On the Structure-aware NLP and Beyond

Survey talk

Hao Fei

June 1, 2022

OUTLINE

- WHAT is syntactic structure?
- WHY integrating structures for NLP?
- HOW to integrate?
- WHAT to do next?

OUTLINE

- **WHAT is syntactic structure?**
- WHY integrating structures for NLP?
- HOW to integrate?
- WHAT to do next?

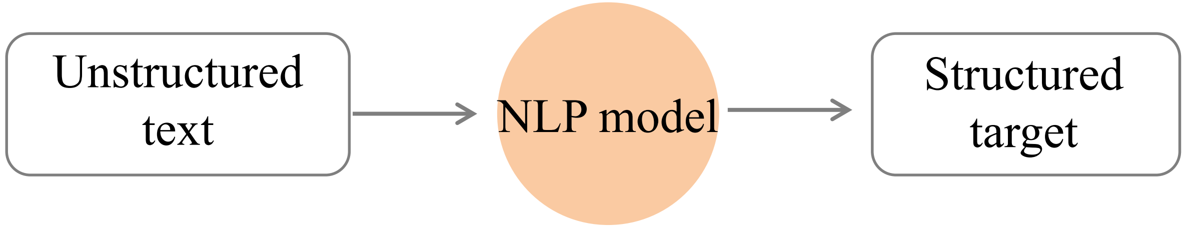
[Structure-aware NLP]

What?

Various NLP applications

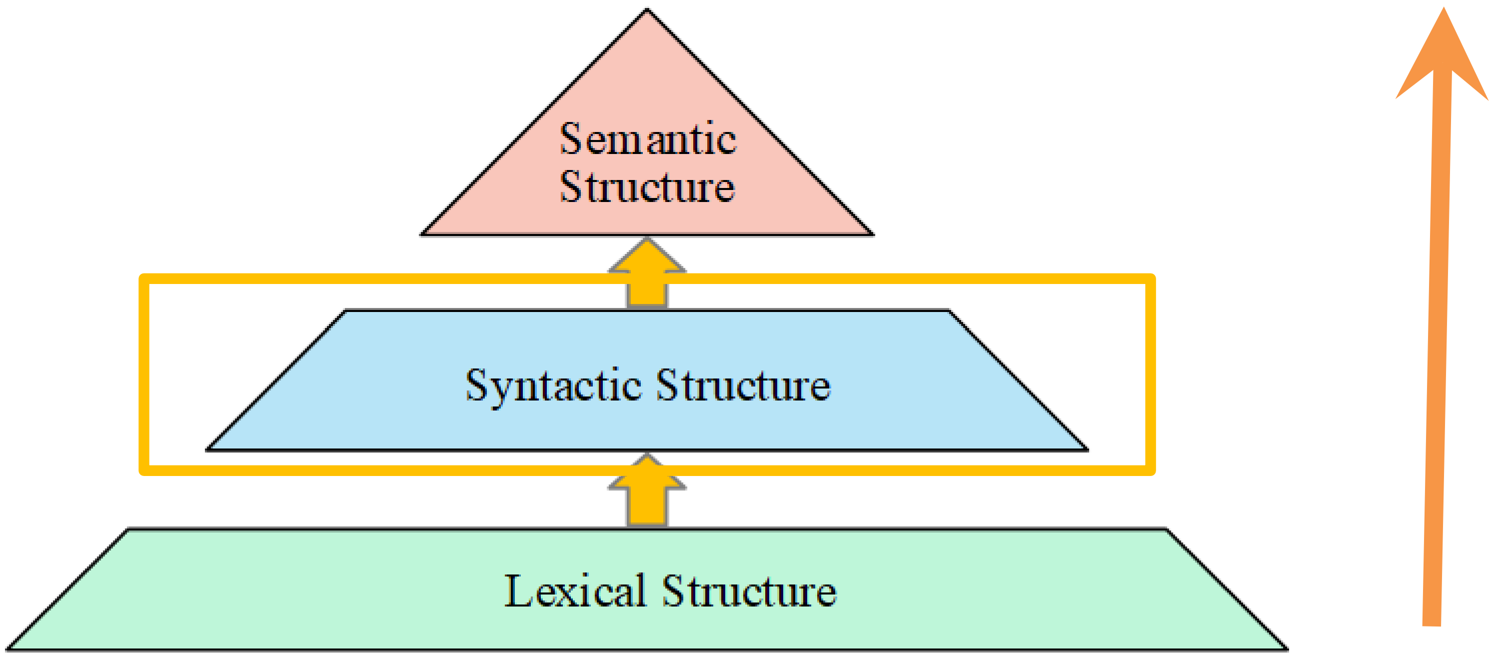
Text classification	NER	Sentiment classification
Relation extraction	Event extraction	Semantic parsing
Neural machine translation	Reading comprehension	Text summarization
Paraphrase	Question answering	……

NLP: *Structure prediction*



[Structure-aware NLP]
What?

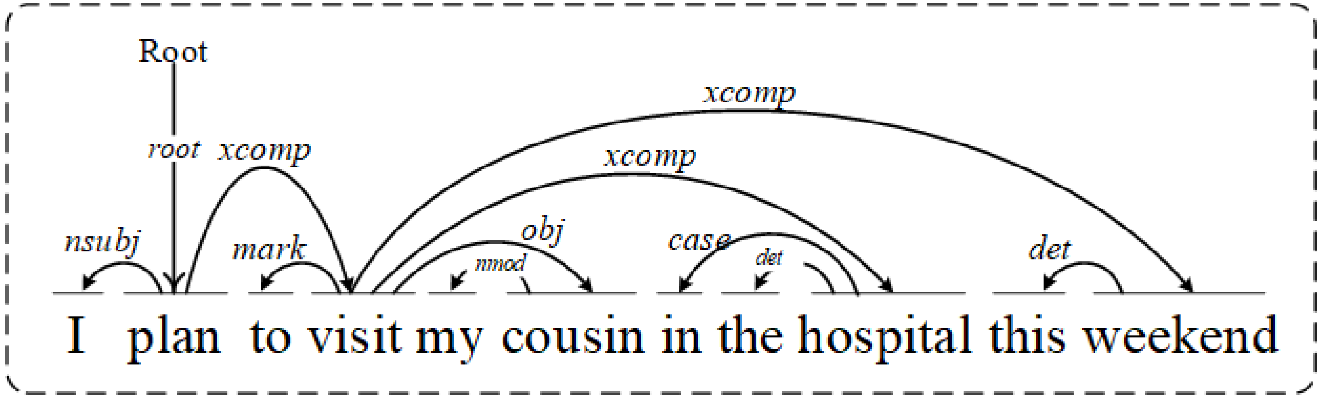
Natural language understanding in a **Three-level Hierarchy**



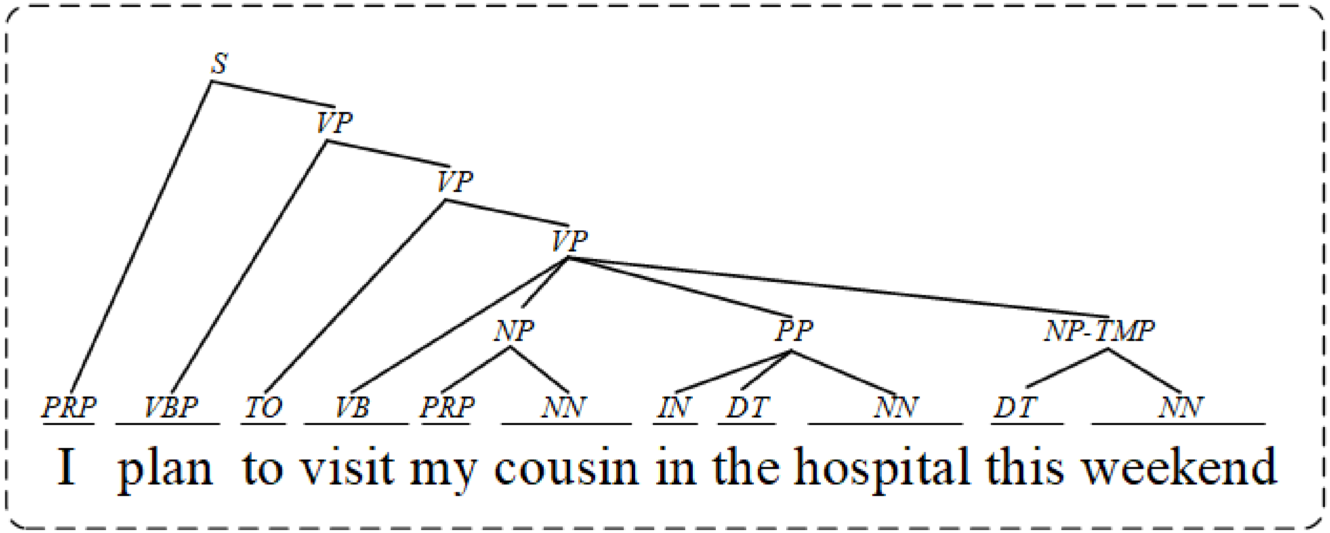
[Structure-aware NLP]

What?

- Syntax structure
  - Syntactic dependency tree
  - Syntactic constituency tree



(*a*) Syntactic dependency tree



(*b*) Syntactic constituency tree

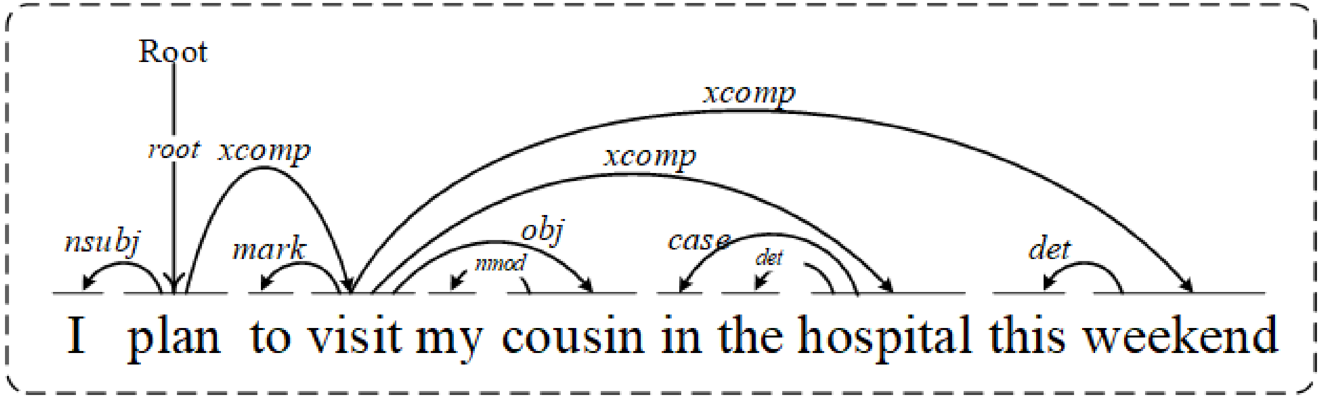
[Structure-aware NLP]

What?

- Syntax structure foundations

  - Syntactic dependency tree

Describe the word-word relations
*in a `**head->dependent**' format*
with specific relation type (label).



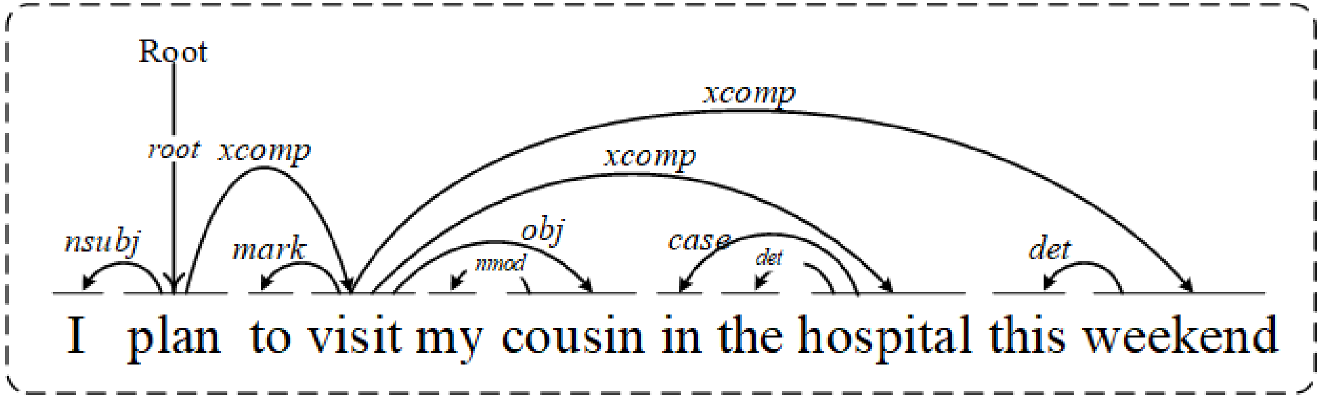
Dep. Label	Description
amod	adjectival modifier
advcl	adverbial clause modifier
advmod	adverb modifier
acomp	adjectival complement
auxpass	passive auxiliary
compound	compound
ccomp	clausal complement
cc	coordination
conj	conjunct
cop	copula
det	determiner
dep	dependent
dobj	direct object
mark	marker
nsubj	nominal subject
nmod	nominal modifier
neg	negation modifier
xcomp	open clausal complement

[Structure-aware NLP]

What?

- Syntax structure foundations
  - Syntactic dependency tree

Describe the word-word relations
*in a `**head->dependent**' format*
with specific relation type (label).



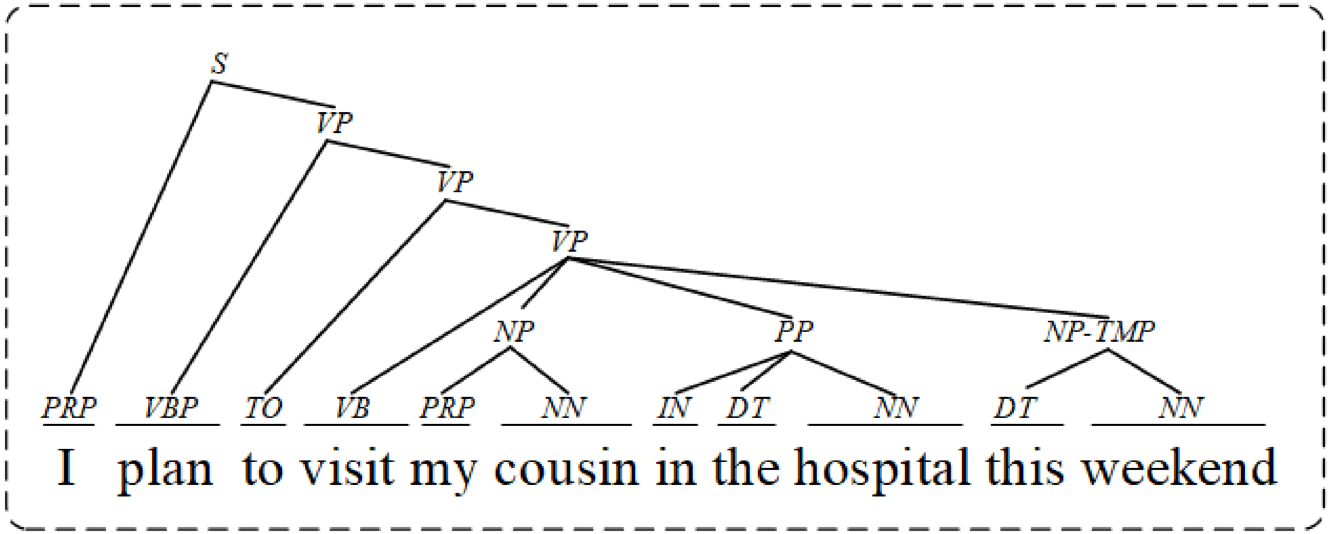
Rules:

- *A dependency tree has only one root with a "root" virtual node word;*
- *In a dependency tree, all nodes except the root node are virtual nodes, and all other nodes are entity words;*
- *Any node in a dependency tree except the root node has a unique parent node that is directly related to it;*
- *A directed acyclic graph composed of a dependency tree as a whole does not cross any dependency arcs.*

What?

- Syntax structure foundations
  - Syntactic constituency tree

*Reveal the **inclusive** and **constituent** relations between **phrases**.*



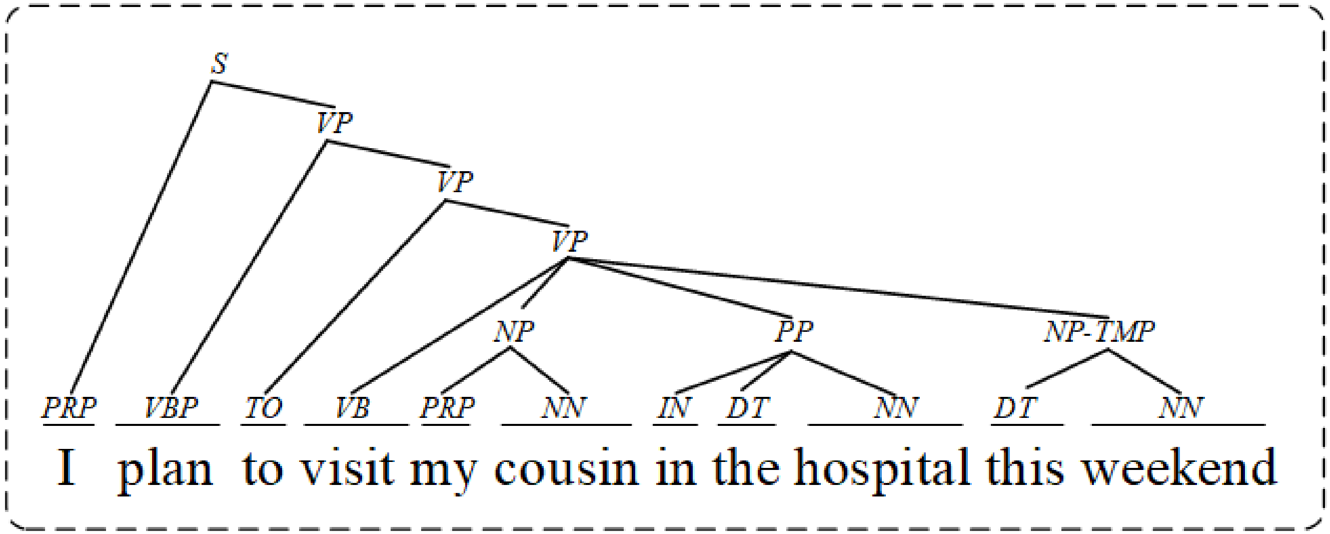
Const. Label	Description
ADJP	Adjective phrase
ADVP	Adverb phrase
NP	Noun phrase
PP	Prepositional phrase
S	Simple declarative clause
SBAR	Subordinate clause
SBARQ	Question introduced by wh-element
SINV	Declarative sentence
SQ	subconstituent of SBARQ
VP	Verb phrase
WHADVP	Wh-adverb phrase
WHNP	Wh-noun phrase
WHPP	Wh-prepositional phrase
X	Unknown or uncertain constituent

[Structure-aware NLP]

What?

- Syntax structure foundations
  - Syntactic constituency tree

*Reveal the **inclusive** and **constituent** relations between **phrases**.*



Rules: Context-free grammar (CFG)

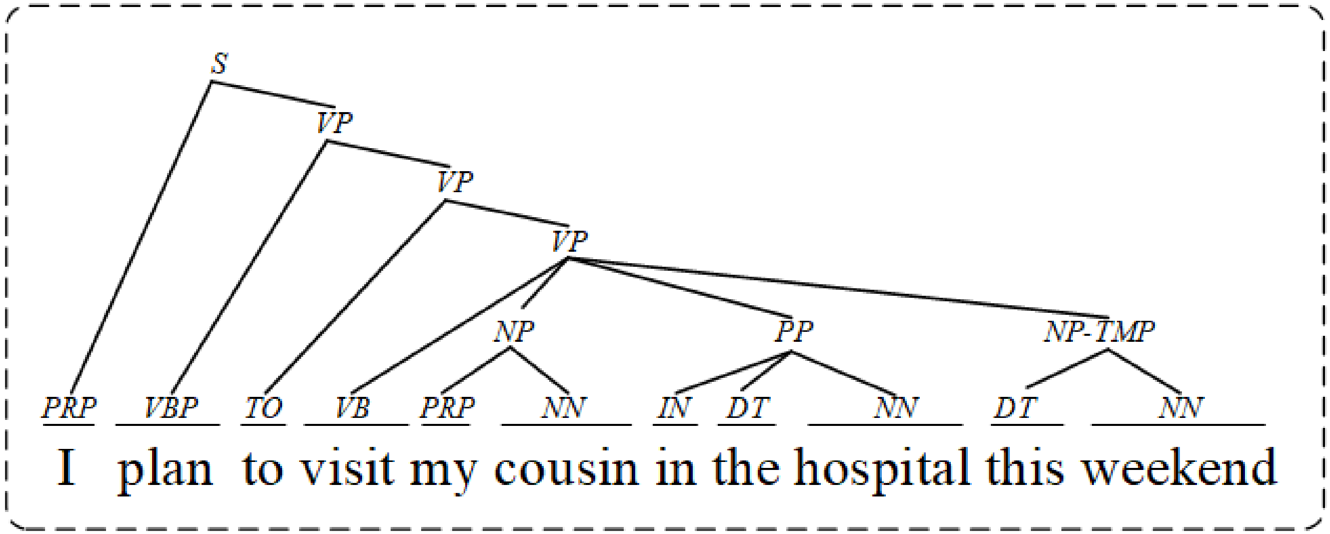
$G = \langle T, N, S, R \rangle$

- T is set of terminals (lexicon)
- N is set of non-terminals For NLP, we usually distinguish out a set $P \subset N$ of *preterminals* which always rewrite as terminals.
- S is start symbol (one of the nonterminals)
- R is rules/productions of the form $X \rightarrow \gamma$, where X is a nonterminal and γ is a sequence of terminals and nonterminals (may be empty).
- A grammar G generates a language L.

What?

- Syntax structure foundations
  - Syntactic constituency tree

*Reveal the **inclusive** and **constituent** relations between **phrases**.*



Rules: Context-free grammar (CFG)

$G = \langle T, N, S, R \rangle$

$T = \{that,\ this,\ a,\ the,\ man,\ book,\ flight,\ meal,\ include,\ read,\ does\}$

$N = \{\text{S, NP, NOM, VP, Det, Noun, Verb, Aux}\}$

$S = S$

$R = \{$

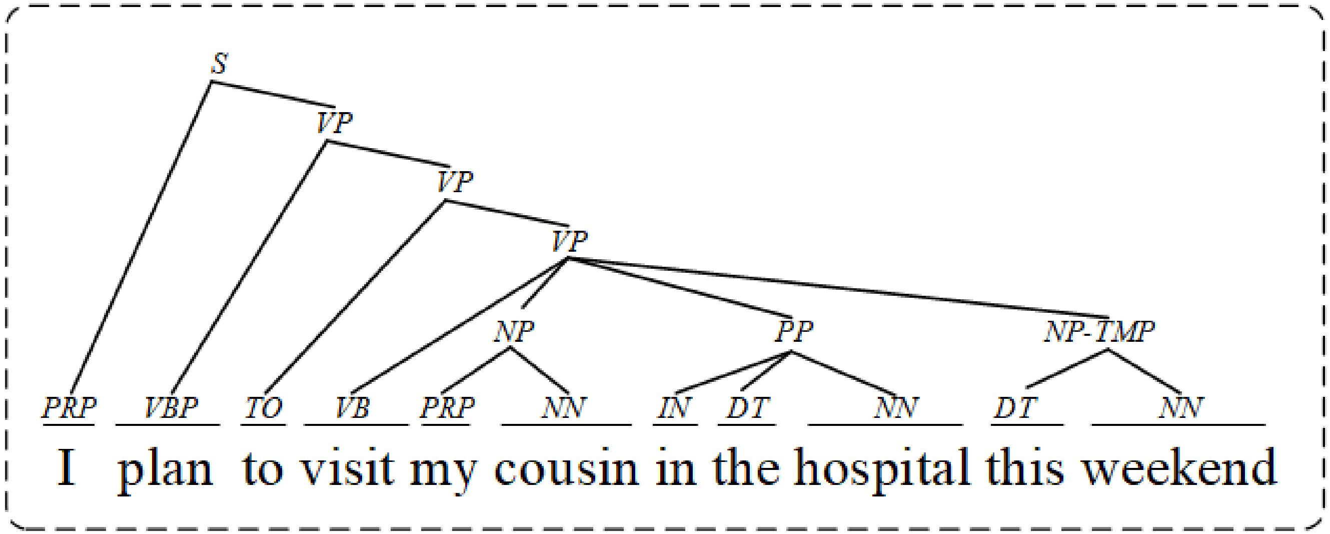
S → NP VP	Det → *that* | *this* | *a* | *the*
S → Aux NP VP	Noun → *book* | *flight* | *meal* | *man*
S → VP	Verb → *book* | *include* | *read*
NP → Det NOM	Aux → *does*
NOM → Noun	
NOM → Noun NOM	
VP → Verb	
VP → Verb NP	

$\}$

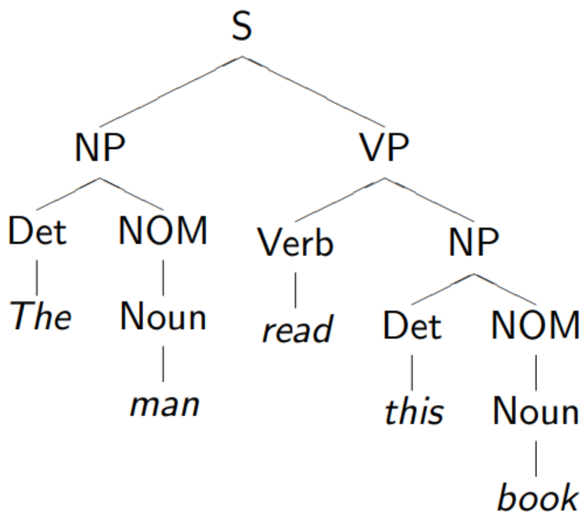
What?

- Syntax structure foundations
  - Syntactic constituency tree

*Reveal the **inclusive** and **constituent** relations between **phrases**.*



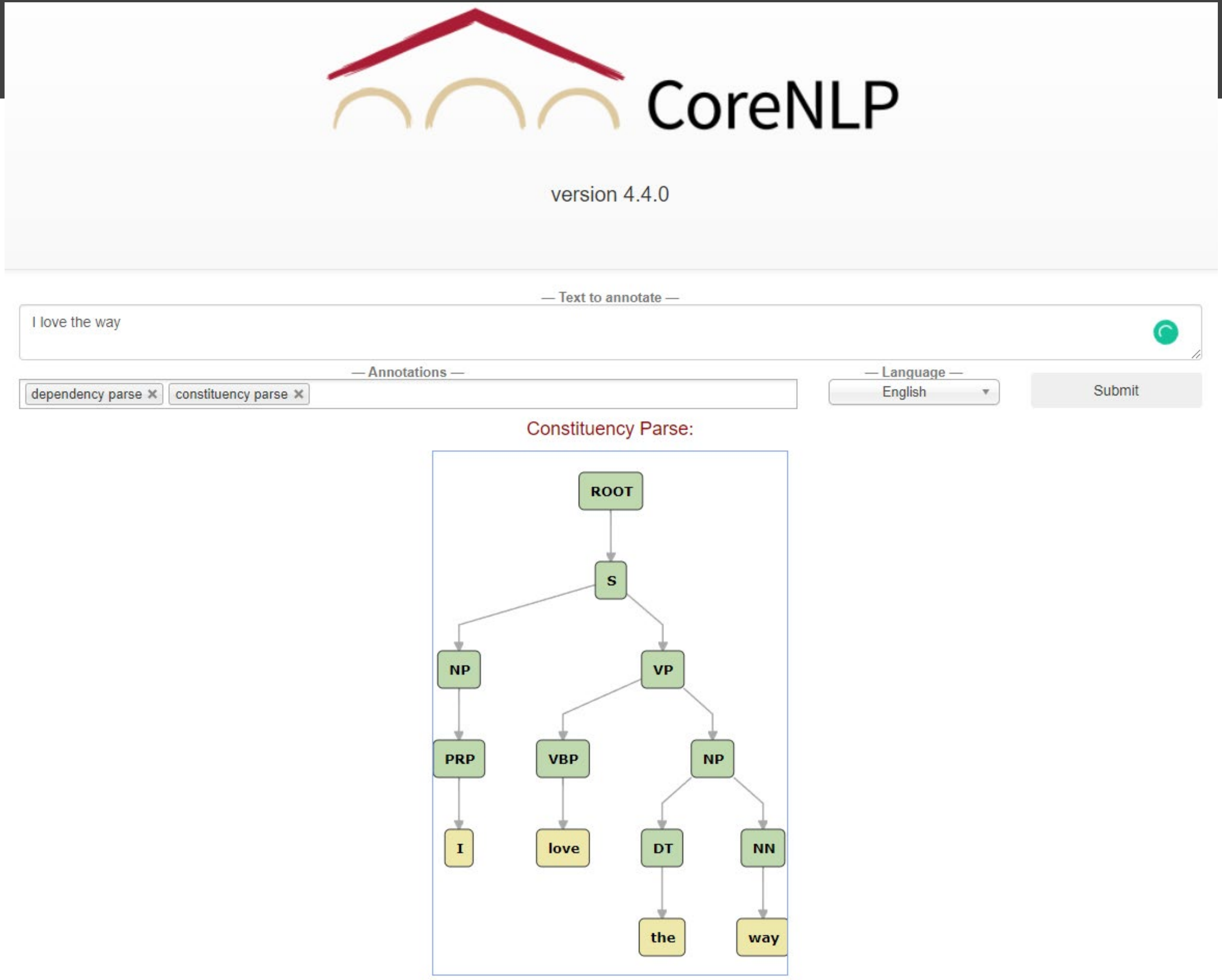
Rules: Context-free grammar (CFG)



[Structure-aware NLP]

What?

- http://corenlp.run/



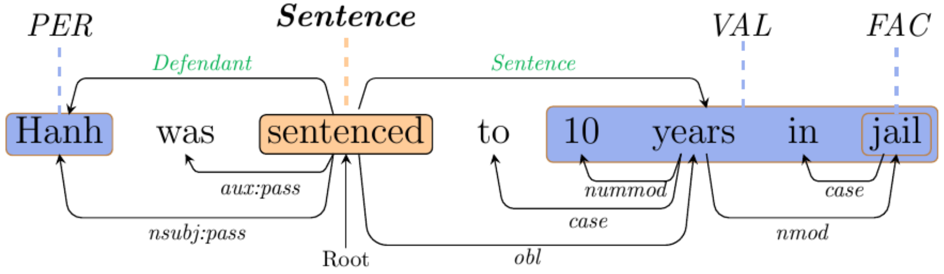
OUTLINE

- WHAT is syntactic structure?
- **WHY integrating structures for NLP?**
- HOW to integrate?
- WHAT to do next?

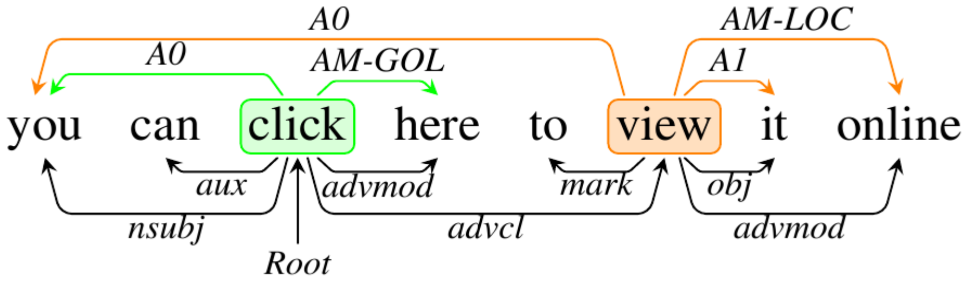
[Structure-aware NLP]
Why?

➢ Aiding NLP with syntax structure

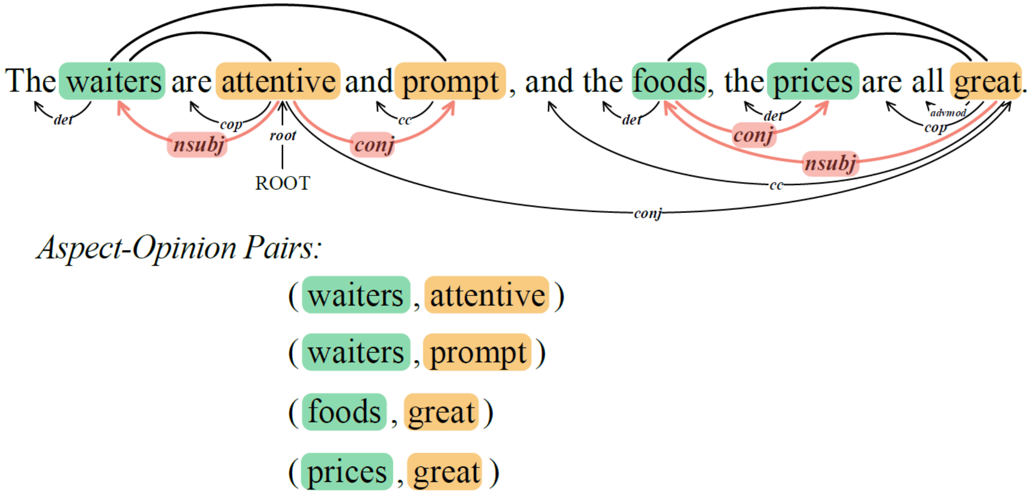
Syntax structure has intrinsically close relationship with semantics, so integrating syntactic structure as external features for improving downstream tasks.

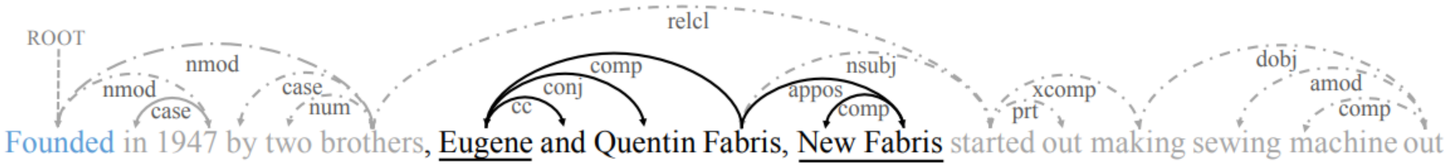


Event extraction



Semantic role labeling





Relation extraction

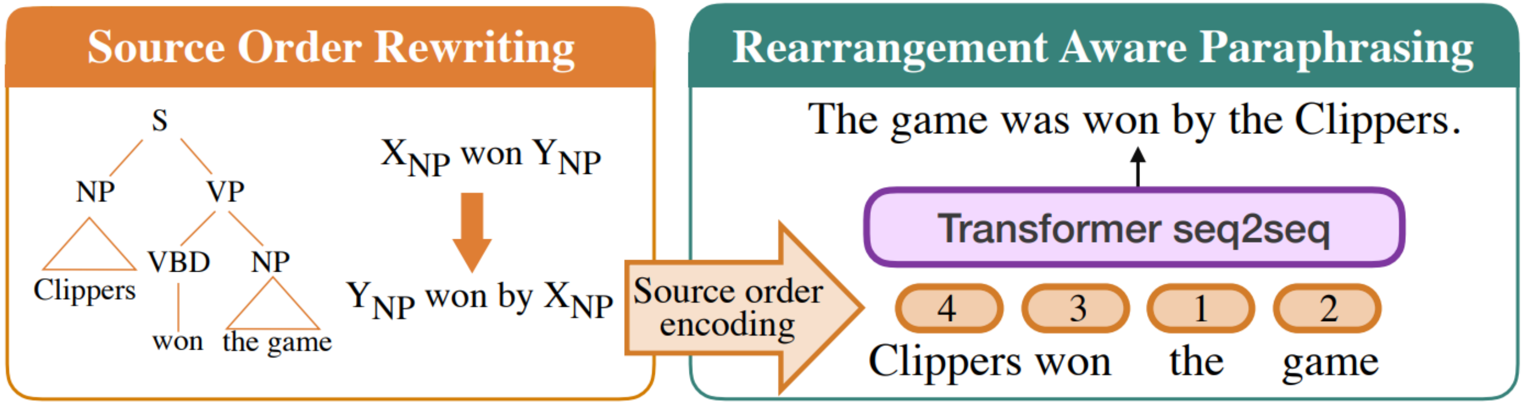
Aspect-opinion pair extraction

[Structure-aware NLP]

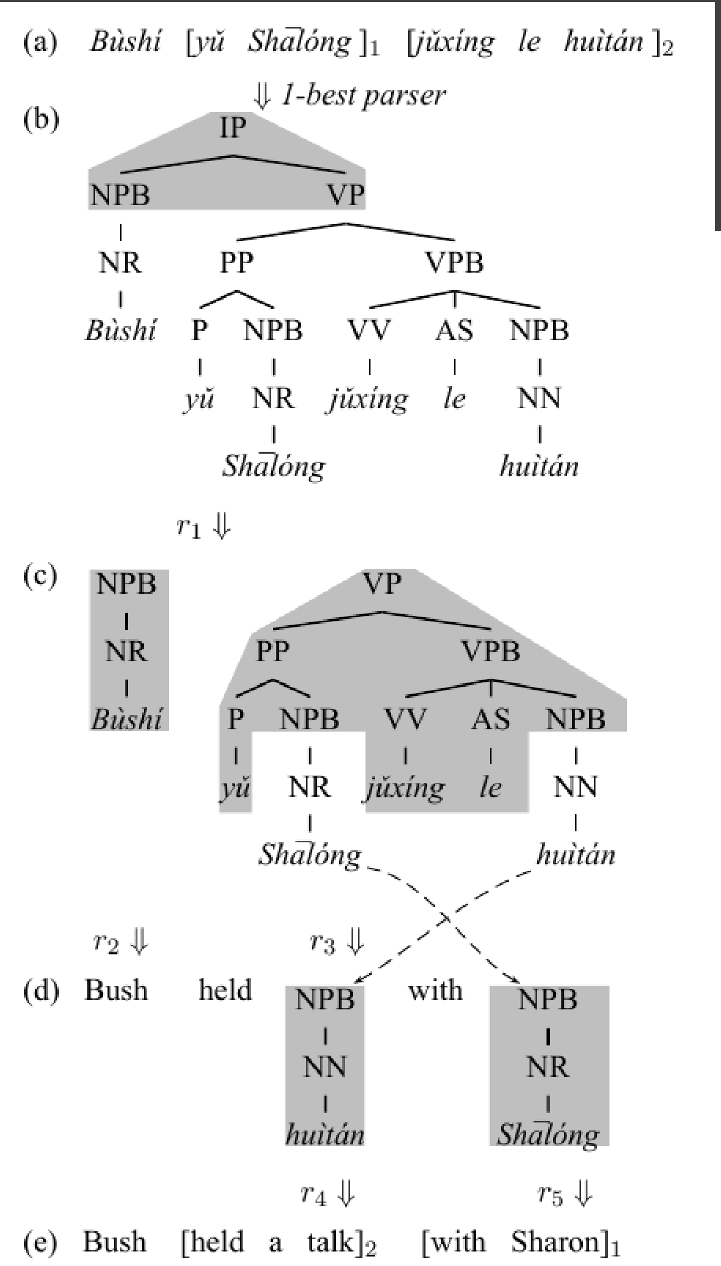
Why?

- Aiding NLP with syntax structure

Syntax structure has intrinsically close relationship with semantics, so integrating syntactic structure as external features for improving downstream tasks.



Paraphrase generation



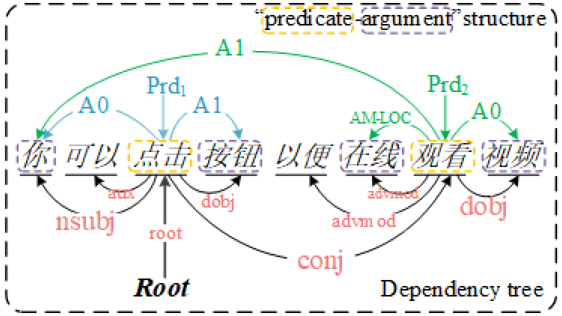
Neural machine translation

[Structure-aware NLP]

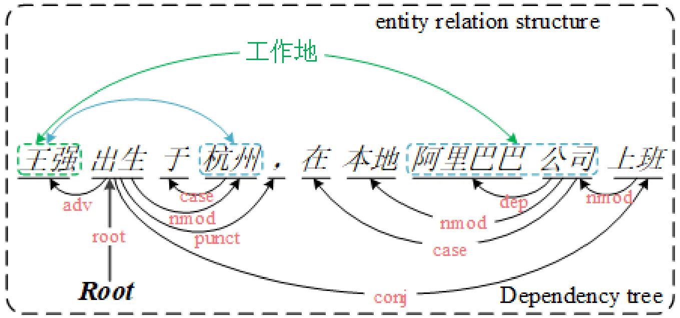
Why?

➢ Aiding NLP with syntax structure

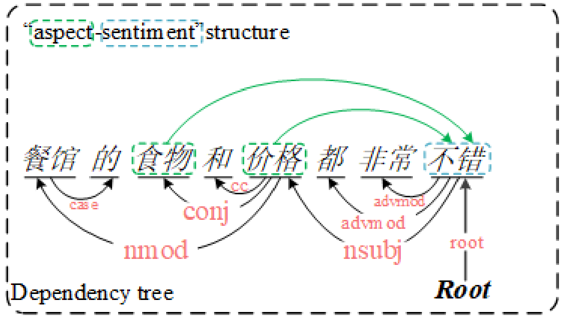
Syntax structure has intrinsically close relationship with semantics, so integrating syntactic structure as external features for improving downstream tasks.



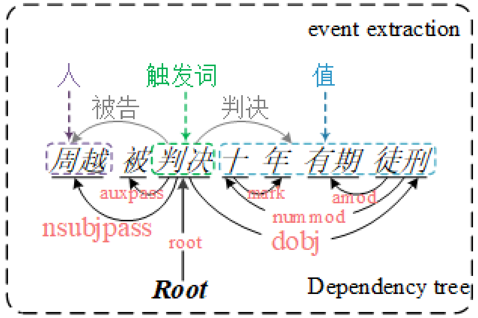
(a) semantic role labeling



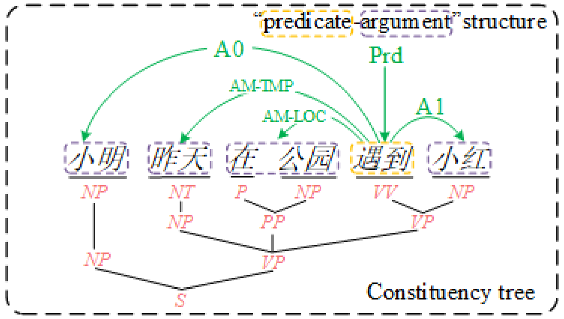
(b) relation extraction



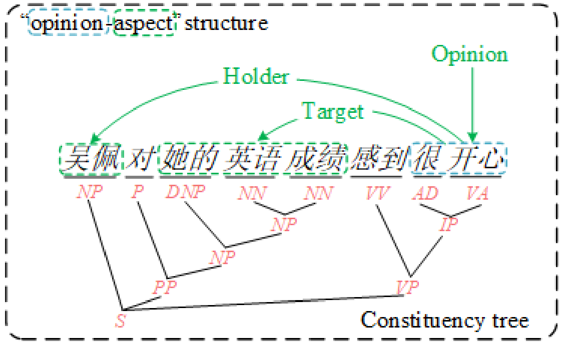
(c) fine-grained sentiment analysis



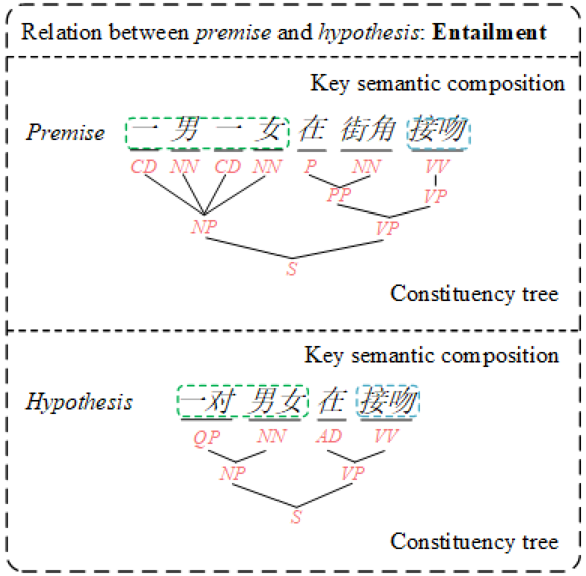
(d) event extraction



(a) semantic role labeling



(b) fine-grained opinion analysis



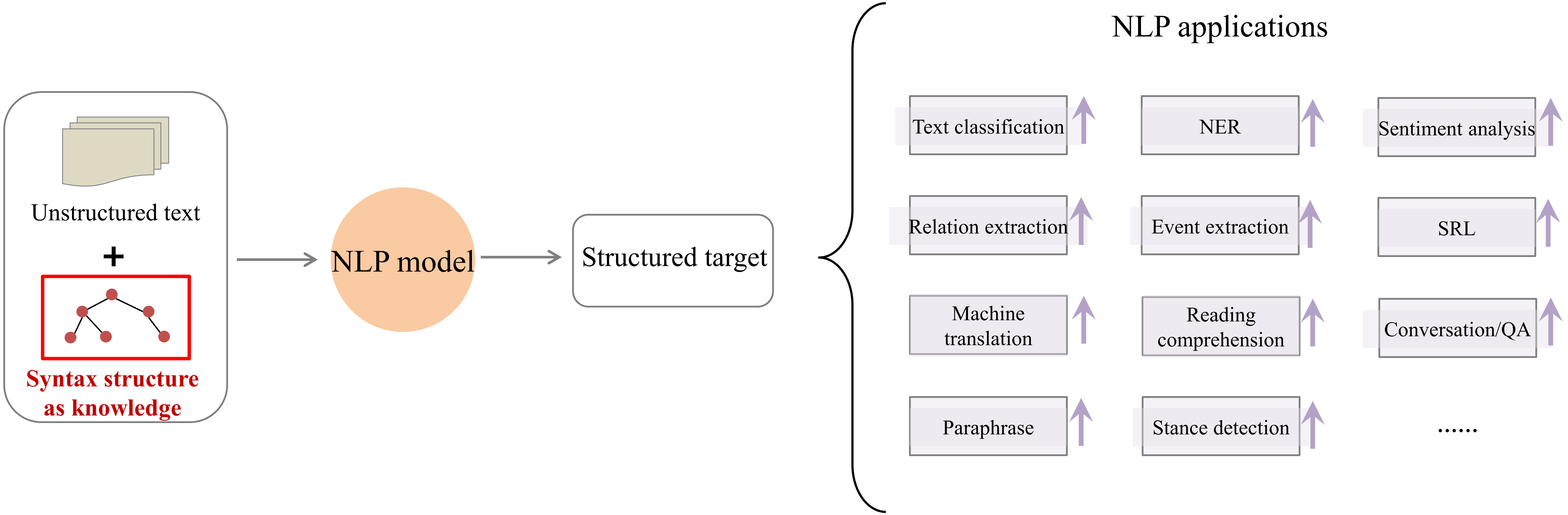
(c) textual entailment recognition

[Structure-aware NLP]

Why?

➢ Aiding NLP with syntax structure

Syntax structure has intrinsically close relationship with semantics, so integrating syntactic structure as external features for improving downstream tasks.



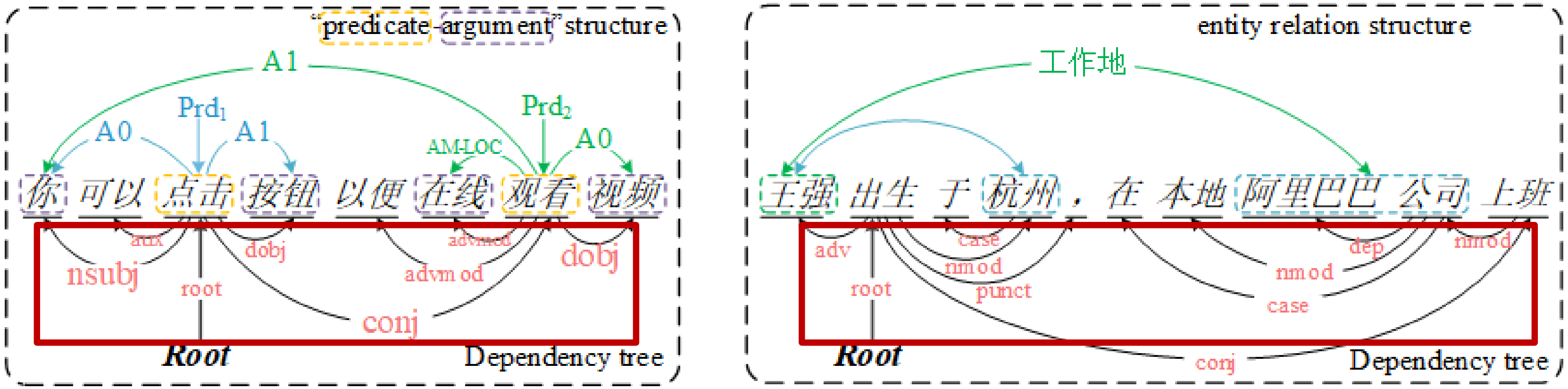
Why?

- Aiding NLP with syntax structure
  - *Enhancing downstream NLP task performances with additional features from low-level linguistic perspective.*
  - *Syntactic structural clues helps better task explainability.*
  - *Syntactic structures, as domain-/language-invariant features, brings up robust transfer learning.*
  - ……

[Structure-aware NLP]

Why?

- Existing challenges
  - Insufficiency and under-optimization of syntax usage
    - Only using the **dependency arcs** between words in the dependency tree;
    - Without considering the use of the **syntactic labels** attached to the dependency arcs.



(a) semantic role labeling (b) relation extraction

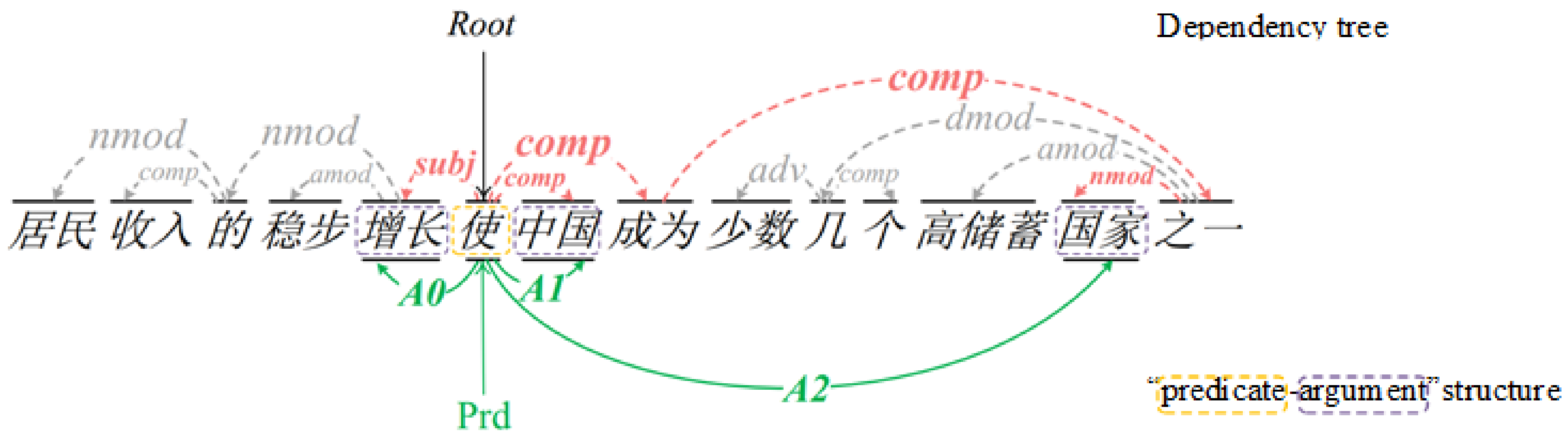
***syntactic label information** attached helps inference equally.*

[Structure-aware NLP]

Why?

➢ Existing challenges

- Insufficiency and under-optimization of syntax usage
  - Not all the entire information within syntactic trees provides positive help:



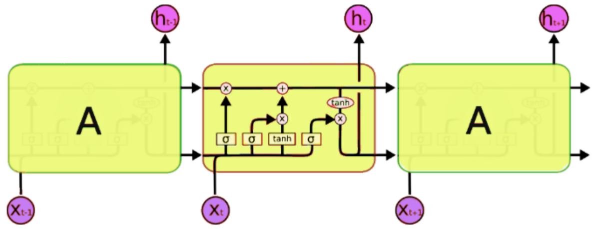
Colored structures: useful
Gray structures: noise

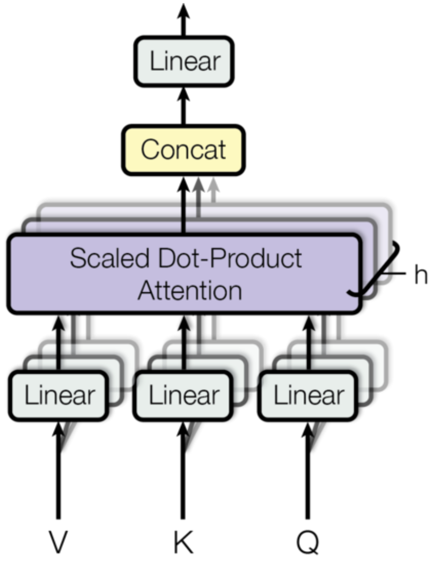
[Structure-aware NLP]

Why?

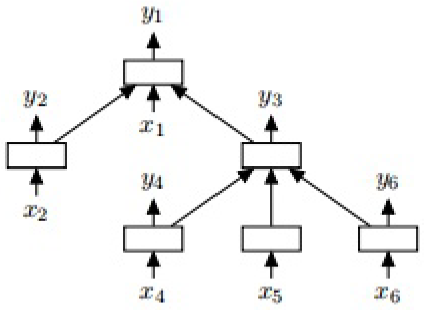
- Existing challenges
  - Shallow structure integration
    - Step1: Encoding syntax with encoders

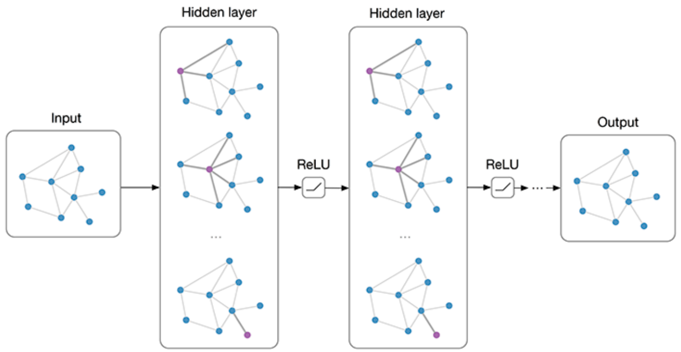
1) **sequential neural** encoders, e.g., LSTM, Attention-based models,





2) **hierarchical encoders**, e.g., TreeLSTM, GCN models, etc.





[Structure-aware NLP]

Why?

- Existing challenges
  - Insufficiently structure integration
    - Step2: *Concatenating* the syntax representation with word embedding as input feature representation.

where the problem come from

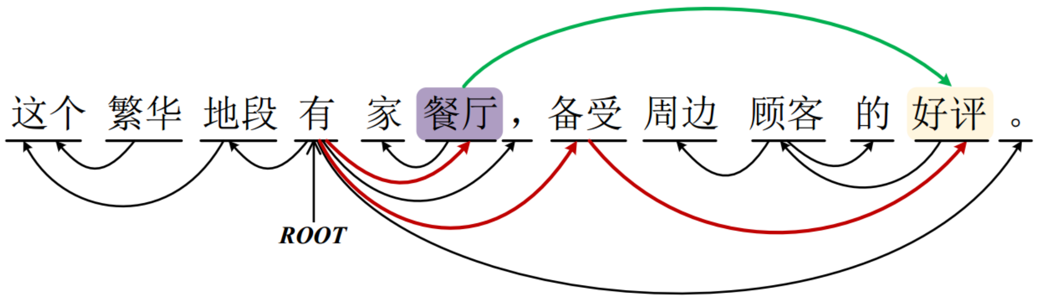
1) *Flattening the hierarchy characteristic of the syntax structure.*

2) *Shallow interaction between syntactics and semantics.*

[Structure-aware NLP]

Why?

- Existing challenges
  - Fixed/rigid syntax structure
    - *Parsing syntax in fixed and rigid tree comes with task-irrelevant substructures, which would deteriorate the efficacy.*



    - *Meanwhile, different NLP tasks rely on distinct bias of structural features.*

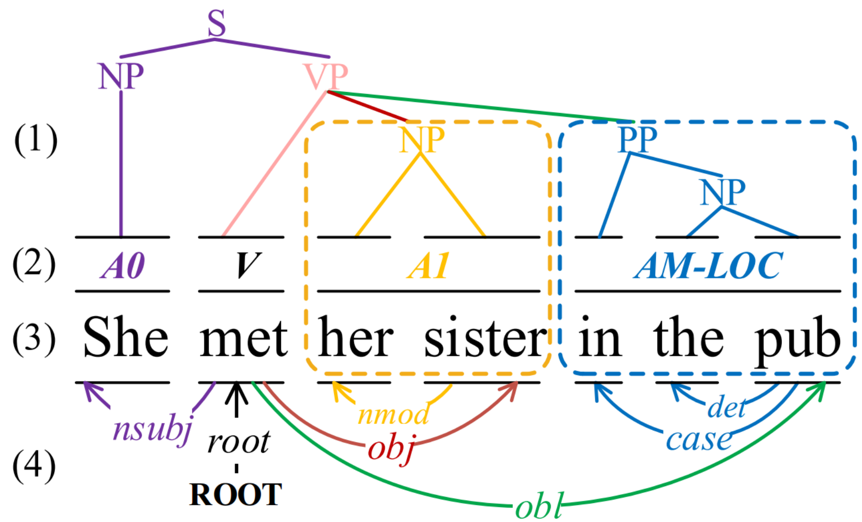
(*a*) parse tree (*b*) balanced tree (*c*) Gumbel tree (*d*) left-branch tree (*e*) right-branch tree

Why?

- Existing challenges
  - Singleton type of syntax structure integration

    1) Making use of only **one singleton syntax** information, i.e., the **dependency syntax**, or sometimes the **constituency tree**.

    2) **Constituent** and **dependency** syntax depict the syntactic structure from different perspectives:



- **Dependency** structure depicts the inter-**relations** between words,
- **Constituency** structure locates more about **span boundaries** of mentions.

[Structure-aware NLP]

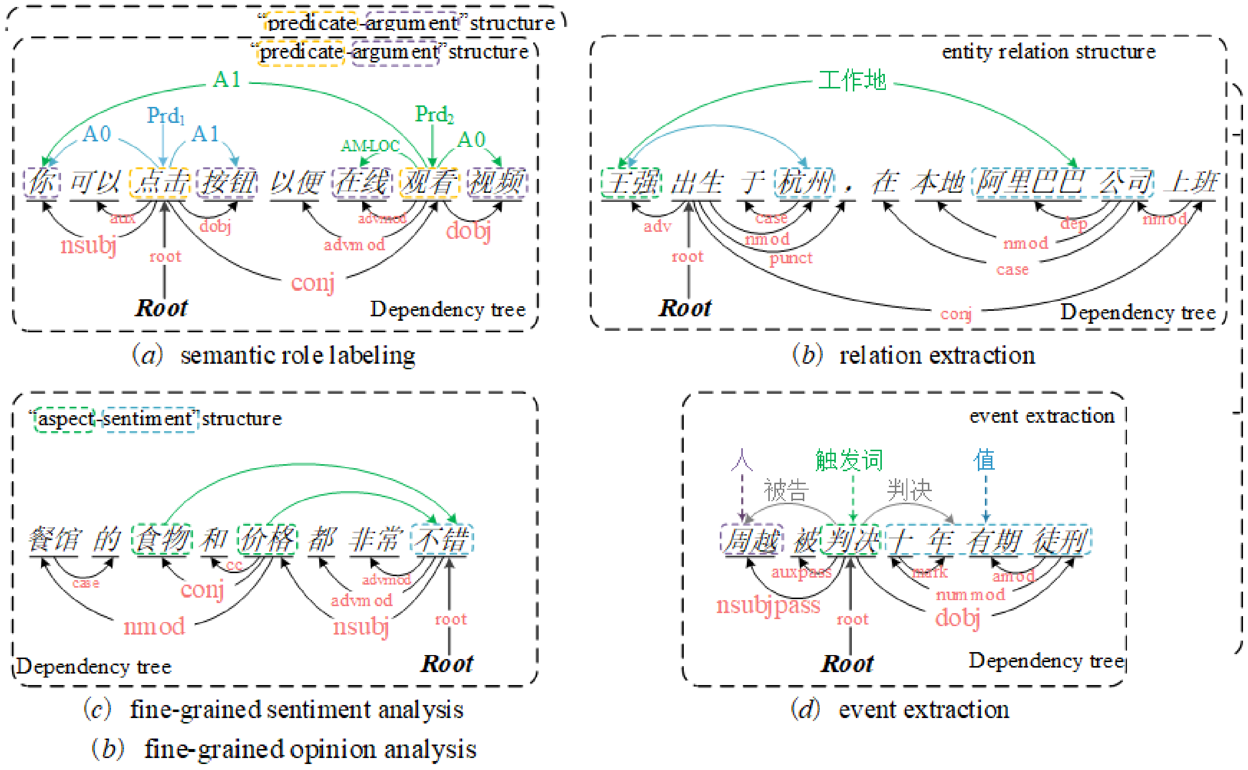
Why?

➢ Existing challenges

- Singleton type of syntax structure integration

  ➢ **Dependency** structure depicts the inter-**relations** between words;

  ➢ **Constituency** structure locates more about **span boundaries** of mentions.



Simultaneously integrating these two heterogeneous representations can bring complementary advantages!

Why?

- Existing challenges
  - Unexploitation of syntax integration on varying scenarios beside sentence
    - *Document-level tasks?*
      - *Multiple sentences, how to model the intrinsic document structure?*
    - *Dialogue-level tasks?*
      - *Multiple separate utterances, how to model the overall structure?*
      - *Multiple speakers, entangled threads.*
      - *Speaker coreference issue.*

OUTLINE

- WHAT is syntactic structure?
- WHY integrating structures for NLP?
- **HOW to integrate?**
- WHAT to do next?

[Structure-aware NLP]
How?

➢ **Preliminary-A**: Syntax structure parsers

- Graph-based parser

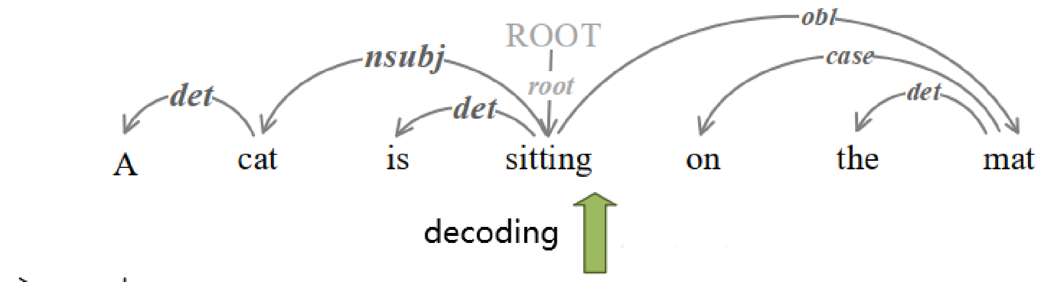
➢ *Parsing task is regarded as the process of building a tree, searching a weighted graph to find the subgraph with the highest score that meets the grammar rules.*

Pros:

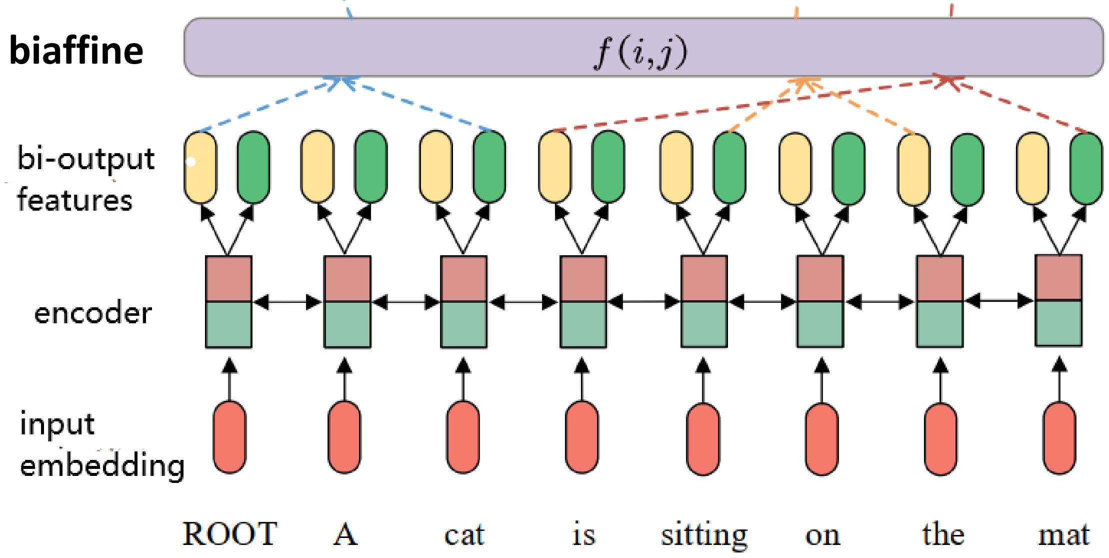
Global-level feature modeling

Cons:

Iteratively enumerating, higher complexity



head \ dep	ROOT	A	cat	is	sitting	on	the	mat
ROOT		0	0	0.1	0.8	0.2	0.1	0.3
A	0.1		0.3	0.4	0.3	0	0.2	0.1
cat	0.2	0.6		0.2	0.1	0.3	0.4	0.2
is	0.4	0.2	0.4		0	0.1	0	0.5
sitting	0	0.5	0.7	0.8		0.5	0.2	0.6
on	0.3	0	0.3	0	0.3		0.5	0.4
the	0.4	0.1	0	0.2	0	0.2		0.2
mat	0.5	0.3	0.4	0.2	0.4	0.7	0.8	



[Structure-aware NLP]

How?

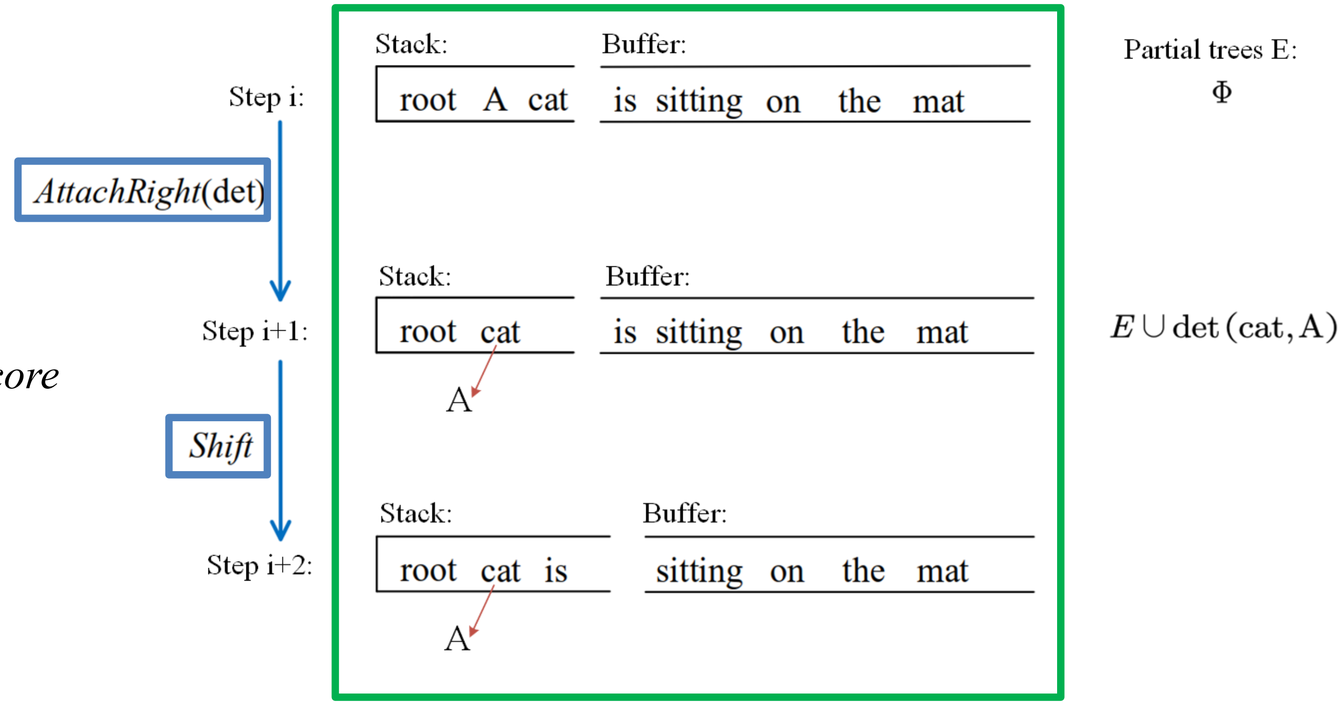
➢ **Preliminary-A**: Syntax structure parsers

- Transition-based parser

  ➢ *The process of a finite state automata.*

  ➢ *Transition process from initial state to terminal state.*

  ➢ *The transition framework consists of two core elements: **Action** and **State**.*



Pros:

Linear decoding, lower complexity

Cons:

Local-level feature modeling

[Structure-aware NLP]

How?

- **Preliminary-A**: Syntax structure parsers

- Transition steps

Sentence : He_1 says_2 the_3 agency_4 seriously_5 needs_6 money_7 to_8 develop_9

Step	Action	σ^o	α^o	λ	σ^r	α^r	β	Ptr	Y
0	-	[]	[]	Null	[]	[]	[1,···,9]		
1	R-START	[]	[]	$(1,1)^r$	[]	[]	[1,···,9]	[1,···,9]	
2	SHIFT	[]	[]	Null	[(1,1)]	[]	[2,···,9]		
3	O-START	[]	[]	$(2,2)^o$	[(1,1)]	[]	[2,···,9]	[2,···,9]	
4	ARC	[]	[]	$(2,2)^o$	[]	[(1,1)]	[2,···,9]		$Y \cup \{\langle(2,2)^o,(1,1)^r(hd)\rangle\}$
5	SHIFT	[(2,2)]	[]	Null	[(1,1)]	[]	[3,···,9]		
6	R-START	[(2,2)]	[]	$(3,4)^r$	[(1,1)]	[]	[3,···,9]	[3,4,···,9]	
7	ARC	[]	[(2,2)]	$(3,4)^r$	[(1,1)]	[]	[3,···,9]		$Y \cup \{\langle(2,2)^o,(3,4)^r(tg)\rangle\}$
8	SHIFT	[(2,2)]	[]	Null	[(1,1),(3,4)]	[]	[4,···,9]		
9	NO-START	[(2,2)]	[]	Null	[(1,1),(3,4)]	[]	[5,···,9]		
10	O-START	[(2,2)]	[]	$(5,6)^o$	[(1,1),(3,4)]	[]	[5,···,9]	[5,6,···,9]	
11	ARC	[(2,2)]	[]	$(5,6)^o$	[(1,1)]	[(3,4]	[5,···,9]		$Y \cup \{\langle(5,6)^o,(3,4)^r(hd)\rangle\}$
12	NO-ARC	[(2,2)]	[]	$(5,6)^o$	[]	[(1,1),(3,4]	[5,···,9]		
13	SHIFT	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4]	[]	[6,···,9]		
14	NO-START	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4]	[]	[7,8,9]		
15	R-START	[(2,2),(5,6)]	[]	$(7,9)^r$	[(1,1),(3,4]	[]	[7,8,9]	[7,8,9]	
16	ARC	[(2,2)]	[(5,6)]	$(7,9)^r$	[(1,1),(3,4]	[]	[7,8,9]		$Y \cup \{\langle(5,6)^o,(7,9)^r(tg)\rangle\}$
17	NO-ARC	[]	[(2,2),(5,6)]	$(7,9)^r$	[(1,1),(3,4]	[]	[7,8,9]		
18	SHIFT	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4,(7,9)]	[]	[8,9]		
19	NO-START	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4,((7,9)]	[]	[9]		
20	NO-START	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4,(7,9)]	[]	[]		

[Structure-aware NLP]

How?

- **Preliminary-A**: Syntax structure parsers

  - Graph-based parser
  - Transition-based parser
  - Table-filling extractor
  - PointerNet extractor
  - MRC extractor
  - Hypergraph extractor
  - ……

Structure parser: Complex information extraction

- *Nested NER*
- *Discontinuous NER*
- *Overlapped RE*
- *Overlapped EE*
- …

[Structure-aware NLP]

How?

➢ **Preliminary-B**: Complete modeling of NLP tasks

(almost) **All NLP tasks**

- Text-pair classification
- Text classification
- Span classification
- Word/token classification
  - Input-output synchronized classification
  - Input-output asynchronized classification

[Structure-aware NLP]

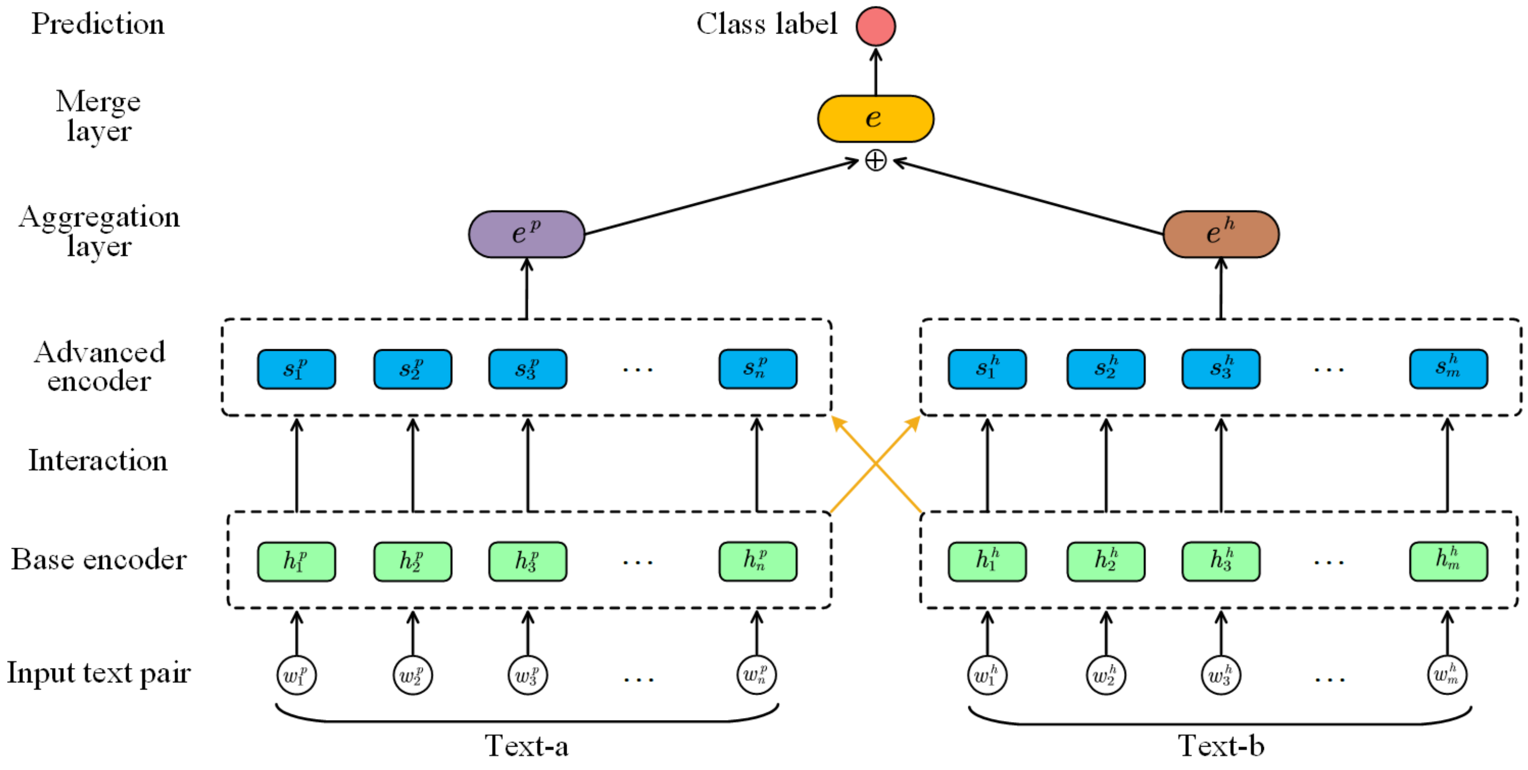
How?

➢ **Preliminary-B**: Complete modeling of NLP tasks

- Text-pair classification

Representative NLP tasks:

1. Recognition of Text Entailment (RTE)
2. Natural language inference
3. Paraphrase Identification
…



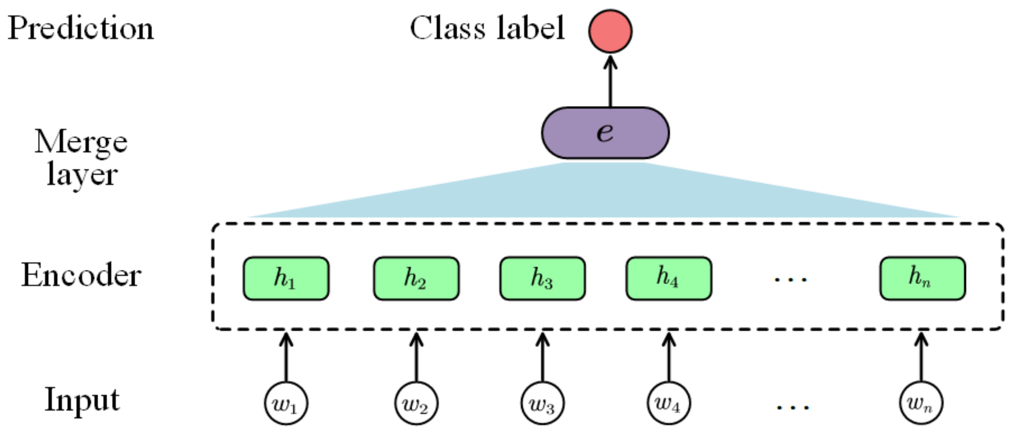
[Structure-aware NLP]
How?

➢ **Preliminary-B**: Complete modeling of NLP tasks

- Text classification

Representative NLP tasks：

1. *Relation classification*
2. *Topic classification*
3. *Sentiment classification*
4. *Question type classification*
5. *Intention classification*
6. *Emotion classification*
7. *Aggressive language classification*
8. …



[Structure-aware
How?

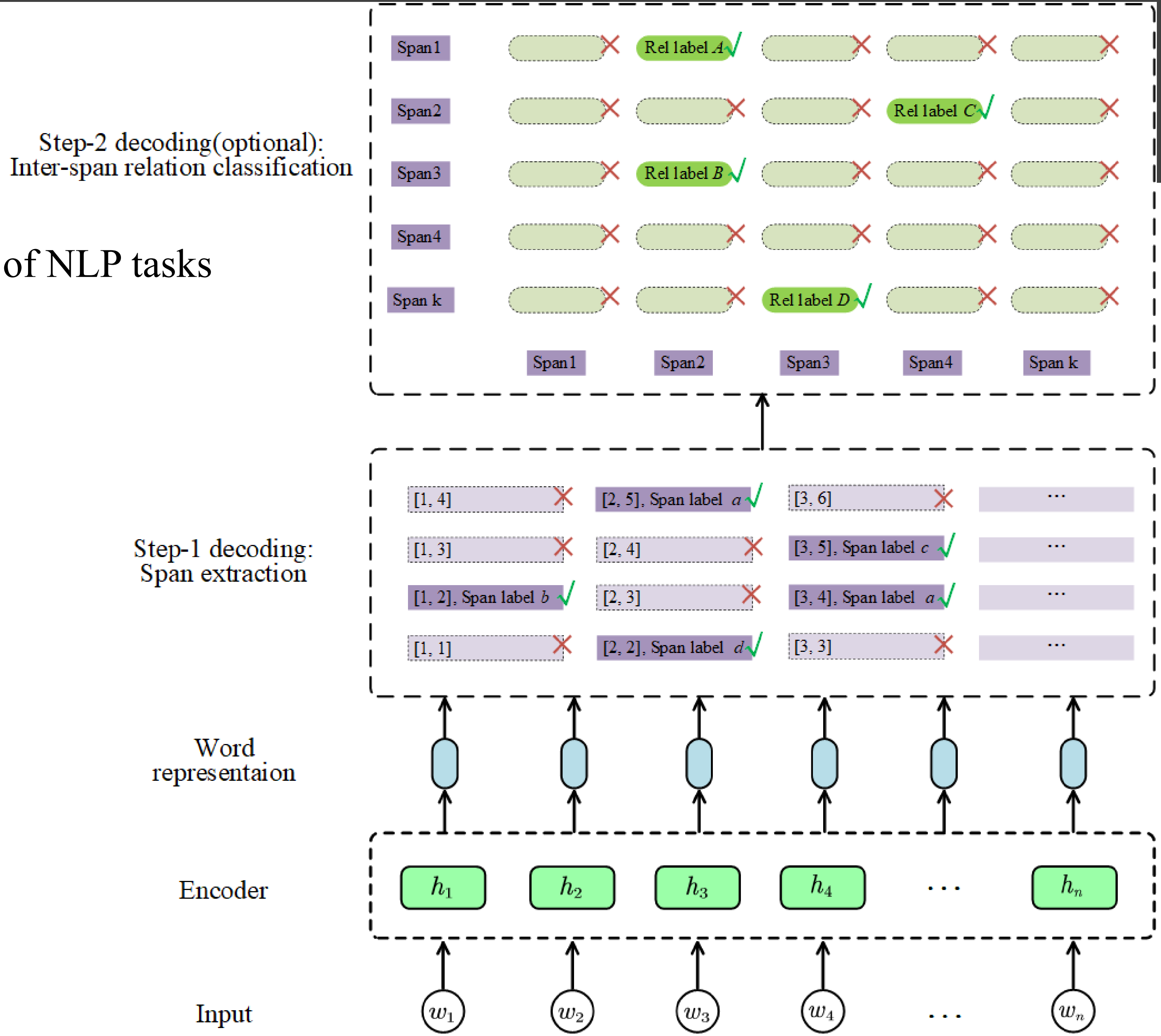
➢ **Preliminary-B**: Complete modeling of NLP tasks

- Span classification

  - **Step-I: span extraction**
  - **Step-II: span-relation classification**

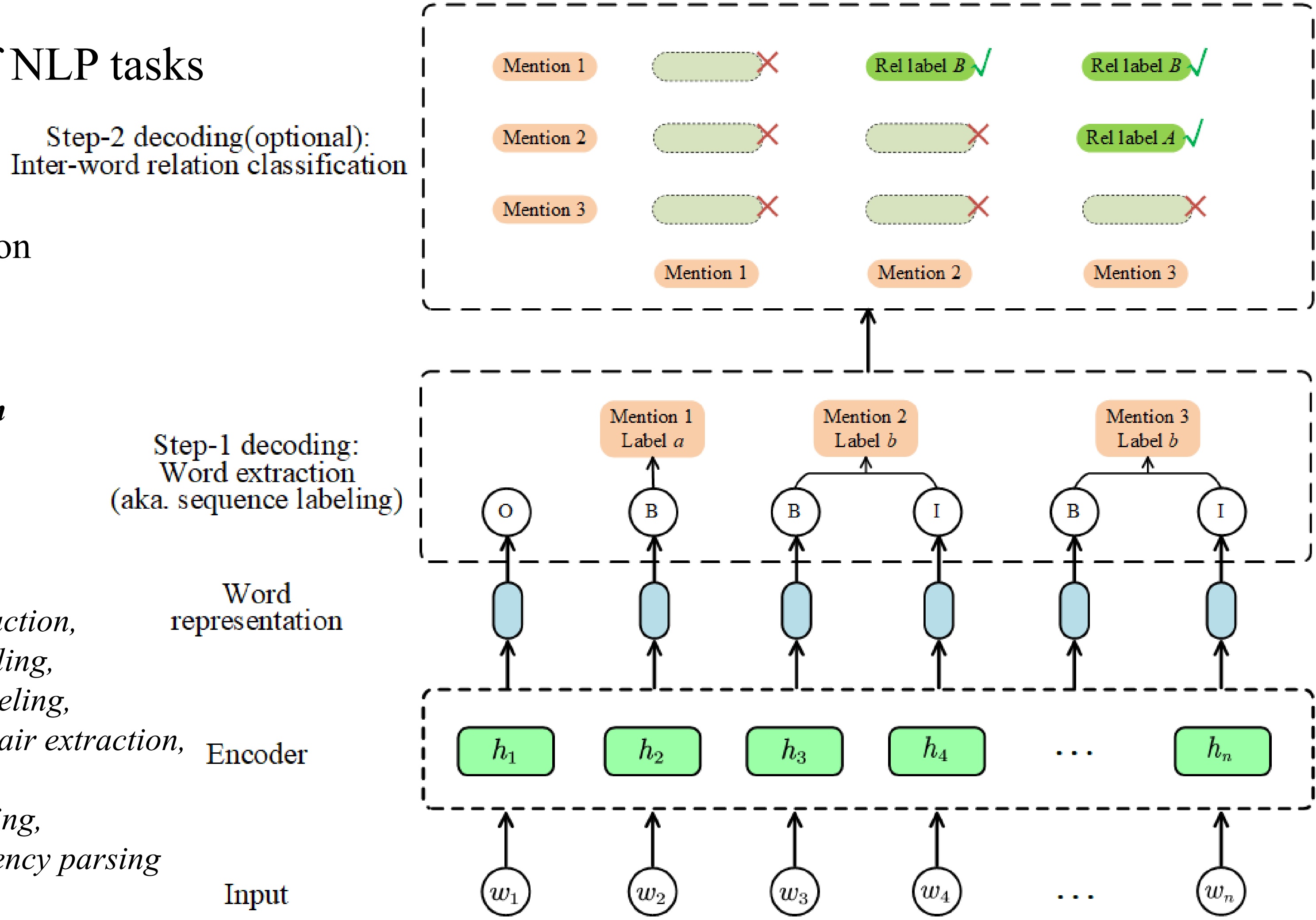
Representative NLP tasks:

1. *Machine reading comprehension*
2. *Extractive automatic summarization*
3. *Nested NER*
4. *Constituency parsing*
5. *Nested RE*
6. *Coreference/anaphora resolution*
7. *...*



[Structure-aware NLP]
How?

- **Preliminary-B**: Complete modeling of NLP tasks
  - Word/token classification
    - Input-output **synchronized** classification
      - ***Step-I：sequence labeling***
      - ***Step-II：mention-relation classification***



Representative NLP tasks：

Step 1:

1. *chunk analysis,*
2. *part of speech tagging,*
3. *named entity recognition,*
4. *Chinese word segmentation,*
5. *fine-grained emotion analysis,*
6. *stance extraction,*
7. *autoregressive language modeling,*
8. …

Step 2:

1. *relationship extraction,*
2. *opinion role labeling,*
3. *semantic role labeling,*
4. *Opinion-aspect pair extraction,*
5. *event extraction,*
6. *dependency parsing,*
7. *semantic dependency parsing*
8. …

How?

➢ **Preliminary-B**: Complete modeling of NLP tasks

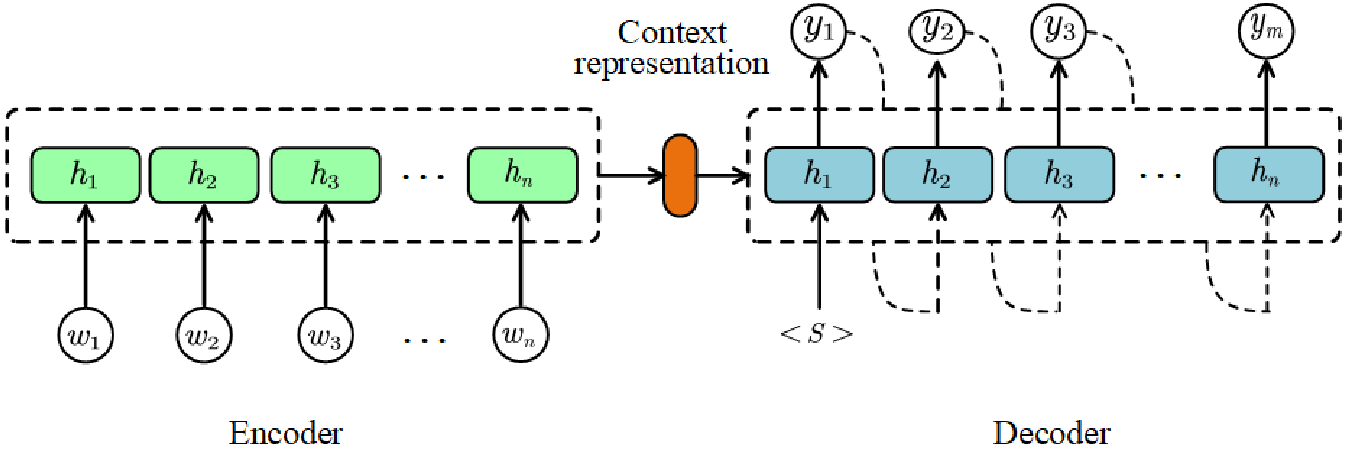
- Word/token classification
  - Input-output **asynchronized** classification

Aka.

➢ *Sequence-to-Sequence framework*
➢ *Encoder-Decoder framework*
➢ *End-to-end framework*

Representative NLP tasks:

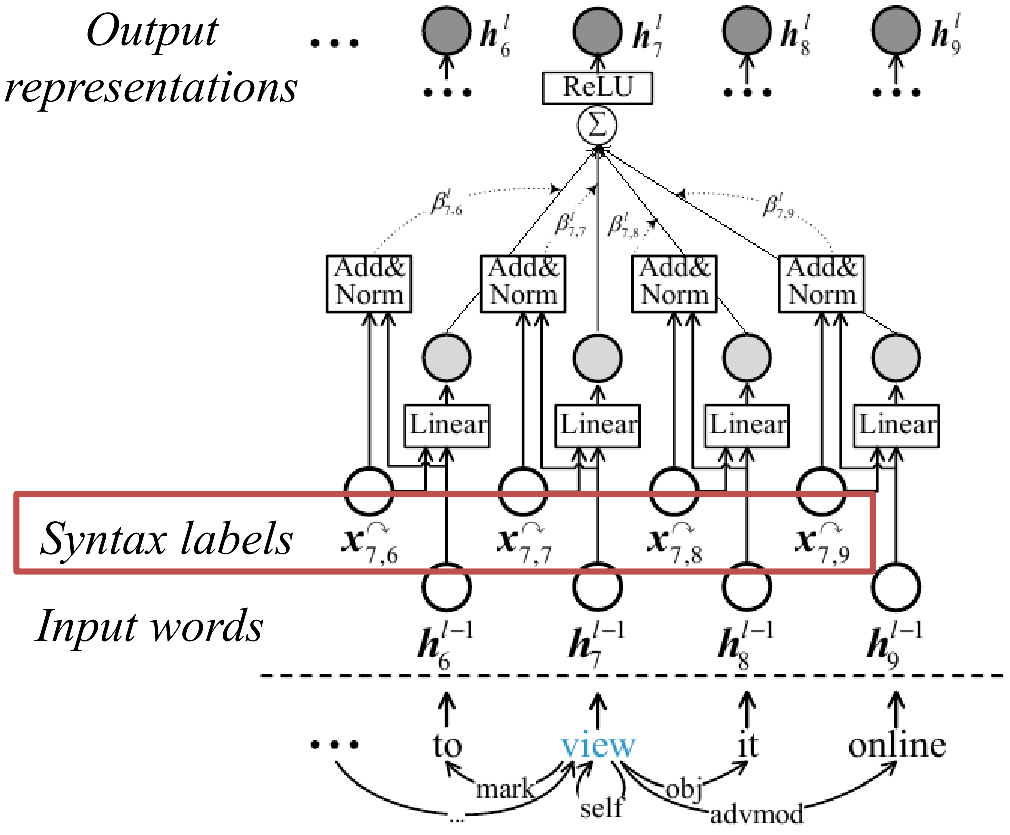
1. *Neural machine translation,*
2. *automatic summarization,*
3. *dialogue system,*
4. *autoregressive language modeling*
5. …



[Structure-aware NLP]

How?

- Case-I: Making use of the syntax label feature
  - Label-aware GCN
    - *GCN backbone*
    - *Simultaneously encoding the syntax label representations*



[1] Hao Fei, Fei Li, Bobo Li, Donghong Ji. Encoder-Decoder Based Unified Semantic Role Labeling with Label-Aware Syntax. AAAI. 2021.

[Structure-aware NLP]

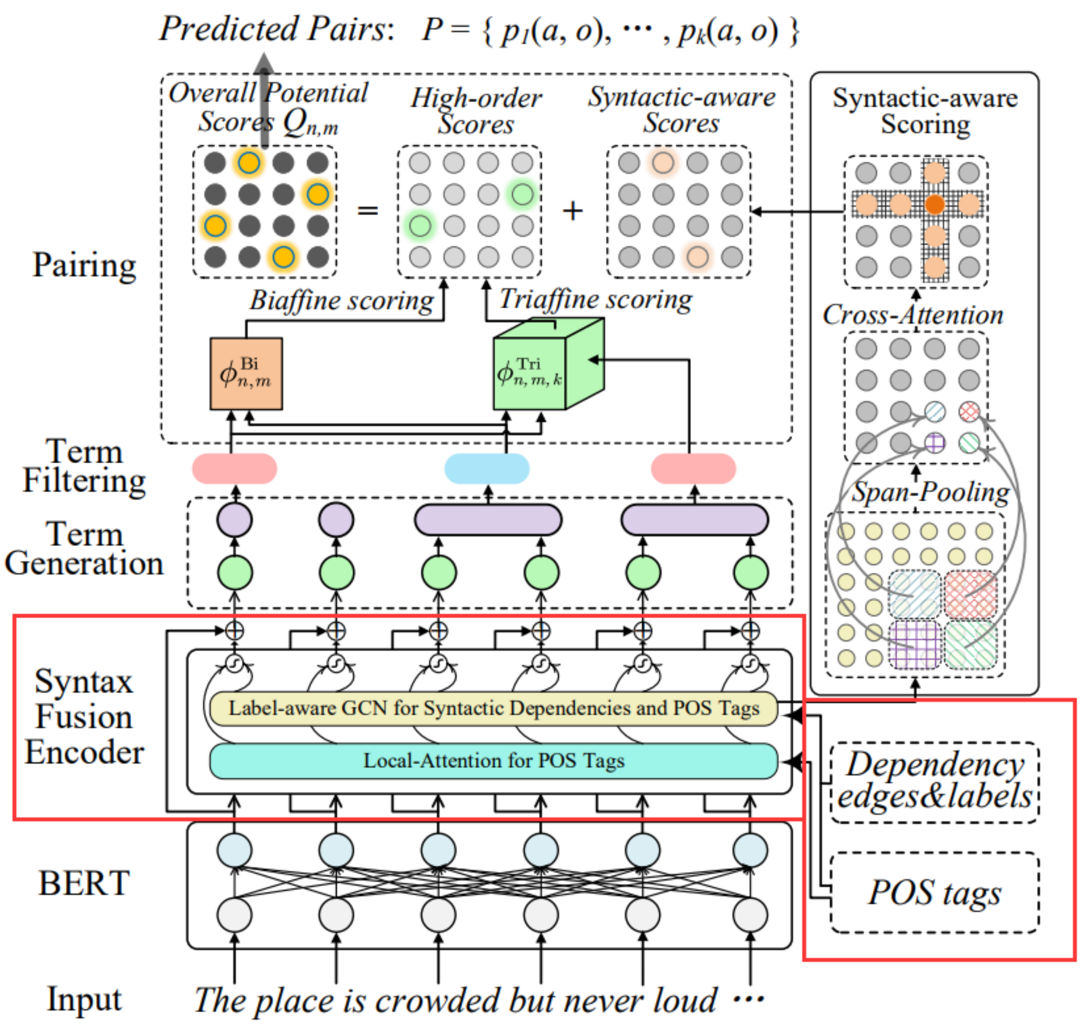
How?

- Case-I: Making use of the syntax label feature
  - Syntax GCN
    - *GCN backbone*
    - *Encoding:*

      1) the dependency edges and

      2) syntax labels

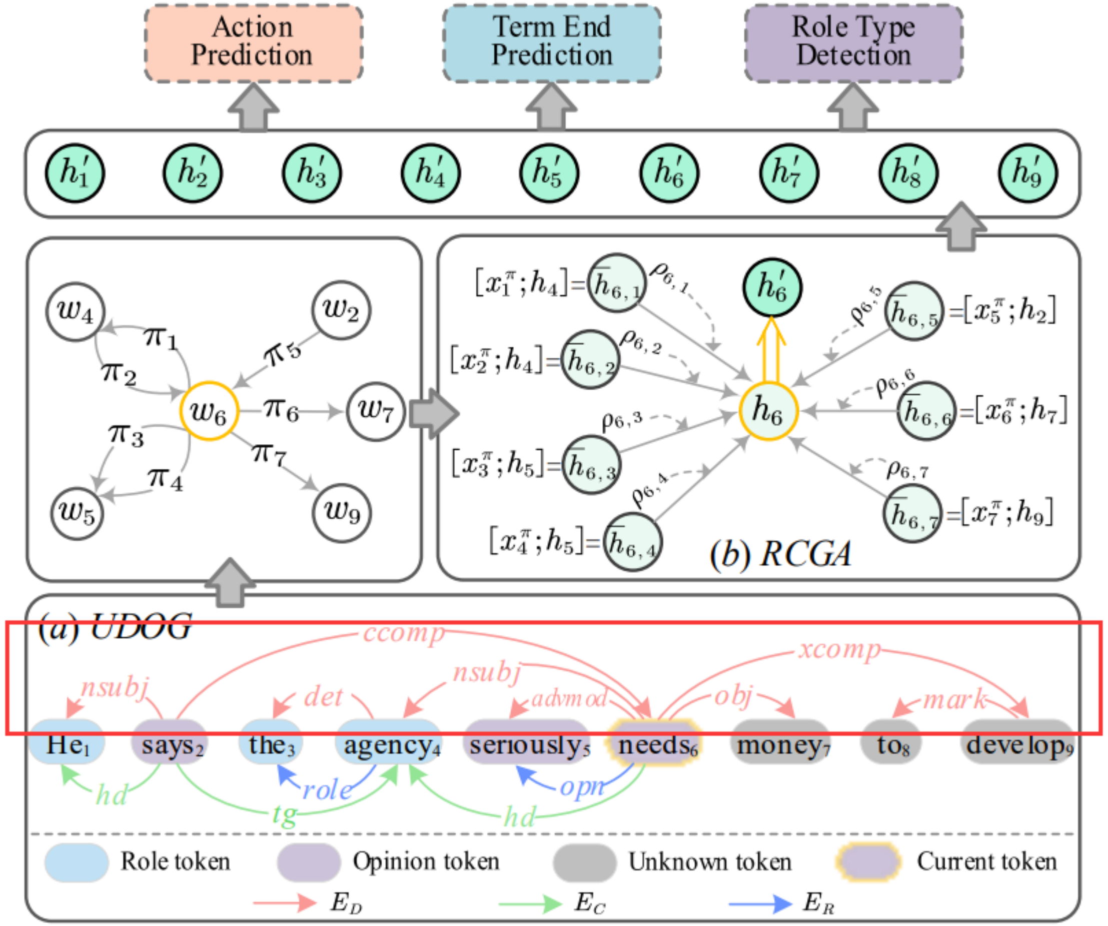
      3) POS tags



[1] Shengqiong Wu, Hao Fei, Yafeng Ren, Donghong Ji, Jingye Li. Learn from Syntax: Improving Pair-wise Aspect and Opinion Terms Extraction with Rich Syntactic Knowledge. IJCAI. 2021.

How?

- Case-I: Making use of the syntax label feature
  - Dependency-aid relation-centered graph aggregator
    - *Graph with multi-relational edges*
    - *Using dependency trees for high-order feature aggregation*

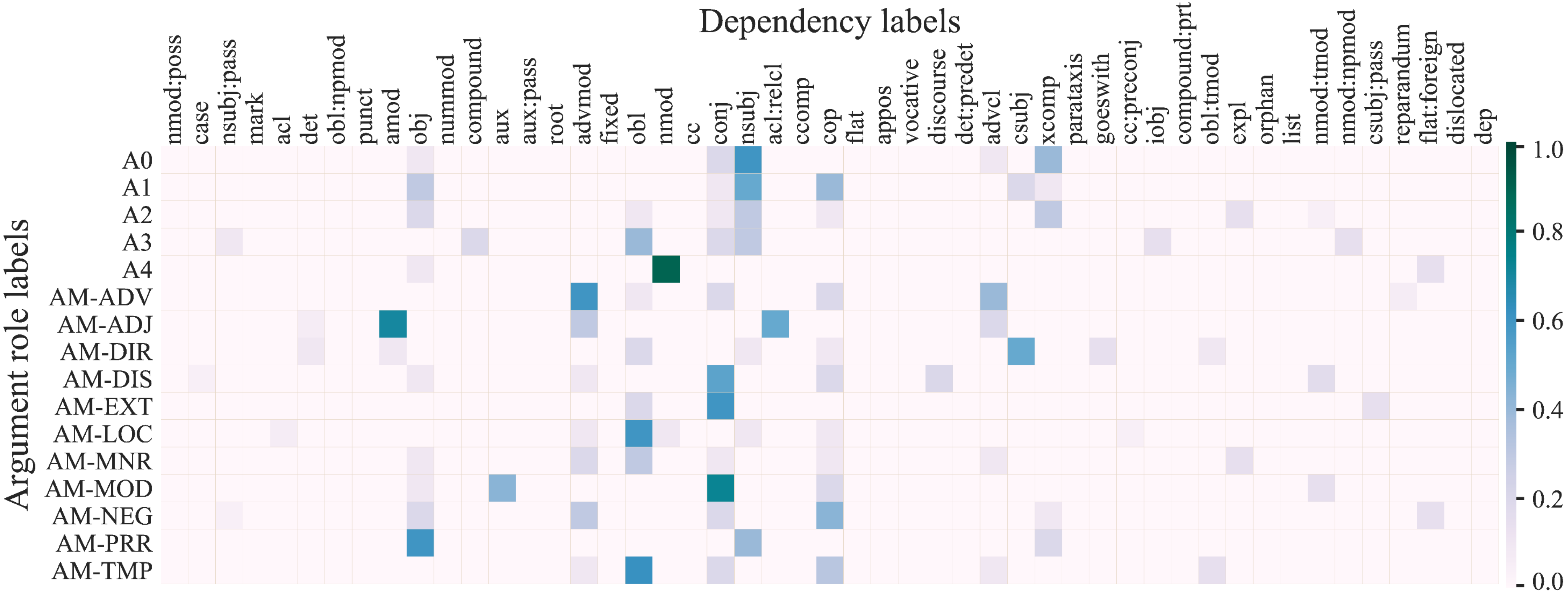


[1] Shengqiong Wu, Hao Fei, Fei Li, Meishan Zhang, etc. Mastering the Explicit Opinion-role Interaction: Syntax-aided Neural Transition System for Unified Opinion Role Labeling. AAAI. 2022.

How?

- ➢ Case-I: Making use of the syntax label feature

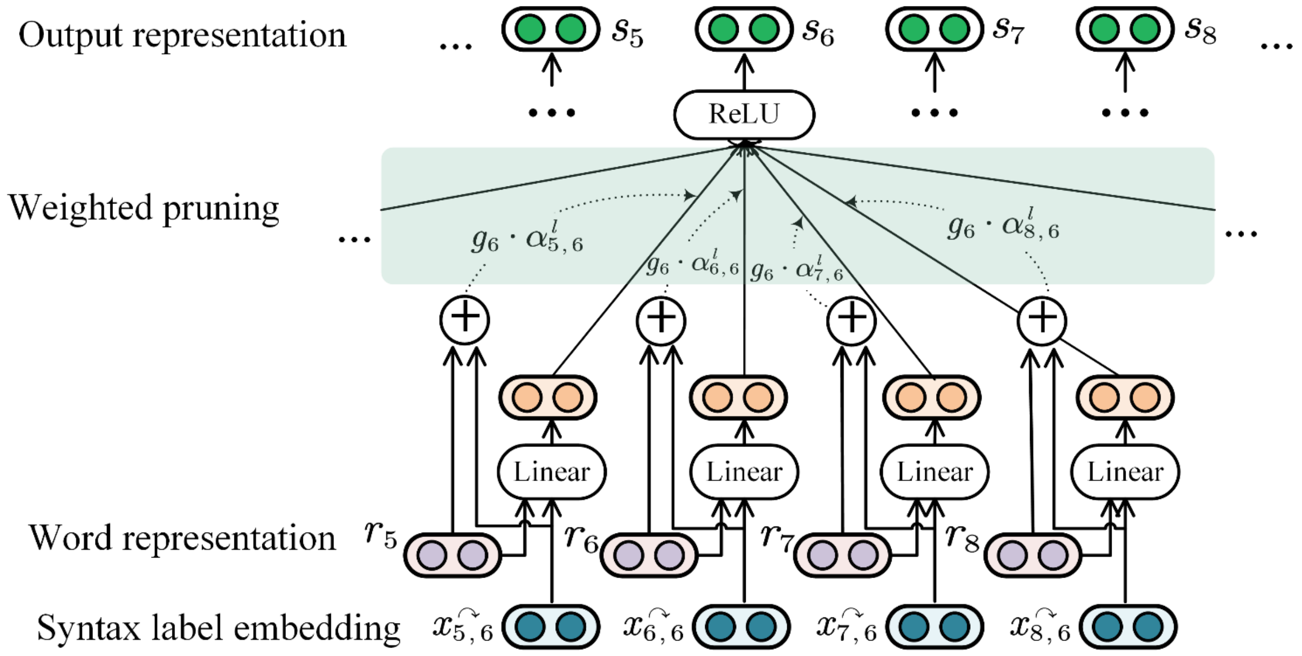
Syntax labels helps produce explainable results



[Structure-aware NLP]

How?

- Case-II: Dynamic structure pruning
  - Dynamic structure pruning mechanism
    - *Based on* Label-aware syntax GCN
    - *Performing structure pruning*

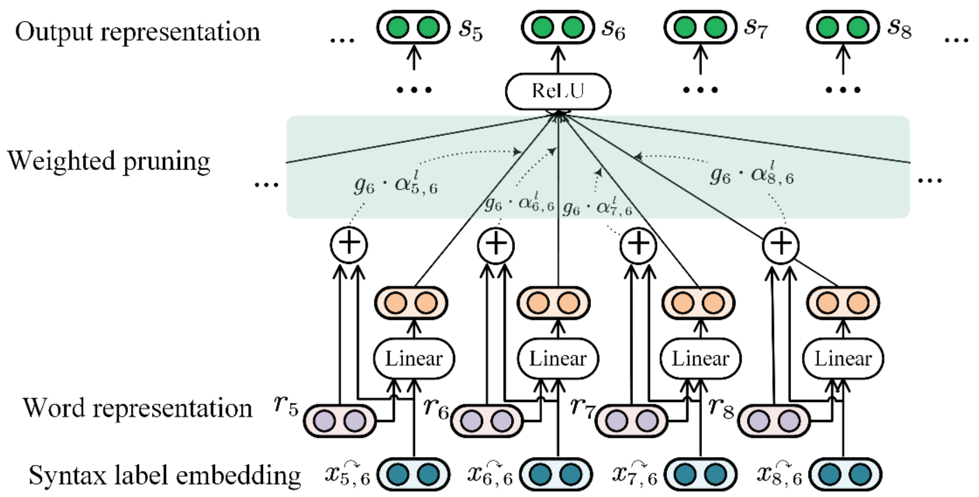


[Structure-aware NLP]

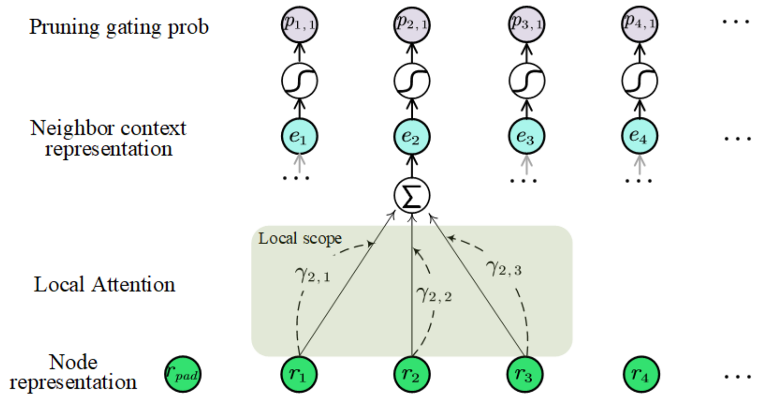
How?

- Case-II: Dynamic structure pruning

  - Dynamic structure pruning mechanism



Step1: Generating pruning gating prob based on local attention module



Step2: Performing discrete transformation on the gating prob via Gumbel-Softmax

$$g_i = \frac{\exp\left(\log p_{i,1} + \epsilon\right)/\tau}{\sum_{t=0,1} \exp\left(\log p_{i,t} + \epsilon\right)/\tau},$$

$$\text{Gumbel}(0,1) = -\log(-\log(\text{Uniform}(0,1))),$$

Step3: Obtaining the weighted pruning values

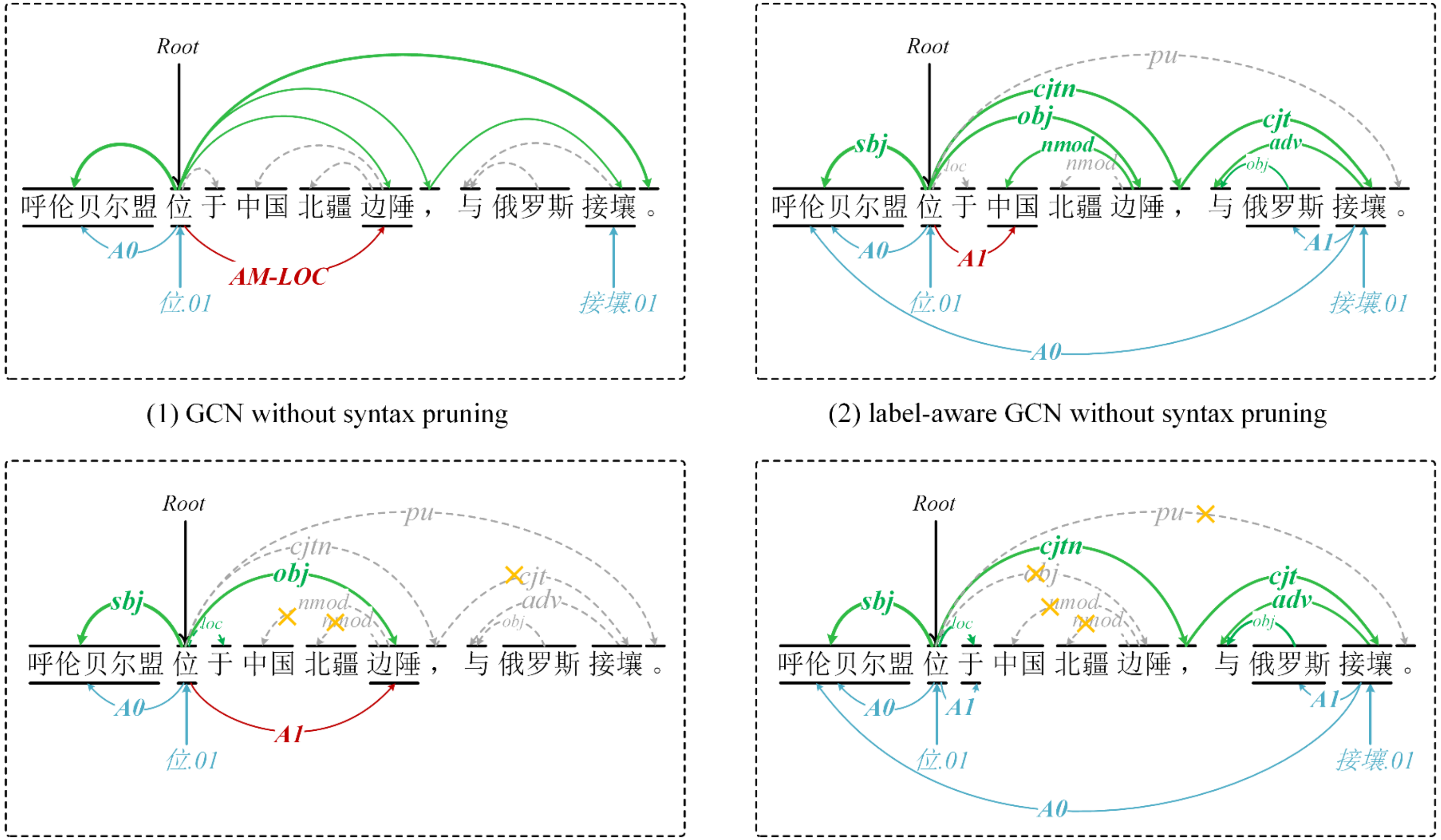
$$\boldsymbol{G} = \{g_i\}_{i=1}^{14}$$

0	0	0	0	1	1	1	1	0	0	0	0	1	1

How?

- Case-II: Dynamic structure pruning

Dynamic structure pruning helps denoising



(1) GCN without syntax pruning

(2) label-aware GCN without syntax pruning

(3) label-aware GCN with static syntax pruning

(4) label-aware GCN with dynamic syntax pruning

How?

- Case-III: Deep Syntax-Semantics Communication
  - Gaps between the syntax and semantics

(a) The same syntactic structure but different semantics.

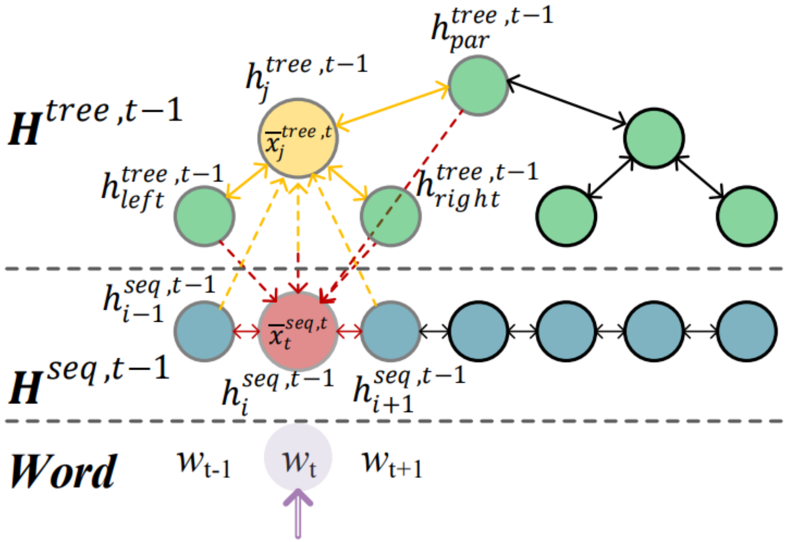
(b) The similar semantics but different syntactic structures.

Insufficient interactions between syntax and semantics!

[1] Hao Fei, Yafeng Ren, Donghong Ji. Improving Text Understanding via Deep Syntax-Semantics Communication. EMNLP (Findings) 2020: 84-93

How?

- Case-III: Deep Syntax-Semantics Communication
  - Local communication
    - between ***<u>syntactic tree encoder</u>*** and ***<u>sequential semantic encoder</u>***.

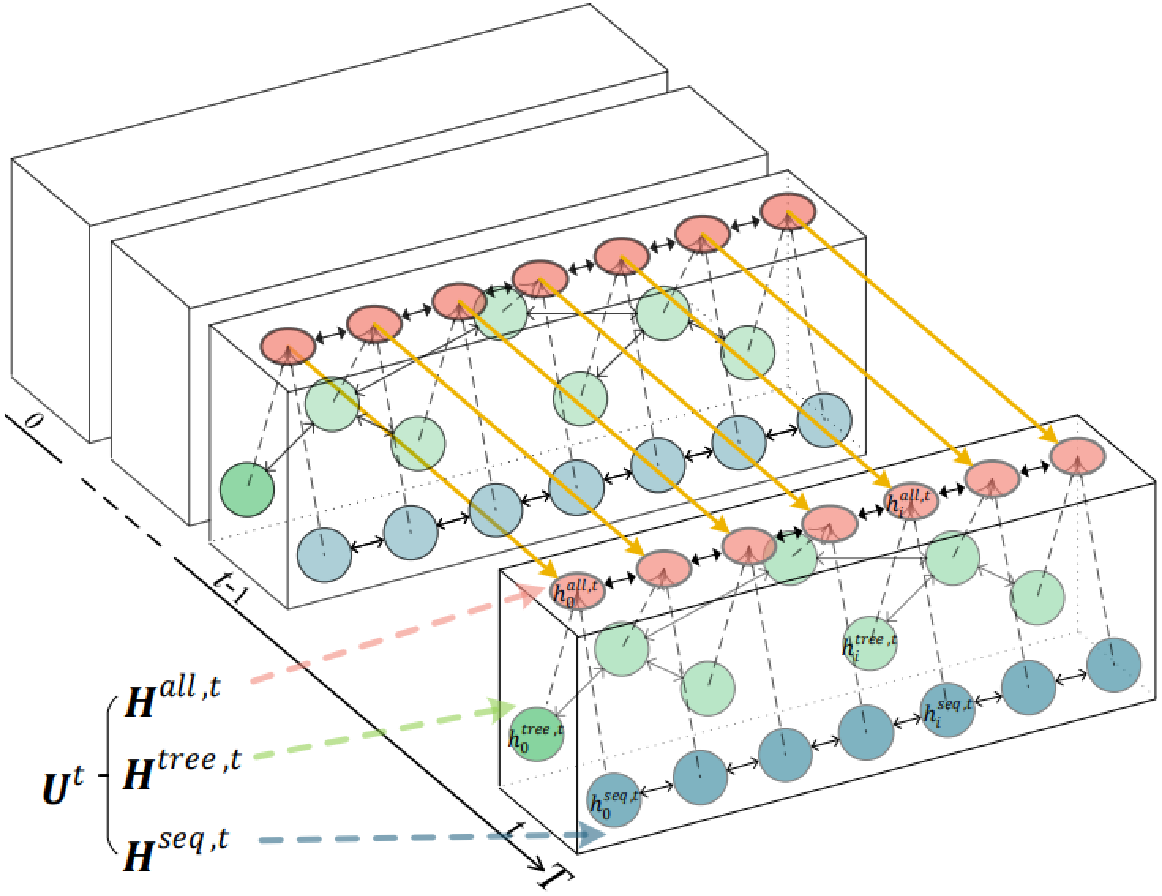


Syntactic tree encoder: GCN/TreeLSTM

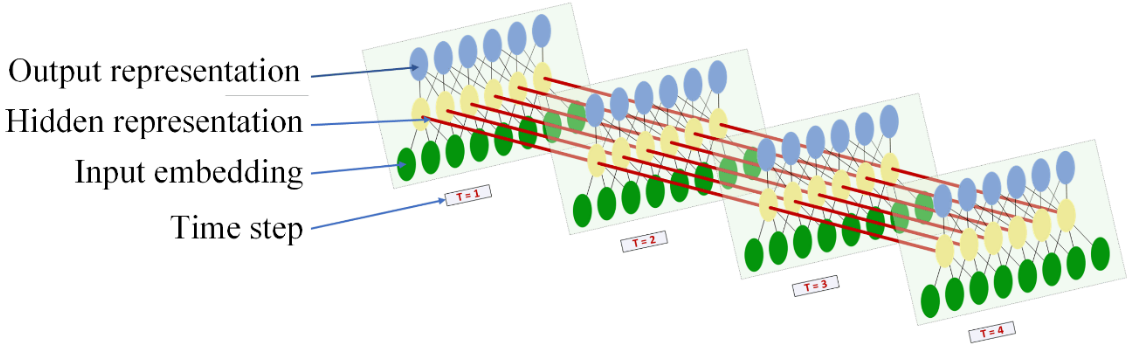
Sequential semantic encoder: LSTM

How?

- Case-III: Deep Syntax-Semantics Communication
  - Global interaction
    - *at the sentence level over **recurrent steps**.*



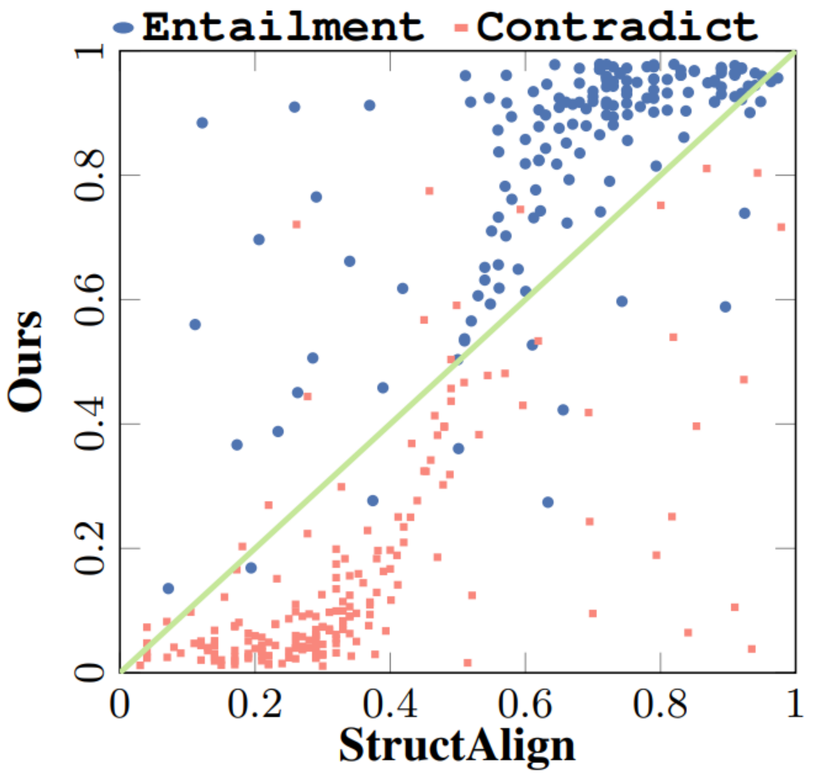
- *Simulating the propagation in RNN*

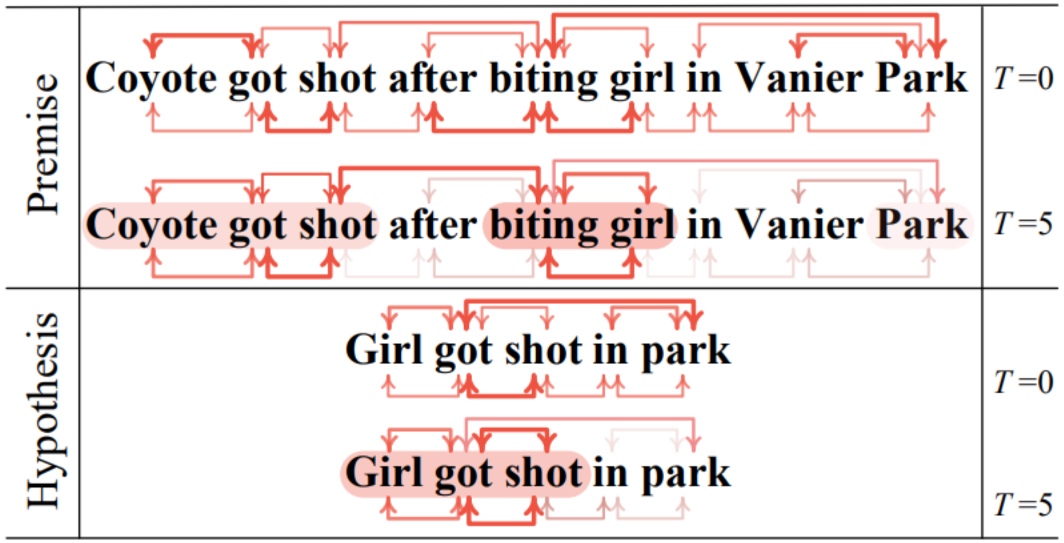


How?

➢ Case-III: Deep Syntax-Semantics Communication

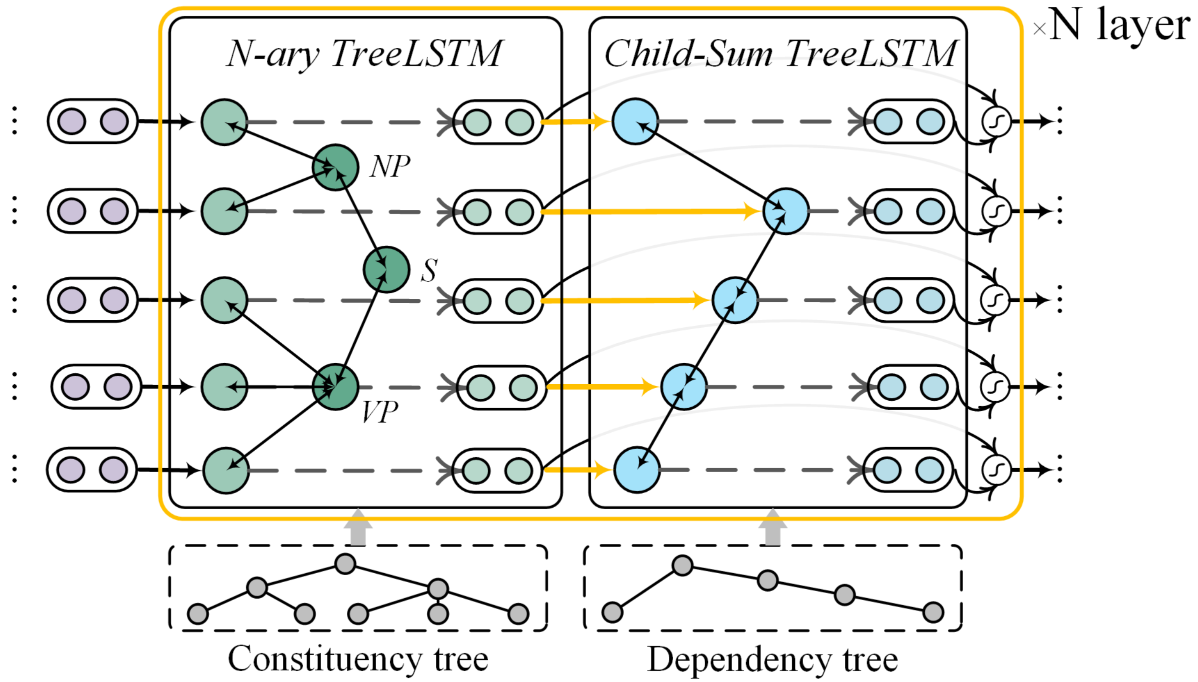
More precise on capturing text semantics





How?

- Case-IV: Heterogeneous syntax integration
  - Explicit heterogeneous syntax fusion
    - *TreeLSTM-based fuser*



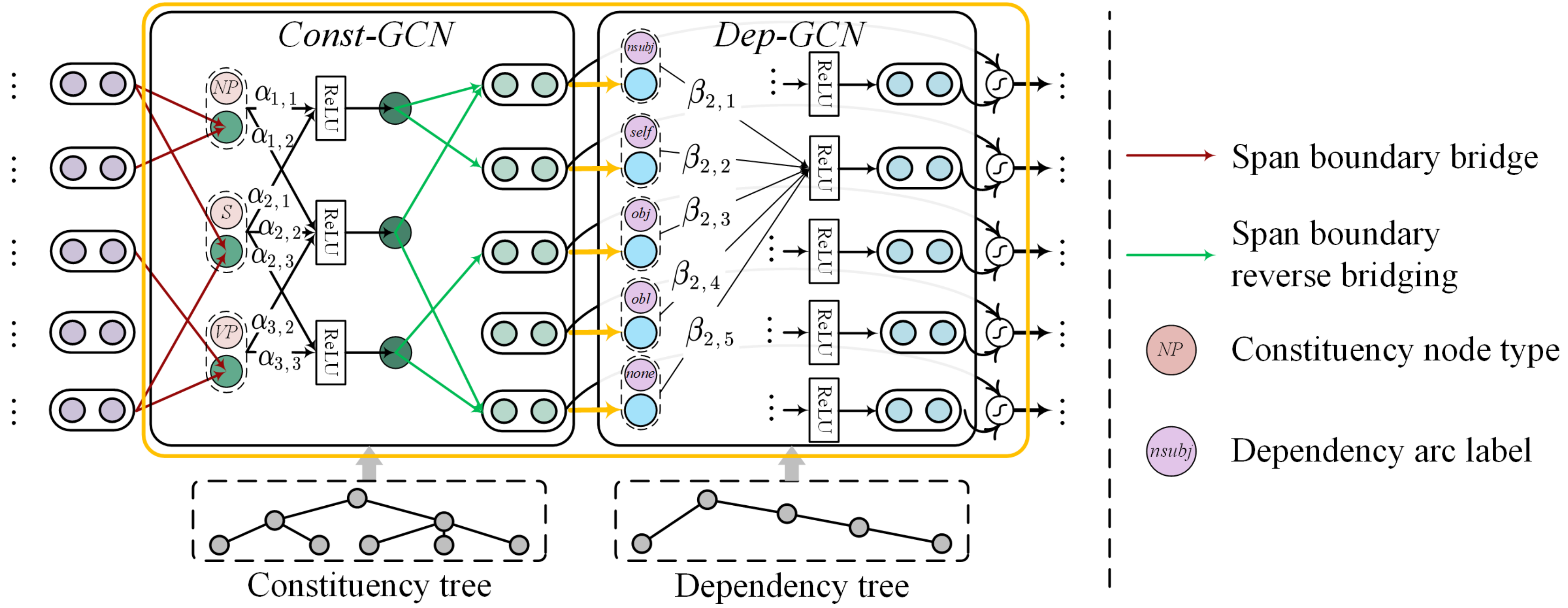
[1] *Hao Fei, Shengqiong Wu, Yafeng Ren, Fei Li, Donghong Ji. Better Combine Them Together! Integrating Syntactic Constituency and Dependency Representations for Semantic Role Labeling. ACL/IJCNLP (Findings) 2021: 549-559*

How?

➢ Case-IV: Heterogeneous syntax integration

- Explicit heterogeneous syntax fusion

➢ *SyntaxGCN-based fuser*

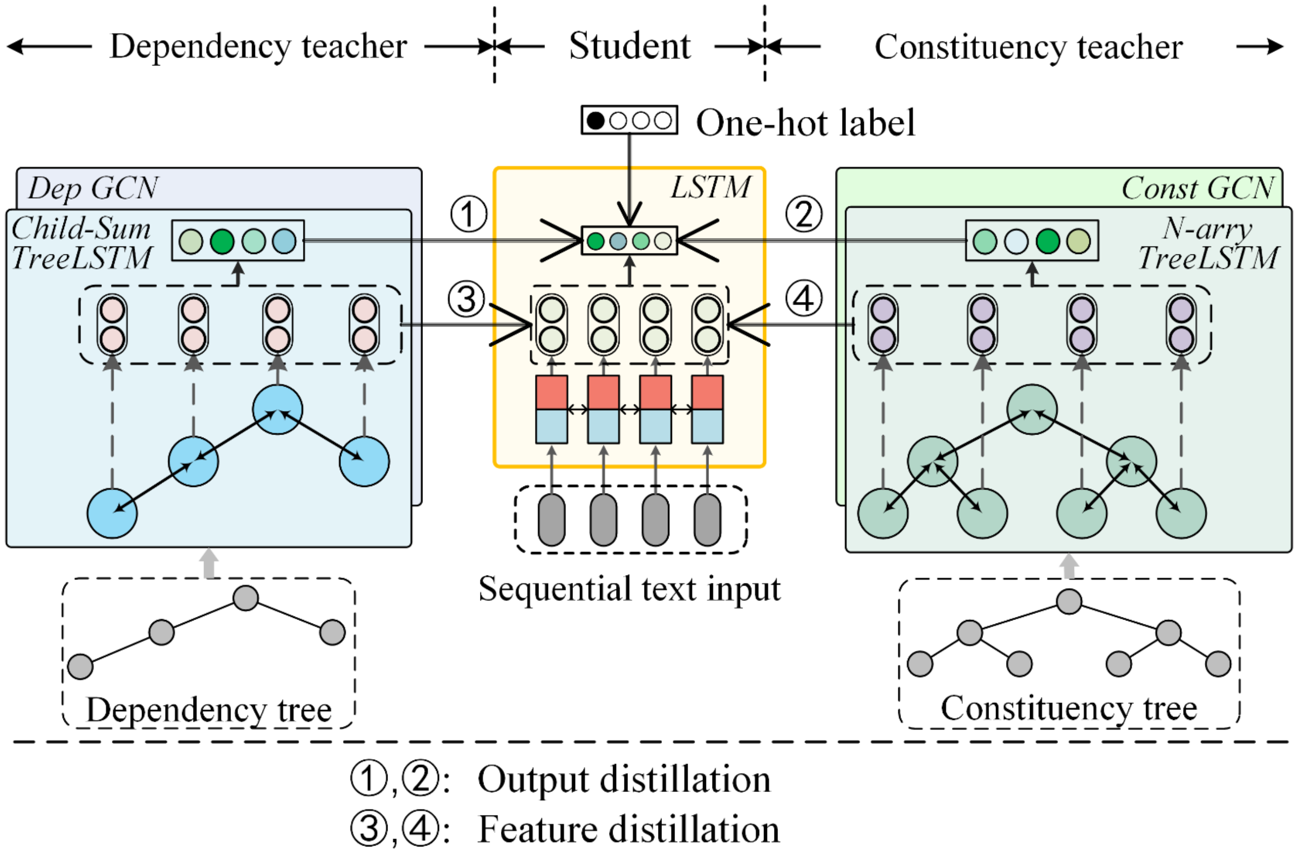


How?

- Case-IV: Heterogeneous syntax integration
  - Implicit heterogeneous syntax fusion

- *Syntax teacher models*
  - Dependency tree:
    - Syntax GCN
    - Child-sum TreeLSTM
  - Constituency tree:
    - Syntax GCN
    - N-arry TreeLSTM

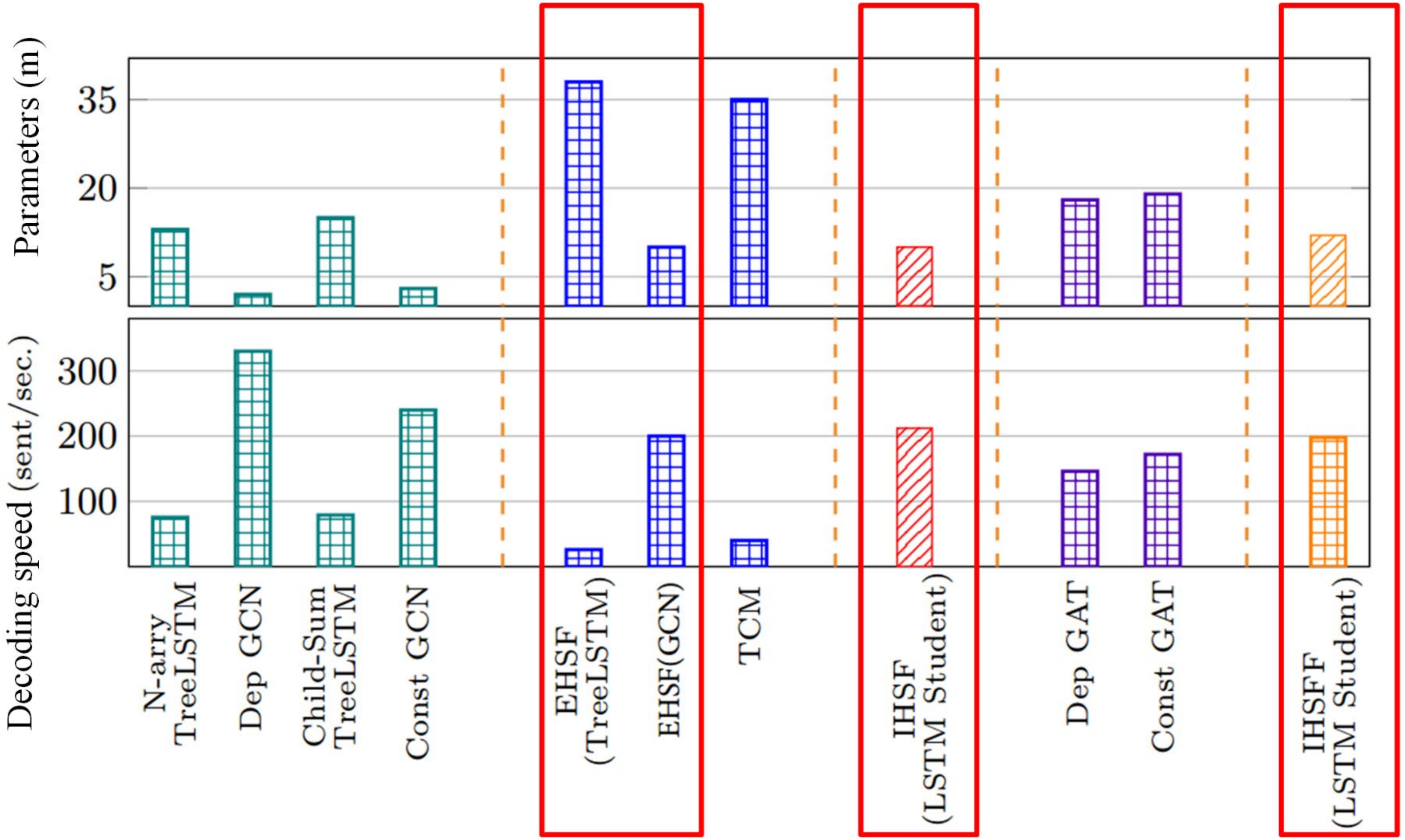
- *Student model*
  - Simply sequential LSTM: linear complexity, faster, lower parameters.



[1] *Hao Fei, Yafeng Ren, Donghong Ji. Mimic and Conquer: Heterogeneous Tree Structure Distillation for Syntactic NLP. EMNLP (Findings) 2020: 183-193*

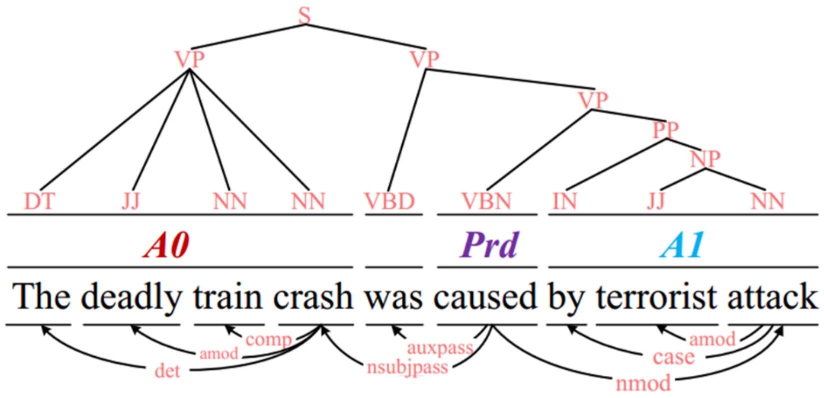
How?

- Case-IV: Heterogeneous syntax integration
  - Efficiency study

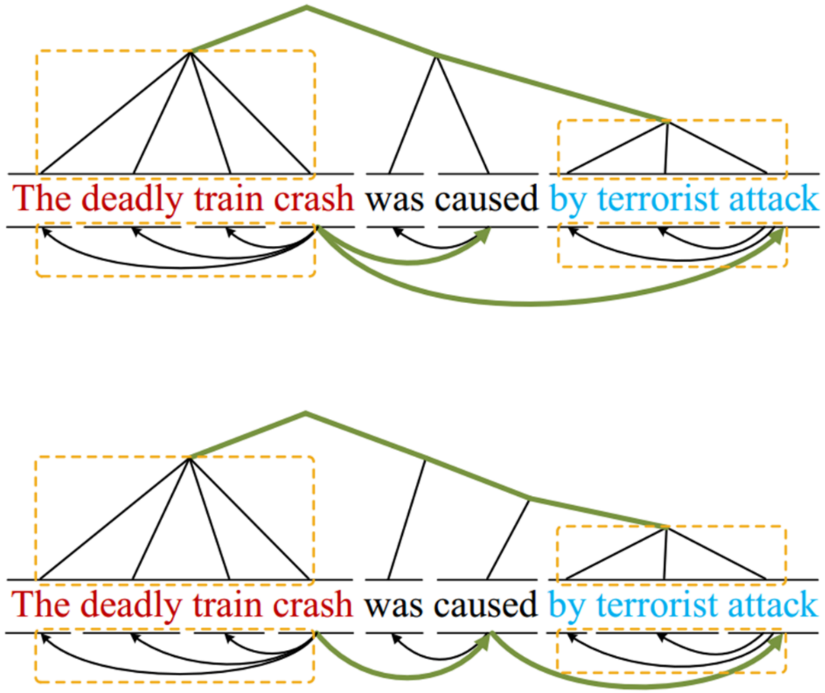


How?

- Case-IV: Heterogeneous syntax integration
  - Visualization



(a) Manually annotated syntax trees

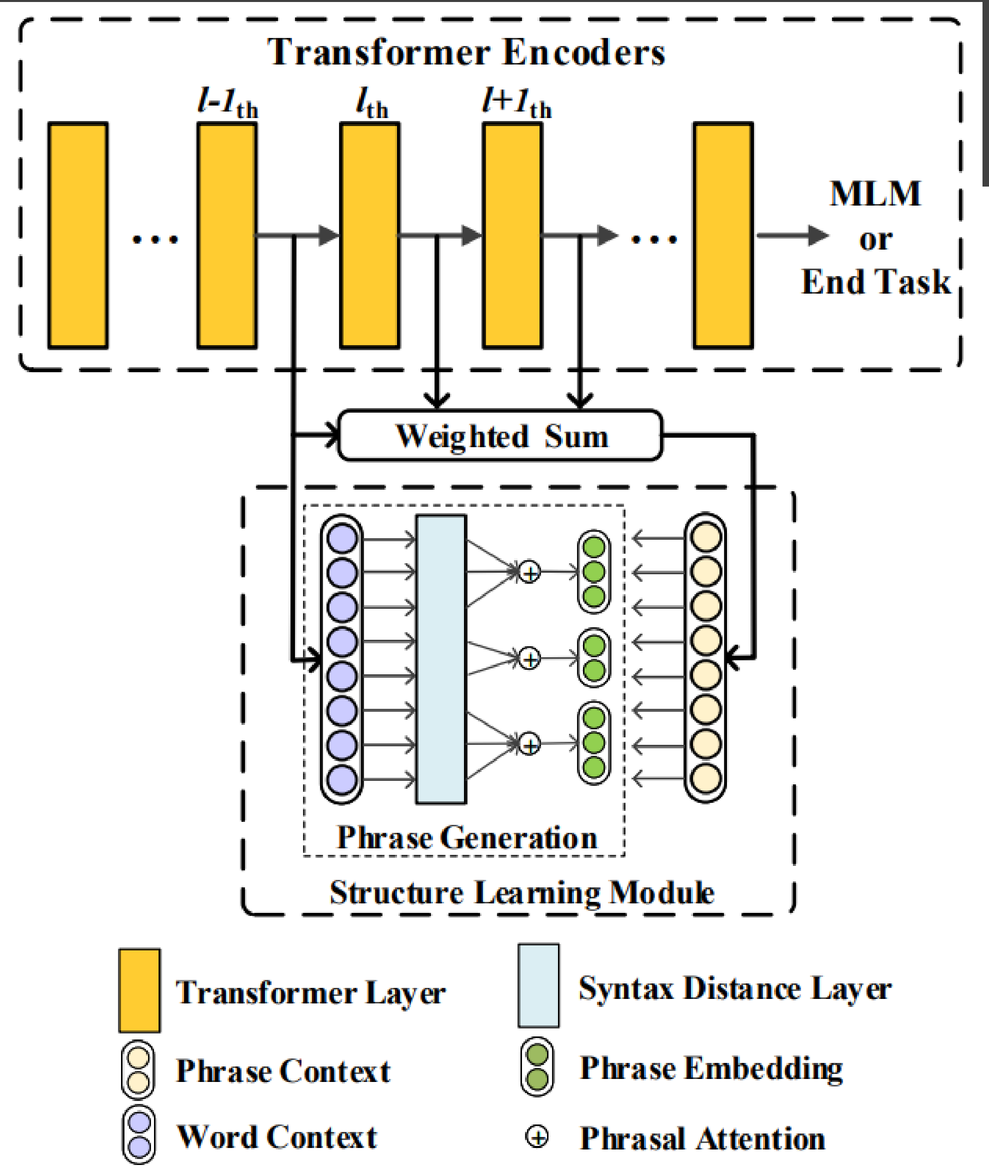


(b) Automatically learned syntax trees by heterogeneous structure fusor

[Structure-aware NLP]

How?

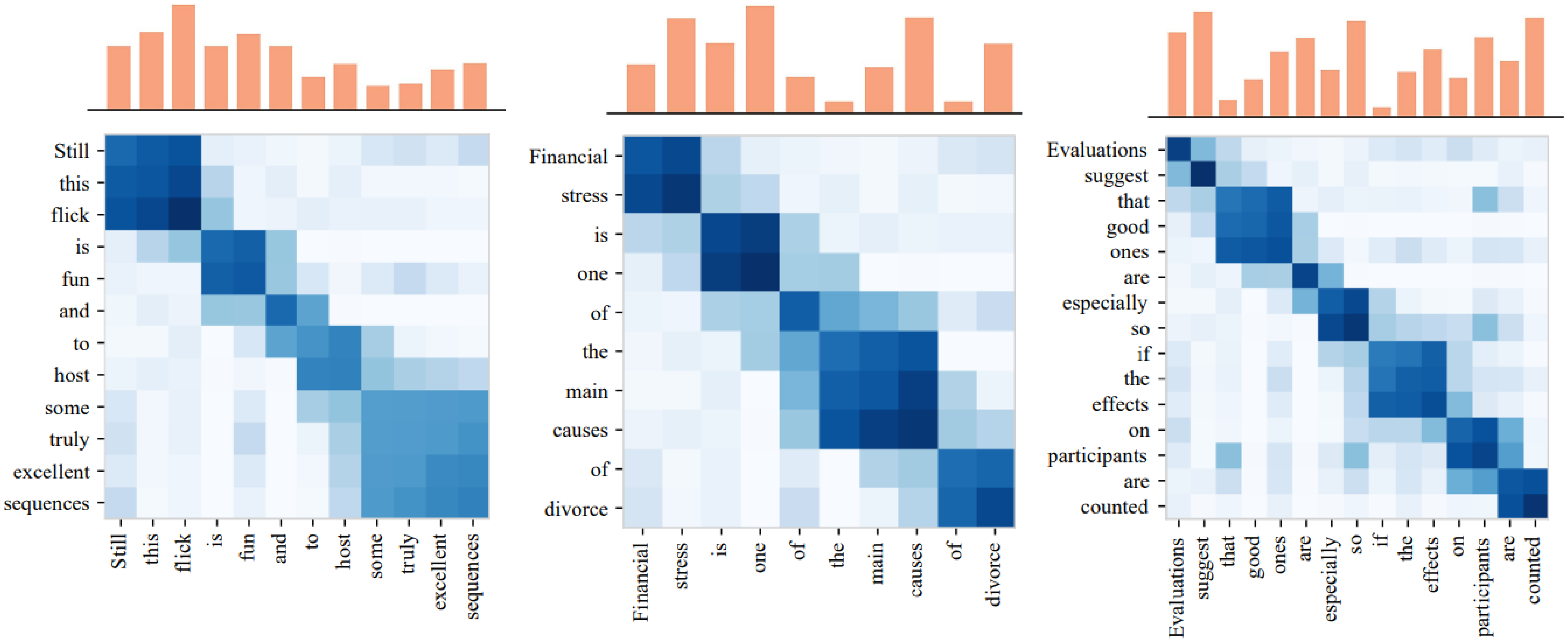
- ➢ Case-V: Syntax integration in LMs
  - Structure-aware Transformer LM
    - ➢ *Integrating heterogeneous syntax*
    - ➢ *Middle-layer syntax-enhanced training strategy*
    - ➢ *Structure-aware fine-tuning with end-task*



[1] Hao Fei, Yafeng Ren, Donghong Ji. Retrofitting Structure-aware Transformer Language Model for End Tasks. EMNLP. 2020: 2151-2161

[Structure-aware NLP]
How?

- Case-V: Syntax integration in LMs
  - Visualization of attention maps of different tasks



(a) SST (b) Rel (c) SRL

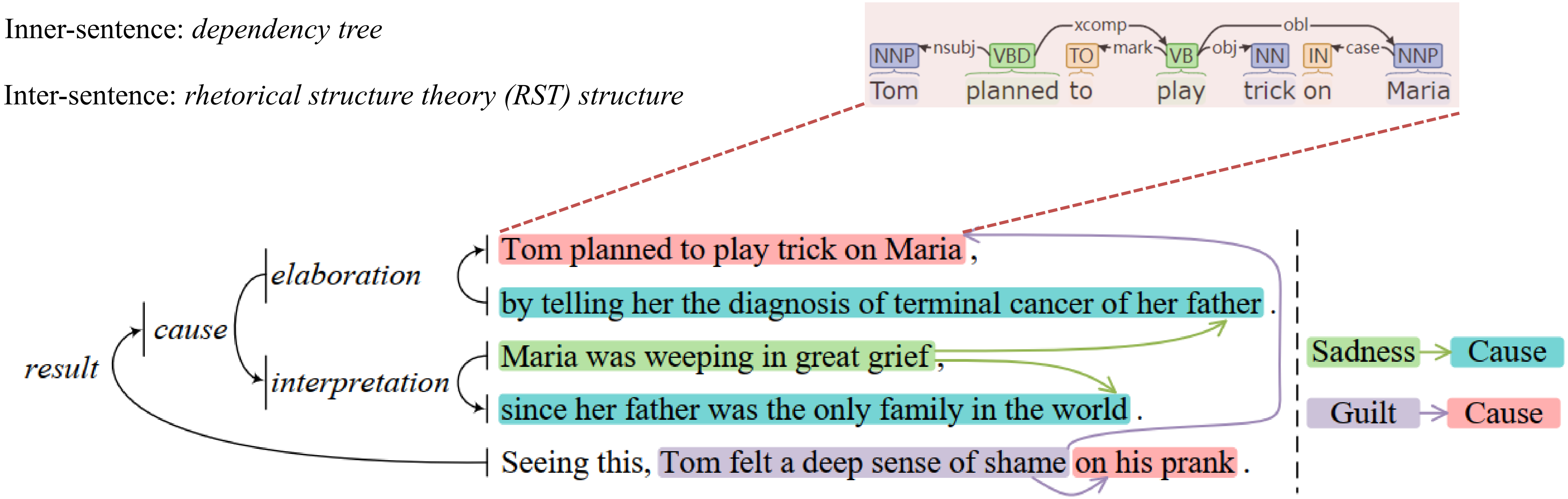
[Structure-aware NLP]

How?

➢ Case-VI: Modeling discourse structure for document-level NLP

- Complete document-level dependency structure

  - Inner-sentence: *dependency tree*
  - Inter-sentence: *rhetorical structure theory (RST) structure*



[1] Hao Fei, etc. Transition-based End-to-end Emotion-Cause Pair Extraction with Implicit Discourse Knowledge. IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, 2022.

How?

➢ Case-VII: Modeling discourse structure for dialogue-level NLP

- Dialogue-level NLP challenges
  - Multi-party dialogue threads are scattered and entangled; There's a logical answering structure between utterances from different speakers (parties).
  - The speaker coreference ambiguity problem.
  - Dependency syntax is an effective feature for sentence; It is intractable to directly apply the syntax structure information for dialogue.

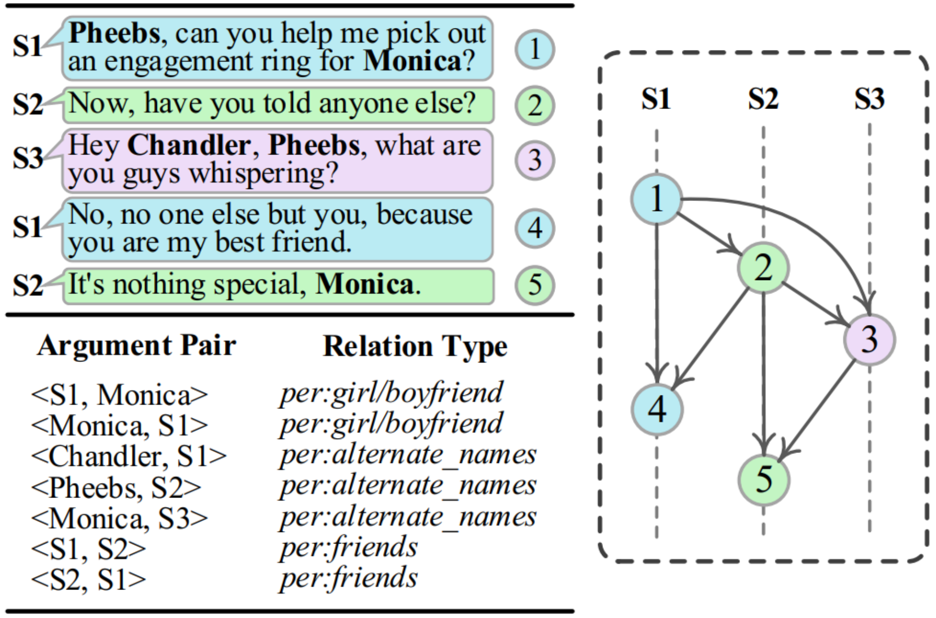


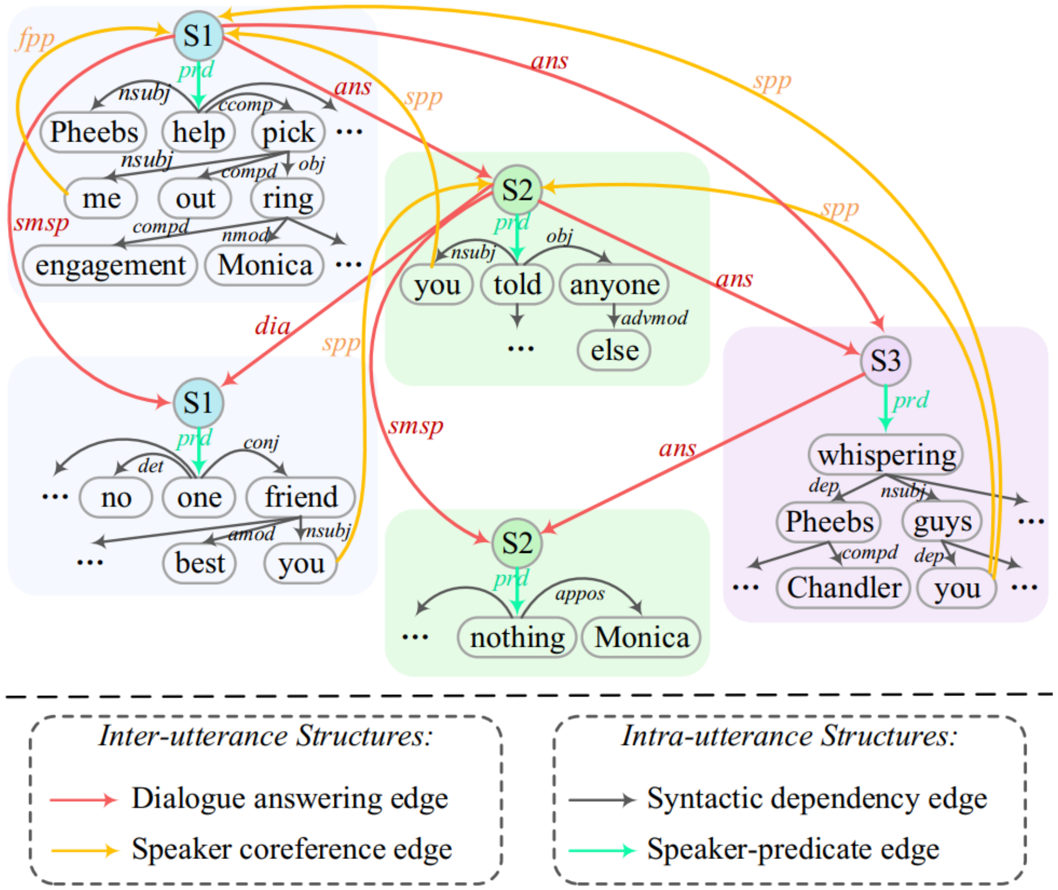
Figure 1: **Left**: dialogue-level relation extraction. **Right**: dialogue-answering structure.

[1] *Hao Fei, Jingye Li, Shengqiong Wu, Chenliang Li, Donghong Ji, Fei Li. Global Inference with Explicit Syntactic and Discourse Structures for Dialogue-Level Relation Extraction. IJCAI. 2022*

[Structure-aware NLP]
How?

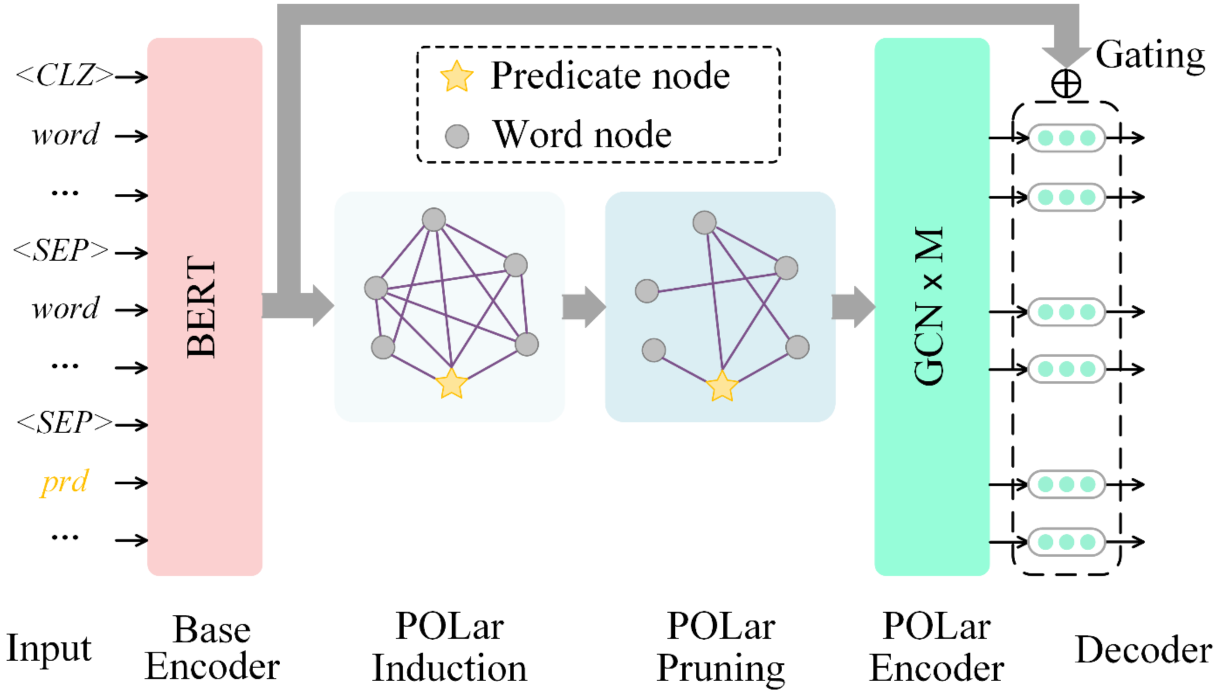
➢ Case-VII: Modeling discourse structure for dialogue-level NLP

- Dialogue-level Mixed Dependency Graph
    - Dialogue answering edge
    - Speaker coreference edge
    - Syntactic dependency edge
    - Speaker-predicate edge



How?

- Case-VIII: Modeling generic latent structure
  - Predicate-oriented latent graph
    - Automatically inducing latent graph structure for end task
    - Latent graph: **HardKuma** distribution theory



[1] *Hao Fei, Shengqiong Wu, Meishan Zhang , Yafeng Ren, Donghong Ji. Conversational Semantic Role Labeling with Predicate-Oriented Latent Graph. IJCAI. 2022*

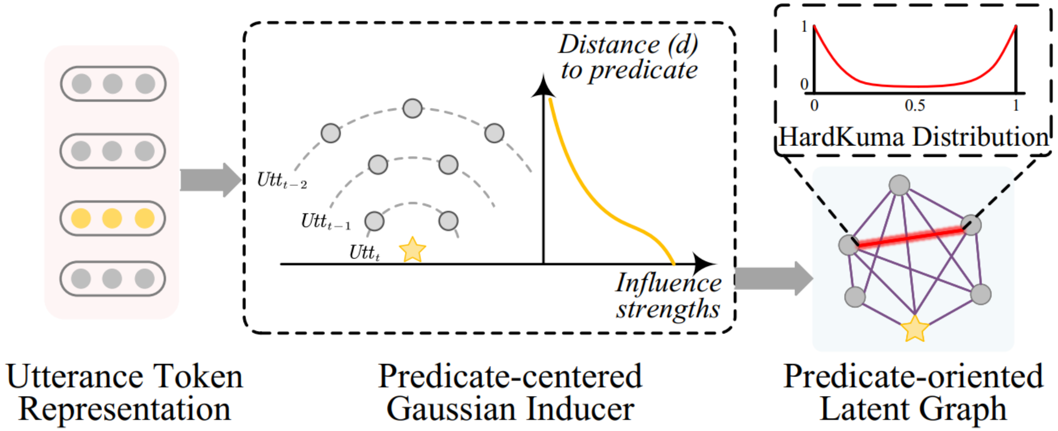
[Structure-aware NLP]

How?

➢ Case-VIII: Modeling generic latent structure

- Predicate-oriented latent graph

  - Predicate-centered Gaussian Inducer:

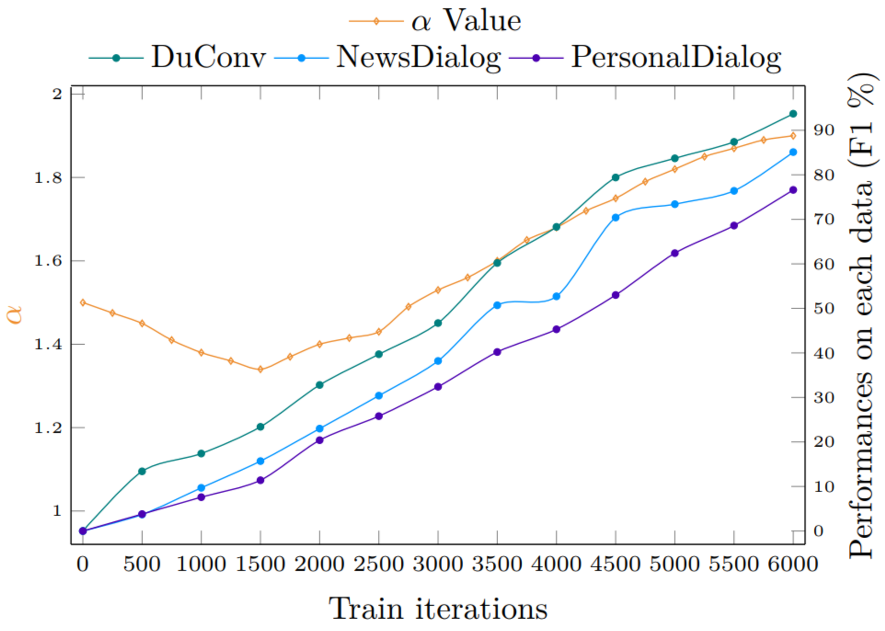


  - Dynamic Structural Pruning:

$$E = \alpha\text{-Entrmax}(E)\,,$$

α is a dynamic parameter controlling the sparsity.

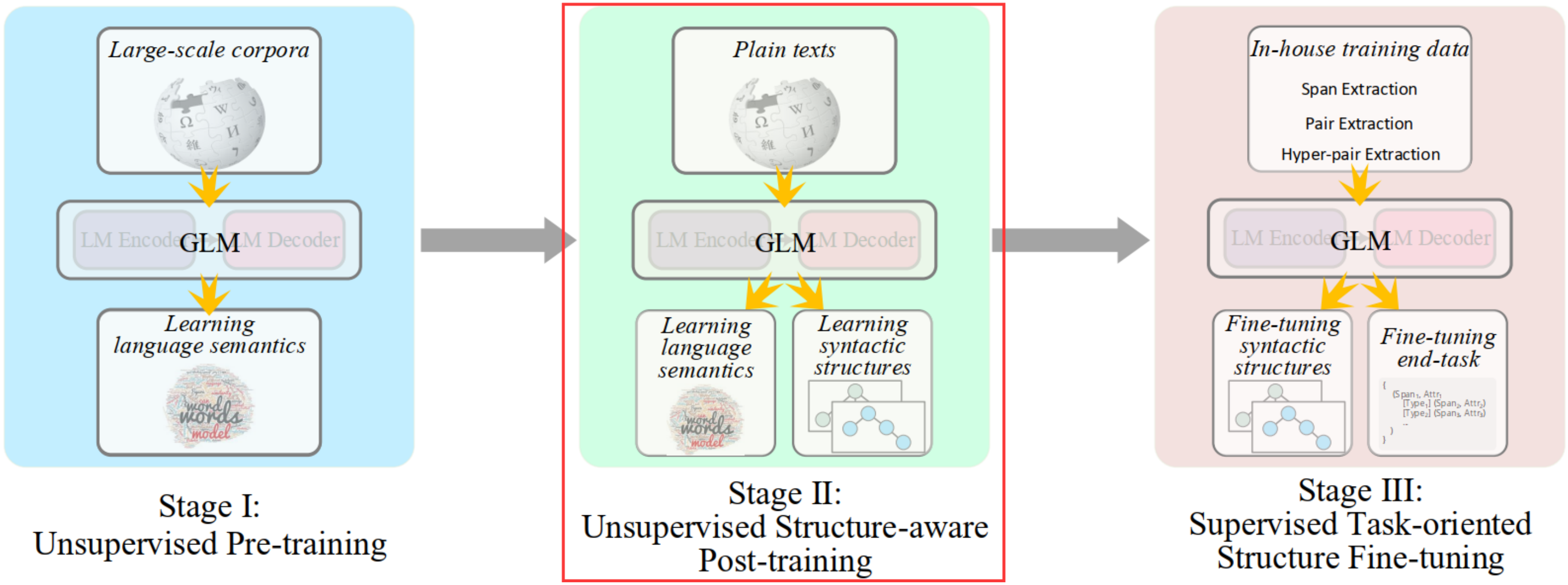
- *When α=2 the Entrmax becomes a* ***Sparsemax*** *mapping,*
- *while α=1 it degenerates into a* ***Softmax*** *mapping*



[Structure-aware NLP]

How?

- Case-VIII: Modeling generic latent structure
  - Latent Adaptive Structure-aware Generative Language Model (GLM) for UIE
    - three-stage training process: structure-aware post-training

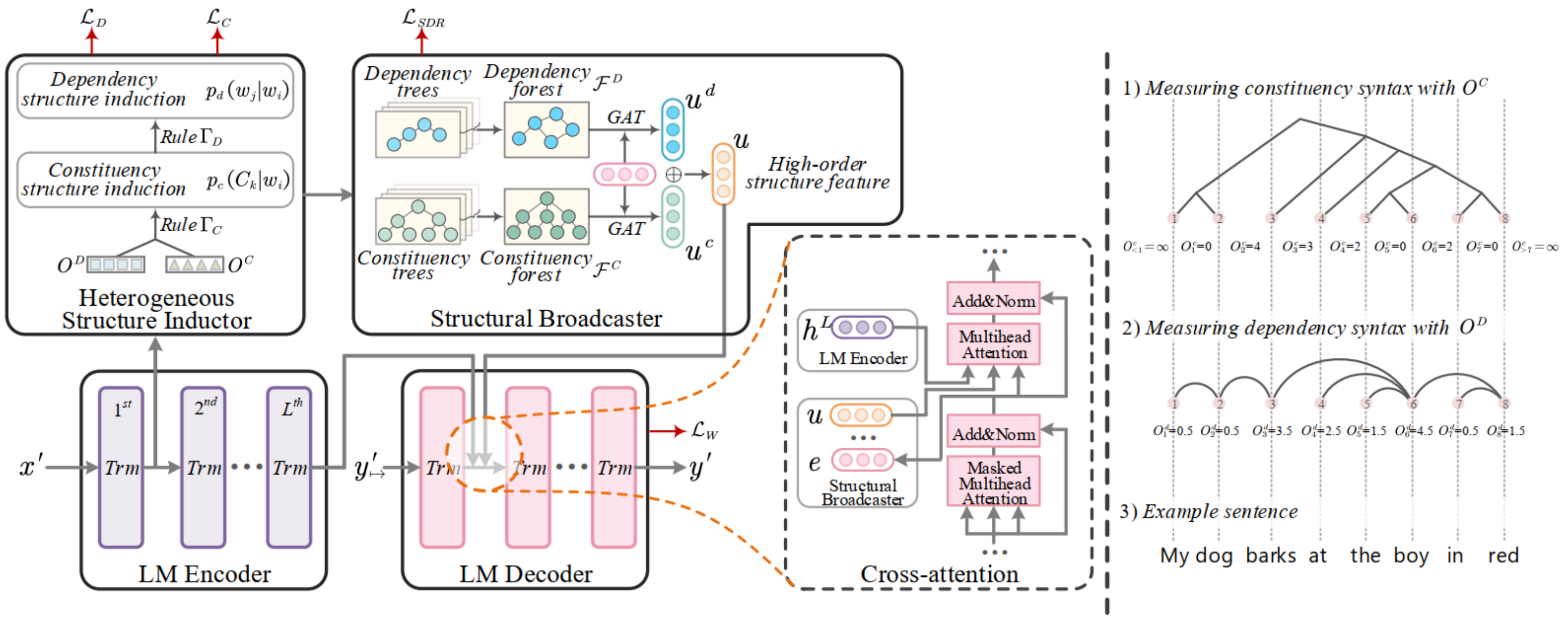


[1] Hao Fei, etc. Unifying Information Extraction with Latent Adaptive Structure-aware Generative Language Model. NIPS2022. submitted.

[Structure-aware NLP]

How?

- Case-VIII: Modeling generic latent structure
  - Latent Adaptive Structure-aware Generative Language Model (GLM) for UIE
    - Unsupervised structure-aware post-training
      - Heterogeneous structure inductor
      - Structural broadcaster



(a) Unsupervised structure-aware post-training

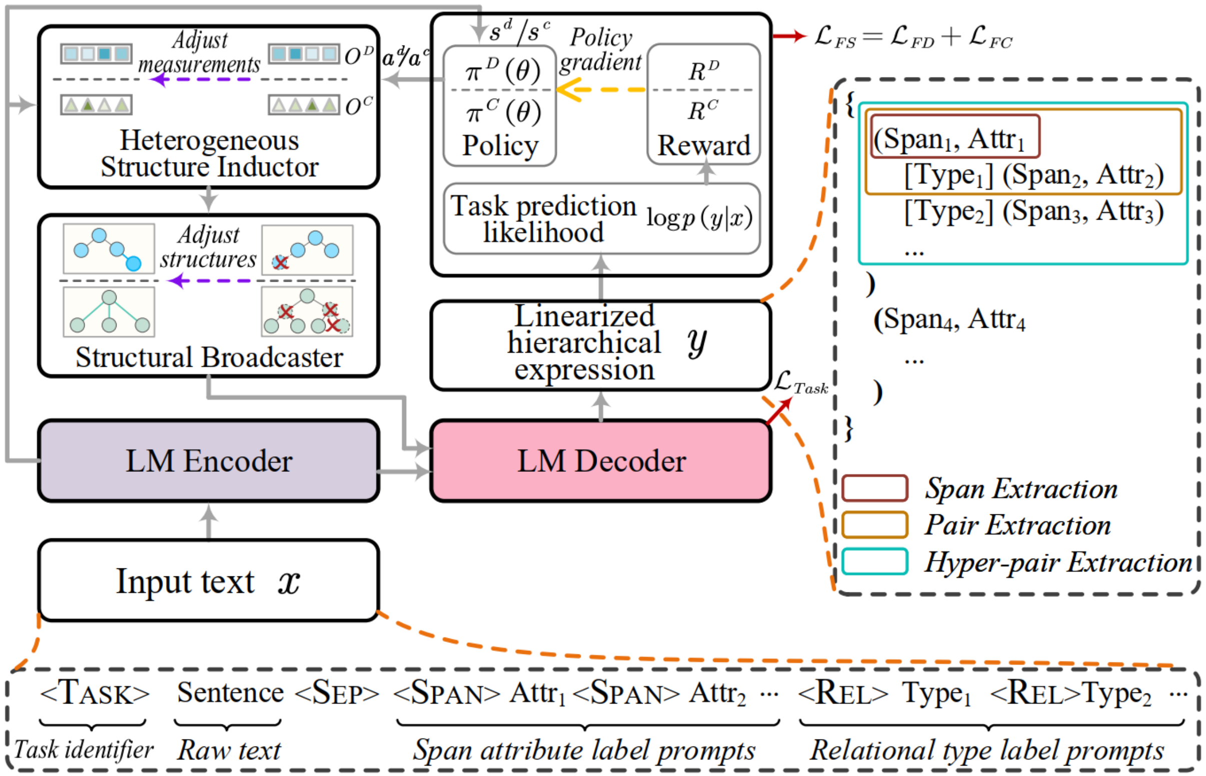
(b) Heterogeneous syntax measurements

How?

➢ Case-VIII: Modeling generic latent structure

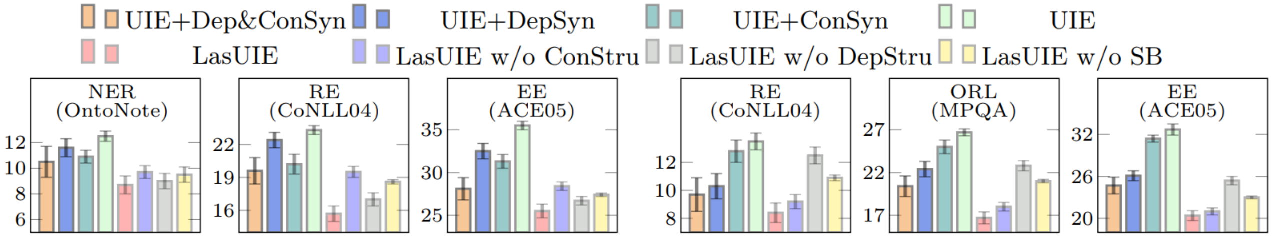
- Latent Adaptive Structure-aware Generative Language Model (GLM) for UIE
  - Task-oriented Structure Fine-tuning

    ✓ narrowing the gaps between the ***induced syntactic*** and ***task-specific*** structures.



How?

- Case-VIII: Modeling generic latent structure
  - Latent Adaptive Structure-aware Generative Language Model (GLM) for UIE
    - Less error on two crux of IE:
      - *long-range dependence issue*
      - *boundary identifying*
    - Structural fine-tuning
      - *Dynamically adjusting the learned structure information in accordance with the end-tasks' need.*



(a) Error rate (%) on boundary recognition (b) Error rate (%) on relation detection

Figure 4: Error rates on boundary recognition and relation detection, respectively.

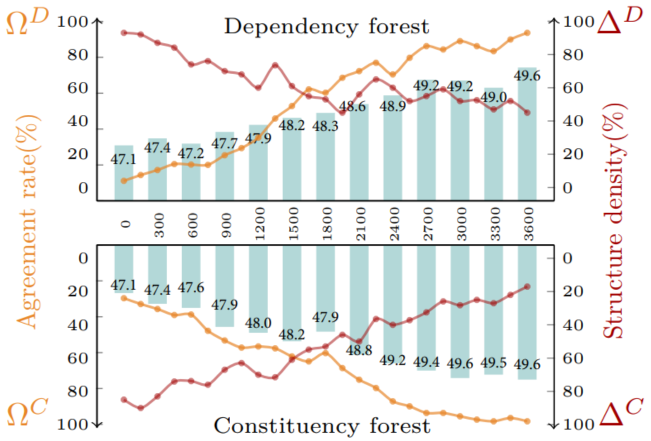


Figure 5: Trajectories of the changing structure agreement rates and densities during task-oriented structure fine-tuning, based on event extraction (ACE05). X-axis is the iteration steps for fine-tuning. Bars means the task performances (F1).

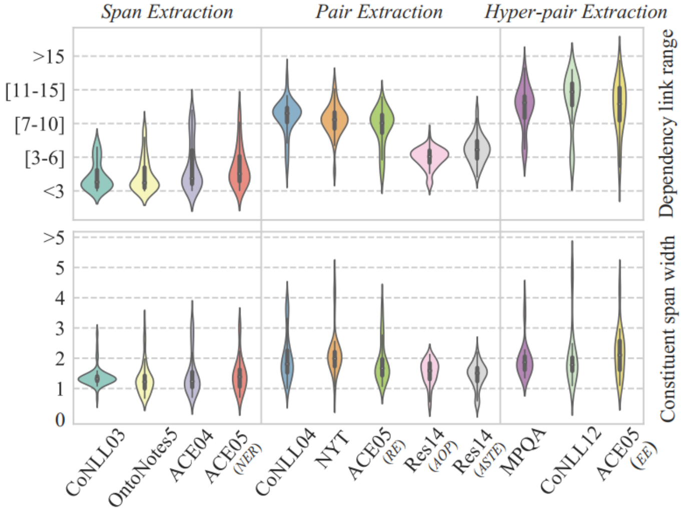


Figure 6: The distributions of the range of word-word dependency link (words) in forest $\mathcal{F}^D$ and the constituency phrasal span width (words) in forest $\mathcal{F}^C$ on each data.

OUTLINE

Structure-aware NLP

- WHAT is syntactic structure?
- WHY integrating structures for NLP?
- HOW to integrate?
- **WHAT to do next?**

What next?

- Summarizing

  - Syntactic structure information offers additional features for NLP semantic understanding from low-level linguistic bias.

  - Effectively modeling the syntactic structure knowledge helps further enhance the utility of structure integration for a wide range of NLP applications. [The focus of this talk]

- *Next, extending the idea of enhancing structural awareness to the applications in other domains besides NLP, i.e., Multimodality modeling.*
  - *Structure-aware Text-Image*
  - *Structure-aware Text-Video*

[Structure-aware multimodal]

What next?

➢ Structure matching in dual learning

- Dual Learning
  - Many NLP/CV/Multimodal tasks appear in dual form.
    - Neural machine translation: [Lan-A -> Lan-B] VS. [Lan-B -> Lan-A]
    - Paraphrase generation: [Target -> Source] VS. [Source -> Target]
    - [Text classification] VS. [Conditioned text generation]
  - Dual learning: modeling the **duality** between the *primal* and *dual* tasks, by minimizing the gap between joint distributions of the two tasks.

- Existing Problem

  Current dual learning scheme fails to explicitly model the ***structural correspondence*** in between.

Duality Scheme	Direction	Representative Application(s)
Text↔Text	⟶ or ⟵	Neural Machine Translation, Paraphrase Generation
Text↔Image	⟶ ⟵	Text-to-Image Synthesis Image Captioning
Text↔Label	⟶ ⟵	Text Classification Conditioned Text Generation
Image↔Label	⟶ ⟵	Image Classification Conditioned Image Generation
Image↔Image	⟶ or ⟵	Image Translation

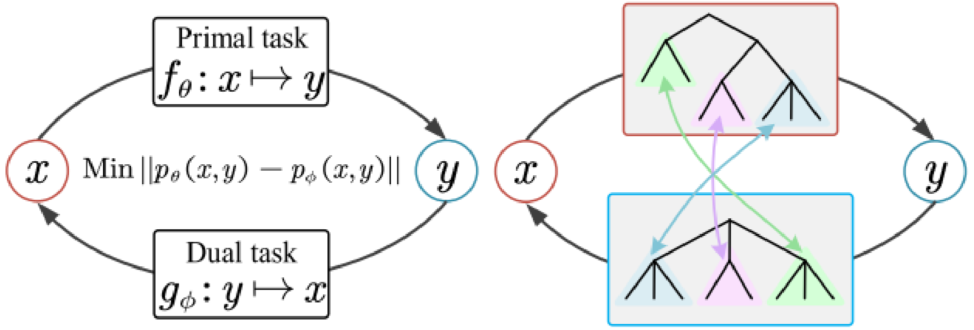


Figure 1. **Left**: dual learning framework. **Right**: dual learning with alignment of structural supervision.

[1] Hao Fei , Shengqiong Wu, Yafeng Ren, Meishan Zhang. Matching Structure for Dual Learning. ICML. 2022

[Structure-aware multimodal]

What next?

- Structure matching in dual learning
  - Dually-Syntactic Structure Matching for **Text-to-Text** Dual Learning

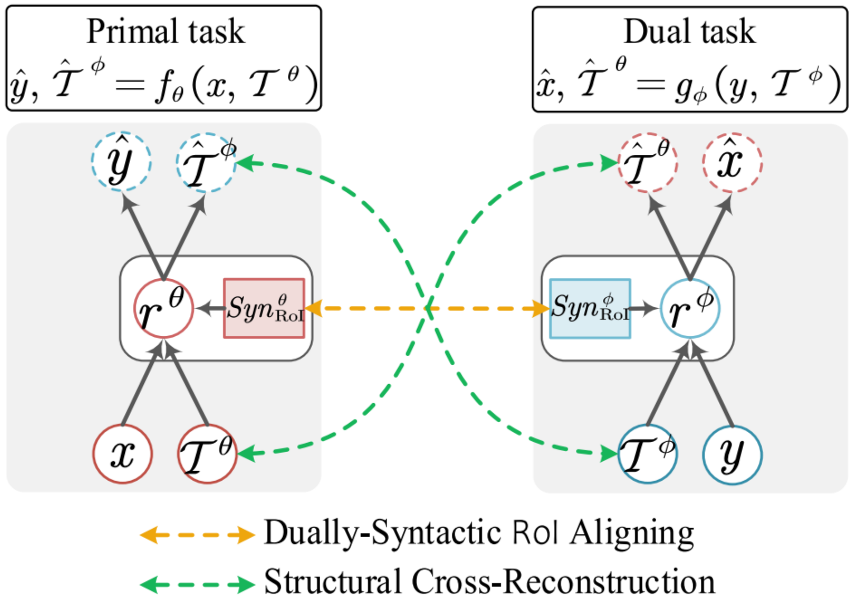
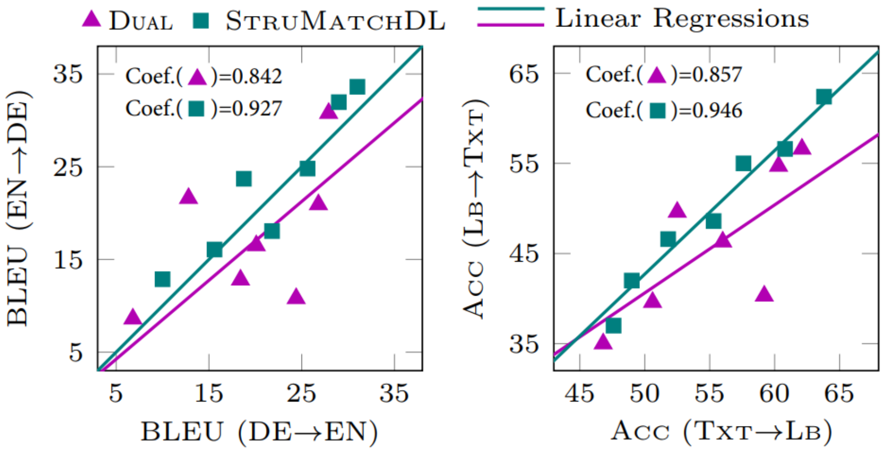


Figure 2. Symmetrically syntactic structure matching for dual learning.

		WMT14 (EN-DE)				WMT14 (EN-FR)			
		EN→DE		EN←DE		EN→FR		EN←FR	
• *Baseline*	B1	28.04	/	30.91	/	39.44	/	35.32	/
	B2	28.22	/	30.72	/	39.68	/	35.90	/
	B3	28.57	/	31.00	/	39.80	/	35.85	/
• *Seq2seq-based*	M1	16.24	/	20.69	/	29.92	/	27.49	/
	M2	17.06	+0.82	21.62	+0.93	31.15	+1.23	28.82	+1.33
	M3	16.81	/	20.81	/	31.99	/	28.35	/
	M4	19.52	+2.71	23.24	+2.43	35.85	+3.86	31.27	+1.92
• *Transformer-based*	M1	25.24	/	28.42	/	37.21	/	32.08	/
	M2	27.07	+1.83	29.84	+1.42	38.73	+1.52	33.95	+1.87
	M3	26.46	/	29.17	/	38.10	/	32.52	/
	M4(RANK)	29.71	+3.25	33.40	+ 4.23	42.28	+4.18	37.09	+4.57
	M4(CL)	30.03	+3.57	33.96	+4.79	42.82	+4.72	37.76	+5.24
	ONLYSYN	27.90	+1.44	30.81	+1.64	39.03	+0.93	34.60	+2.08
	-SALN	28.23	+1.77	31.15	+1.98	39.55	+1.45	35.07	+2.55
	-SYREC	29.56	+3.10	32.68	+3.51	41.17	+3.07	36.34	+3.82

Table 1. Results (BLEU scores) on NMT. Two colors indicate the coupled tasks, respectively. Color depth highlights the significances of the results improvements. '+' means the improvement over the counterpart without using structure knowledge (e.g., M2-M1, M4-M3).



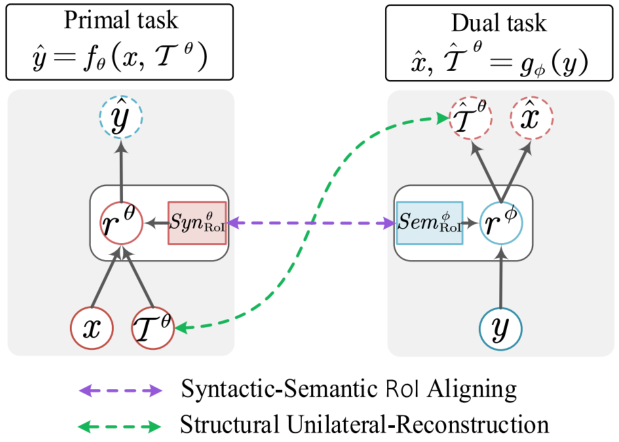
(a) WMT14 (EN-DE)

(b) Yelp2014

[Structure-aware mul

What next?

➢ Syntactic-Semantic Structure Matching for **Text-to-nonText** Dual Learning



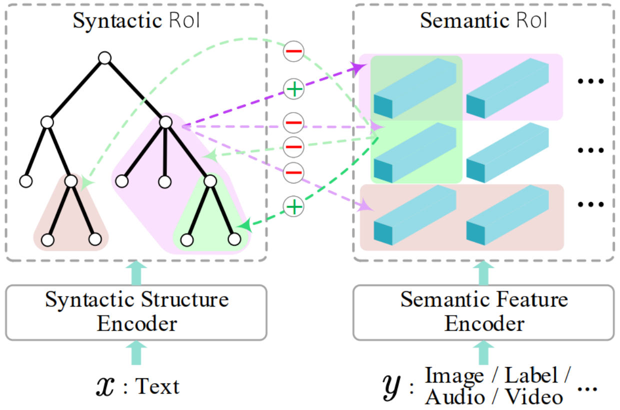


Figure 4. Syntactic-semantic RoI alignment via contrastive representation learning.

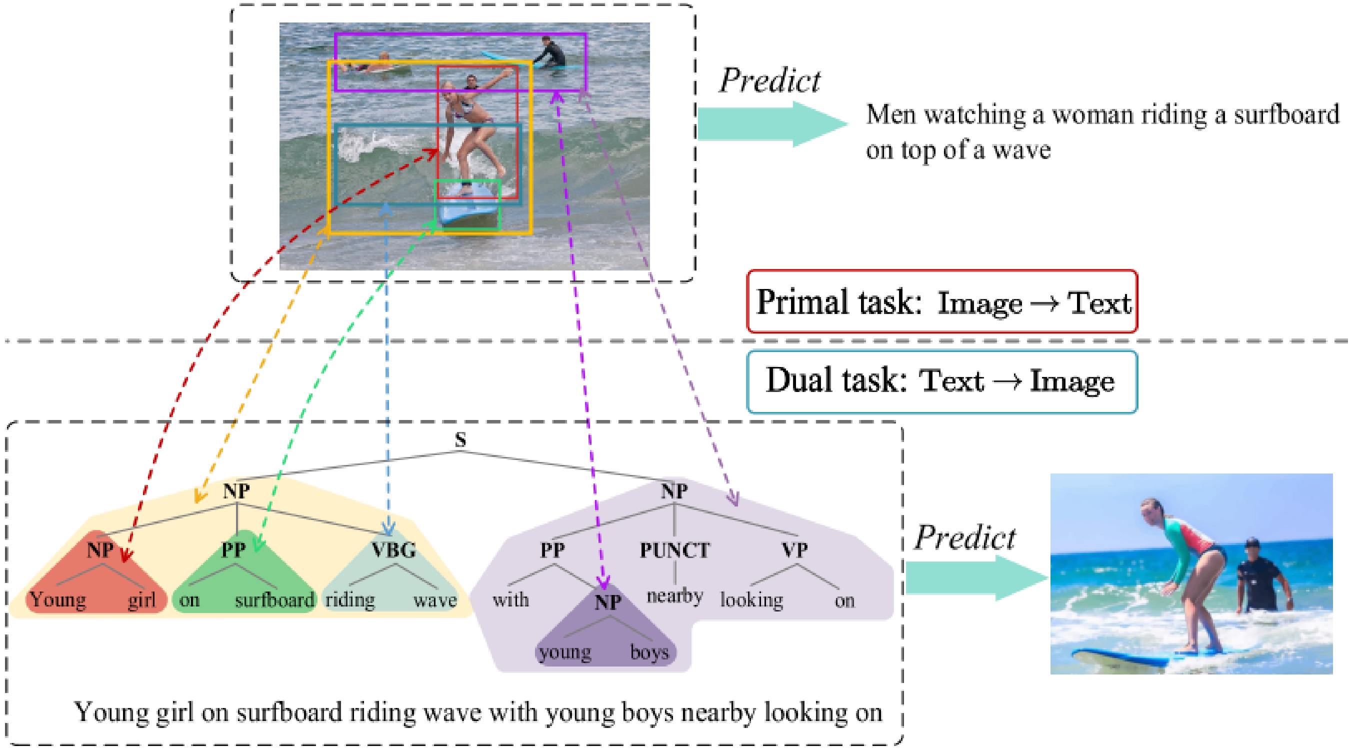
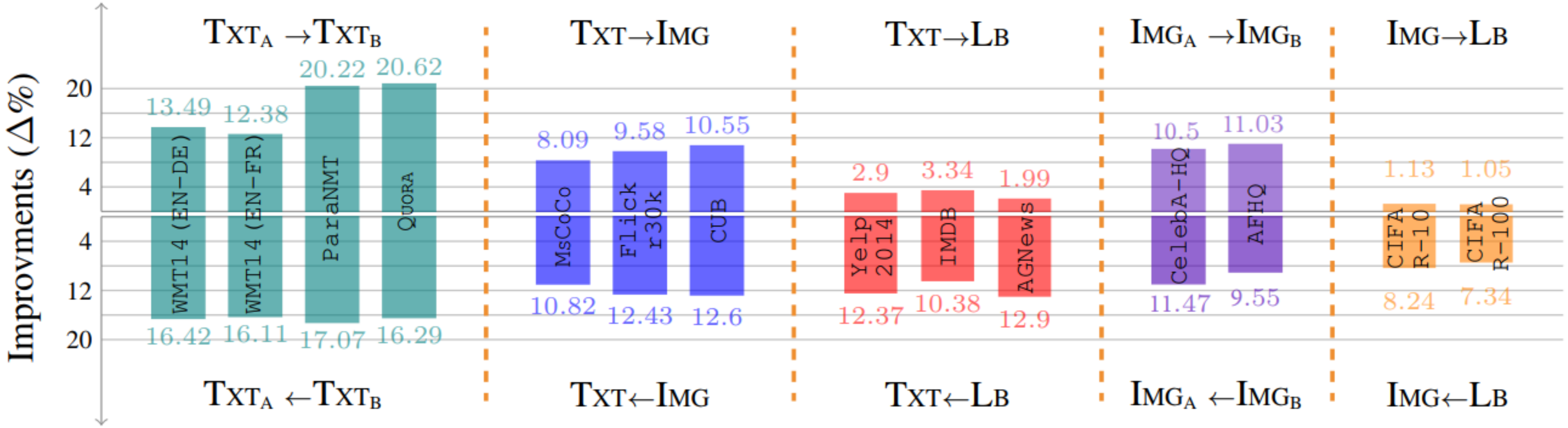


Figure 16. Case study for Text↔Image application.



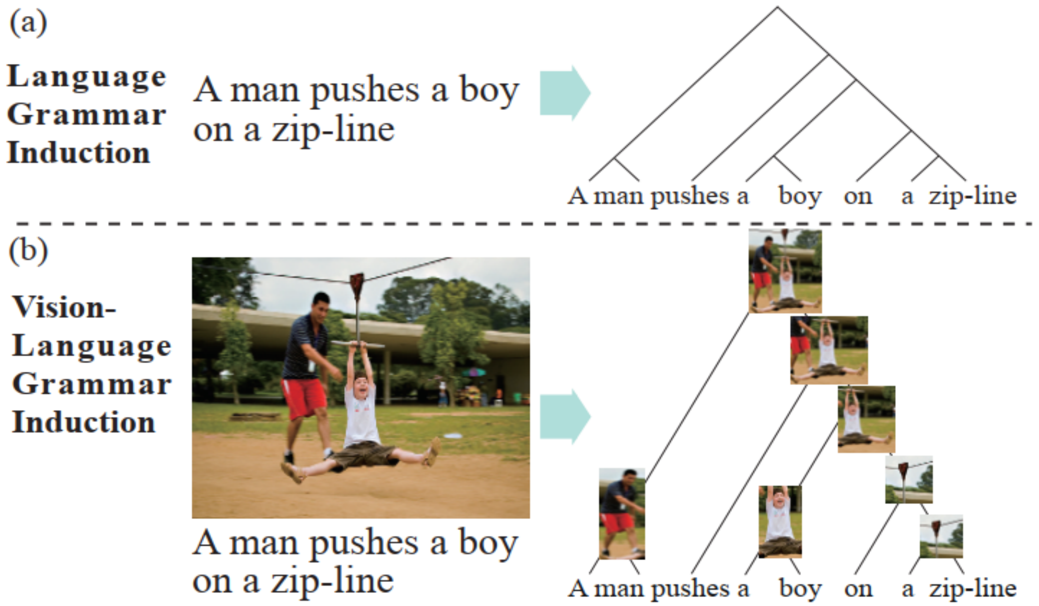
[Structure-aware multimodal]

What next?

➢ On-Going work-I: Unsupervised vision-language grammar induction

- Two main challenges

  - Context-dependent semantic representation learning.
  - Fine-grained vision-language alignment for all levels of the hierarchical structure.



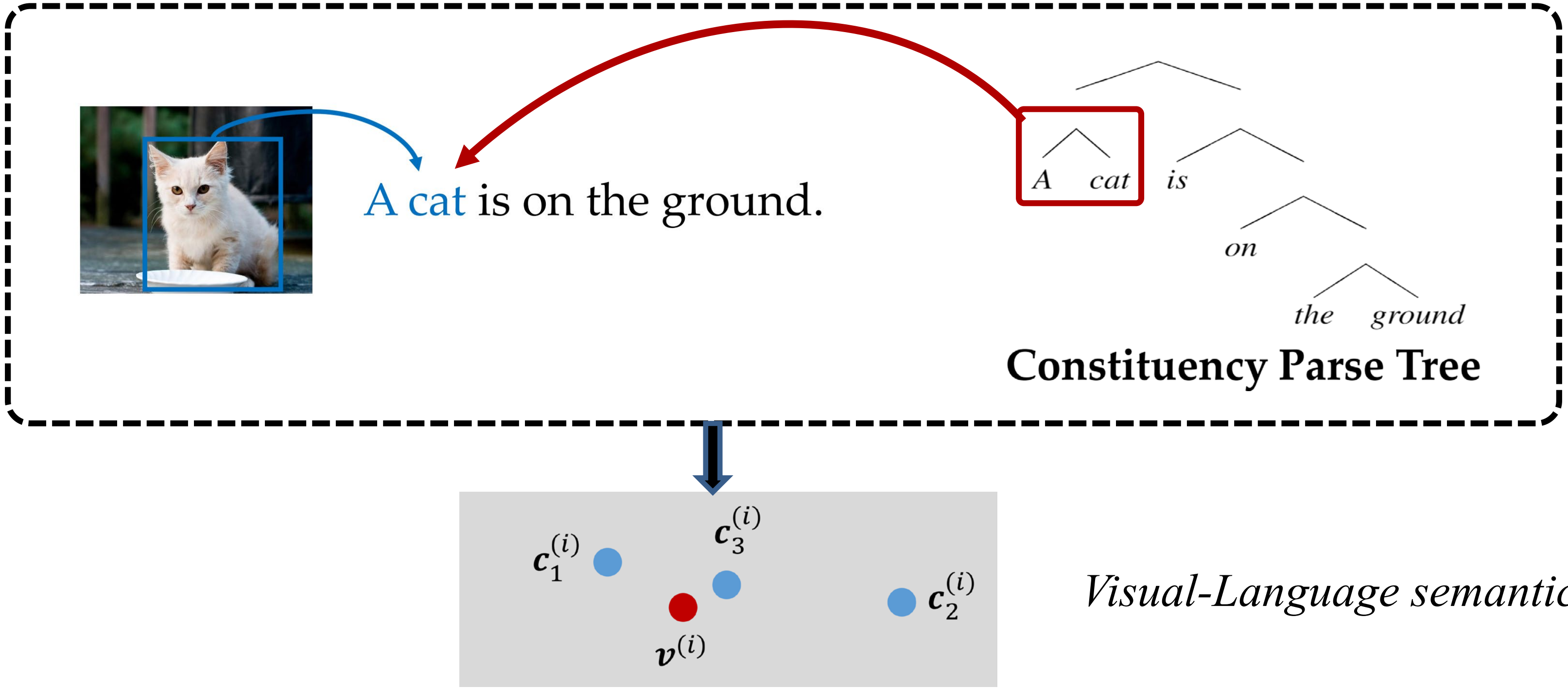
✓ *Pre-training multimodal language model via unsupervised learning*

[Structure-aware multimodal]
What next?

➢ On-Going work-II: Deep Text-Image structure alignment

- Image side: scene graph/objective proposals
- Text side: constituency-dependency structure

➢ *Improving the explainability of the multi-modal system*



Visual-Language semantics alignment

What next?

- On-Going work-III: Deep Text-Video structure alignment
  - Video side: event proposals?
  - Text side: event/semantic structure

Static

One single image:

Rachel and **Rose**.

The **alignments** between the **visual object** and **text entities**

Dynamic

Video:

Rachel explains to **her dad** on the **phone** why she can't marry her **fiancé**.

A0 prd A1

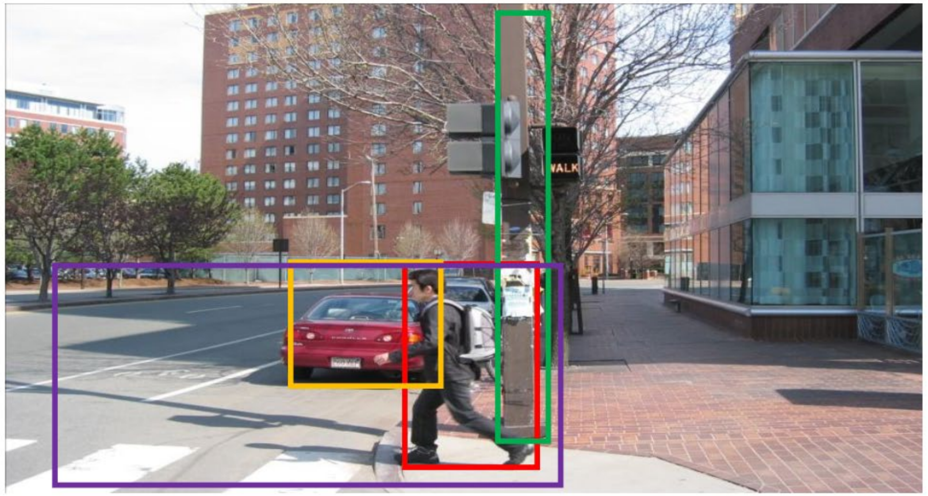
Rachel **explains** to her dad …

- **Static entities alignments**
- **Dynamic event alignments**

[Structure-aware multimodal]

What next?

- On-Going work-IV: Multimodal spatial semantics understanding
  - Spatial relation detection/Visual spatial description
    - Given an image and two objects inside it, VSD produces one description focusing on the spatial perspective between the two objects.
    - Created benchmark datasets:
      - 30K images;
      - 100K sentences.



	Condition	Target Text
Image Captioning	– –	*A man is walking past a car A*
SRL-based Captioning	walk; <Arg>, <Loc>	*man is walking cross a street*
Visual Question Answering	*What color is the car?*	*The car is red*
Our Work: VSD	<man, car>	*A man is walking behind a red car.*
	<car, pole>	*A red car is parked to the left of a pole.*

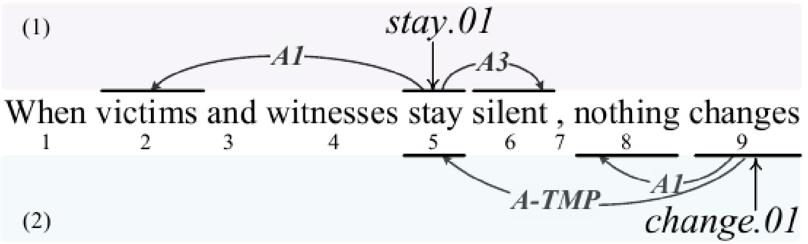
[Structure-aware multimodal]

What next?

- On-Going work-V: Multimodal semantic role labeling

  - Semantic role labeling (SRL)

    'who did what to whom, when and where'



  - Multimodal semantic role labeling/Spatial Role Labeling

SemEval-2015 Task 8

Trajector *Landmark*

a flag on top of the building

Sp Indicator

<*flag*, *on top of*, *building*>

Sp Relation

[**Arriving**$_{m1}$] [**in**$_{ms1}$] the [**town of Juanjui**$_{pl1}$], near the [**park**$_{pl2}$], [**I**$_{se1}$] learned that my map had lied to me.

<MOTION id=m1 extent="Arriving" motion type=PATH motion class=REACH motion sense=LITERAL>
<MOTION SIGNAL id=ms1 extent="in" motion signal type=PATH>
<PLACE id=pl1 extent="town of Juanjui" form=NAM countable=TRUE dimensionality=AREA>
<PLACE id=pl2 extent="park" form=NAM countable=TRUE dimensionality=AREA>
<SPATIAL ENTITY id=se1 extent="I" form=NOM countable=TRUE dimensionality=VOLUME>
<MOVELINK id=mvl1 trigger=m1 goal=pl1 mover=se1 goal reached=TRUE motion signalID=ms1>

[Structure-aware multimodal]

What next?

Structure awareness has more potential…

Thank You!

Q&A