Event extraction using iterative optimization

Aju Thalappillil Scaria

Rishita Anubhai

Rose Marie Philip

Project goal

In response to a specific antigen and to immune cell signals (not shown), one B cell divides and forms a clone of cells. The remaining B cells, which have antigen receptors specific for other antigens, do not respond. The clone of cells formed by the selected B cell gives rise to memory B cells and antibody-secreting plasma cells.
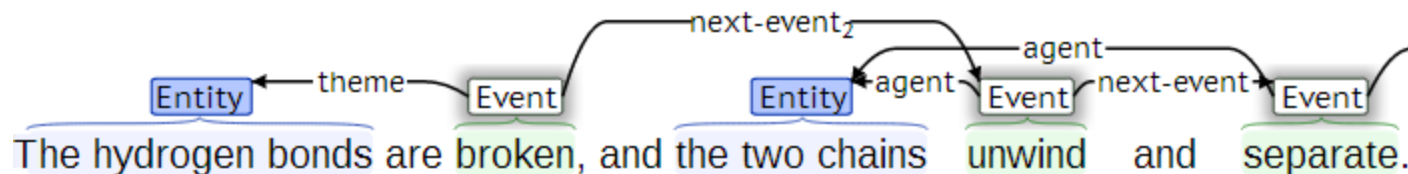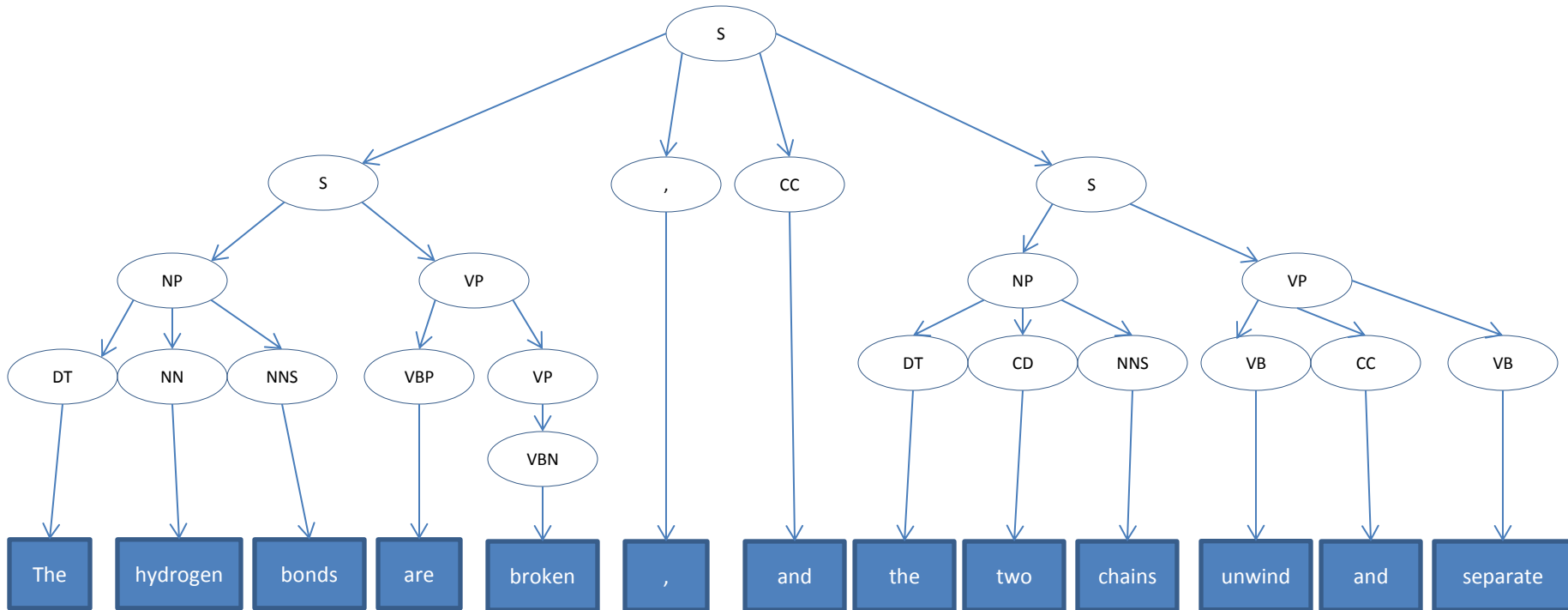
Model
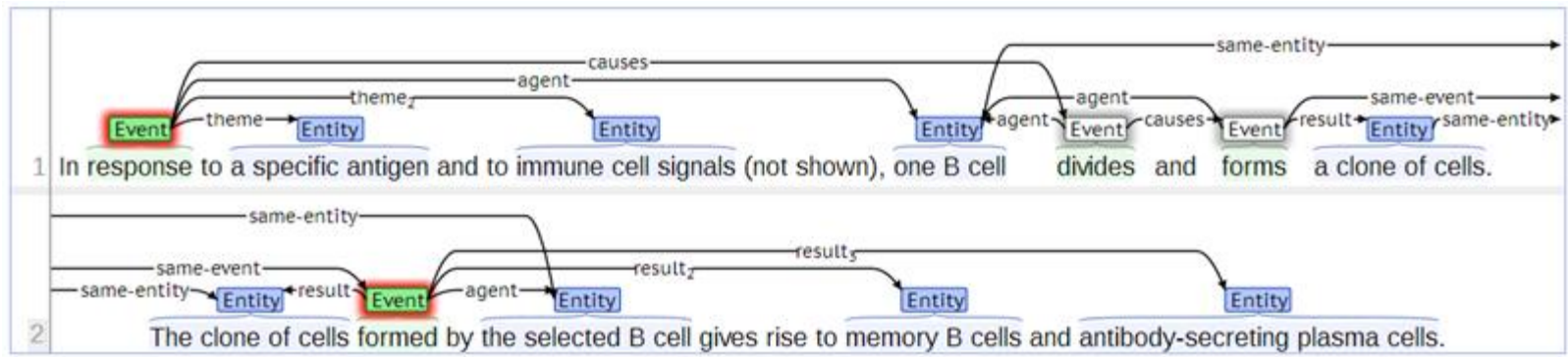
Stages

- Three high level stages
  - Event/trigger prediction
  - Entity/argument identification for triggers
  - Semantic role labeling the entities identified
- MaxEnt based classifier for prediction
- Features
  - Lexical
  - Dependency tree based
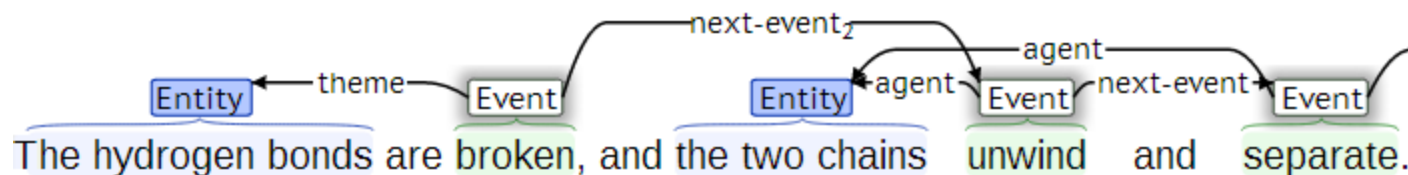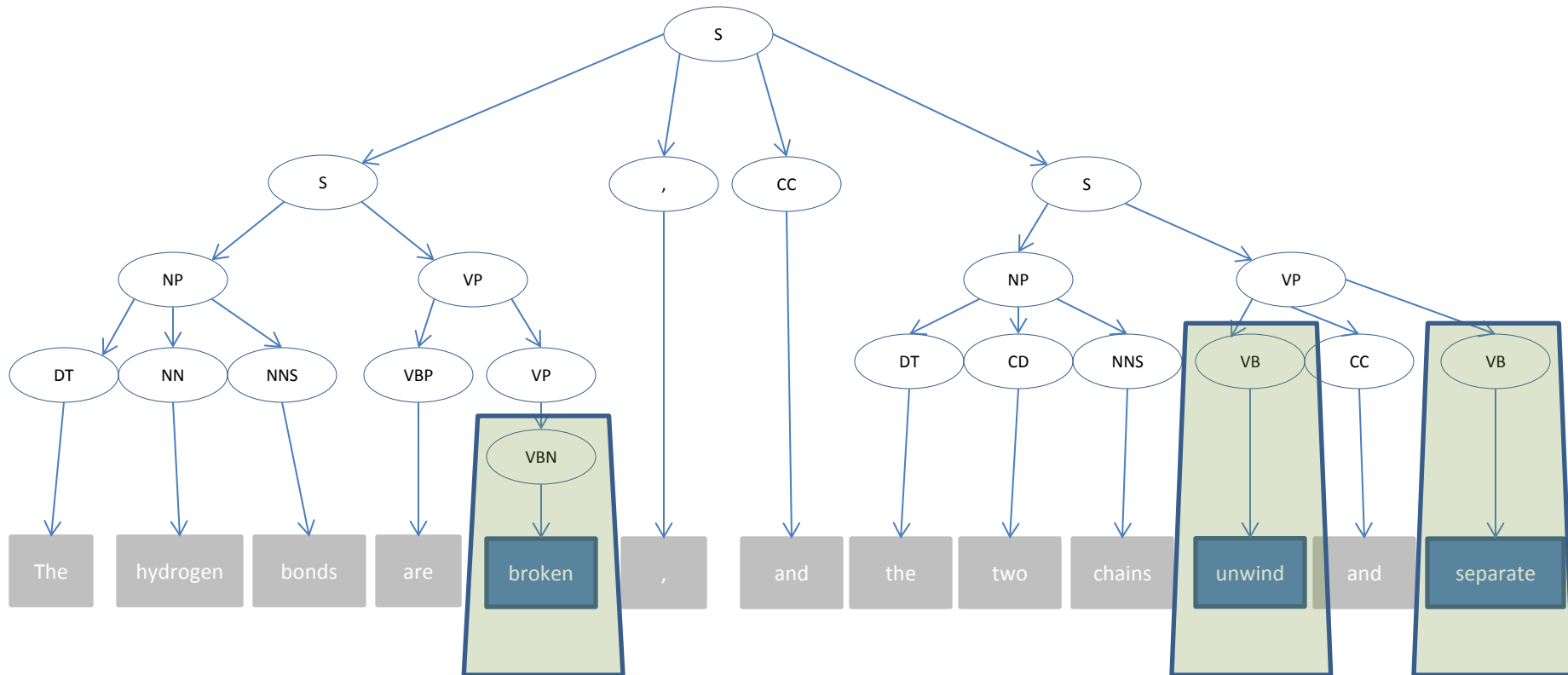 - Parse tree based

We use Stanford CoreNLP Toolkit

Representation

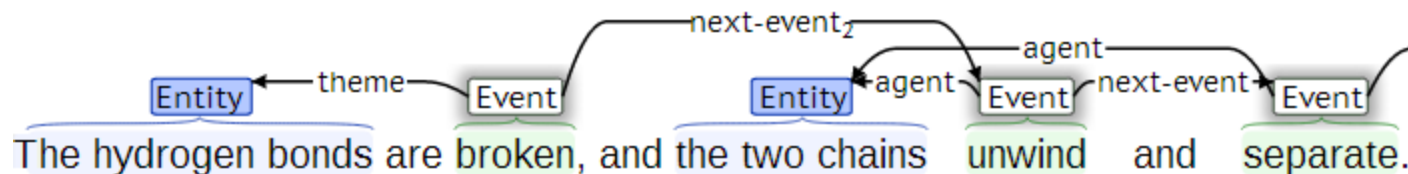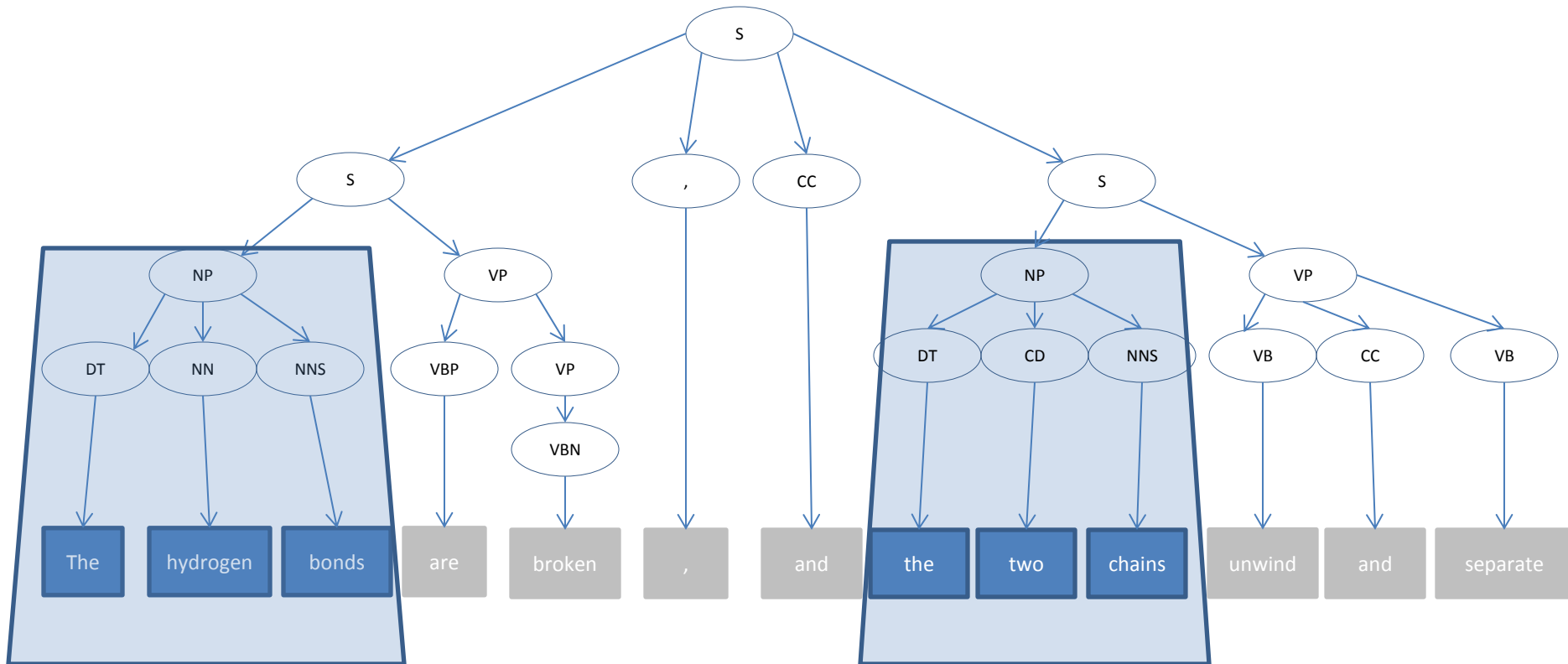The hydrogen bonds are broken, and the two chains unwind and separate.

Representation – Event triggers

The hydrogen bonds are broken, and the two chains unwind and separate.
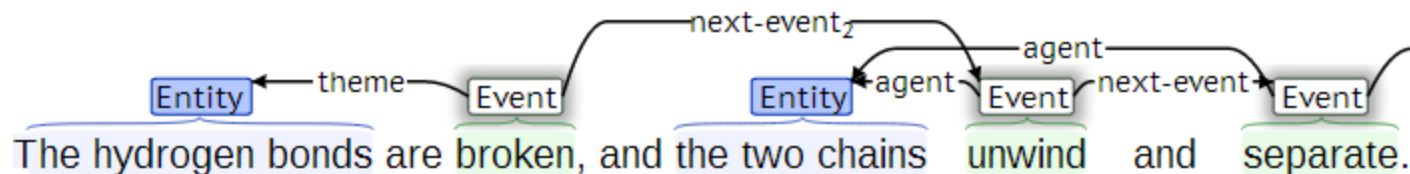
Representation - Entities
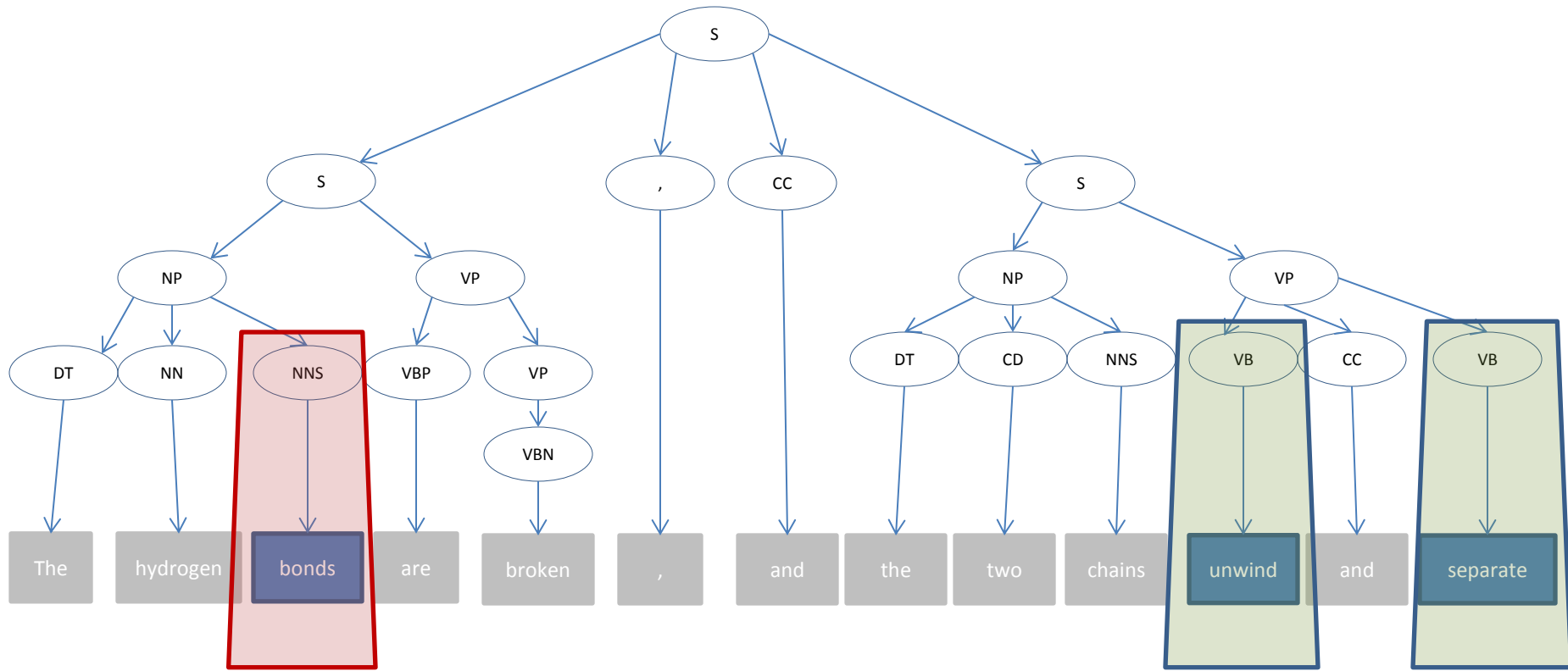
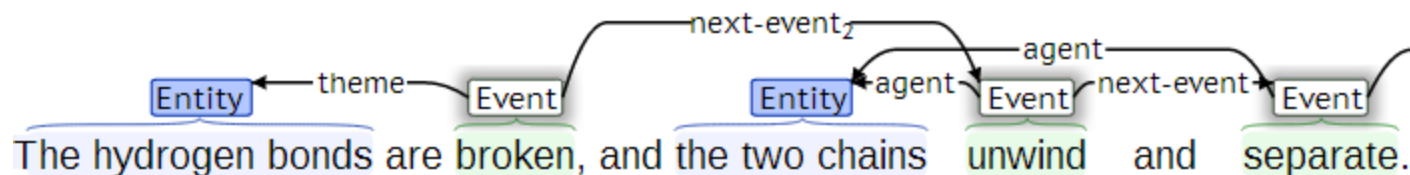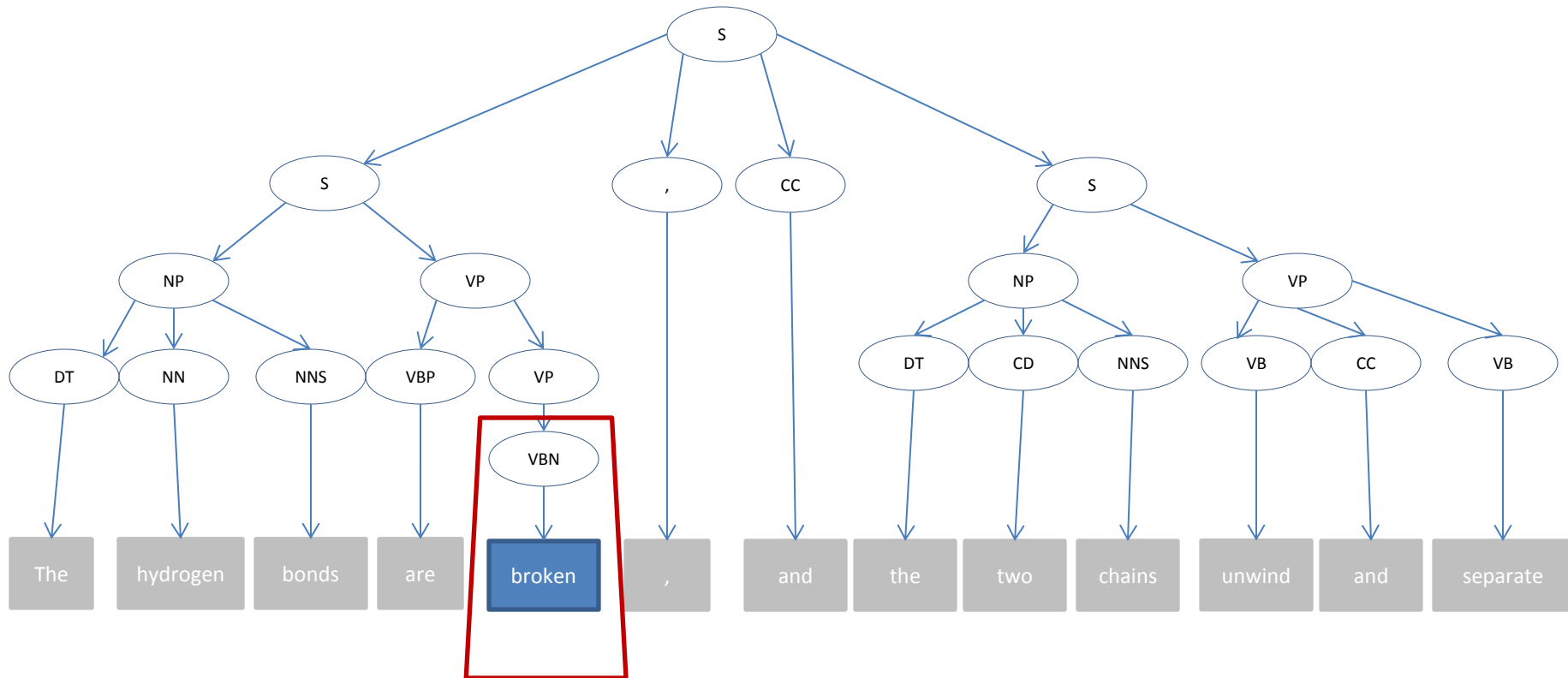The hydrogen bonds are broken, and the two chains unwind and separate.

MODELS

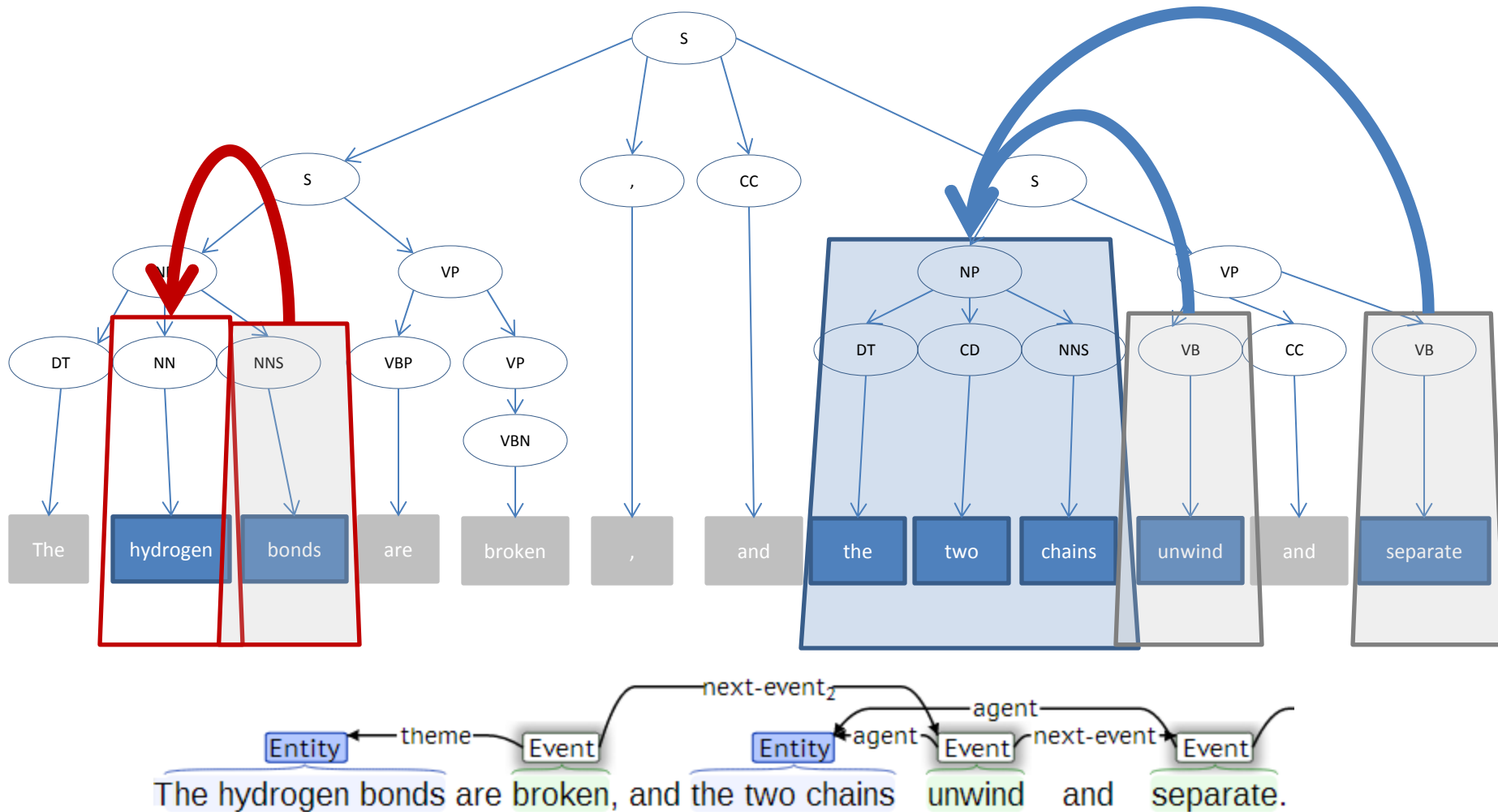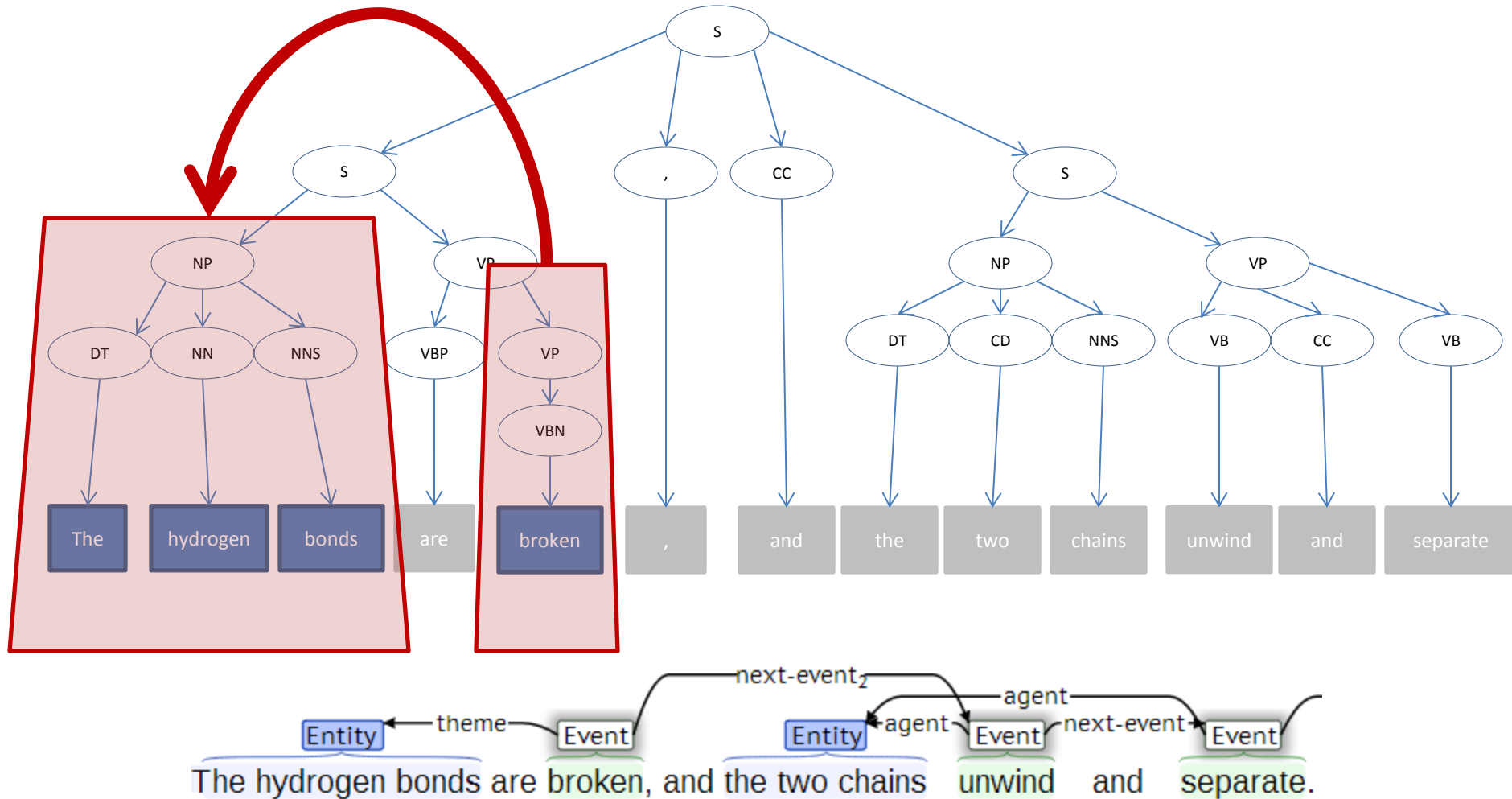# Event trigger prediction

$$P(word \in \{TRIGGER\} \mid sentence)$$

Event trigger prediction

Event argument (entity) prediction

- For each trigger

$$P(phrase = argument \mid trigger, sentence)$$

- Non overlapping constraint

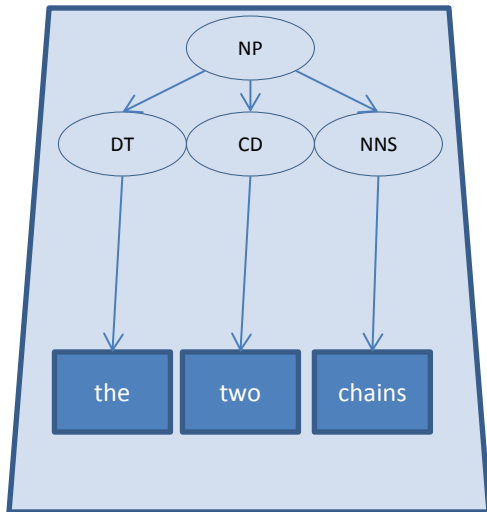 – Dynamic program

Argument prediction for trigger

Argument prediction for trigger

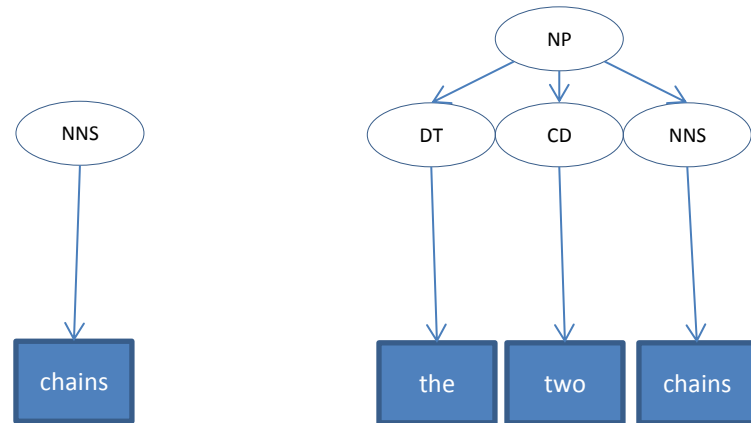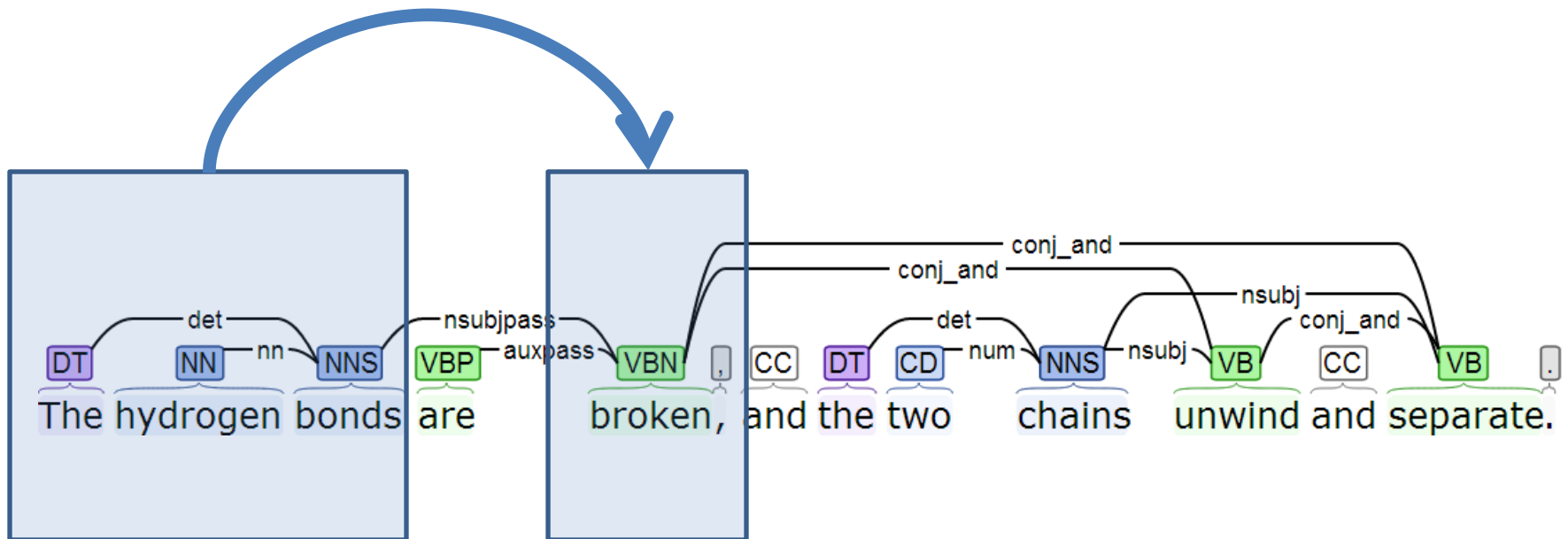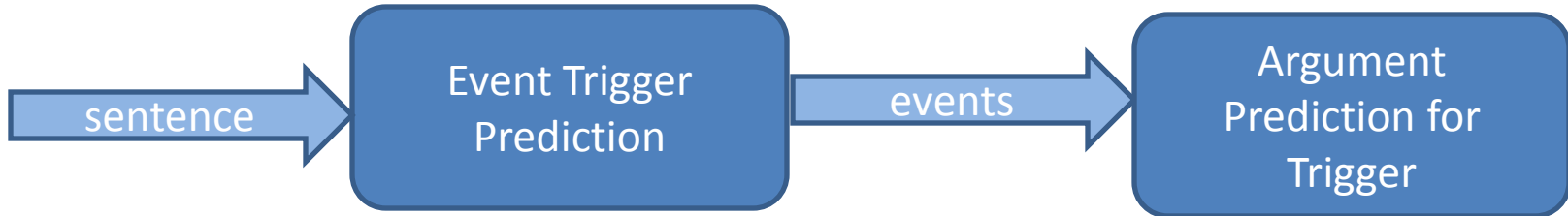# Dynamic program

Actual
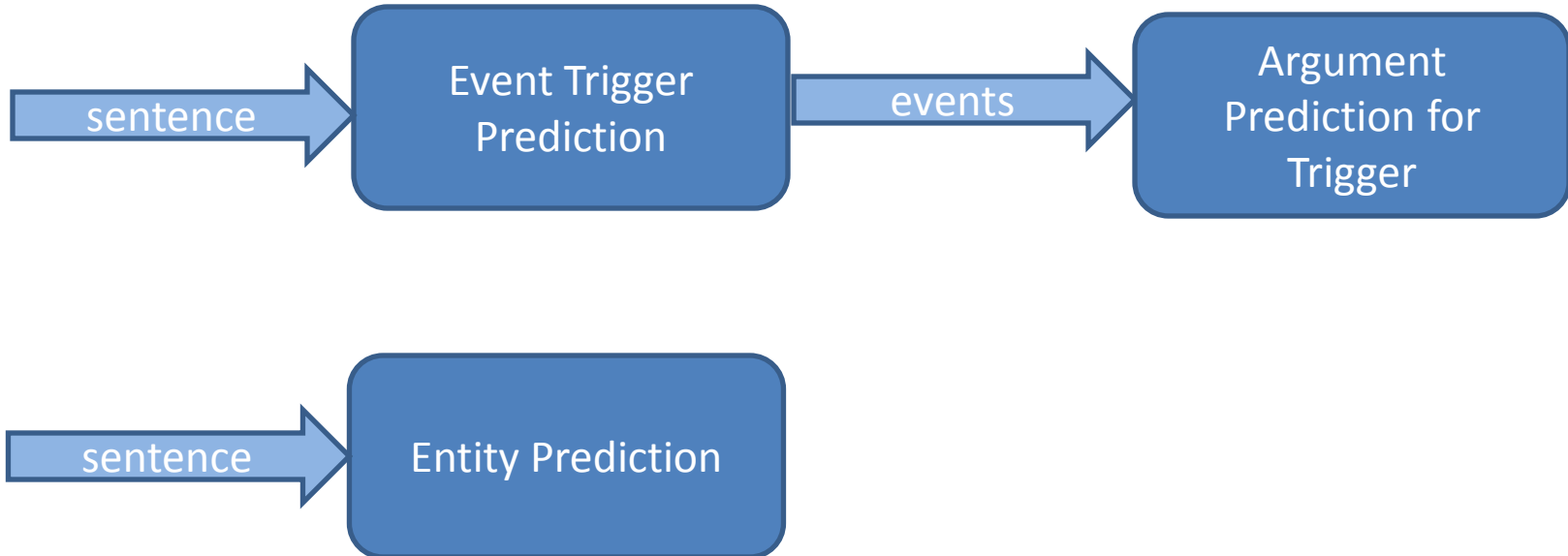
Predicts both

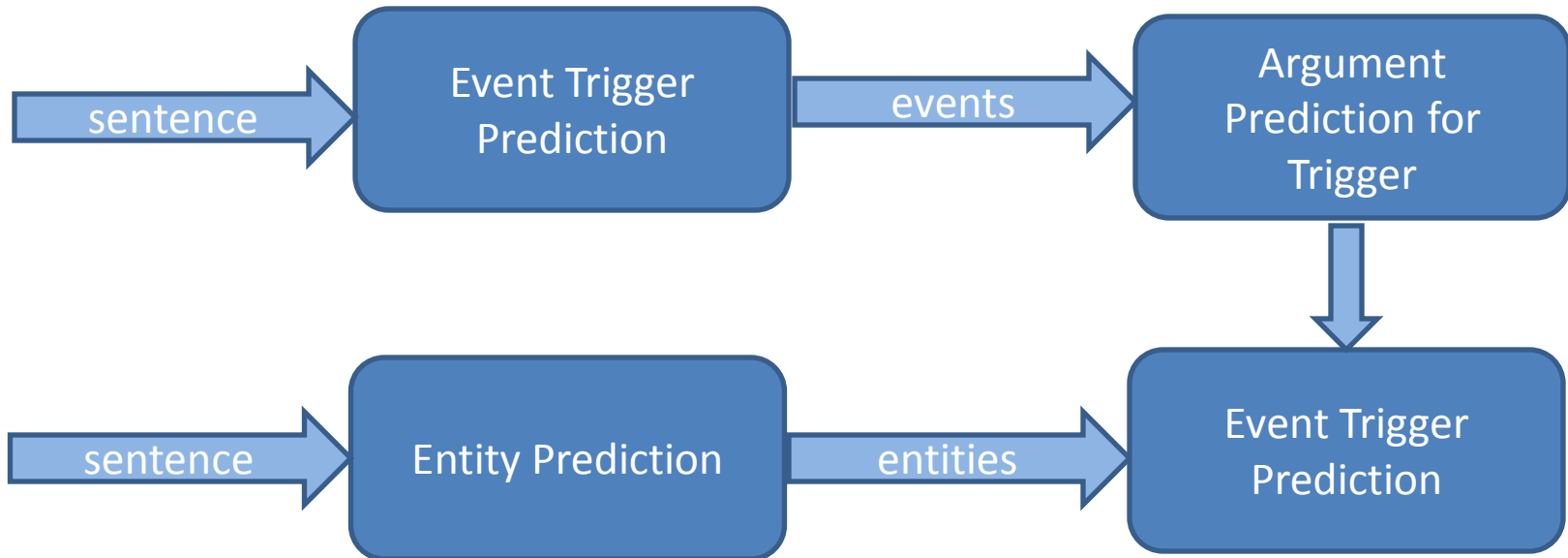# Can we use entities to predict triggers?
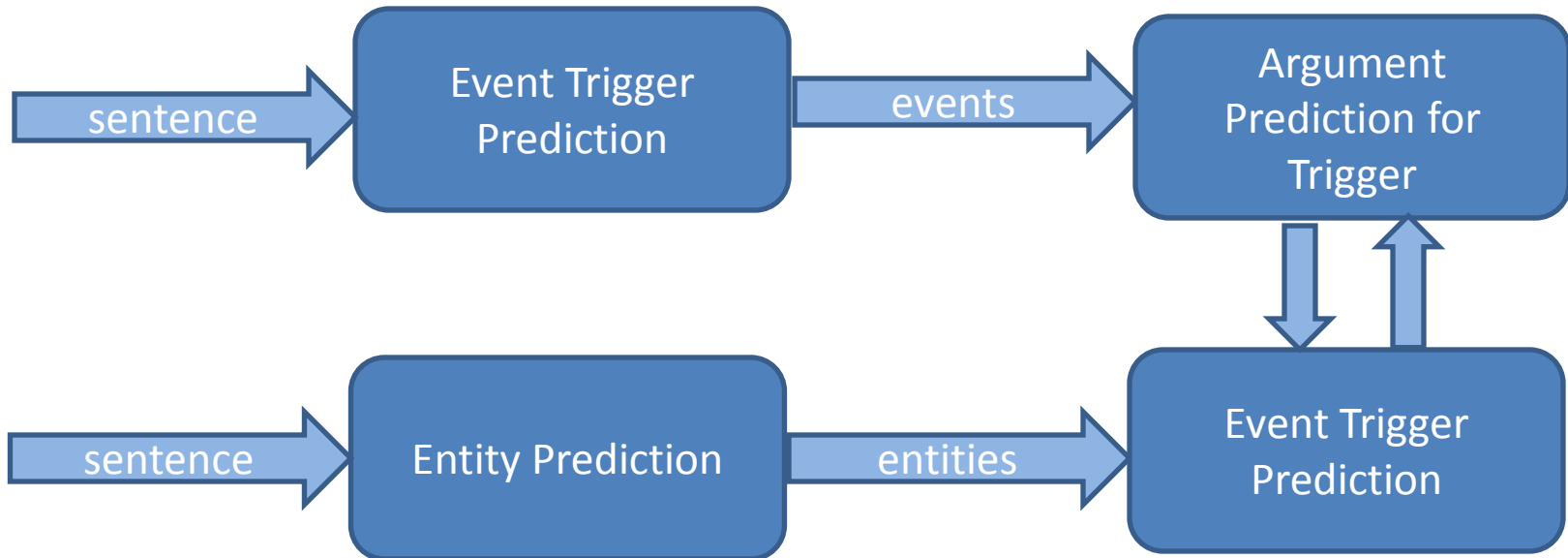
Dependency parse

Iterative optimization

sentence → **Event Trigger Prediction** → events → **Argument Prediction for Trigger**
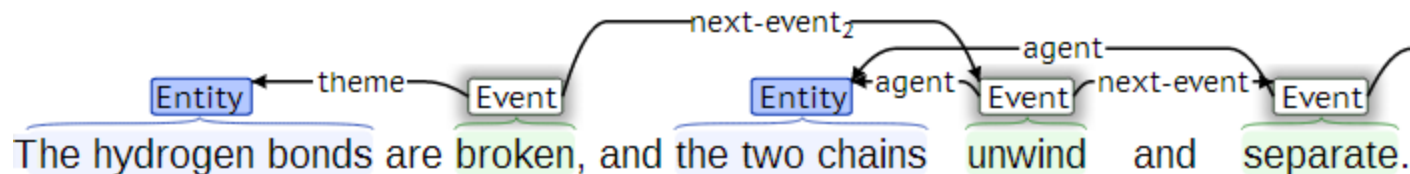
Iterative optimization
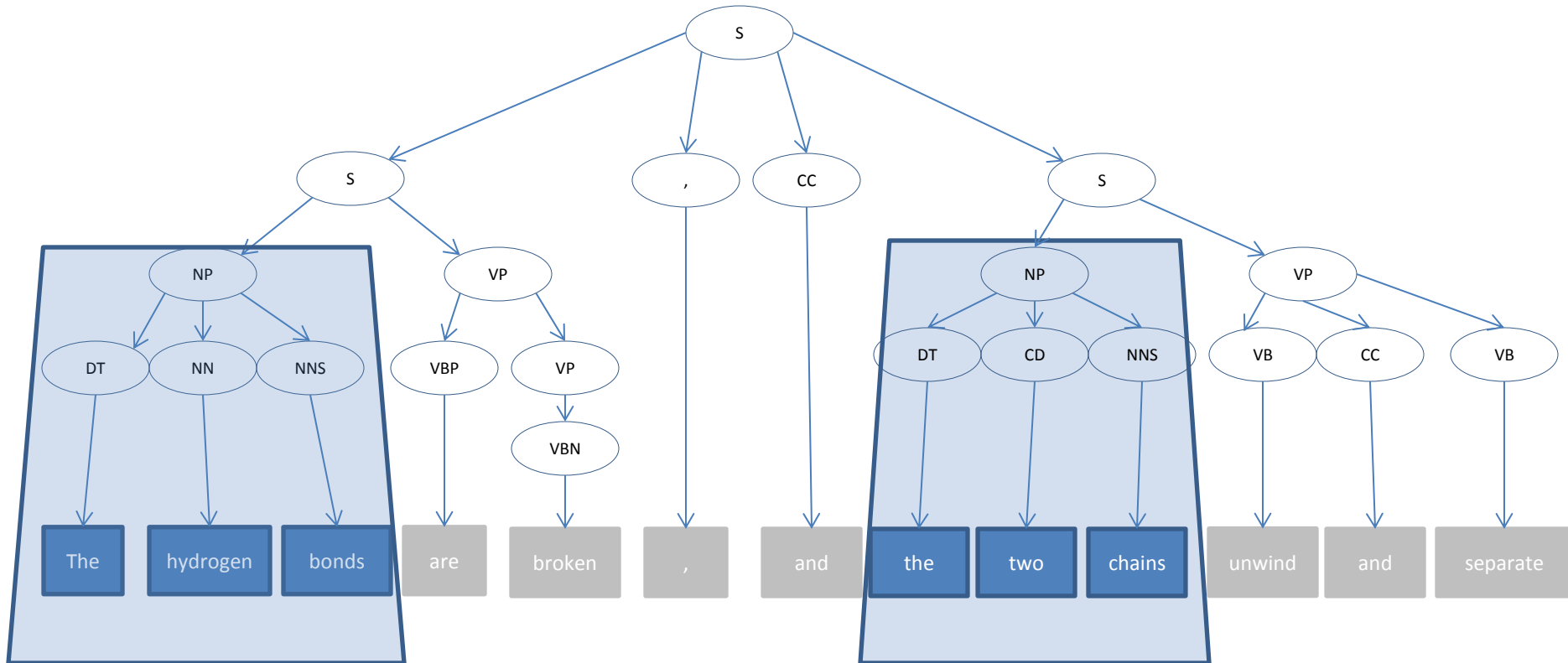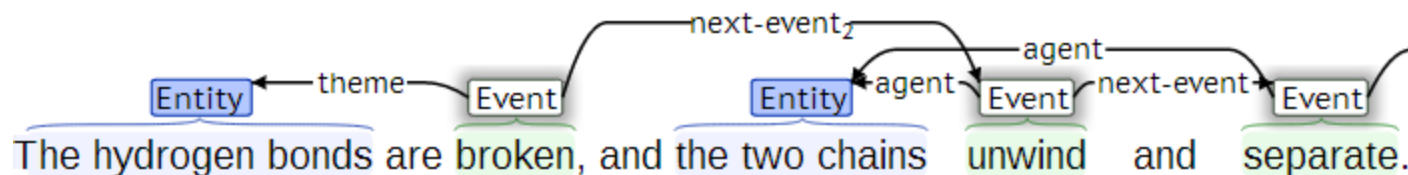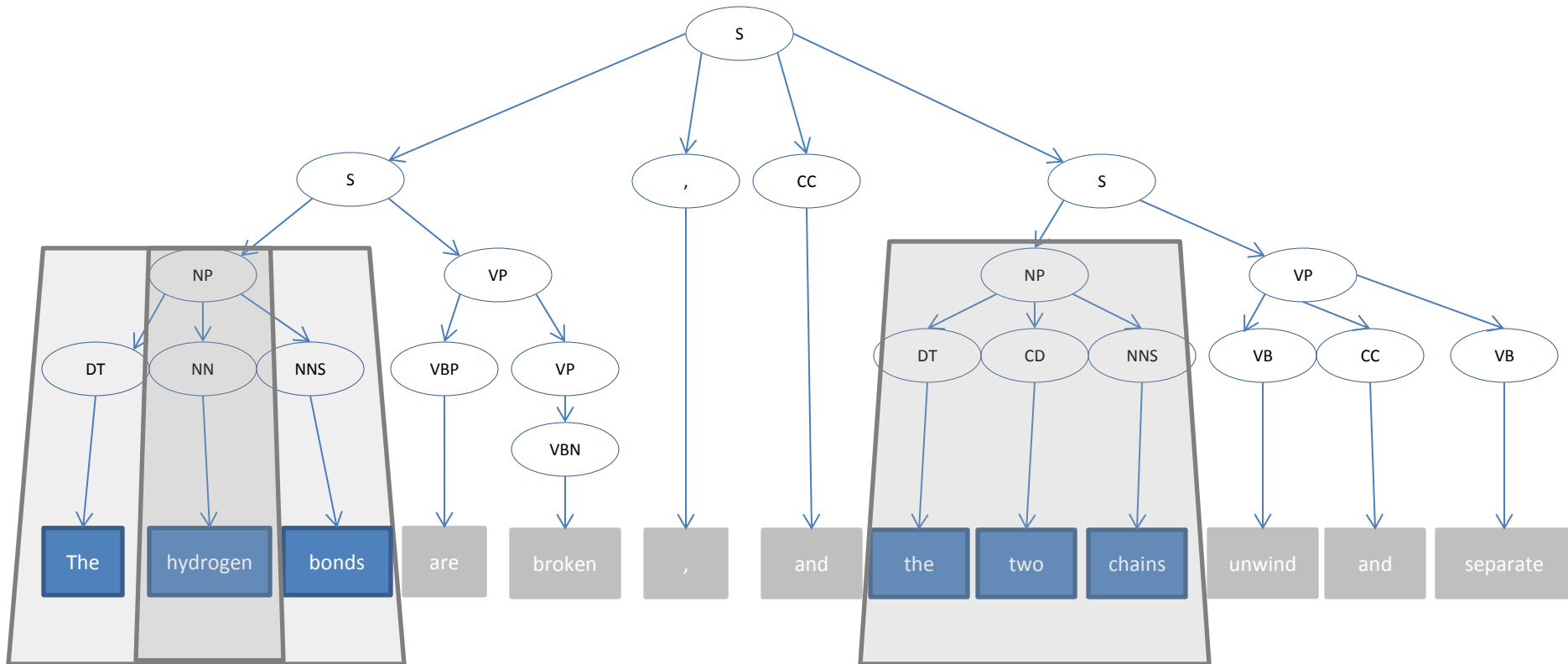
Iterative optimization

Iterative optimization

Entity Prediction

$$P(phrase = argument \mid sentence)$$

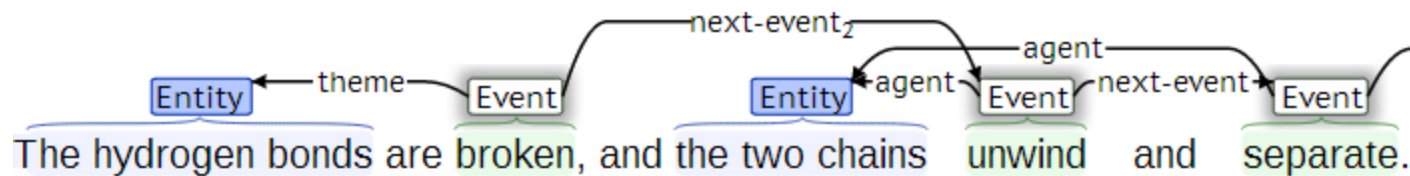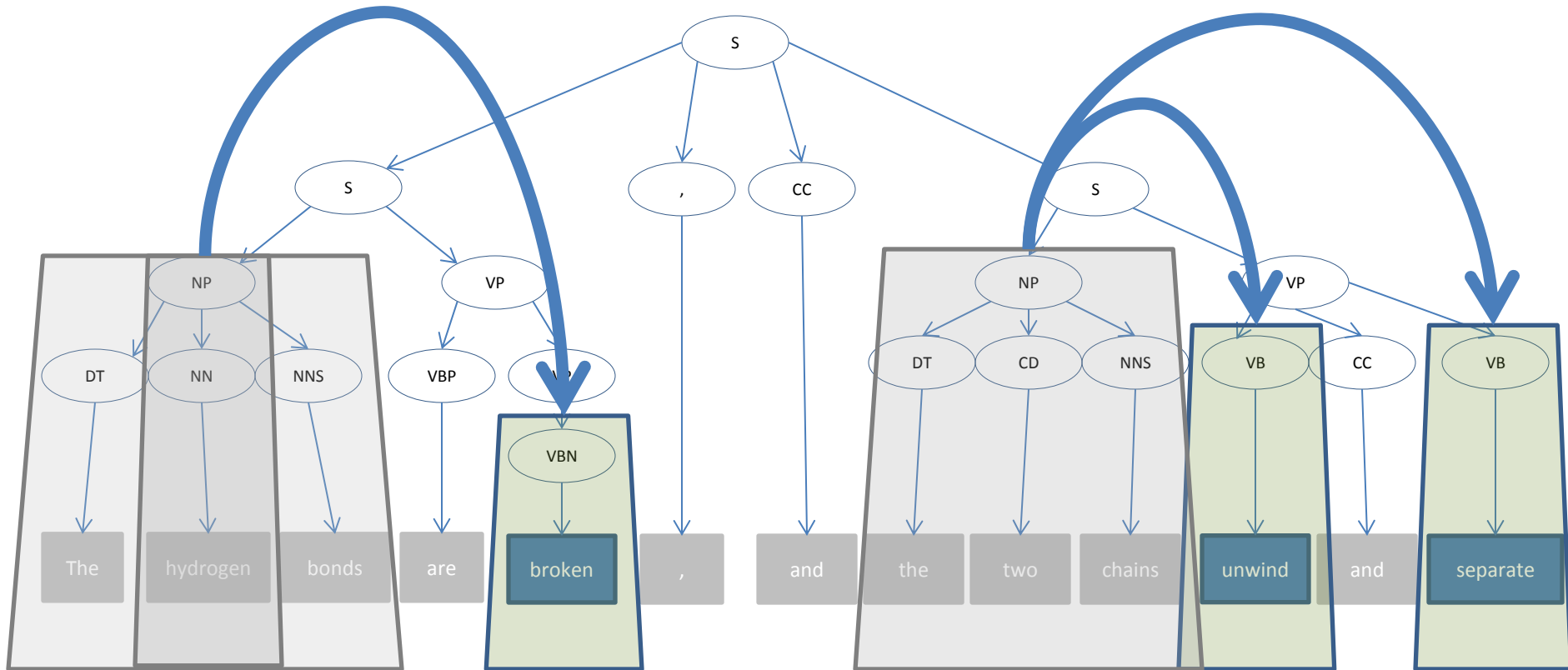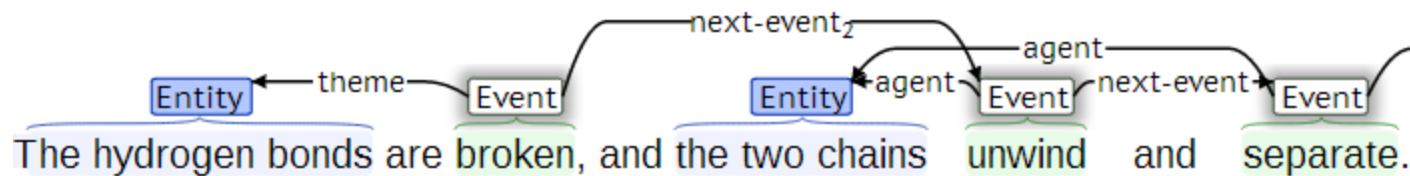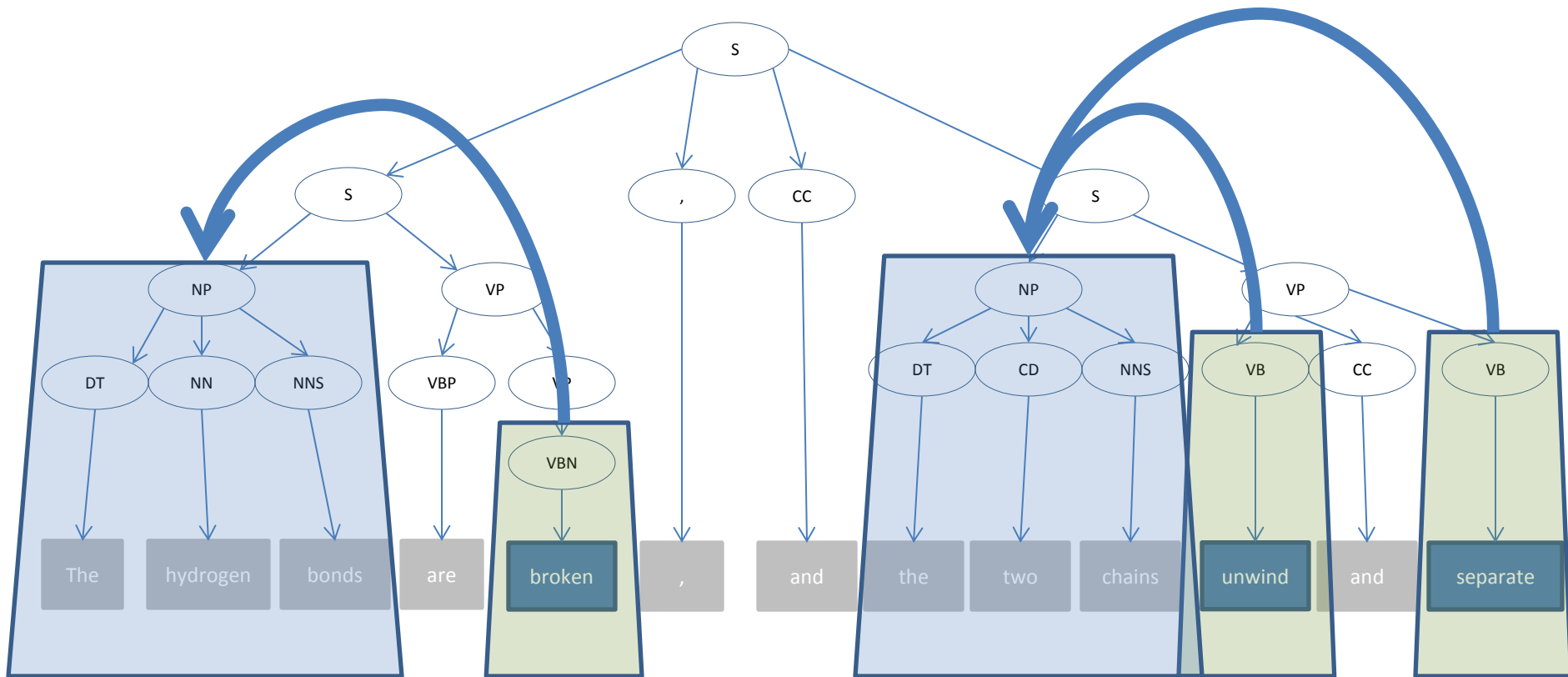# Iterative optimization

$$P(word \in \{TRIGGER\} \mid \{ENTITIES\}, sentence)$$

Iterative optimization

Iterative optimization

Results

Event trigger prediction

Type	Precision	Recall	F1
Baseline	0.47	0.73	0.57
MaxEnt_Basic	0.69	0.66	0.67
MaxEnt_Iterative	0.72	0.70	0.71

Event argument prediction

Type	Precision	Recall	F1
Baseline	0.44	0.53	0.48
MaxEnt_Basic	0.56	0.45	0.50
MaxEnt_Iterative	0.55	0.50	0.52

Next steps

- Improve performance of classifiers
  - Tune features
- Semantic role labeling
  - Multiclass MaxEnt
  - Re-ranking
- Joint models