Nested Coordination in Universal Dependencies

Adam Przepiórkowski[1,2,3] and Agnieszka Patejuk[1,4]

[1]Institute of Computer Science, **Polish Academy of Sciences**
[2]Institute of Philosophy, **University of Warsaw**
[3]Wolfson College, **University of Oxford**
[4]Centre for Linguistics and Philology, **University of Oxford**

Syntax Fest 2019
Paris, 30 August 2019

Problem – UD

1. Tom and Jerry and Scooby-Doo (ternary)
2. [Tom and Jerry] and Scooby-Doo (binary)
3. Tom and [Jerry and Scooby-Doo] (binary)

Universal Dependencies (UD):

1, 2

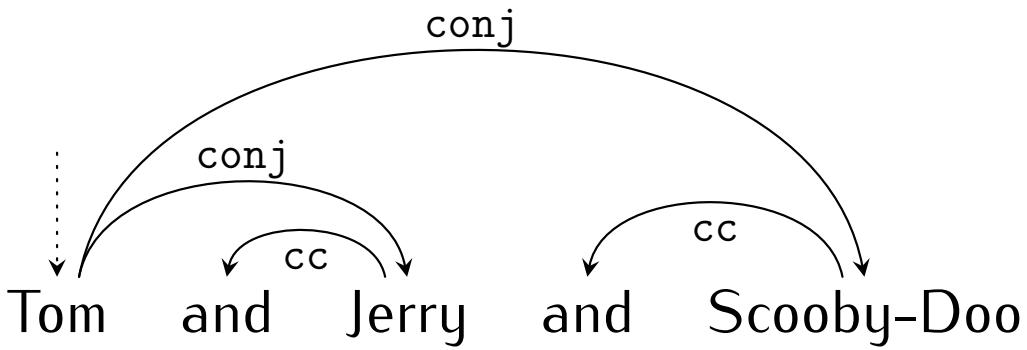
3

Problem – UD

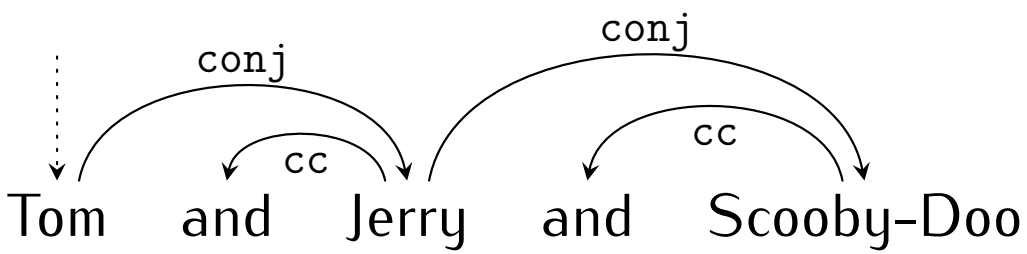
1 Tom and Jerry and Scooby-Doo (ternary)
2 [Tom and Jerry] and Scooby-Doo (binary)
3 Tom and [Jerry and Scooby-Doo] (binary)

Universal Dependencies (UD):

1, 2



3

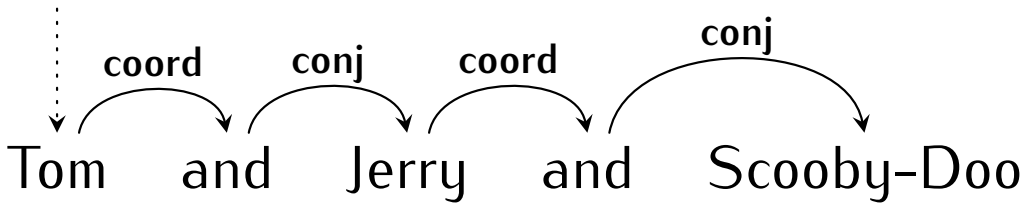


Problem – MTT

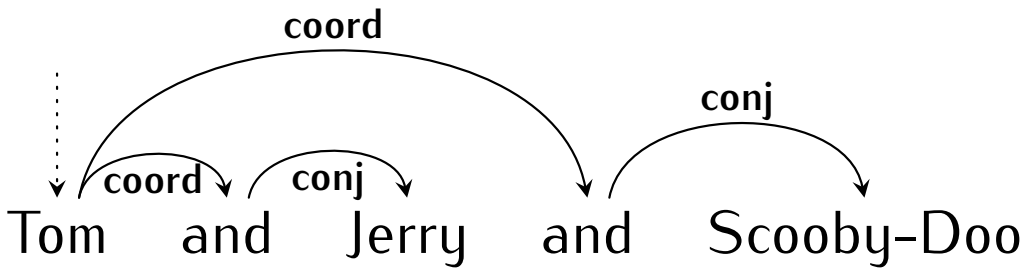
1. Tom and Jerry and Scooby-Doo
2. [Tom and Jerry] and Scooby-Doo
3. Tom and [Jerry and Scooby-Doo]

Igor Mel'čuk's Meaning–Text Theory (MTT):

1, 3



2

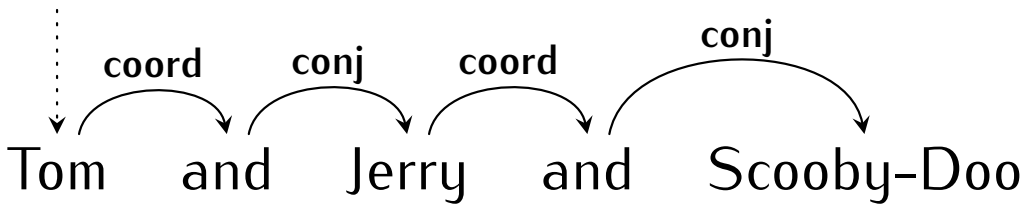


Solution – MTT

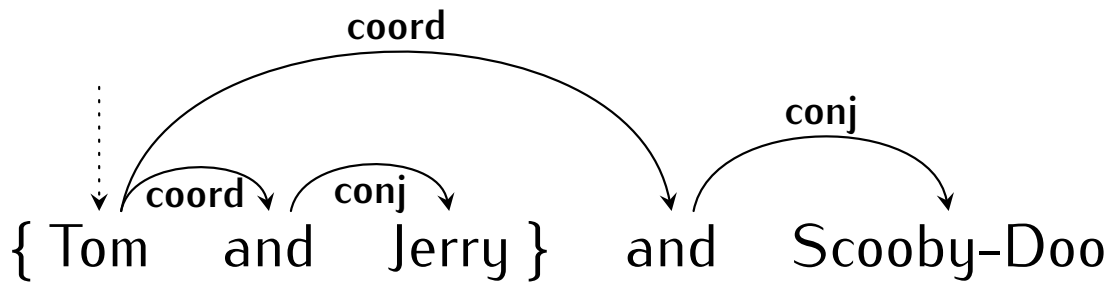
1. Tom and Jerry and Scooby-Doo
2. [Tom and Jerry] and Scooby-Doo
3. Tom and [Jerry and Scooby-Doo]

Groupings in MTT:

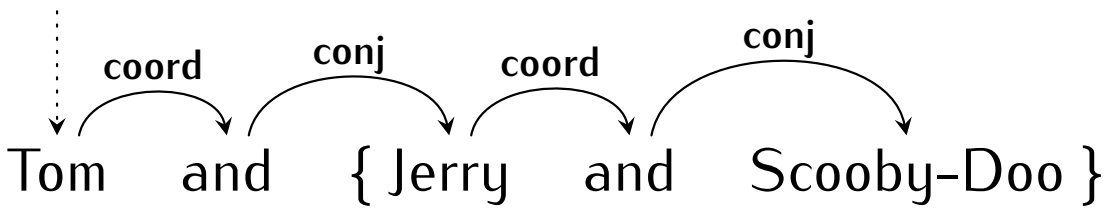
1



2



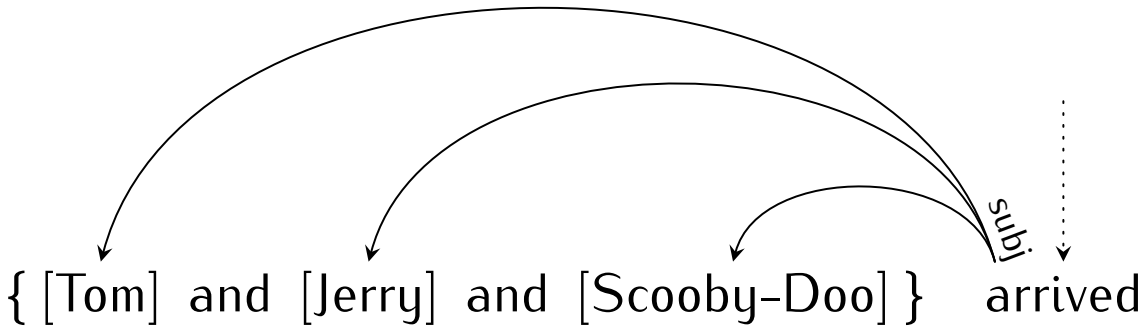
3



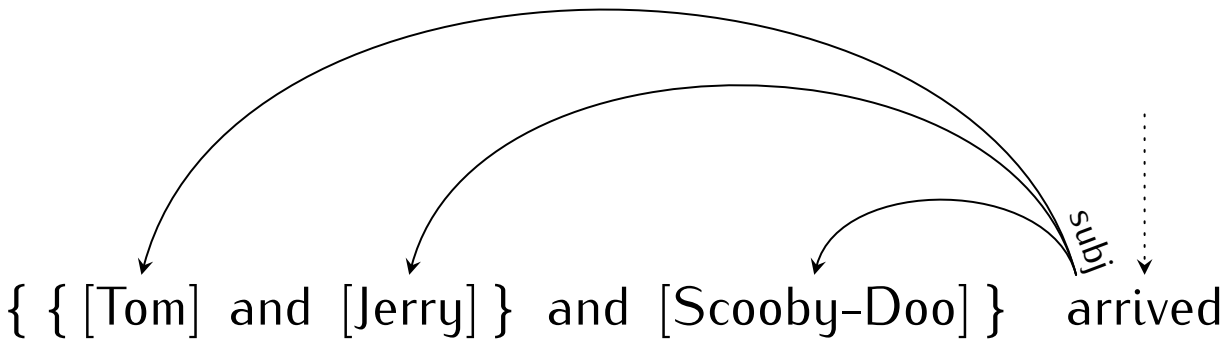
Solution – WG

Constituents in Dick Hudson's Word Grammar:

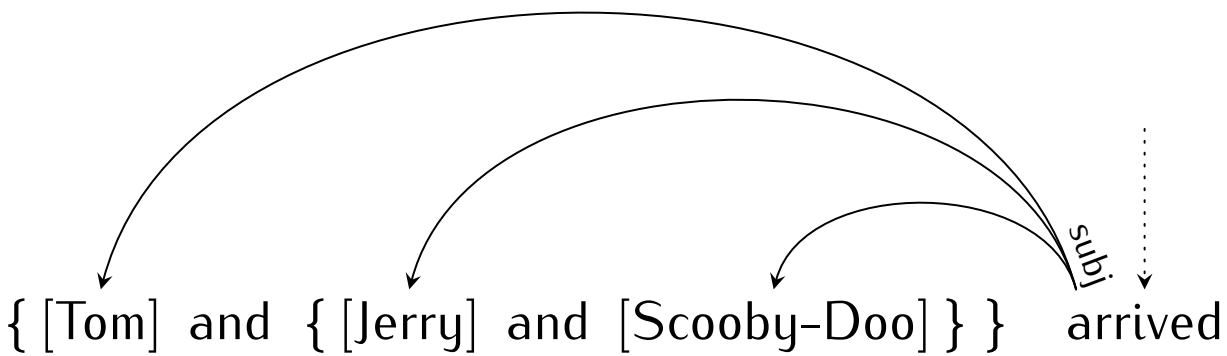
1



2



3



No problem for Prague

1. Tom and Jerry and Scooby-Doo
2. [Tom and Jerry] and Scooby-Doo
3. Tom and [Jerry and Scooby-Doo]

Prague-style:

1

2

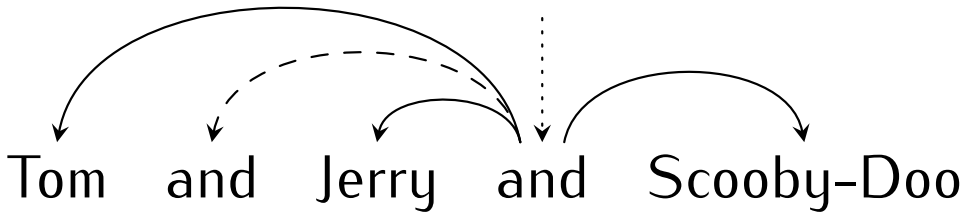
3

No problem for Prague

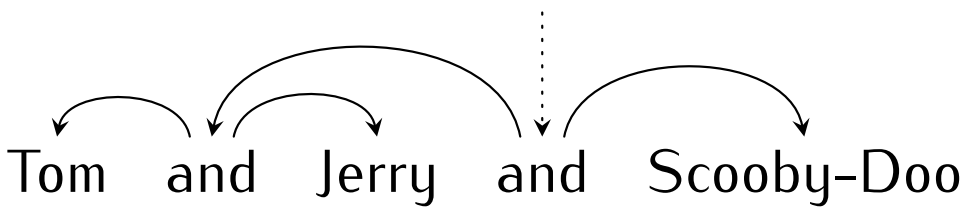
1. Tom and Jerry and Scooby-Doo
2. [Tom and Jerry] and Scooby-Doo
3. Tom and [Jerry and Scooby-Doo]

Prague-style:

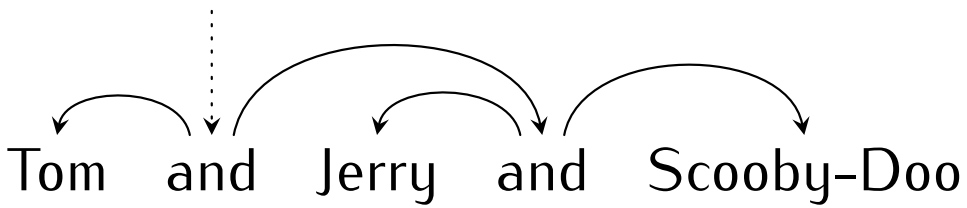
1



2



3



Problem – UD (summary)

But Prague-style analysis is **theoretically problematic** and rejected by many linguists of different theoretical persuasions (Mel'čuk and Pertsov 1987: 65, Hudson 1988: 314–315, Gerdes and Kahane 2015: 102–105; also Borsley 2005).

Summary of the problem:

- UD does not distinguish between certain nestings of coordination,
- MTT and WG use mechanisms unavailable in UD (groupings, constituents),
- not a technical problem for Prague-style analysis of coordination,
- but such an analysis is rejected by many theoretical linguists.

This talk: conservative solutions within standard UD.

Problem – UD (summary)

But Prague-style analysis is **theoretically problematic** and rejected by many linguists of different theoretical persuasions (Mel'čuk and Pertsov 1987: 65, Hudson 1988: 314–315, Gerdes and Kahane 2015: 102–105; also Borsley 2005).

Summary of the problem:

- UD does not distinguish between certain nestings of coordination,
- MTT and WG use mechanisms unavailable in UD (groupings, constituents),
- not a technical problem for Prague-style analysis of coordination,
- but such an analysis is rejected by many theoretical linguists.

This talk: conservative solutions within standard UD.

Problem – UD (summary)

But Prague-style analysis is **theoretically problematic** and rejected by many linguists of different theoretical persuasions (Mel'čuk and Pertsov 1987: 65, Hudson 1988: 314–315, Gerdes and Kahane 2015: 102–105; also Borsley 2005).

Summary of the problem:

- **UD does not distinguish between certain nestings of coordination**,
- MTT and WG use mechanisms unavailable in UD (groupings, constituents),
- **not a technical problem for Prague**-style analysis of coordination,
- but such an analysis is rejected by many theoretical linguists.

This talk: **conservative solutions within standard UD**.

Problem – UD (summary)

But Prague-style analysis is **theoretically problematic** and rejected by many linguists of different theoretical persuasions (Mel'čuk and Pertsov 1987: 65, Hudson 1988: 314–315, Gerdes and Kahane 2015: 102–105; also Borsley 2005).

Summary of the problem:

- **UD does not distinguish between certain nestings of coordination**,
- **MTT and WG use mechanisms unavailable in UD** (groupings, constituents),
- **not a technical problem for Prague**-style analysis of coordination,
- but such an analysis is rejected by many theoretical linguists.

This talk: **conservative solutions within standard UD**.

Problem – UD (summary)

But Prague-style analysis is **theoretically problematic** and rejected by many linguists of different theoretical persuasions (Mel'čuk and Pertsov 1987: 65, Hudson 1988: 314–315, Gerdes and Kahane 2015: 102–105; also Borsley 2005).

Summary of the problem:

- **UD does not distinguish between certain nestings of coordination**,
- **MTT and WG use mechanisms unavailable in UD** (groupings, constituents),
- **not a technical problem for Prague**-style analysis of coordination,
- but such an analysis is rejected by many theoretical linguists.

This talk: **conservative solutions within standard UD.**

Problem – UD (summary)

But Prague-style analysis is **theoretically problematic** and rejected by many linguists of different theoretical persuasions (Mel'čuk and Pertsov 1987: 65, Hudson 1988: 314–315, Gerdes and Kahane 2015: 102–105; also Borsley 2005).

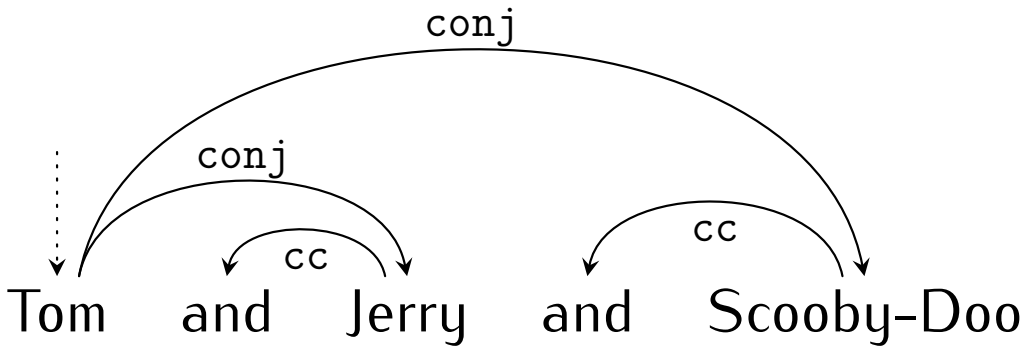
Summary of the problem:

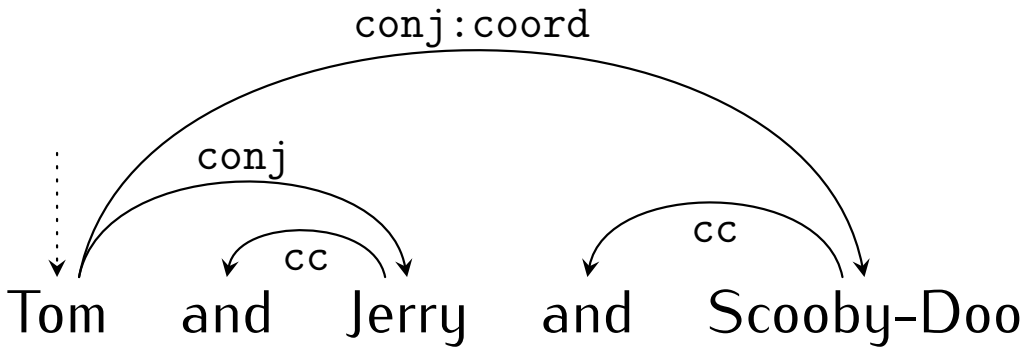
- **UD does not distinguish between certain nestings of coordination**,
- **MTT and WG use mechanisms unavailable in UD** (groupings, constituents),
- **not a technical problem for Prague**-style analysis of coordination,
- but such an analysis is rejected by many theoretical linguists.

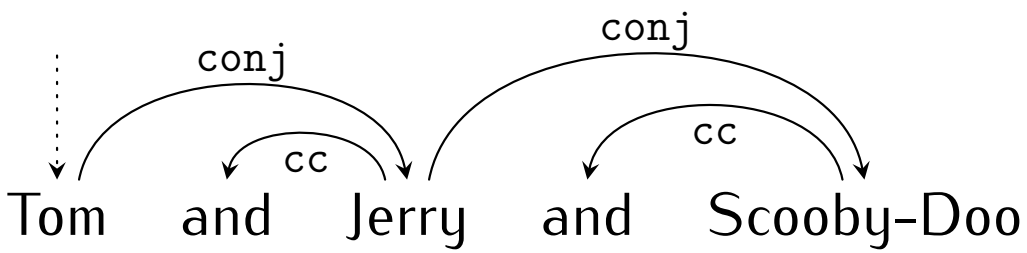
This talk: **conservative solutions within standard UD**.

Enriching labels 1

Enriching dependency labels via **subtyping**:

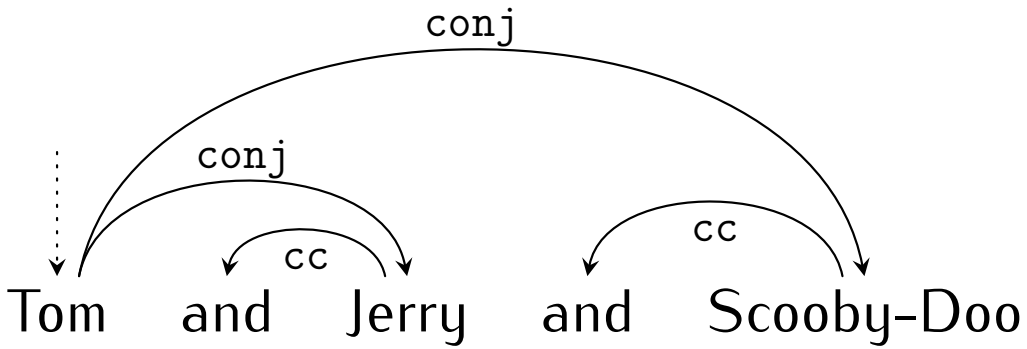
- 

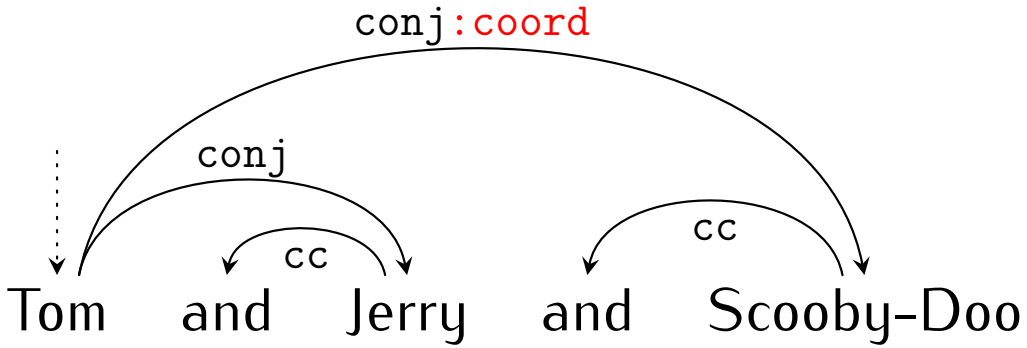
- 

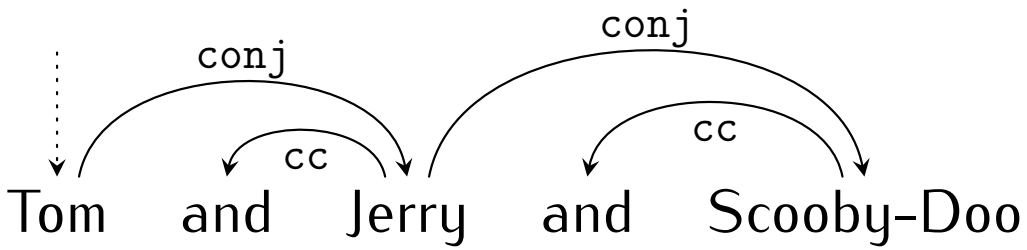
- 

Enriching labels 1

Enriching dependency labels via **subtyping**:

- 

- 

- 

Enriching labels 2

Similar solution considered in Mel'čuk 1988: 30 and Mel'čuk 2009: 93–94 and earlier. **Rejected** as 'highly unnatural' and leading to the doubling of dependency labels.

A problem for UD – **no theoretical limit** to the number of subtypes (Schuster *et al.* 2017: 130–131):

- [[Tom and Jerry] and Spike] and Scooby-Doo
-

Enriching labels 2

Similar solution considered in Mel'čuk 1988: 30 and Mel'čuk 2009: 93–94 and earlier. **Rejected** as 'highly unnatural' and leading to the doubling of dependency labels.

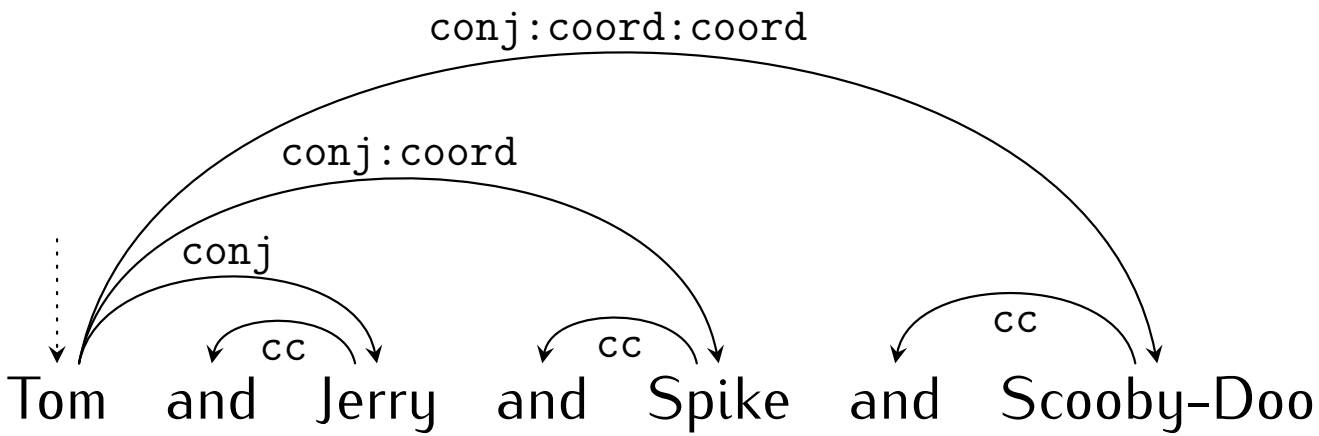
A problem for UD – **no theoretical limit** to the number of subtypes (Schuster *et al.* 2017: 130–131):

- [[Tom and Jerry] and Spike] and Scooby-Doo
-

Enriching labels 2

Similar solution considered in Mel'čuk 1988: 30 and Mel'čuk 2009: 93–94 and earlier. **Rejected** as 'highly unnatural' and leading to the doubling of dependency labels.

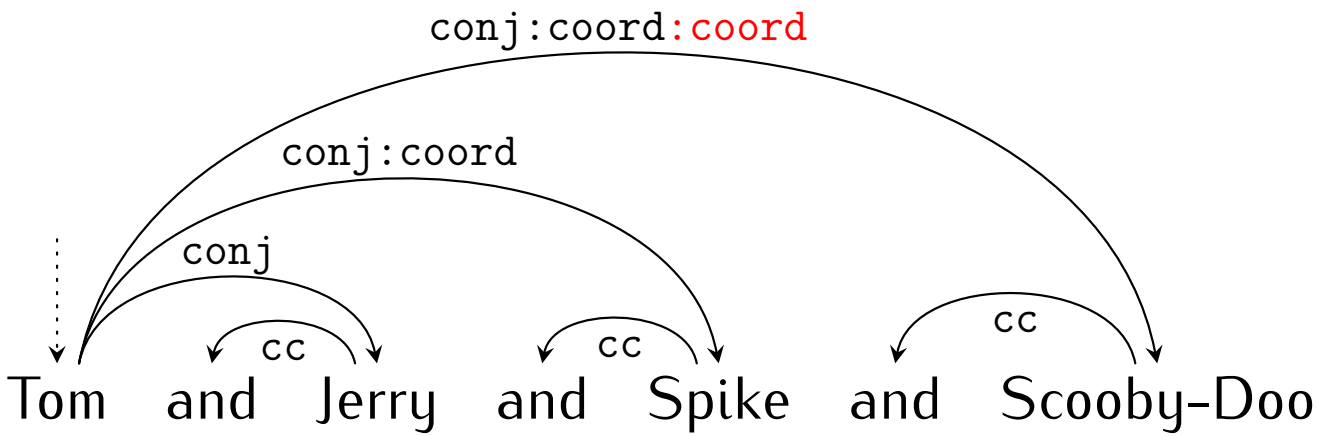
A problem for UD – **no theoretical limit** to the number of subtypes (Schuster *et al.* 2017: 130–131):

- [[Tom and Jerry] and Spike] and Scooby-Doo
- 

Enriching labels 2

Similar solution considered in Mel'čuk 1988: 30 and Mel'čuk 2009: 93–94 and earlier. **Rejected** as 'highly unnatural' and leading to the doubling of dependency labels.

A problem for UD – **no theoretical limit** to the number of subtypes (Schuster *et al.* 2017: 130–131):

- [[Tom and Jerry] and Spike] and Scooby-Doo
- 

Conjuncts as co-heads in enhanced representations 1

Retain basic tree representation, add **different enhanced representation**:

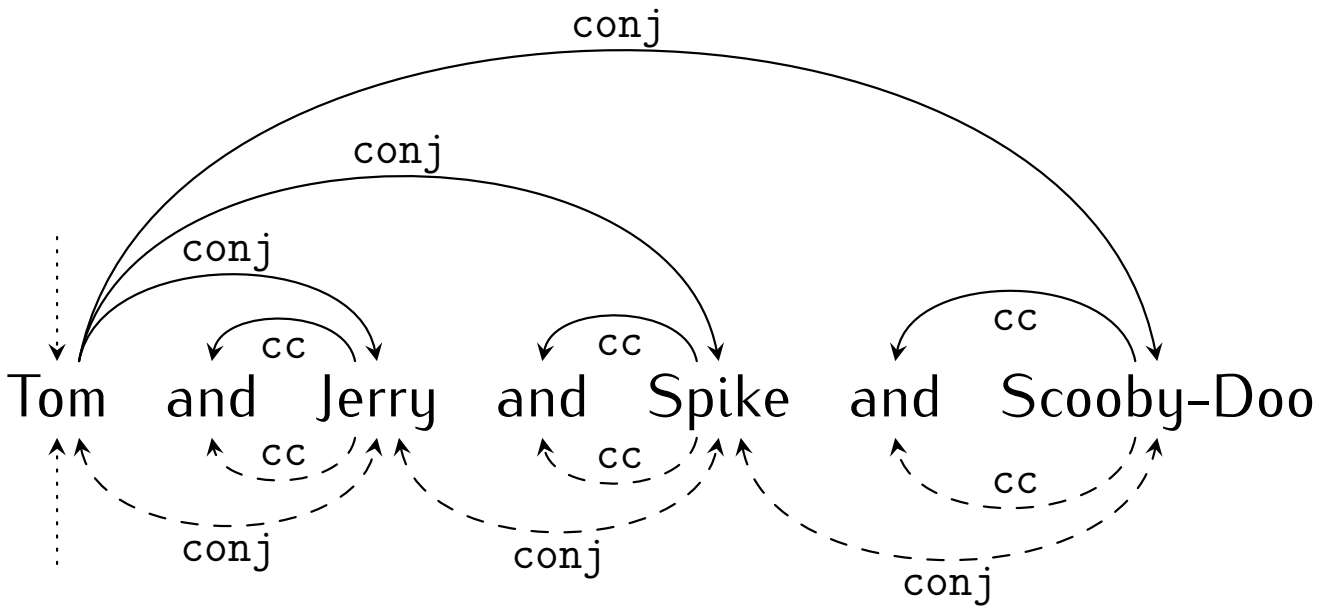
- Tom and Jerry and Spike and Scooby-Doo: (no nesting)

Idea: pairs of dependencies (in both directions) between neighbouring conjuncts.

Conjuncts as co-heads in enhanced representations 1

Retain basic tree representation, add **different enhanced representation**:

- Tom and Jerry and Spike and Scooby-Doo: (no nesting)

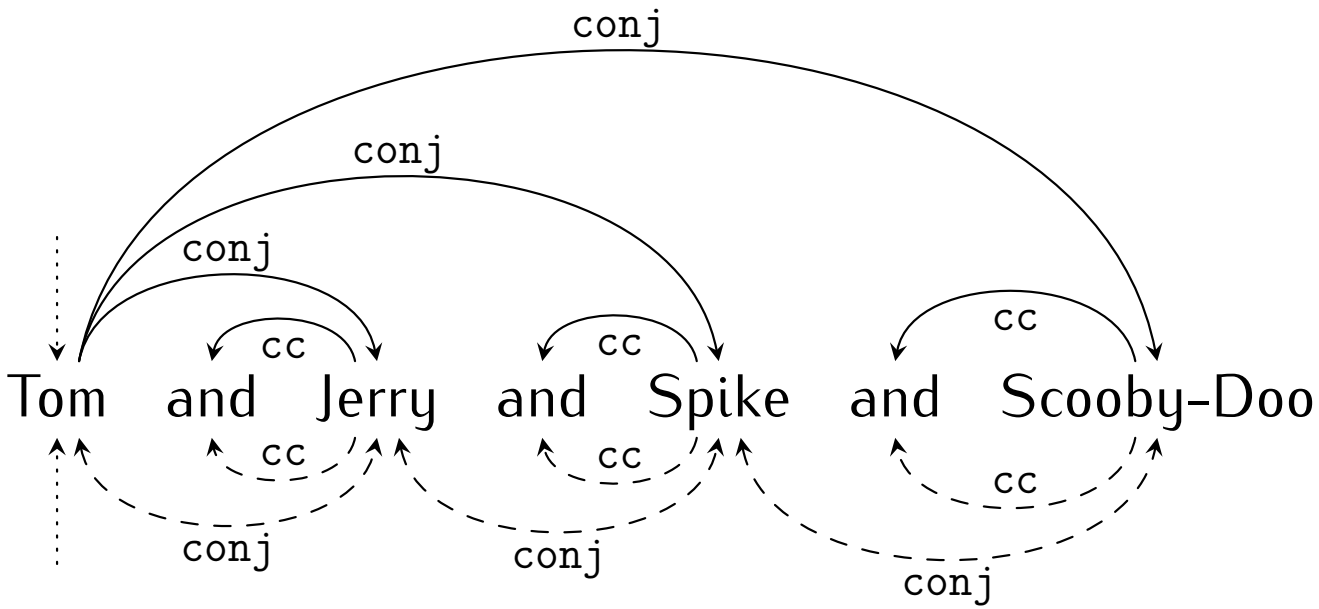


Idea: pairs of dependencies (in both directions) between neighbouring conjuncts.

Conjuncts as co-heads in enhanced representations 1

Retain basic tree representation, add **different enhanced representation**:

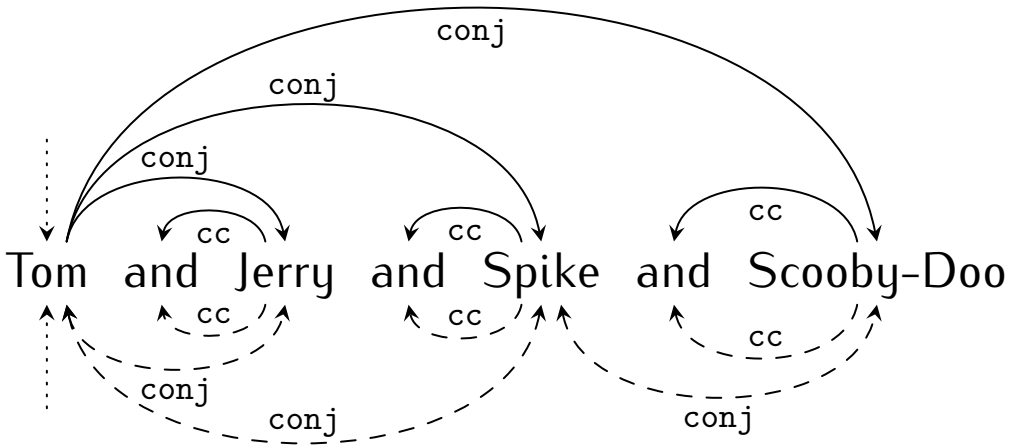
- Tom and Jerry and Spike and Scooby-Doo: (no nesting)



Idea: pairs of dependencies (in both directions) between neighbouring conjuncts.

Conjuncts as co-heads in enhanced representations 2

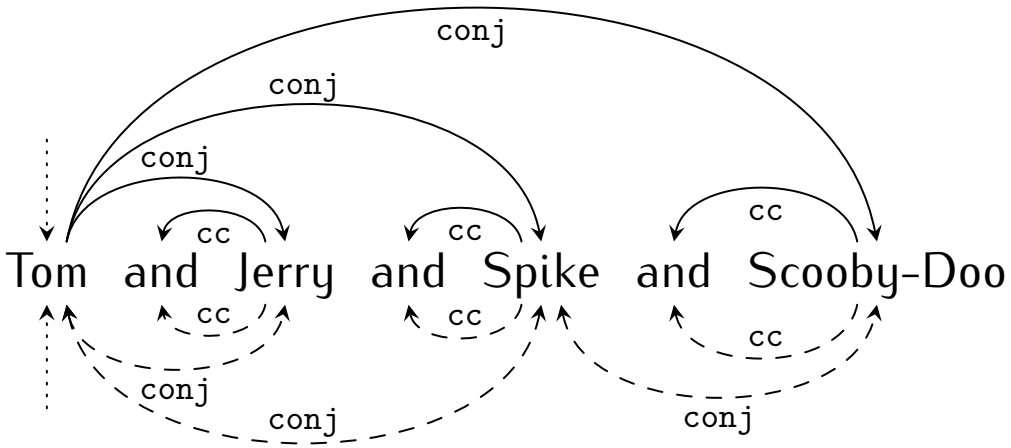
- [Tom and Jerry] and Spike and Scooby-Doo:



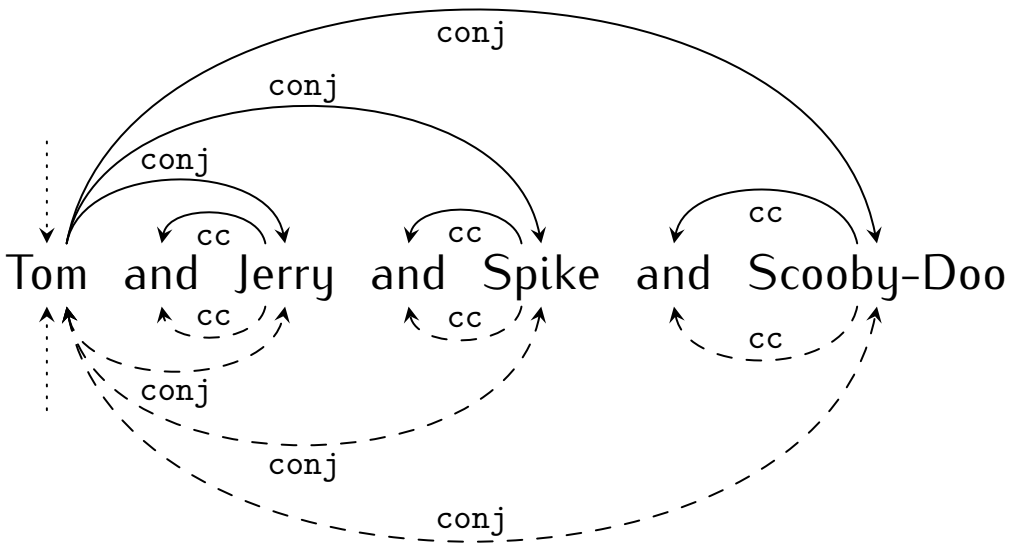
- [[Tom and Jerry] and Spike] and Scooby-Doo:

Conjuncts as co-heads in enhanced representations 2

- [Tom and Jerry] and Spike and Scooby-Doo:



- [[Tom and Jerry] and Spike] and Scooby-Doo:



Conjuncts as co-heads in enhanced representations 2

Pros:

- it can be shown that **representations of different nestings differ**
- enhanced graph implements the common idea that **conjuncts are co-heads**
  - in dependency approaches: Tesnière 1959 (similar sentiments expressed by Hudson)
  - in constituency approaches: Gazdar *et al.* 1985 (similar sentiments in some HPSG work)
  - in combined approaches: Kahane 1997, Kahane and Mazziotta 2015

Cons:

- **basic trees and enhanced graphs must be inspected together** to reveal full structure
- although this may be rectified by copying basic tree to enhanced graphs (and modifying labels accordingly)

Conjuncts as co-heads in enhanced representations 2

Pros:

- it can be shown that **representations of different nestings differ**
- enhanced graph implements the common idea that **conjuncts are co-heads**
  - in dependency approaches: Tesnière 1959 (similar sentiments expressed by Hudson)
  - in constituency approaches: Gazdar *et al.* 1985 (similar sentiments in some HPSG work)
  - in combined approaches: Kahane 1997, Kahane and Mazziotta 2015

Cons:

- **basic trees and enhanced graphs must be inspected together** to reveal full structure
- although this may be rectified by copying basic tree to enhanced graphs (and modifying labels accordingly)

Conjuncts as co-heads in enhanced representations 2

Pros:

- it can be shown that **representations of different nestings differ**
- enhanced graph implements the common idea that **conjuncts are co-heads**
  - in dependency approaches: Tesnière 1959 (similar sentiments expressed by Hudson)
  - in constituency approaches: Gazdar *et al.* 1985 (similar sentiments in some HPSG work)
  - in combined approaches: Kahane 1997, Kahane and Mazziotta 2015

Cons:

- **basic trees and enhanced graphs must be inspected together** to reveal full structure
- although this may be rectified by copying basic tree to enhanced graphs (and modifying labels accordingly)

Conjuncts as co-heads in enhanced representations 2

Pros:

- it can be shown that **representations of different nestings differ**
- enhanced graph implements the common idea that **conjuncts are co-heads**
  - in dependency approaches: Tesnière 1959 (similar sentiments expressed by Hudson)
  - in constituency approaches: Gazdar *et al.* 1985 (similar sentiments in some HPSG work)
  - in combined approaches: Kahane 1997, Kahane and Mazziotta 2015

Cons:

- **basic trees and enhanced graphs must be inspected together** to reveal full structure
- although this may be rectified by copying basic tree to enhanced graphs (and modifying labels accordingly)

Conjuncts as co-heads in enhanced representations 2

Pros:

- it can be shown that **representations of different nestings differ**
- enhanced graph implements the common idea that **conjuncts are co-heads**
  - in dependency approaches: Tesnière 1959 (similar sentiments expressed by Hudson)
  - in constituency approaches: Gazdar *et al.* 1985 (similar sentiments in some HPSG work)
  - in combined approaches: Kahane 1997, Kahane and Mazziotta 2015

Cons:

- **basic trees and enhanced graphs must be inspected together** to reveal full structure
- although this may be rectified by copying basic tree to enhanced graphs (and modifying labels accordingly)

Conjuncts as co-heads in enhanced representations 2

Pros:

- it can be shown that **representations of different nestings differ**
- enhanced graph implements the common idea that **conjuncts are co-heads**
  - in dependency approaches: Tesnière 1959 (similar sentiments expressed by Hudson)
  - in constituency approaches: Gazdar *et al.* 1985 (similar sentiments in some HPSG work)
  - in combined approaches: Kahane 1997, Kahane and Mazziotta 2015

Cons:

- **basic trees and enhanced graphs must be inspected together** to reveal full structure
- although this may be rectified by copying basic tree to enhanced graphs (and modifying labels accordingly)

Distribution in enhanced representations 1

Recall **standard UD treatment of coordination**:

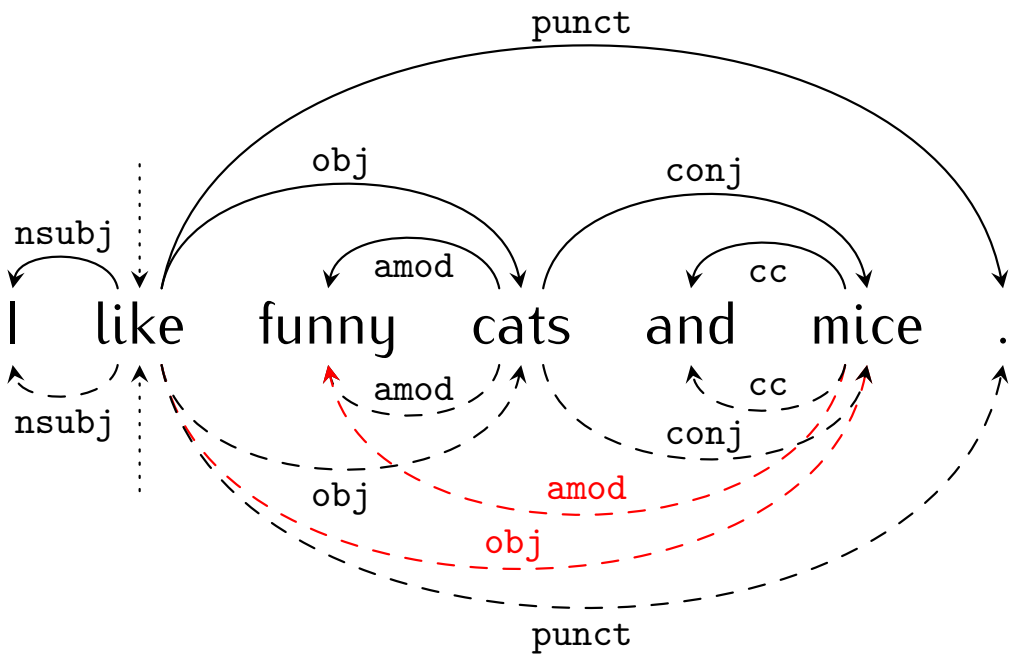
- I like funny cats and mice.
-

In **enhanced representation**:

- distribute dependencies **to** coordinate structure (cf. obj above),
- distribute dependencies **from** coordinate structure (cf. amod above).

Distribution in enhanced representations 1

Recall **standard UD treatment of coordination**:

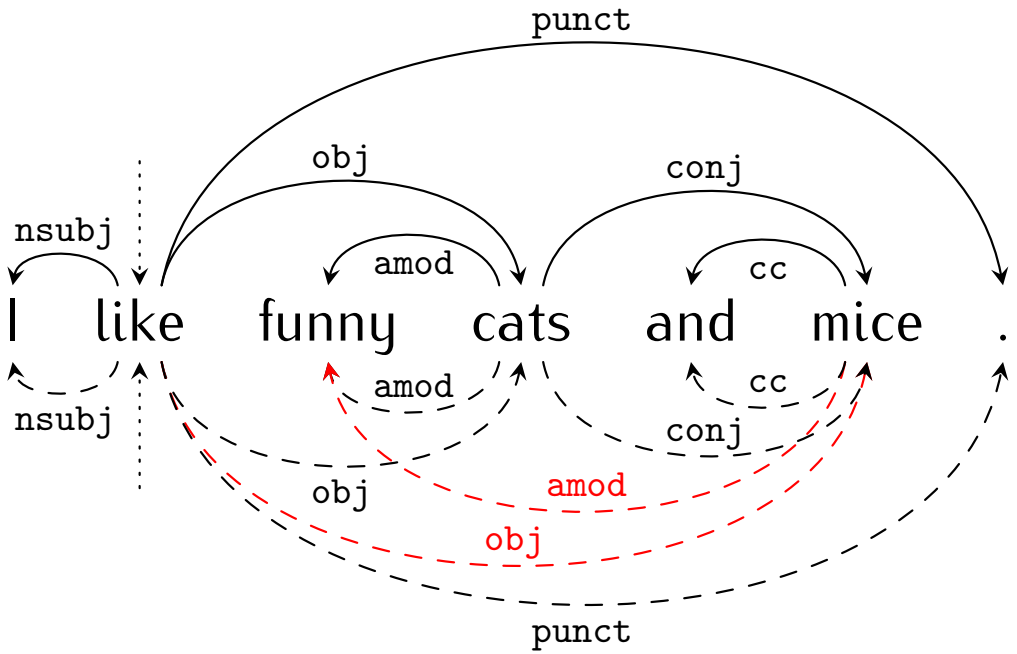
- I like funny cats and mice.
- 

In **enhanced representation**:

- distribute dependencies **to** coordinate structure (cf. obj above),
- distribute dependencies **from** coordinate structure (cf. amod above).

Distribution in enhanced representations 1

Recall **standard UD treatment of coordination**:

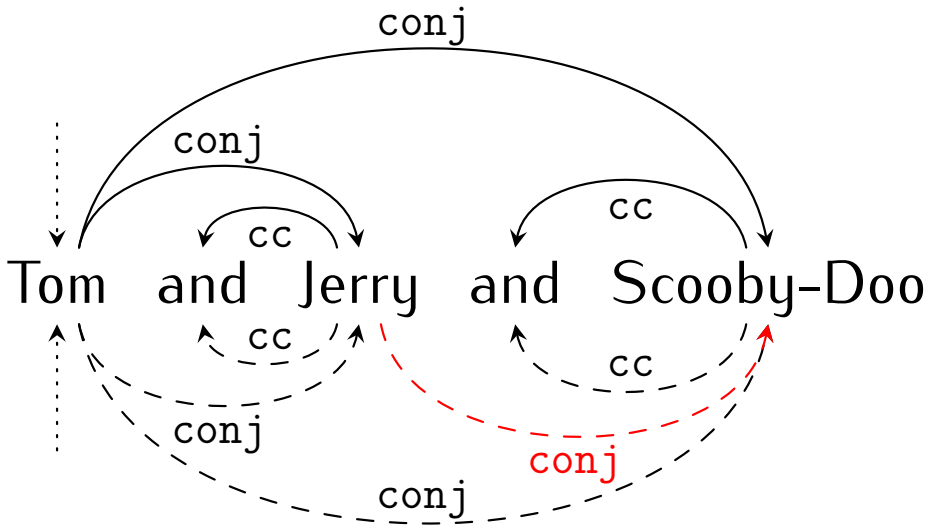
- I like funny cats and mice.
- 

In **enhanced representation**:

- distribute dependencies **to** coordinate structure (cf. obj above),
- distribute dependencies **from** coordinate structure (cf. amod above).

Distribution in enhanced representations 2

- [Tom and Jerry] and Scooby-Doo:



- Tom and [Jerry and Scooby-Doo]:

Pros:

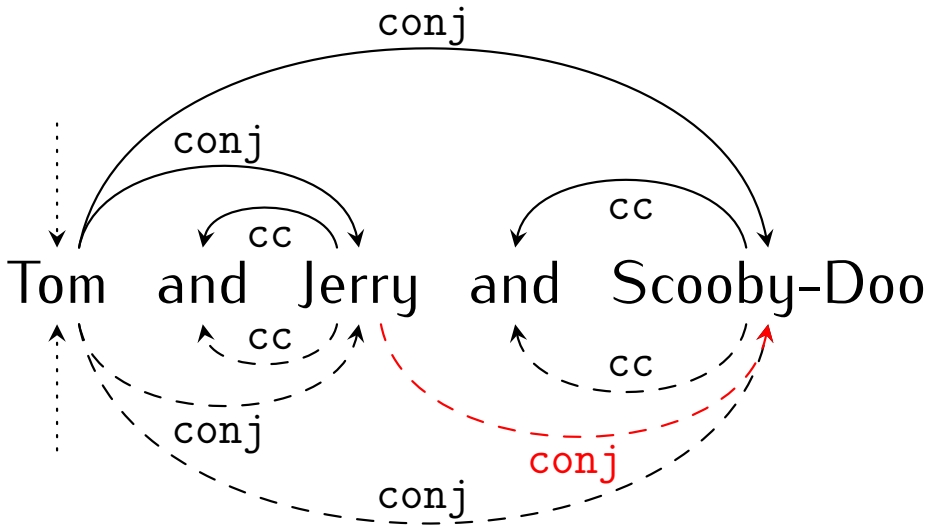
- maximally conservative solution

Cons:

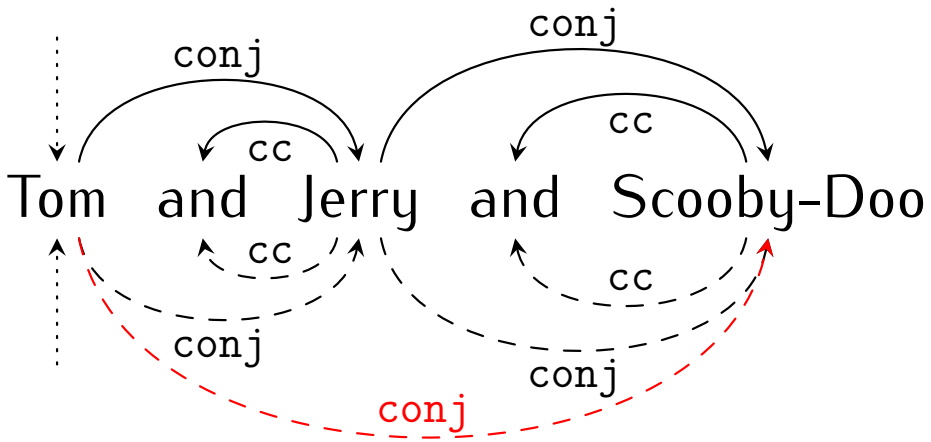
- requires inspecting both levels (basic trees and enhanced graphs)
- this time, this can't be rectified so easily
- (does not encode the co-heads idea)

Distribution in enhanced representations 2

- [Tom and Jerry] and Scooby-Doo:



- Tom and [Jerry and Scooby-Doo]:



Pros:

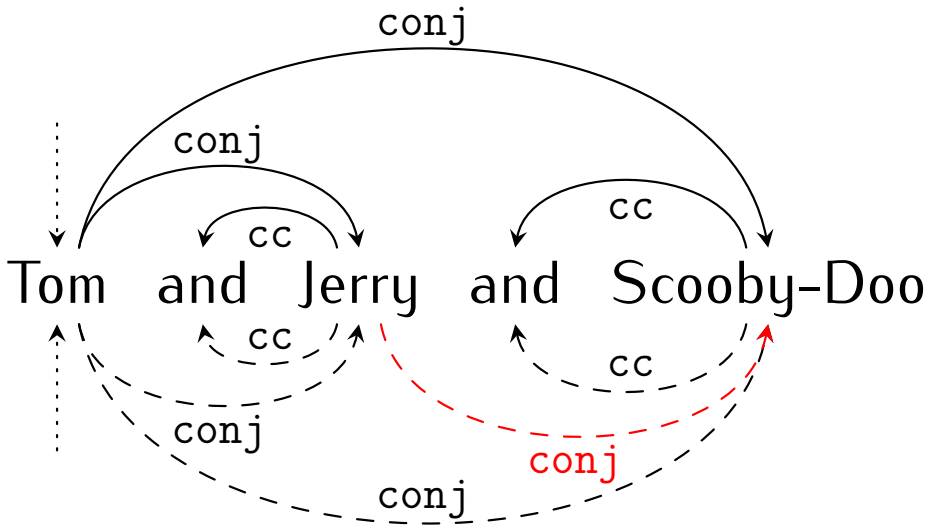
- maximally conservative solution

Cons:

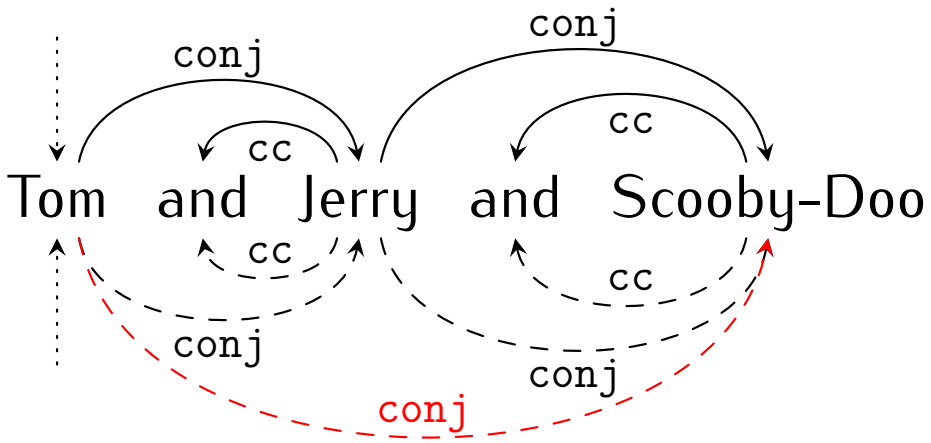
- requires inspecting both levels (basic trees and enhanced graphs)
- this time, this can't be rectified so easily
- (does not encode the co-heads idea)

Distribution in enhanced representations 2

- [Tom and Jerry] and Scooby-Doo:



- Tom and [Jerry and Scooby-Doo]:



Pros:

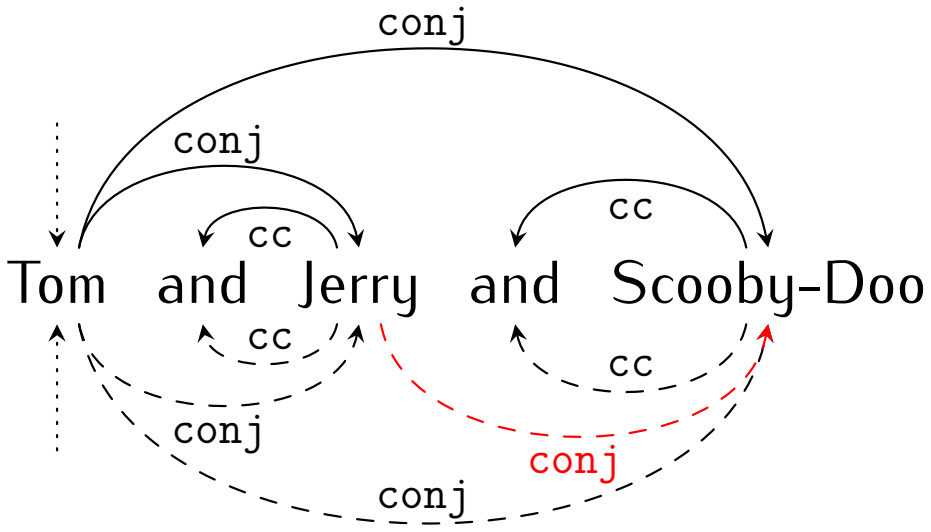
- **maximally conservative** solution

Cons:

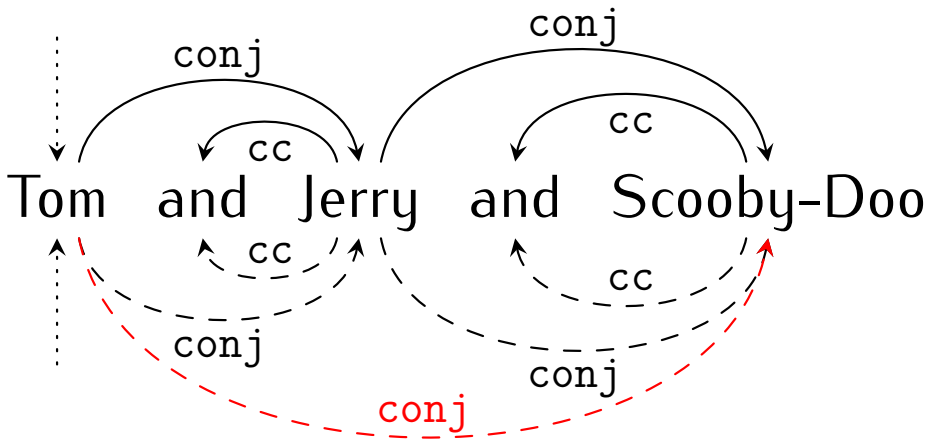
- **requires inspecting both levels** (basic trees and enhanced graphs)
- this time, this can't be rectified so easily
- (**does not encode the co-heads idea**)

Distribution in enhanced representations 2

- [Tom and Jerry] and Scooby-Doo:



- Tom and [Jerry and Scooby-Doo]:



Pros:

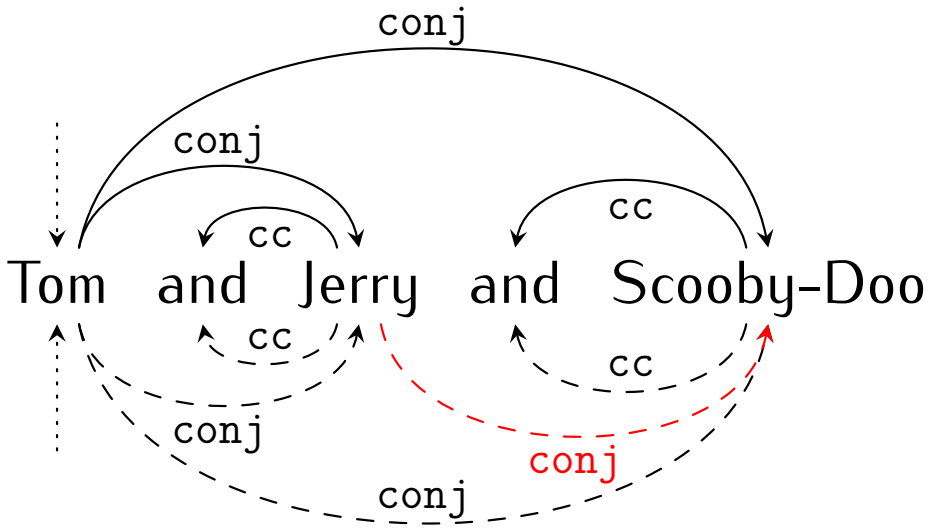
- **maximally conservative** solution

Cons:

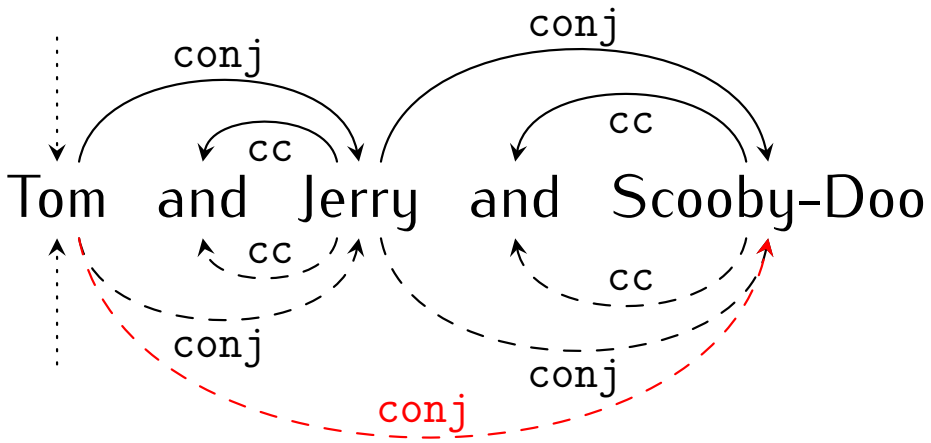
- **requires inspecting both levels** (basic trees and enhanced graphs)
- this time, this can't be rectified so easily
- (does not encode the co-heads idea)

Distribution in enhanced representations 2

- [Tom and Jerry] and Scooby-Doo:



- Tom and [Jerry and Scooby-Doo]:



Pros:

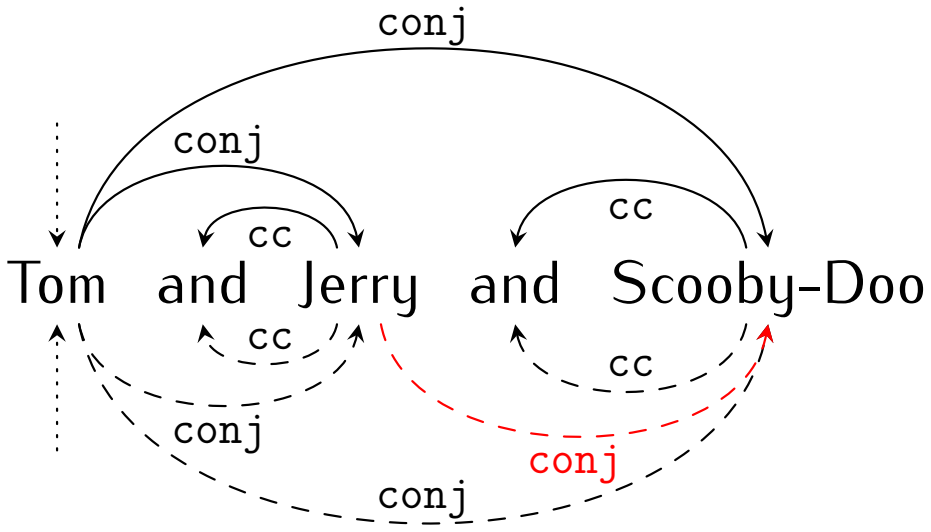
- **maximally conservative** solution

Cons:

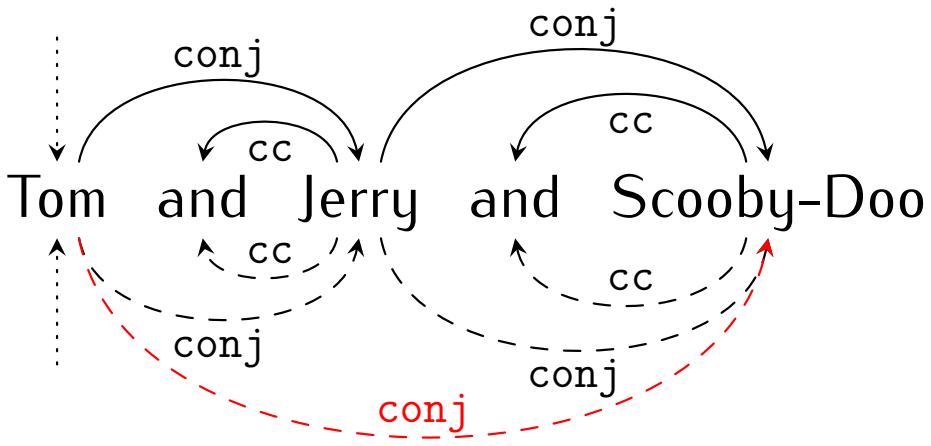
- **requires inspecting both levels** (basic trees and enhanced graphs)
- this time, this can't be rectified so easily
- (does not encode the co-heads idea)

Distribution in enhanced representations 2

- [Tom and Jerry] and Scooby-Doo:



- Tom and [Jerry and Scooby-Doo]:



Pros:

- **maximally conservative** solution

Cons:

- **requires inspecting both levels** (basic trees and enhanced graphs)
- this time, this can't be rectified so easily
- (**does not encode the co-heads idea**)

Distribution in enhanced representations 3

Better solution: only distribute dependencies **from** coordinate structures:

- Tom and [[Jerry and Spike] and Scooby-Doo]:

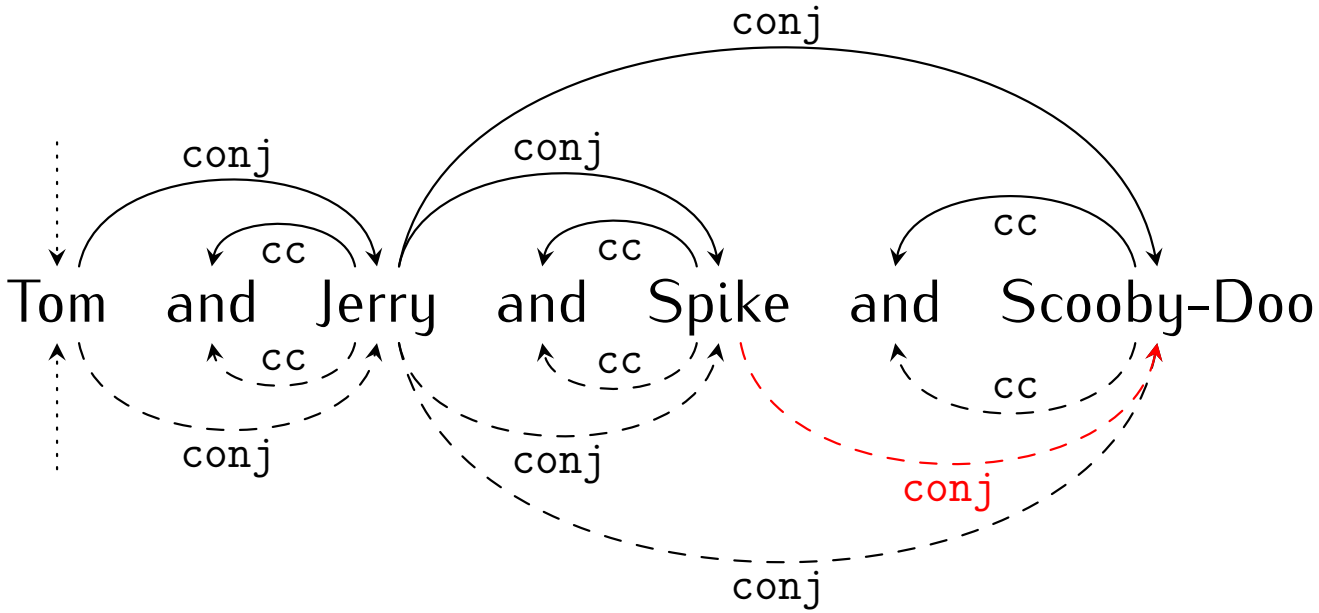
Pros:

- **different nestings distinguished in enhanced representations** (alone)
- another argument for getting rid of the problematic aspect of UD approach to distribution in coordinate structures (two outgoing `obl` dependencies – two dependents or one?)

Distribution in enhanced representations 3

Better solution: only distribute dependencies **from** coordinate structures:

- Tom and [[Jerry and Spike] and Scooby-Doo]:



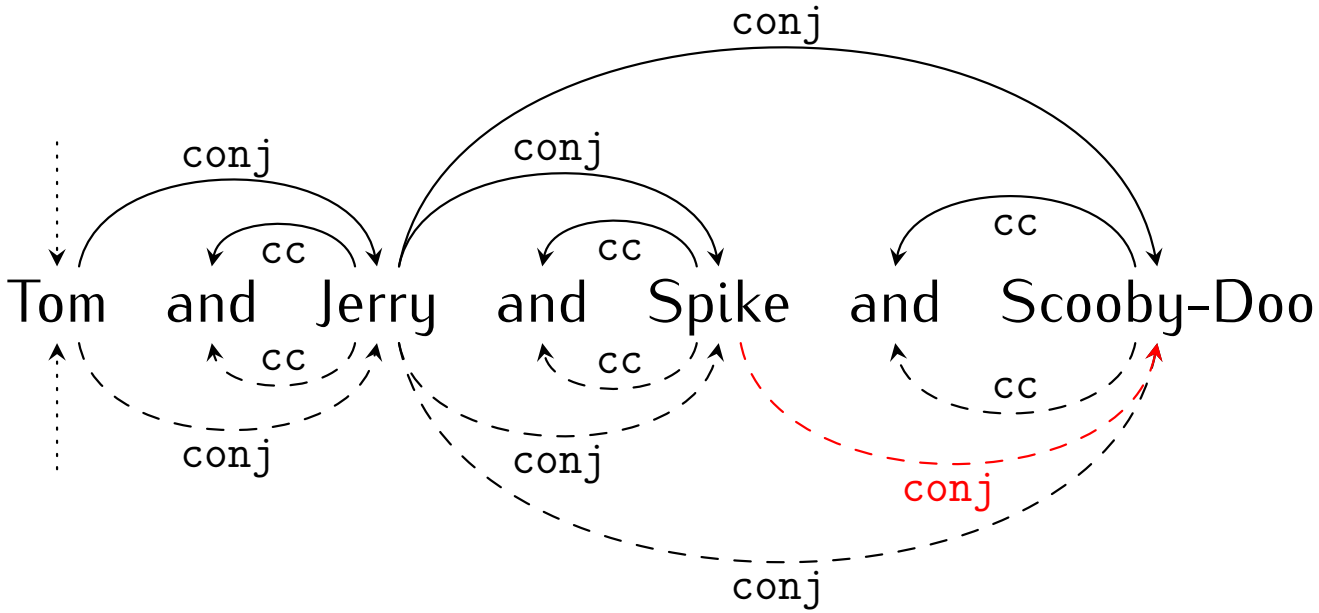
Pros:

- **different nestings distinguished in enhanced representations** (alone)
- another argument for getting rid of the problematic aspect of UD approach to distribution in coordinate structures (two outgoing `obl` dependencies – two dependents or one?)

Distribution in enhanced representations 3

Better solution: only distribute dependencies **from** coordinate structures:

- Tom and [[Jerry and Spike] and Scooby-Doo]:



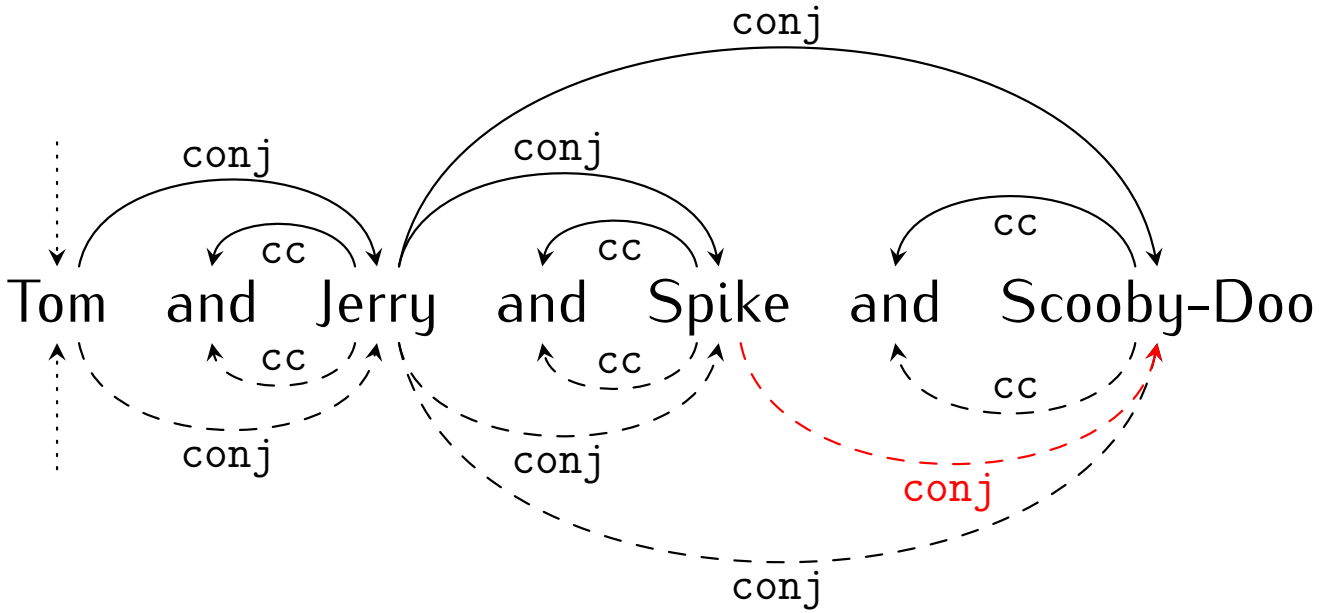
Pros:

- **different nestings distinguished in enhanced representations** (alone)
- another argument for getting rid of the problematic aspect of UD approach to distribution in coordinate structures (two outgoing `obl` dependencies – two dependents or one?)

Distribution in enhanced representations 3

Better solution: only distribute dependencies **from** coordinate structures:

- Tom and [[Jerry and Spike] and Scooby-Doo]:



Pros:

- **different nestings distinguished in enhanced representations** (alone)
- another argument for getting rid of the problematic aspect of UD approach to distribution in coordinate structures (two outgoing `obl` dependencies – two dependents or one?)

Solutions – summary

Solution	Pros	Cons
subtypes	• works at the level of basic trees	• indefinite number of labels • 'highly unnatural' (Mel'čuk)
co-heads	• conjuncts as co-heads • may be made to work at the enhanced level	• basic version requires inspecting both levels
distribution	• works at the level of enhanced graphs • relatively conservative • argument for getting rid of distribution **to** coordination	• (does not encode co-heads)

Solutions – summary

Solution	Pros	Cons
subtypes	• works at the level of basic trees	• indefinite number of labels • 'highly unnatural' (Mel'čuk)
co-heads	• conjuncts as co-heads • may be made to work at the enhanced level	• basic version requires inspecting both levels
distribution	• works at the level of enhanced graphs • relatively conservative • argument for getting rid of distribution **to** coordination	• (does not encode co-heads)

Thank you
for your attention!

Fun fact

Fun fact: how many nestings for n conjuncts?

Fun fact

Fun fact: how many nestings for n conjuncts?

conjuncts	2	3	4	5	...	10	...
nestings	1	3	11	45	...	103,049	...

Fun fact

Fun fact: how many nestings for n conjuncts?

conjuncts	2	3	4	5	...	10	...
nestings	1	3	11	45	...	103,049	...

- **little Schröder numbers** (Schröder–Hipparchus numbers, super-Catalan numbers)

Fun fact

Fun fact: how many nestings for n conjuncts?

conjuncts	2	3	4	5	...	10	...
nestings	1	3	11	45	...	103,049	...

- **little Schröder numbers** (Schröder–Hipparchus numbers, super-Catalan numbers)
- sequence A001003 in the On-line Encyclopedia of Integer Sequences

Fun fact

Fun fact: how many nestings for n conjuncts?

conjuncts	2	3	4	5	...	10	...
nestings	1	3	11	45	...	103,049	...

- **little Schröder numbers** (Schröder–Hipparchus numbers, super-Catalan numbers)
- sequence A001003 in the On-line Encyclopedia of Integer Sequences
- the value for 10 calculated already by Hipparchus of Nicaea, c. 190 – c. 120 BC

Fun fact

Fun fact: how many nestings for n conjuncts?

conjuncts	2	3	4	5	...	10	...
nestings	1	3	11	45	...	103,049	...

- **little Schröder numbers** (Schröder–Hipparchus numbers, super-Catalan numbers)
- sequence A001003 in the On-line Encyclopedia of Integer Sequences
- the value for 10 calculated already by Hipparchus of Nicaea, c. 190 – c. 120 BC
- see Stanley 1997 for the history of these numbers, and their other interpretations

Borsley, R. D. (2005). Against ConjP. *Lingua*, **115**(4), 461–482.

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A. (1985). *Generalized Phrase Structure Grammar*. Blackwell / Harvard University Press.

Gerdes, K. and Kahane, S. (2015). Non-constituent coordination and other coordinative constructions as dependency graphs. In E. Hajičová and J. Nivre, eds., *Proceedings of the Third International Conference on Dependency Linguistics (DepLing 2015)*, pp. 101–110.

Hudson, R. (1988). Coordination and grammatical relations. *Journal of Linguistics*, **24**(2), 303–342.

Kahane, S. (1997). Bubble trees and syntactic representations. In T. Becker and H.-U. Krieger, eds., *Proceedings of Mathematics of Language 5*, pp. 70–76.

Kahane, S. and Mazziotta, N. (2015). Syntactic polygraphs: A formalism extending both constituency and dependency. In M. Kuhlmann, M. Kanazawa, and G. M. Kobele, eds., *Proceedings of Mathematics of Language 14*, pp. 152–164.

Mel'čuk, I. (1988). *Dependency Syntax: Theory and Practice*. The SUNY Press.

Mel'čuk, I. (2009). Dependency in natural language. In A. Polguère and I. Mel'čuk, eds., *Dependency in Linguistic Description*, pp. 1–110. John Benjamins.

Mel'čuk, I. and Pertsov, N. (1987). *Surface Syntax of English. A Formal Model within the Meaning–Text Framework*. John Benjamins.

Popel, M., Mareček, D., Štěpánek, J., Zeman, D., and Žabokrtský, Z. (2013). Coordination structures in dependency treebanks. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 517–527.

Schuster, S., Lamm, M., and Manning, C. D. (2017). Gapping constructions in Universal Dependencies v2. In M.-C. de Marneffe, J. Nivre, and S. Schuster, eds., *Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017)*, pp. 123–132. Association for Computational Linguistics.

Stanley, R. P. (1997). Hipparchus, Plutarch, Schröder, and Hough. *The American Mathematical Monthly*, **104**(4), 344–350.

Tesnière, L. (1959). *Éléments de Syntaxe Structurale*. Klincksieck.