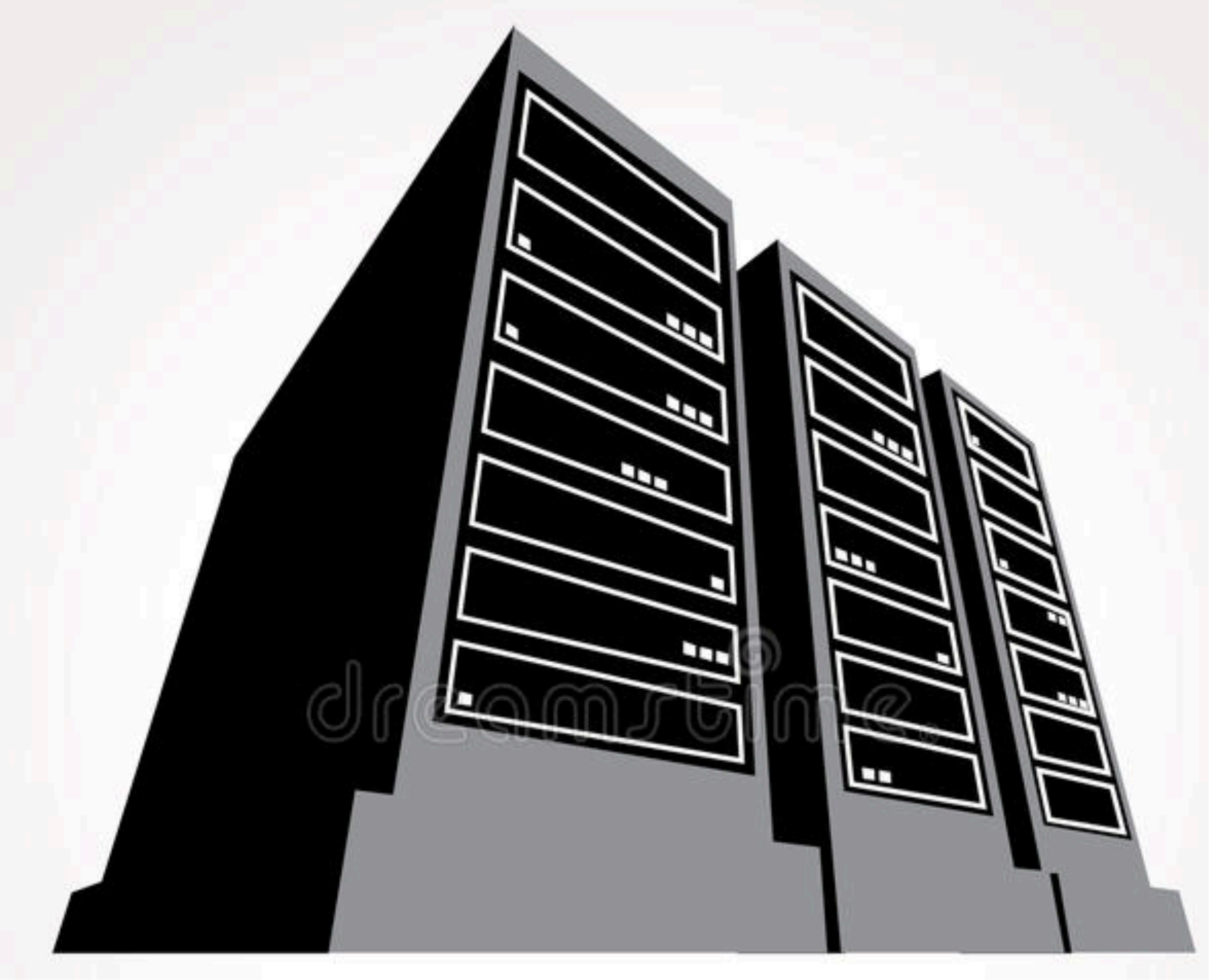
Semantic Role Labeling as Syntactic Dependency Parsing

By Tianze Shi, Igor Malioutov and Ozan Iroy

	Pattern	Example	Percentage English	Percentage Chinese
(D)	pred arg	nsubj: *She* *designed* the bridge …	87.5%	82.7%
(C)	arg pred	nsubj, xcomp: *She* wanted to *design* the bridge …	6.1%	10.4%
(R)	arg pred	rcmod: The *bridge*, which is *designed* by her, …	4.7%	5.7%
	arg pred	nsubj, xcomp, conj: *She* wanted to design and *build* the bridge …	1.1%	1.0%
	Others		0.5%	0.2%

Table 1: The most common structural relations in the training data between the *predicates* (pred) and the *arguments* (arg). Appendix §C and §D include more examples as well as Chinese data.

Supervised Learning



1. Feature-based

2. Neural Algorithm

Feature-based

Standard structure - based on parsing

- 1. **Pruning**: simple heuristics to prune unlikely constituents.
- 2. **Identification**: a binary classification of each node as an argument to be labeled or a None
- 3. **Classification**: a 1-of-N classification all the constituents that were labeled as arguments by the previous stage

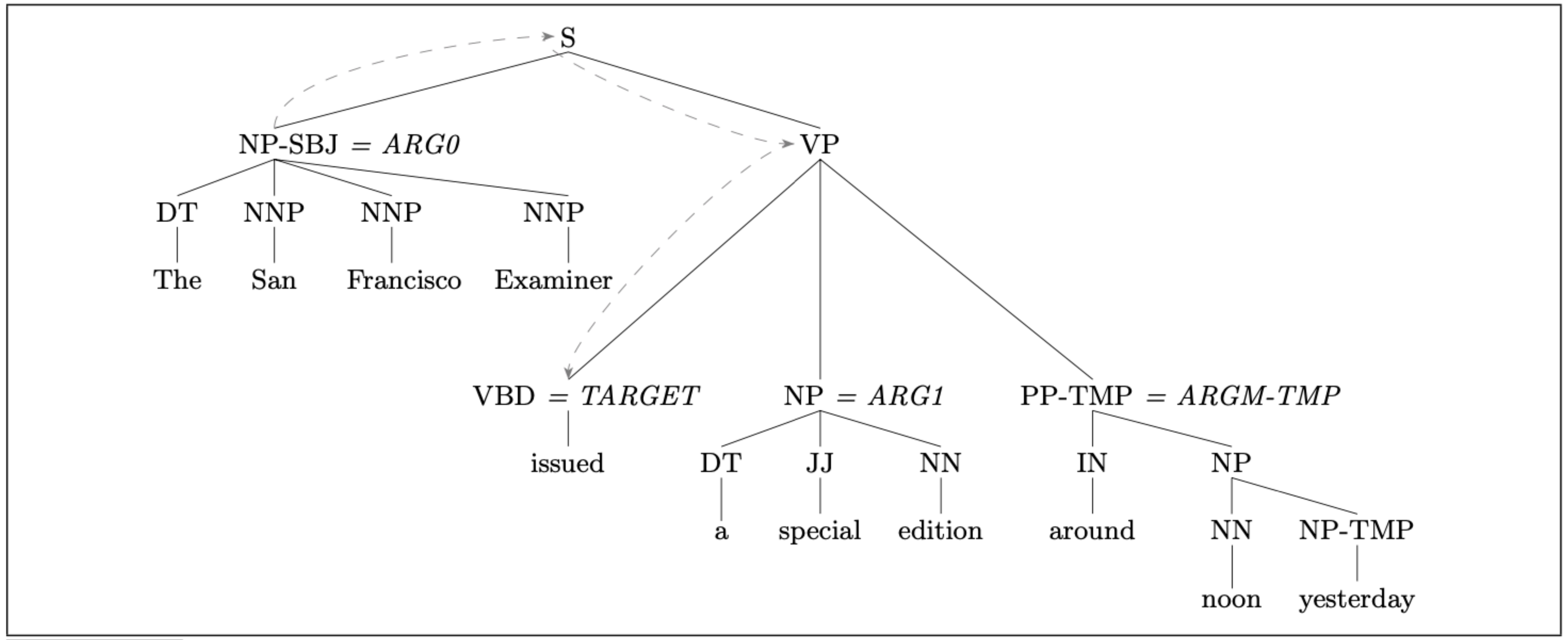


Figure 19.5 Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the **path** feature NP↑S↓VP↓VBD for ARG0, the NP-SBJ constituent *The San Francisco Examiner.*

Neural-based

Standard structure

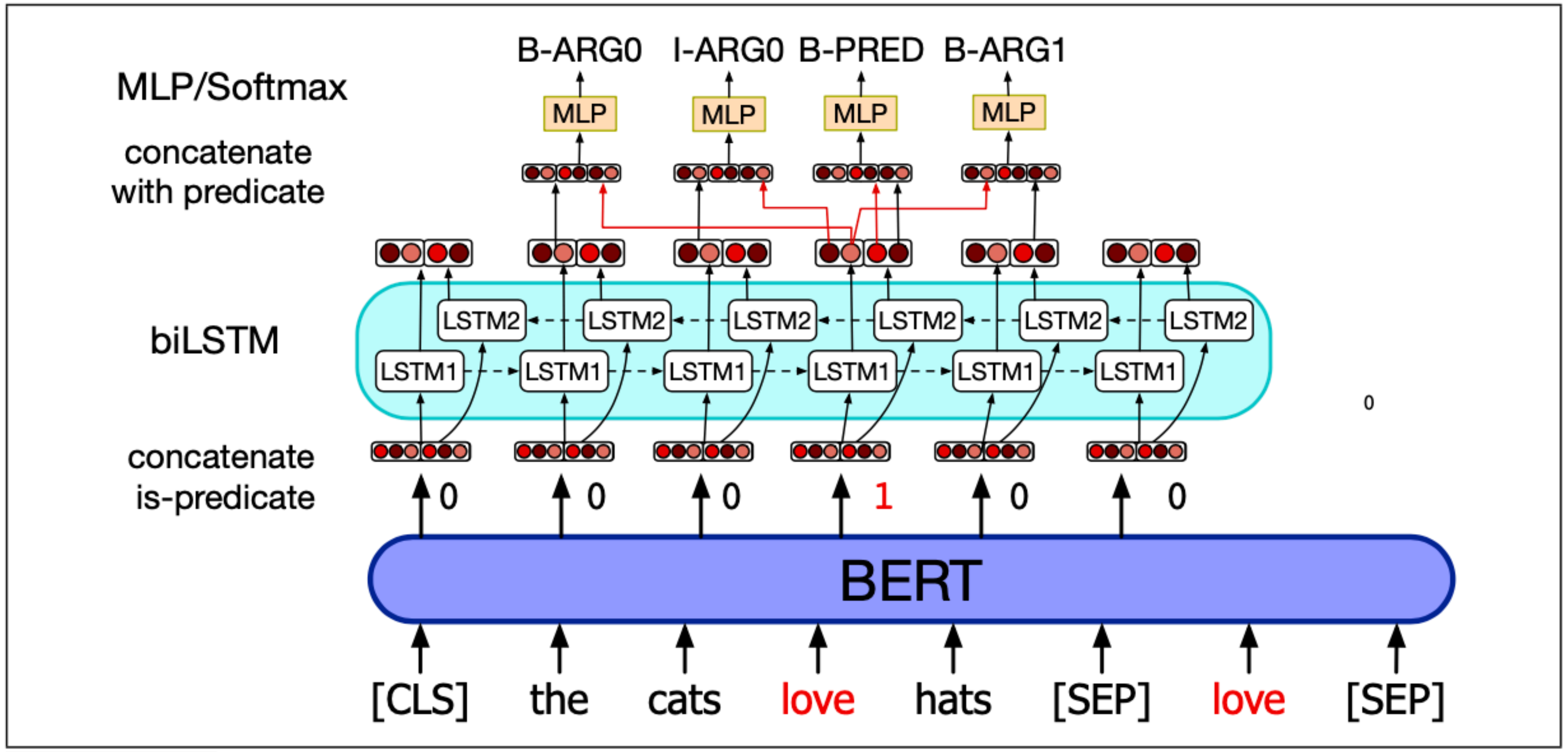


Figure 19.6 A BERT + biLSTM approach to semantic role labeling. The input sentence is followed by [SEP] and an extra input for the predicate, in this case *love*. The BERT outputs are concatenated to an indicator variable which is 1 for the predicate and 0 for all other words , passed through a biLSTM, and then the output embedding at each token position is concatenated with the embedding for the predicate, and passed through a single-layer MLP. After Shi and Lin (2019) and He et al. (2017).

Multilingual

X-SRL: A Parallel Cross-Lingual Semantic Role Labeling Dataset