Universal Conceptual Cognitive Annotation (UCCA)

Abend and Rappoport,2013

Hanzhang Xia

University of Edinburgh

March 4, 2016

Outline

- Introduction
- UCCA Scheme
- Building a UCCA-Annotated Corpus
- Conclusion
- Evaluation

Introduction: Motivation

- Syntax: only indirectly reflect semantic distinctions

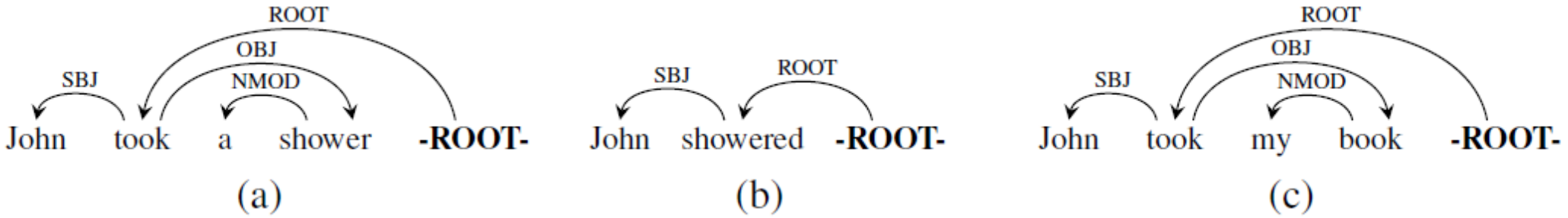

Figure 1. CoNLL-style dependency annotations.

Introduction: Motivation

- Syntax: only indirectly reflect semantic distinctions

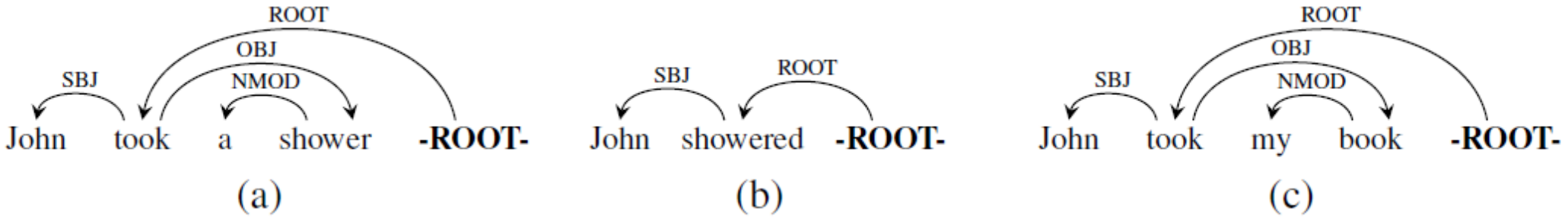

Figure 1. CoNLL-style dependency annotations.

- Applications such as Machine Translation, Question Answering

Introduction: UCCA

Universal Conceptual Cognitive Annotation (UCCA)

- *Universal*: catches a rich set of semantic distinctions
- *Conceptual*: contrasts with “syntactic”
- *Cognitive*: theory (Basic Linguistic Theory, Cognitive Linguistics)

Introduction: UCCA

Universal Conceptual Cognitive Annotation (UCCA)

- Built as a multi-layered structure
- This paper: focus on foundational layer (coarse-grained)

UCCA: Categories

Abb.	Category	Short Definition
	Scene Elements	
P	**Process**	The main relation of a Scene that evolves in time (usually an action or movement).
S	**State**	The main relation of a Scene that does not evolve in time.
A	**Participant**	A participant in a Scene in a broad sense (including locations, abstract entities and Scenes serving as arguments).
D	**Adverbial**	A secondary relation in a Scene (including temporal relations).
	Elements of Non-Scene Units	
C	**Center**	Necessary for the conceptualization of the parent unit.
E	**Elaborator**	A non-Scene relation which applies to a single Center.
N	**Connector**	A non-Scene relation which applies to two or more Centers, highlighting a common feature.
R	**Relator**	All other types of non-Scene relations. Two varieties: (1) Rs that relate a C to some super-ordinate relation, and (2) Rs that relate two Cs pertaining to different aspects of the parent unit.
	Inter-Scene Relations	
H	**Parallel Scene**	A Scene linked to other Scenes by regular linkage (e.g., temporal, logical, purposive).
L	**Linker**	A relation between two or more Hs (e.g., "when", "if", "in order to").
G	**Ground**	A relation between the speech event and the uttered Scene (e.g., "surprisingly", "in my opinion").
	Other	
F	**Function**	Does not introduce a relation or participant. Required by the structural pattern it appears in.

Table 1: The complete set of categories in UCCA's foundational layer.

UCCA: Categories

John(A) saw(P) the film(A) yesterday(D).
John(A) loves(S) banana(A).

Abb.	Category	Short Definition
	Scene Elements	
P	**Process**	The main relation of a Scene that evolves in time (usually an action or movement).
S	**State**	The main relation of a Scene that does not evolve in time.
A	**Participant**	A participant in a Scene in a broad sense (including locations, abstract entities and Scenes serving as arguments).
D	**Adverbial**	A secondary relation in a Scene (including temporal relations).
	Elements of Non-Scene Units	
C	**Center**	Necessary for the conceptualization of the parent unit.
E	**Elaborator**	A non-Scene relation which applies to a single Center.
N	**Connector**	A non-Scene relation which applies to two or more Centers, highlighting a common feature.
R	**Relator**	All other types of non-Scene relations. Two varieties: (1) Rs that relate a C to some super-ordinate relation, and (2) Rs that relate two Cs pertaining to different aspects of the parent unit.
	Inter-Scene Relations	
H	**Parallel Scene**	A Scene linked to other Scenes by regular linkage (e.g., temporal, logical, purposive).
L	**Linker**	A relation between two or more Hs (e.g., "when", "if", "in order to").
G	**Ground**	A relation between the speech event and the uttered Scene (e.g., "surprisingly", "in my opinion").
	Other	
F	**Function**	Does not introduce a relation or participant. Required by the structural pattern it appears in.

Table 1: The complete set of categories in UCCA's foundational layer.

UCCA: Categories

Abb.	Category	Short Definition
	Scene Elements	
P	**Process**	The main relation of a Scene that evolves in time (usually an action or movement).
S	**State**	The main relation of a Scene that does not evolve in time.
A	**Participant**	A participant in a Scene in a broad sense (including locations, abstract entities and Scenes serving as arguments).
D	**Adverbial**	A secondary relation in a Scene (including temporal relations).
	Elements of Non-Scene Units	
C	**Center**	Necessary for the conceptualization of the parent unit.
E	**Elaborator**	A non-Scene relation which applies to a single Center.
N	**Connector**	A non-Scene relation which applies to two or more Centers, highlighting a common feature.
R	**Relator**	All other types of non-Scene relations. Two varieties: (1) Rs that relate a C to some super-ordinate relation, and (2) Rs that relate two Cs pertaining to different aspects of the parent unit.
	Inter-Scene Relations	
H	**Parallel Scene**	A Scene linked to other Scenes by regular linkage (e.g., temporal, logical, purposive).
L	**Linker**	A relation between two or more Hs (e.g., "when", "if", "in order to").
G	**Ground**	A relation between the speech event and the uttered Scene (e.g., "surprisingly", "in my opinion").
	Other	
F	**Function**	Does not introduce a relation or participant. Required by the structural pattern it appears in.

Table 1: The complete set of categories in UCCA's foundational layer.

big(E) dogs(C)

and (N)

in, of (R)

It(F) is weird that he disappeared.

UCCA: Examples

- Directed acyclic graphs (DAGs)

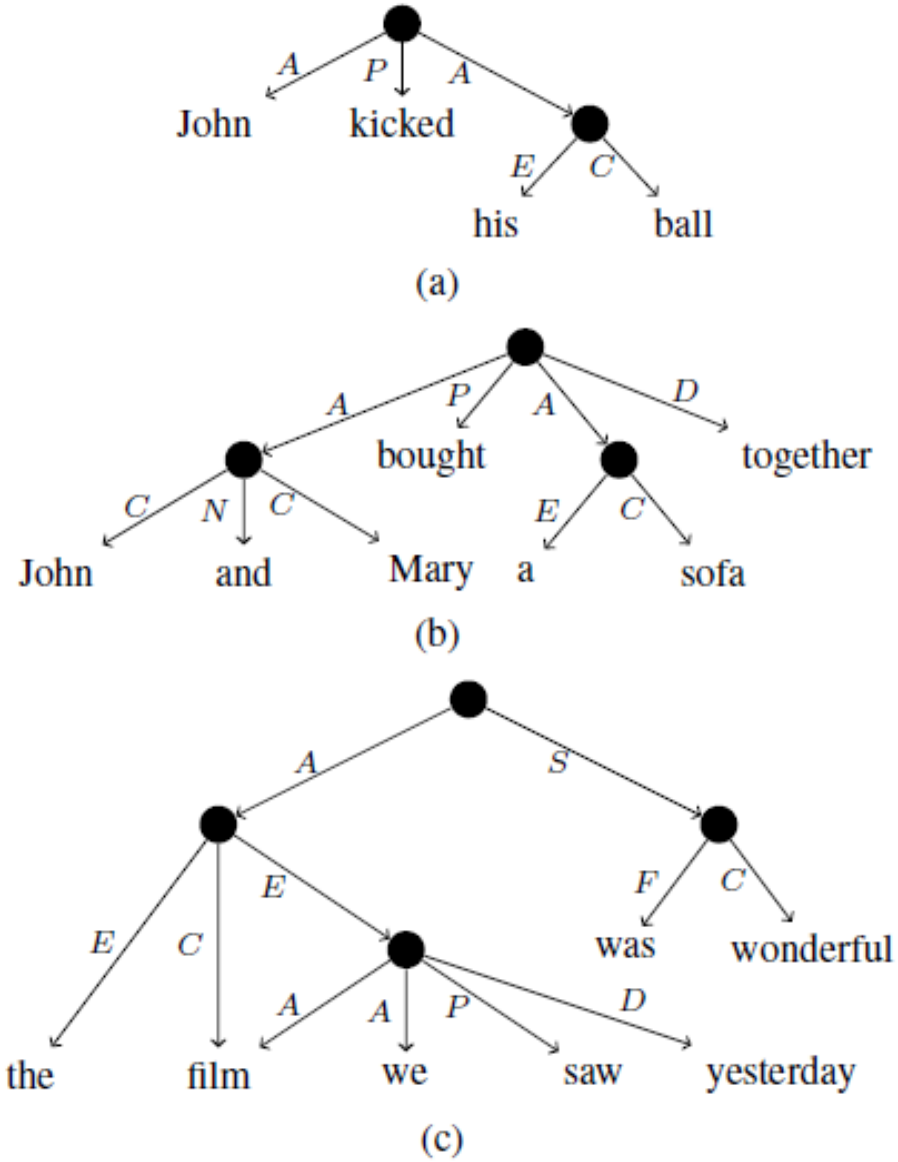

Figure 2: Examples of UCCA annotation graphs.

UCCA: Examples

- Directed acyclic graphs (DAGs):
- Terminals
- Units
- Edges

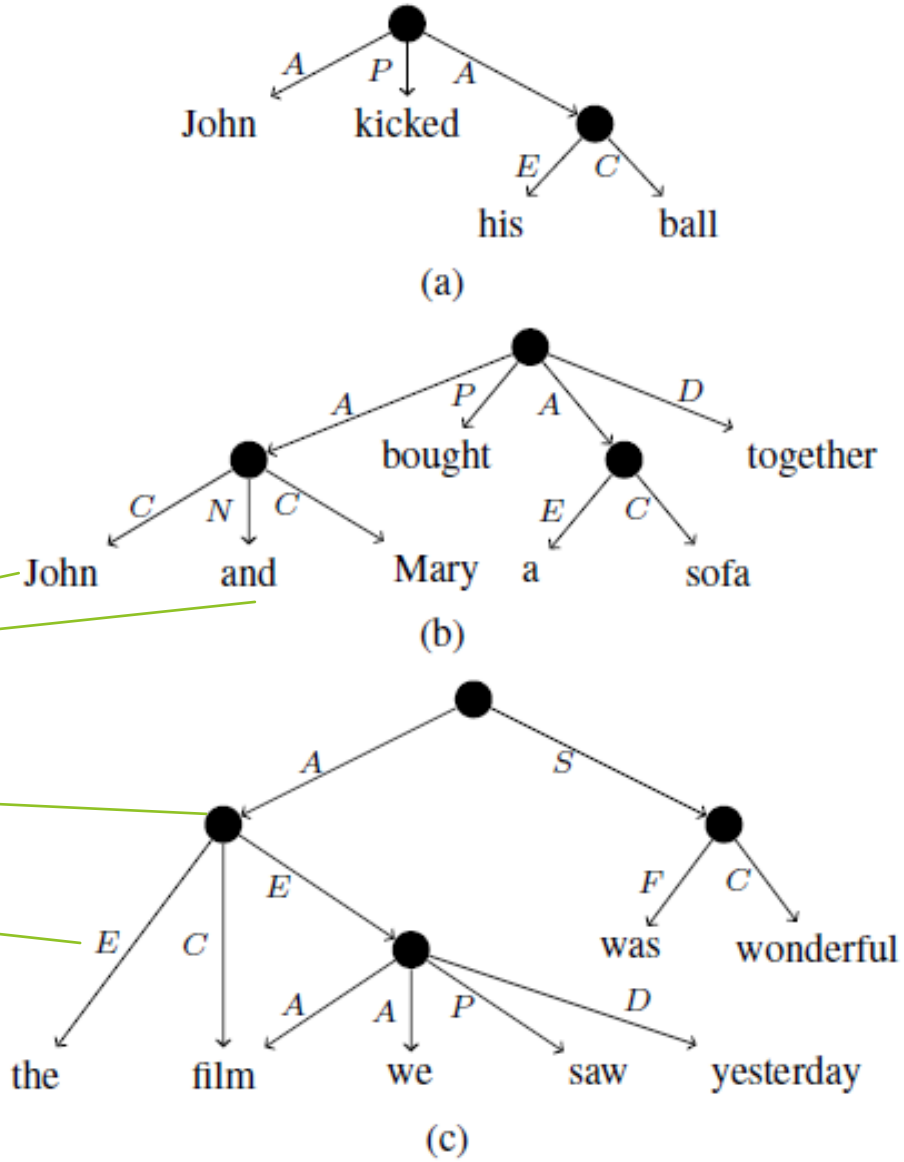

Figure 2: Examples of UCCA annotation graphs.

UCCA: Examples

- Directed acyclic graphs (DAGs):

- Units: either

(i) a terminal

(ii) several elements jointly viewed as a single entity

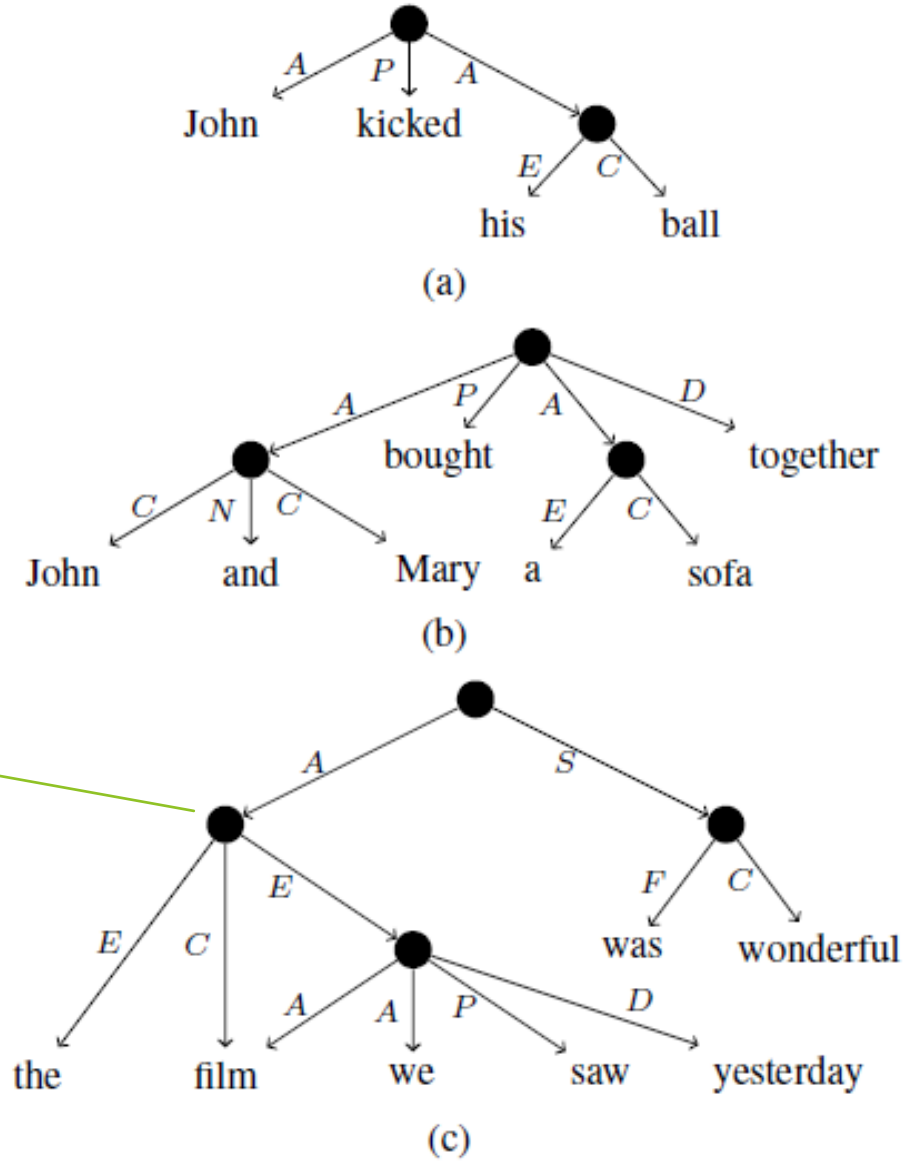

Figure 2: Examples of UCCA annotation graphs.

UCCA: Example Sentence

- *John encouraged the studio to accept his demands.*

- 1. John$_{A}$ encouraged$_{P}$ [the$_{E}$ stuidio$_{C}$]$_{A}$ [to$_{R}$] [accept his demands]$_{C}$]$_{A}$
- 2. [the studio]$_{A}$…accept$_{P}$ [his demands]$_{A}$
- 3. his$_{A}$ demands$_{P}$ **IMP**$_{A}$

UCCA: Complex Examples

- Units participating in multiple relations

e.g., *John asked Mary to join him*.

- Implicit units

e.g., *(For people), playing games is fun.*

- Inter-Scene relations

e.g., *John said [[he must leave]$_S$]$_A$].*

UCCA: Multi-layered Structure

- Additional layers added to refine relations

e.g.

- linkage: temporal, purposive, causal
- Co-reference layer: *John kicked his ball*.
- ……

A UCCA-Annotated Corpus

- Annotated text: English Wikipedia articles for celebrities
- 56890 tokens in 148 annotated passages

- Annotators: 4 annotators with different levels in linguistics
- Training annotators: 30-40 hours

A UCCA-Annotated Corpus

- Inter-annotator agreement

Passage#	1	2	3	4	5	6
ITA	67.3	74.1	71.2	73.5	77.8	81.1
vs. Gold	72.4	76.7	75.5	75.7	79.5	84.2

Table 2. average F-scores. ITA: comparing the annotations of the different annotators among themselves. vs.Gold: comparing them to a gold standard

A UCCA-Annotated Corpus

- Inter-annotator agreement

Passage#	1	2	3	4	5	6
ITA	67.3	74.1	71.2	73.5	77.8	81.1
vs. Gold	72.4	76.7	75.5	75.7	79.5	84.2

Table 2. average F-scores. ITA: comparing the annotations of the different annotators among themselves. vs.Gold: comparing them to a gold standard

- Do not need proficient annotators!

UCCA: Disagreement examples

- Elaborators? Centers?

[truck]*E* [company]*C*

[The Fox drama]*E?C?* [Glory days]*C?E?*

- Scenes? Non-scenes?

[John's [portrayal]*P?C?* of the character]*A* has been described as…

UCCA: benefits

- Relative insensitivity to syntactic forms
- Can be applied cross domains and languages
- Multi-layer: more fine-grained representations
- No proficient annotators needed

Conclusion

- UCCA: a multi-layered framework for semantic representation
- The Foundational Layer
- Annotation
- Advantages (insensitivity to syntactic variation, across domains and languages, no proficient annotators, can be more fine-grained)

Evaluation

- A good try of meaning representation by using cognitive categories

- Multi-layers: how many?
- Across languages: new annotations needed
- Annotators?

Reference

- Abend, O., & Rappoport, A. (2013). Universal Conceptual Cognitive Annotation (UCCA). In *ACL (1)* (pp. 228-238).