Symbolic AI

Andre Freitas

Photo by Vasilyev Alexandr

Acknowledgements

- Based on the slides of:
  - J. Sowa, Existential & Conceptual Graphs
  - Cristiano Broccias, Cognitive Lexical Semantics
  - Taboada, Introduction to Rhetorical Structure Theory
  - CMSC 473/673 (UMBC), Semantic Roles and Frames

This Lecture

- Representing complex statements.
  - We will focus on events
- Representing discourse elements.

Representing Events

- How do we represent time and temporal relationships between events?
  - It seems only yesterday that Martha Stewart was in prison but now she has a popular TV show. There is no justice.

- Where do we get temporal information?
  - Verb tense
  - Temporal expressions
  - Sequence of presentation

Representing Events

- Temporal, tense logic.
  - I arrived in New York.
  - I am arriving in New York.
  - I will arrive in New York.

$$\exists w\, ISA(w, Arriving)$$
$$\wedge Arriver(w, Speaker) \wedge Destination(w, NewYork)$$

- The temporal information provided by the tense of the verbs can be exploited by predicating additional information about the **event variable *w***.

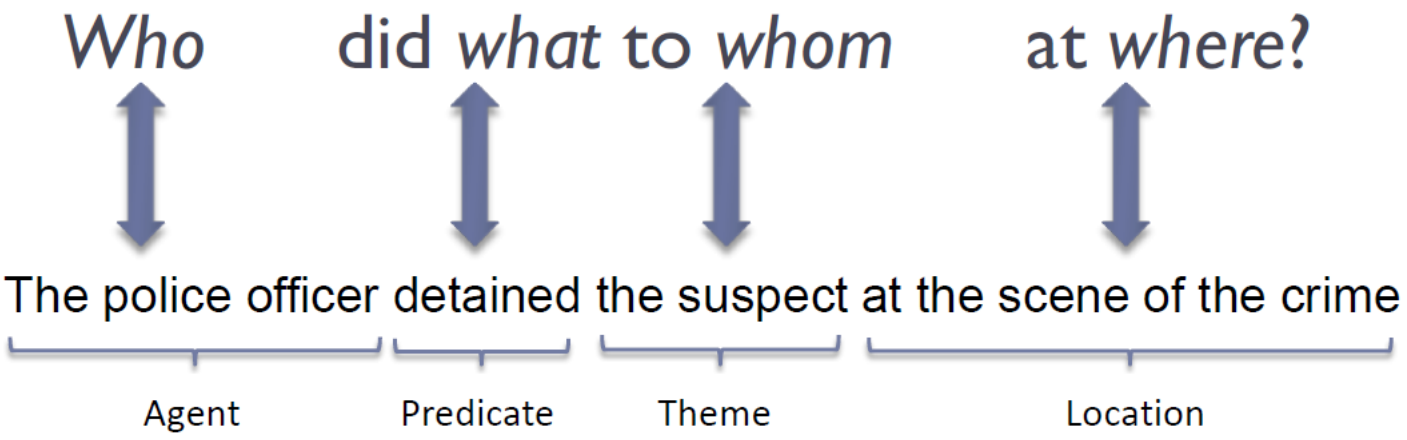
Representing Events

We can add temporal variables representing the interval corresponding to the event, the end point of the event, and temporal predicates relating this end point to the current time as indicated by the tense of the verb.

$$\exists i, e, w, t\ ISA(w, Arriving)$$
$$\wedge Arriver(w, Speaker) \wedge Destination(w, NewYork)$$
$$IntervalOf(w, i) \wedge EndPoint(i, e) \wedge Precedes(e, Now)$$

$$\exists i, e, w, t\ ISA(w, Arriving)$$
$$\wedge Arriver(w, Speaker) \wedge Destination(w, NewYork)$$
$$IntervalOf(w, i) \wedge MemberOf(i, Now)$$

Semantic Roles: Frame!



Thematic Roles

Sasha broke the window

Pat opened the door

Subjects of break and open:
Breaker and **Opener**

Specific to each event

Thematic Roles

Sasha broke the window

Pat opened the door

Subjects of break and open:
Breaker and **Opener**

Specific to each event

Breaker and **Opener** have something in common!

- Volitional actors
- Often animate
- Direct causal responsibility for their events

Thematic roles are a way to capture this semantic commonality between *Breakers* and *Eaters*.

Thematic Roles

Sasha broke the window

Pat opened the door

Subjects of break and open: **Breaker** and **Opener**

Specific to each event

Breaker and **Opener** have something in common!

- Volitional actors
- Often animate
- Direct causal responsibility for their events

Thematic roles are a way to capture this semantic commonality between *Breakers* and *Eaters*.

They are both AGENTS.

The *BrokenThing* and *OpenedThing*, are THEMES.

- prototypically inanimate objects affected in some way by the action

Modern formulation from
Fillmore (1966,1968), Gruber (1965)

Fillmore influenced by Lucien Tesnière's (1959) *Éléments de Syntaxe Structurale,* the book that introduced dependency grammar

Typical Thematic Roles

Thematic Role	Definition	Example
AGENT	The volitional causer of an event	*The waiter* spilled the soup.
EXPERIENCER	The experiencer of an event	*John* has a headache.
FORCE	The non-volitional causer of the event	*The wind* blows debris from the mall into our yards.
THEME	The participant most directly affected by an event	Only after Benjamin Franklin broke *the ice*...
RESULT	The end product of an event	The city built a *regulation-size baseball diamond*...
CONTENT	The proposition or content of a propositional event	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	An instrument used in an event	He poached catfish, stunning them *with a shocking device*...
BENEFICIARY	The beneficiary of an event	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	The origin of the object of a transfer event	I flew in *from Boston*.
GOAL	The destination of an object of a transfer event	I drove *to Portland*.

Verb Alternations (Diathesis Alternations)

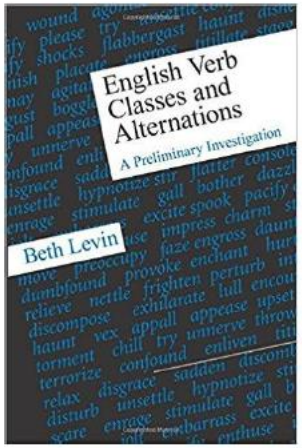
Doris gave the book to Cary.
AGENT THEME GOAL

Doris gave Cary the book.
AGENT GOAL THEME

Break: AGENT, INSTRUMENT, or THEME as subject

Give: THEME and GOAL in either order

Levin (1993): 47 semantic classes ("**Levin classes**") for 3100 English verbs and alternations. In online resource **VerbNet**.



Alternative to Thematic Roles

PropBank

1. **Fewer roles**: generalized semantic roles, defined as prototypes (Dowty 1991)

   PROTO-AGENT

   PROTO-PATIENT

FrameNet

2. **More roles**: Define roles specific to a group of predicates

PropBank Frame Files

agree.01

Arg0: Agreer
Arg1: Proposition
Arg2: Other entity agreeing

Ex1: [$_{\text{Arg0}}$ The group] *agreed* [$_{\text{Arg1}}$ it wouldn't make an offer].
Ex2: [$_{\text{ArgM-TMP}}$ Usually] [$_{\text{Arg0}}$ John] *agrees* [$_{\text{Arg2}}$ with Mary] [$_{\text{Arg1}}$ on everything].

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106

View Commonalities Across Sentences

increase.01 "go up incrementally"

Arg0: causer of increase
Arg1: thing increasing
Arg2: amount increased by, EXT, or MNR
Arg3: start point
Arg4: end point

[$_{\text{Arg0}}$ Big Fruit Co.] increased [$_{\text{Arg1}}$ the price of bananas].
[$_{\text{Arg1}}$ The price of bananas] was increased again [$_{\text{Arg0}}$ by Big Fruit Co.]
[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

Frege's Begriffsschrift for the Same Sentence

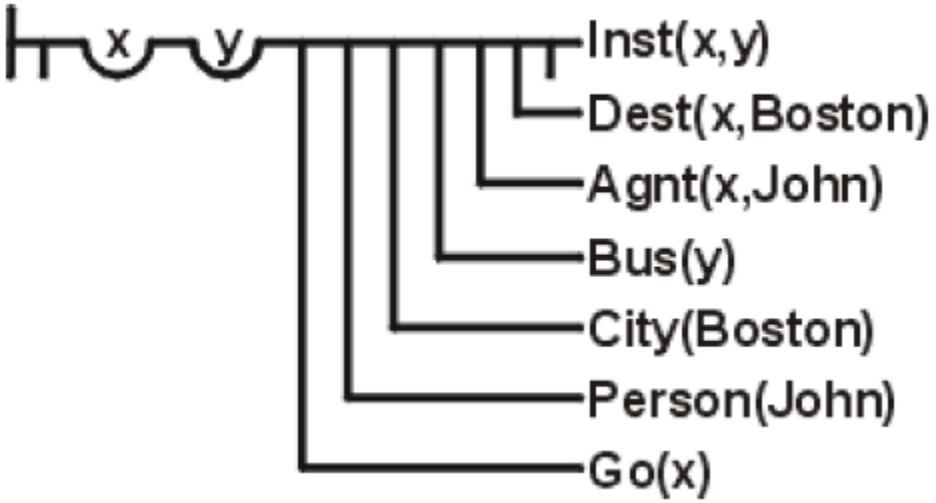
- Peirce's algebraic notation (1885):

$$\Sigma_x\Sigma_y\,(\mathrm{Go}(x) \bullet \mathrm{Person}(\mathrm{John}) \bullet \mathrm{City}(\mathrm{Boston}) \bullet \mathrm{Bus}(y) \bullet \mathrm{Agnt}(x,\mathrm{John}) \bullet \mathrm{Dest}(x,\mathrm{Boston}) \bullet \mathrm{Inst}(x,y))$$

- With Peano's choice of symbols:

$$(\exists x)(\exists y)(\mathrm{Go}(x) \wedge \mathrm{Person}(\mathrm{John}) \wedge \mathrm{City}(\mathrm{Boston}) \wedge \mathrm{Bus}(y) \wedge \mathrm{Agnt}(x,\mathrm{John}) \wedge \mathrm{Dest}(x,\mathrm{Boston}) \wedge \mathrm{Inst}(x,y))$$

Frege's Begriffsschrift for the Same Sentence



- Translation to Peirce-Peano notation:

$$\sim(\forall x)(\forall y)(\mathrm{Go}(x) \supset (\mathrm{Person}(\mathrm{John}) \supset (\mathrm{City}(\mathrm{Boston}) \supset (\mathrm{Bus}(y) \supset (\mathrm{Agnt}(x,\mathrm{John}) \supset (\mathrm{Dest}(x,\mathrm{Boston}) \supset \sim \mathrm{Inst}(x,y)))))))$$

Frege's Begriffsschrift for the Same Sentence

- Translation to Peirce-Peano notation:

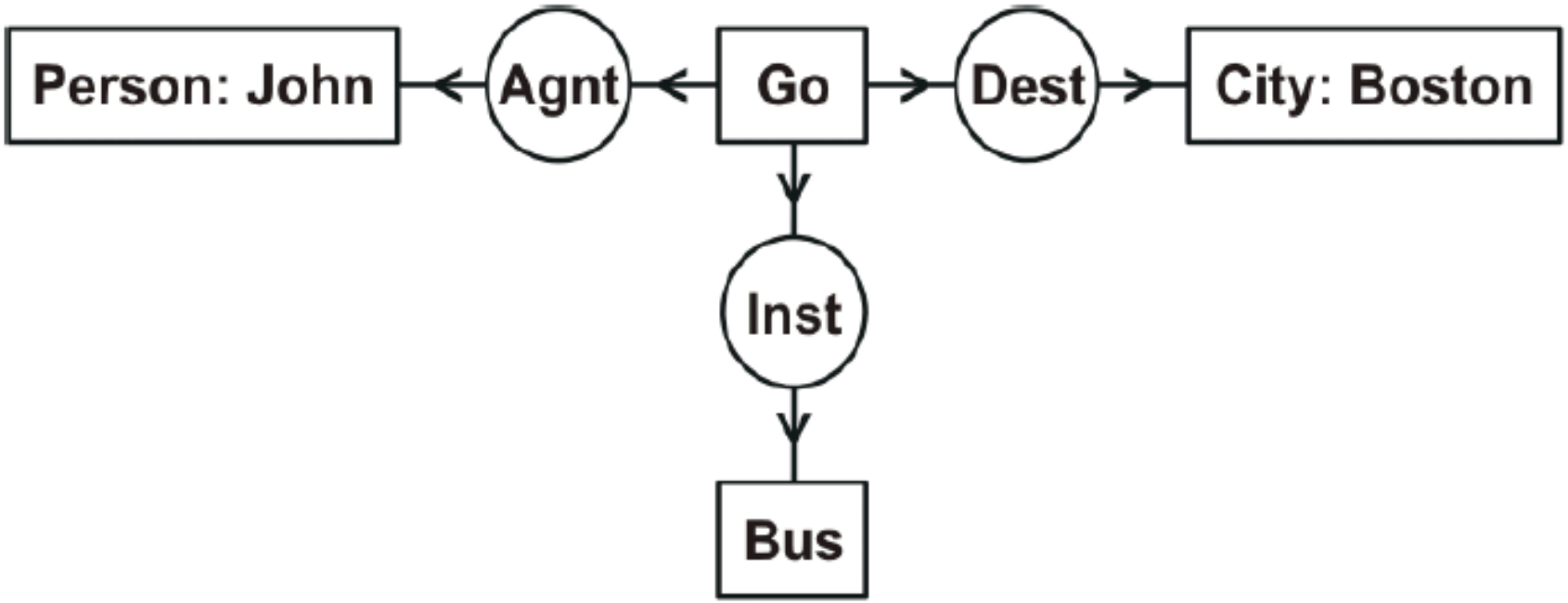
$$\sim(\forall x)(\forall y)(\mathrm{Go}(x) \supset (\mathrm{Person}(\mathrm{John}) \supset (\mathrm{City}(\mathrm{Boston}) \supset (\mathrm{Bus}(y) \supset (\mathrm{Agnt}(x,\mathrm{John}) \supset (\mathrm{Dest}(x,\mathrm{Boston}) \supset \sim\mathrm{Inst}(x,y)))))))$$

- Equivalent in English:

It is false that for every x and y, if x is an instance of going then if John is a person then if Boston is a city then if y is a bus then if the agent of x is John then if the destination of x is Boston then the instrument of x is not y.

J. Sowa

Sentence: *John is going to Boston by bus*

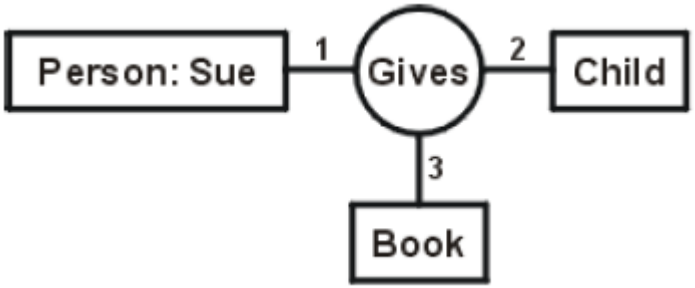


Conceptual Graphs

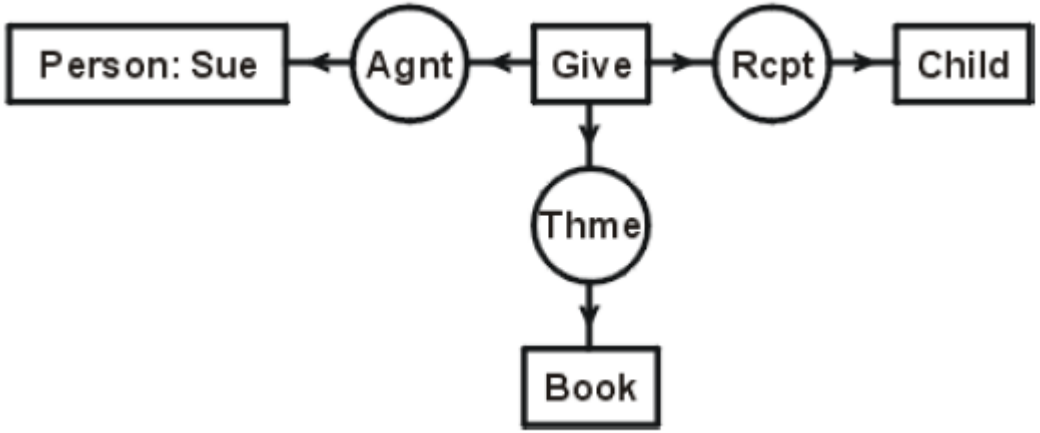
J. Sowa

Existential Graphs

- Existential Graphs vs Conceptual Graphs

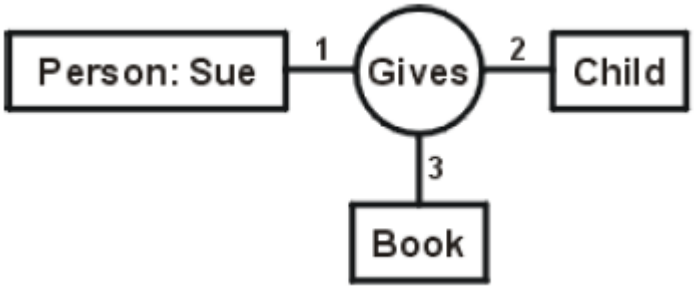


- The concept type **Give** is, in Peirce's terminology, a *hypostatic abstraction* of the relation type **Gives**.
- The idea of representing a verb by an entity that can be related by quantified variables is what Davidson called *event semantics*.

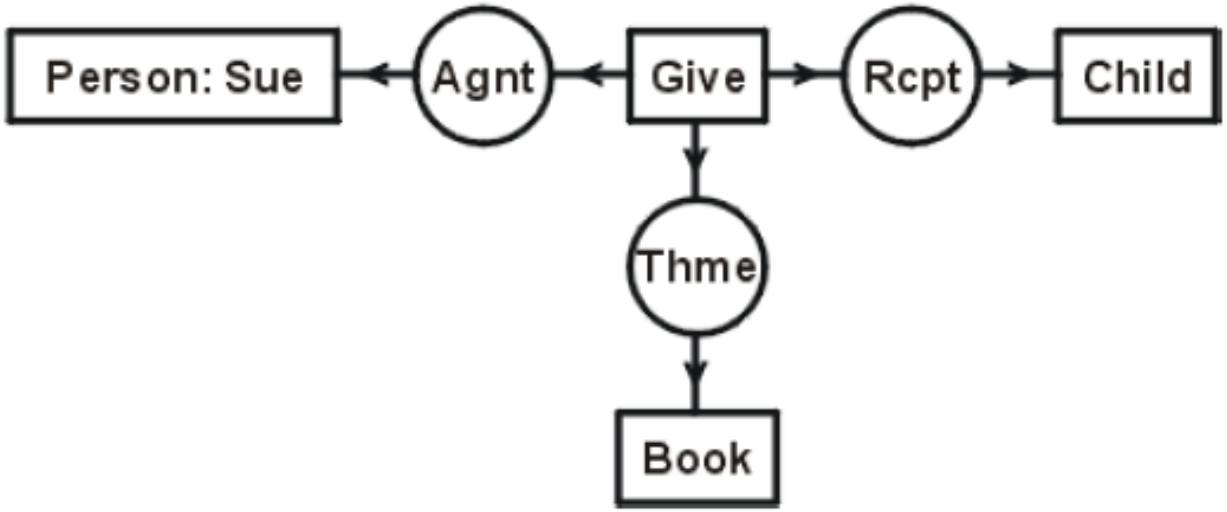


J. Sowa

- The equivalent operation can be performed in the algebraic notation, but its effect on the structure is harder to see and to express in a systematic generalization.



$$(\exists x)(\exists y)(\text{Person}(\text{Sue}) \wedge \text{Child}(x) \wedge \text{Book}(y) \wedge \text{Gives}(\text{Sue},x,y)).$$



- For the CG above, the triadic connection is represented by five occurrences of the variable z, three of which correspond to the three arcs attached to the concept **[Give].**

- The conceptual relations **(Rcpt)** for recipient and **(Thme)** for theme are translated to dyadic relations in predicate calculus:

$$(\exists x)(\exists y)(\exists z)(\text{Person}(\text{Sue}) \wedge \text{Child}(x) \wedge \text{Book}(y) \wedge \text{Give}(z) \wedge \text{Agnt}(z,\text{Sue}) \wedge \text{Rcpt}(z,x) \wedge \text{Thme}(z,y))$$

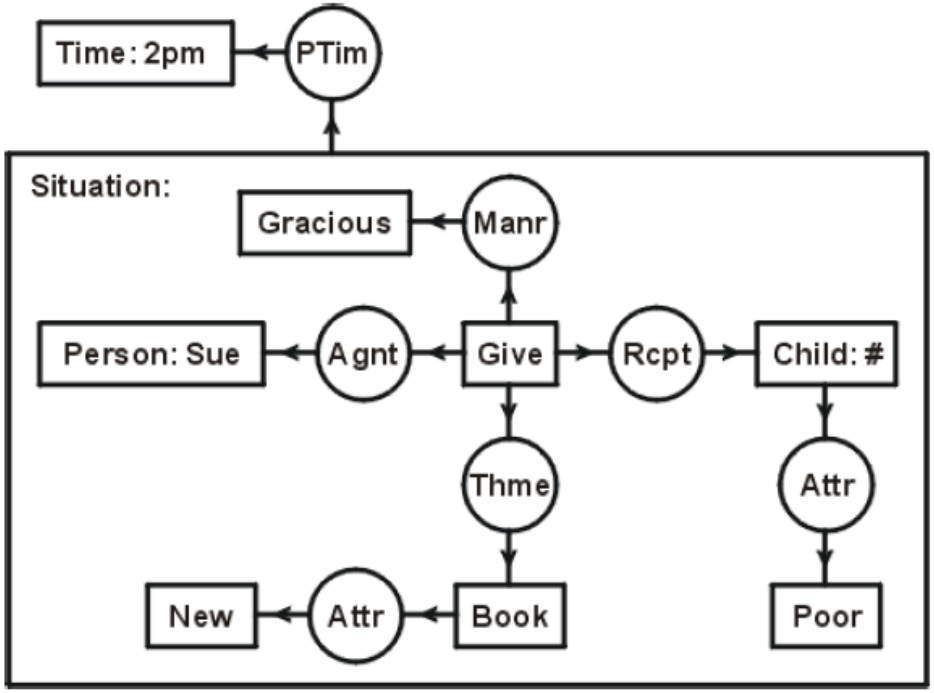
EG & CG vs Algebraic Notation

- As this example illustrates, the graph notation directly shows the *topology* of the logic, which is determined by the connectivity of the nodes and the cycles in the graph.

- That same topology is present in the algebraic formulas, but it is obscured by the notation for variables and quantifiers.

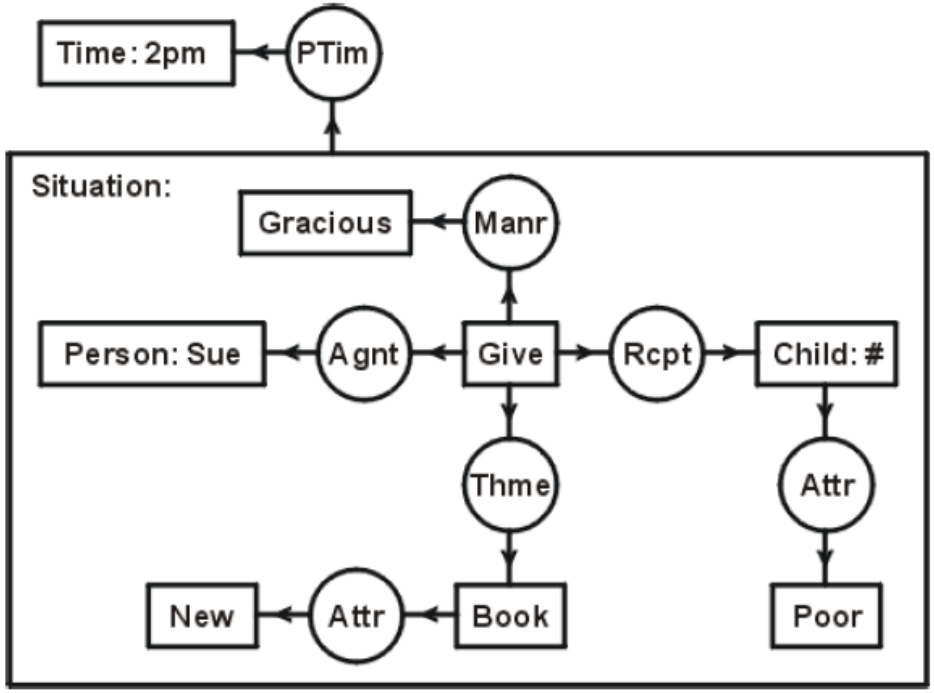
- By showing the connections directly, the graph notation in either CG or EG form enables efficient graph operations that are difficult or impossible to apply to the formulas without first converting them to an equivalent graph.

Sentence: *At 2 pm, Sue graciously gave the poor child a new book.*



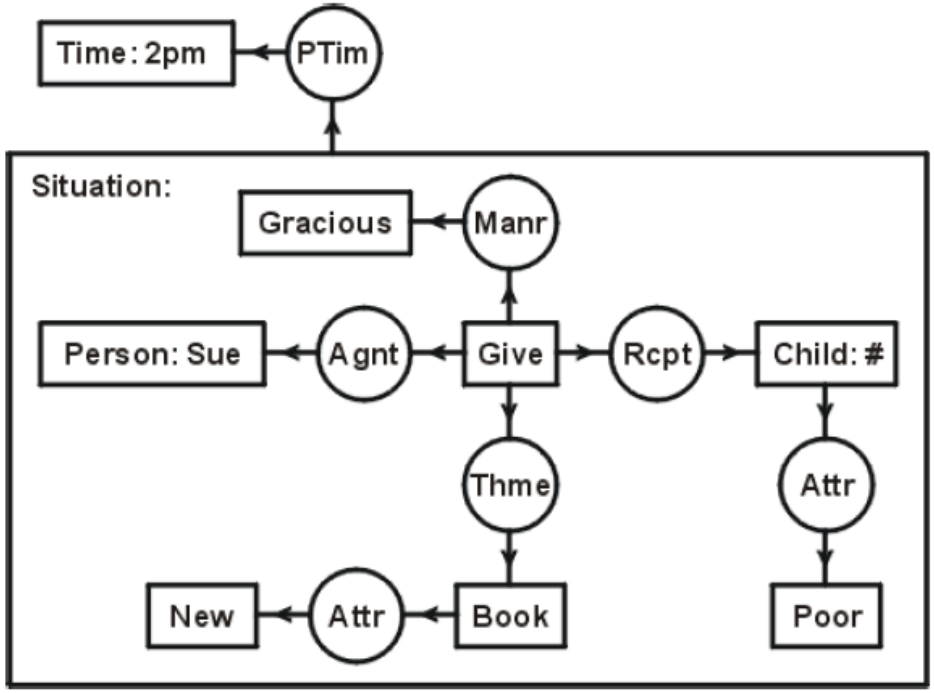
$(\exists s)(\text{Situation}(s) \wedge \text{Time}(2pm) \wedge \text{PTim}(s,2pm) \wedge \text{dscr}(s,$
$(\exists y)(\exists z)(\exists u)(\exists v)(\exists w)(\text{Person}(\text{Sue}) \wedge \text{Child}(\text{Bob}) \wedge$
$\text{Book}(y) \wedge \text{Give}(z) \wedge \text{Gracious}(u) \wedge \text{Poor}(v) \wedge$
$\text{New}(w) \wedge \text{Manr}(z,u) \wedge \text{Attr}(\text{Bob},v) \wedge \text{Attr}(y,w) \wedge$
$\text{Agnt}(z,\text{Sue}) \wedge \text{Rcpt}(z,\text{Bob}) \wedge \text{Thme}(z,y)))).$

J. Sowa



J. Sowa

- The concept of type Situation with a nested CG represents a situation described by that CG.
- In the algebraic formula, the relation dscr(s,p) is used to state that a situation s is described by a proposition p.
- The relation **(Ptim)** shows the point in time of that situation.
- The relations **(Manr)** and **(Attr)** represent the manner and attribute relations that are linked to the hypostatic abstractions **[Gracious]** and **[Poor]**, which were derived from an adverb and an adjective in the original sentence.
- Those concepts represent instances of graciousness and poverty, and the graphs allow additional connections to those nodes to represent phrases such as *very graciously* or *poor as a church mouse*.

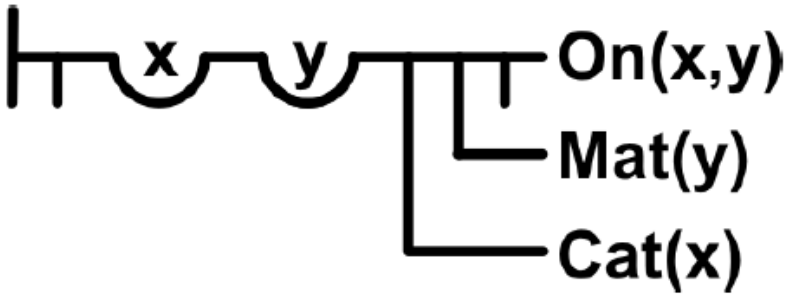


J. Sowa

- The symbol **#** in the concept **[Child: #]** represents the indexical effect of the phrase *the child*.
- Before the CG can be translated to other versions of logic, the indexical must be resolved to some individual in the context, either in the discourse or in the surrounding environment.
- In the algebraic formula, the symbol **#** is replaced by the name Bob.
- Hans Kamp developed discourse representation theory as a method of resolving such references. It turns out that the notation Kamp developed has context boxes that are isomorphic to the ovals of Peirce's existential graphs. By following Peirce's structures, the CG boxes turned out to be nested in the same ways as Kamp's.

How to say "A cat is on a mat."

Gottlob Frege (1879):



Charles Sanders Peirce (1885): $\Sigma_x \Sigma_y \mathrm{Cat}_x \bullet \mathrm{Mat}_y \bullet \mathrm{On}_{x,y}$

Giuseppe Peano (1895): $\exists x \exists y\ \mathrm{Cat}(x) \wedge \mathrm{Mat}(y) \wedge \mathrm{On}(x,y)$

Charles Sanders Peirce (1897): Cat — On — Mat

J. Sowa

First-Order Logic

J. Sowa

- Shaded ovals are sufficient to express full FOL:

Existence: —

Negation: 

Relations: **Cat- -On- -Under- -With- -Mat**

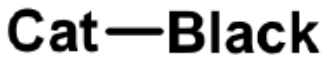
***A cat is on a mat*:** **Cat—On—Mat**

***Something is under a mat*:** **—Under—Mat**

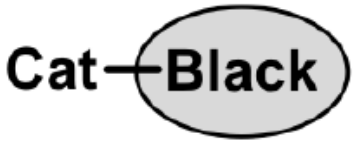
***Some cat is not on a mat*:** **Cat—(On—Mat)**

***Some cat is on something that is not a mat*:** **Cat—On—(Mat)**

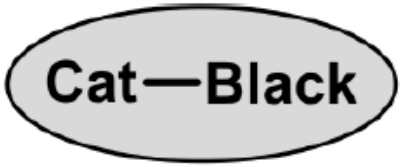
The Scope of Quantifiers



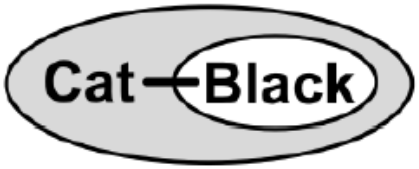
Some cat is black.



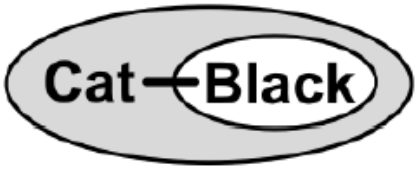
Some cat is not black.



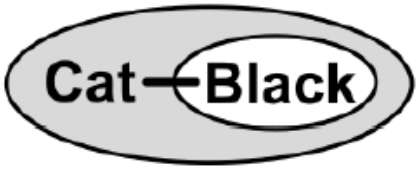
No cat is black.



It is false that some cat is not black.

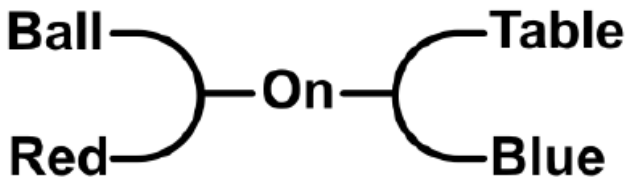


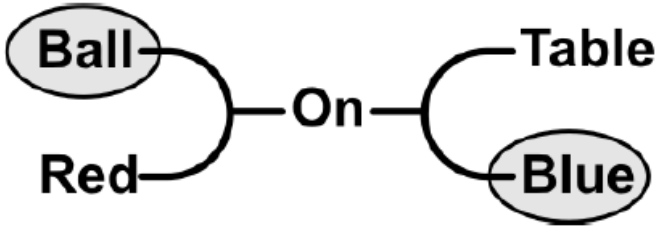
If there is a cat, then it is black.



Every cat is black.

Translating EGs to and from English





Left graph:

A red ball is on a blue table.

Some ball that is red is on some table that is blue.

Right graph:

Something red that is not a ball is on a table that is not blue.

A red non-ball is on a non-blue table.

On some non-blue table, there is something red that is not a ball.

J. Sowa

Existential Graph Interchange Format

A subset of the Conceptual Graph Interchange Format (CGIF):

Existence: — **[*x]**

Negation:  **~[ ]**

Relations: **(Cat ?x) (On ?x ?y) (Under ?x ?y) (Mat ?y)**

***A cat is on a mat*:** **[*x] [*y] (Cat ?x) (On ?x ?y) (Mat ?y)**

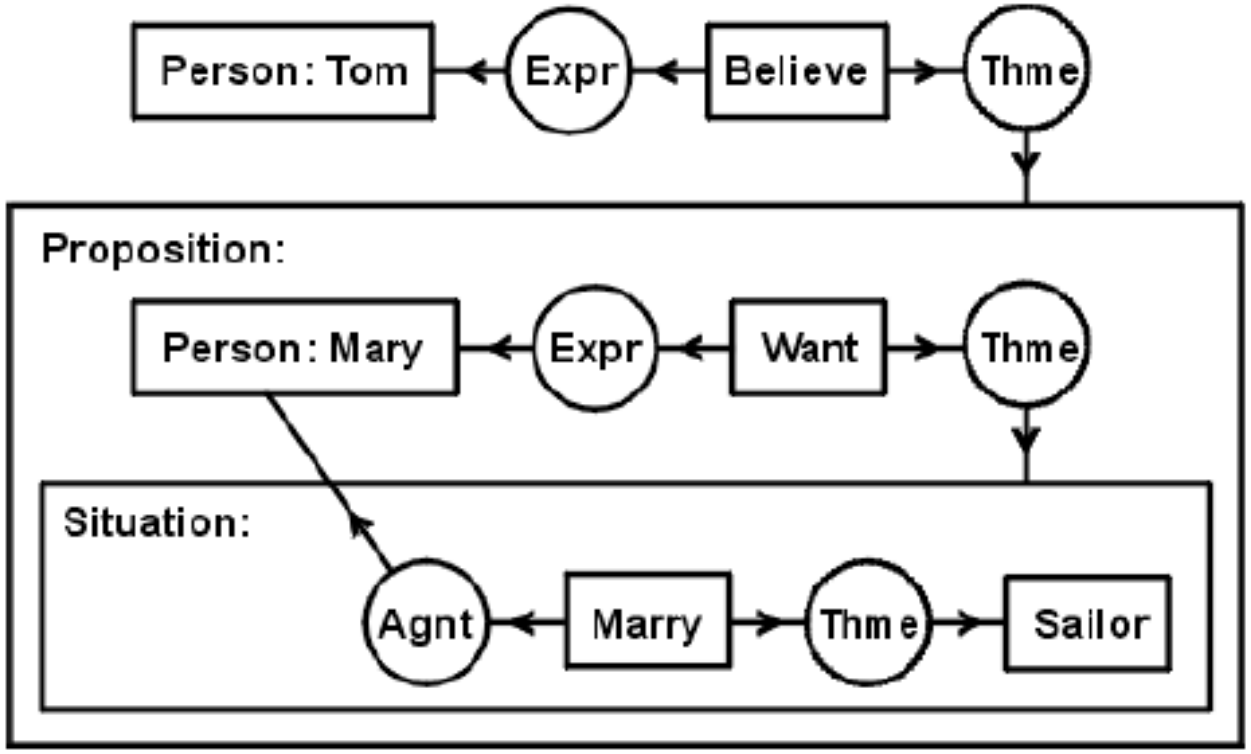
***Something is under a mat*:** **[*x] [*y] (Under ?x ?y) (Mat ?y)**

***Some cat is not on a mat*:** **[*x] (Cat ?x) ~[[*y] (On ?x ?y) (Mat ?y)]**

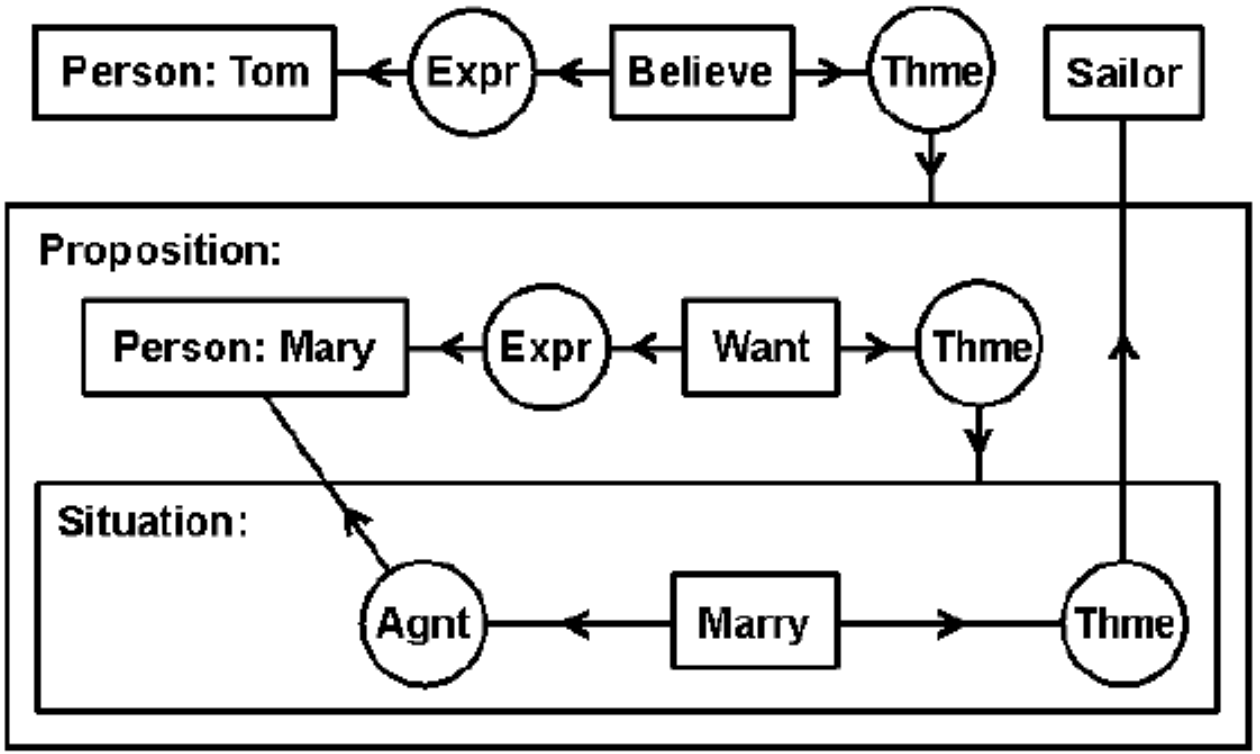
***Some cat is on something that is not a mat*:**
[*x] [*y] (Cat ?x) (On ?x ?y) ~[(Mat ?y)]

J. Sowa

Example: *Tom believes Mary wants to marry a sailor.*



Example: *There is a sailor that Tom believes Mary wants to marry.*



J. Sowa

Interpreting Discourse

Interpreting Discourse

- Discourse is a sequence of sentences.

- When we look at discourse, interesting **challenges arise**:
  - Interpreting co-references/anaphoras (pronominal resolution).
  - Representing discourse relations between propositions.

Discourse Representation Theory (DRS)

A woman walks. **She** smokes.

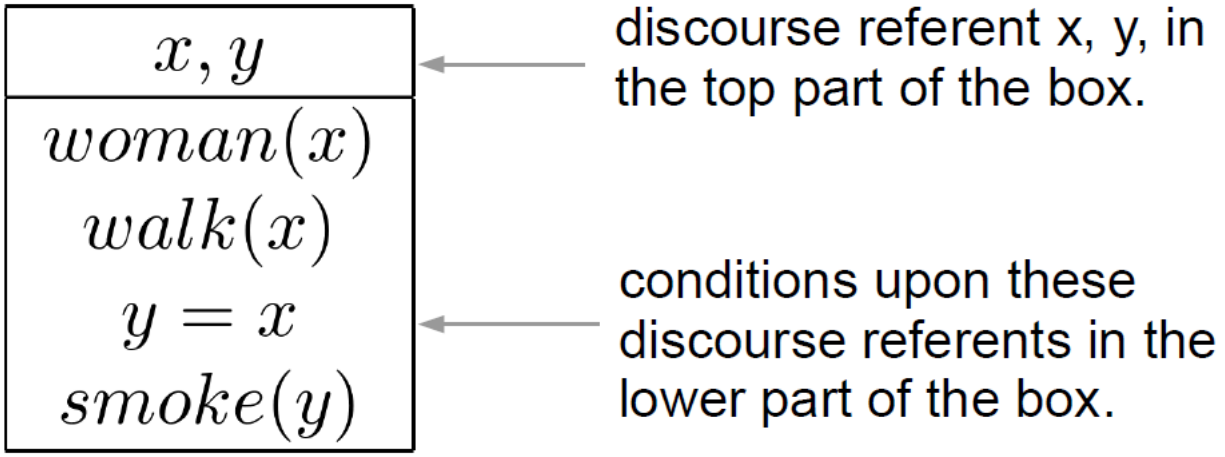
$\exists x(woman(x) \wedge walk(x))$ $smoke(x)$

$\exists x(woman(x) \wedge walk(x) \wedge smoke(x))$

Need to Expand the scope of the existential quantifier.

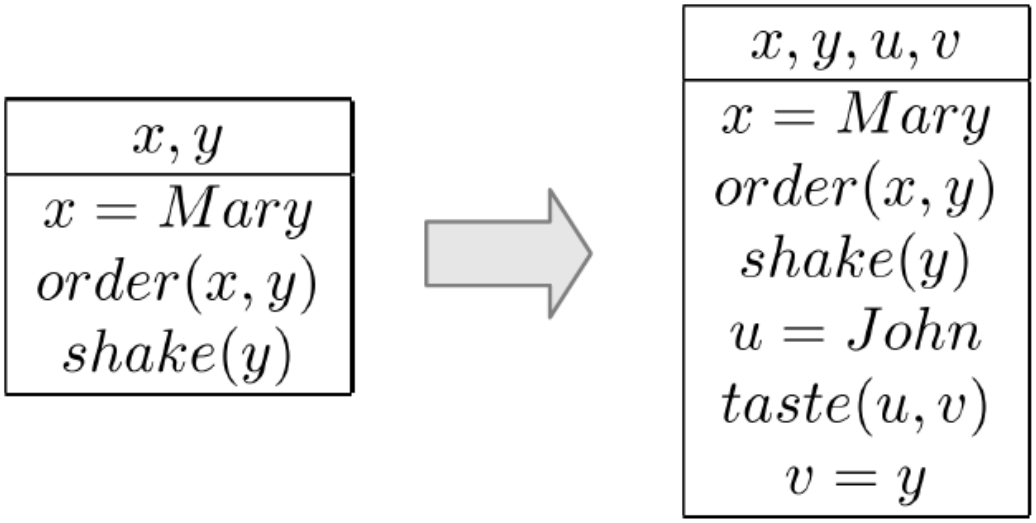
A woman walks. **She** smokes.

- **A woman** walks. **She** smokes.



Discourse Structure and Accessibility

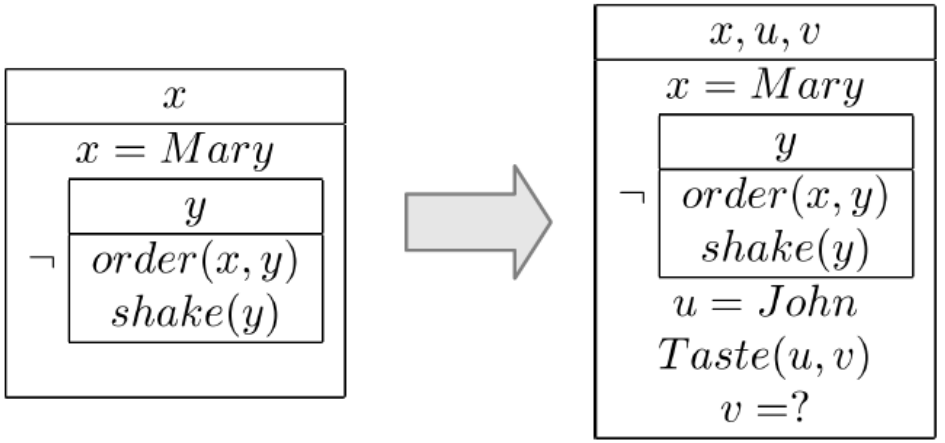
- Mary ordered a milk shake, John tasted **it**.



- The discourse referent **y is accessible** for discourse referent **v**.
- An anaphoric link between **it** and **milk shake** is allowed.

Discourse Structure and Accessibility

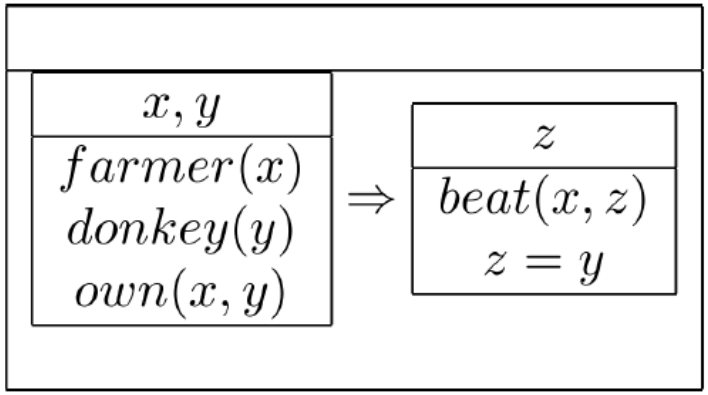
- Mary **did not** ordered a milk shake. John tasted **it**.



- When we introduced **negation**, an anaphoric link is blocked.
- Hence, **y is not accessible** for **v**

Discourse Structure and Accessibility

- Every farmer that owns **a donkey** beats **it.**



- Following the definition of accessibility, the discourse referent **y** introduced by a donkey is available as antecedent.

- A link is established by the DRS-condition **z = y.**

Rhetorical Structure Theory (RST)

Principles

- Coherent texts consist of minimal units, which are linked to each other, recursively, through rhetorical relations
  - Rhetorical relations also known, in other theories, as coherence or discourse relations

- Coherent texts do not show gaps or non-sequiturs
  - Therefore, there must be some relation holding among the different parts of the text

Taboada

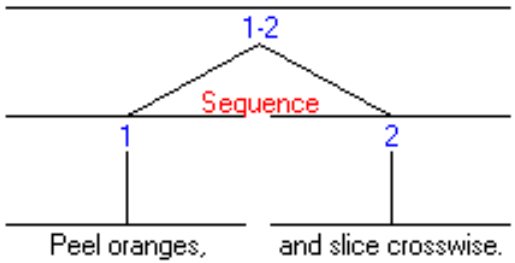
Components

- Units of discourse
  - Texts can be segmented into minimal units, or spans.
- Nuclearity
  - Some spans are more central to the text's purpose (nuclei), whereas others are secondary (satellites).
  - Based on hypotactic and paratactic relations in language.
- Relations among spans
  - Spans are joined into discourse relations.
- Hierarchy/recursion
  - Spans that are in a discourse relation may enter into new relations.

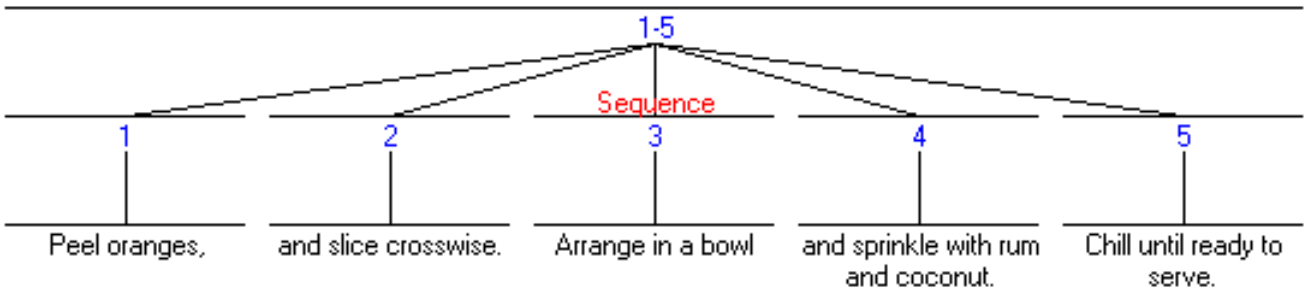
Taboada

Paratactic (coordinate)

- At the sub-sentential level (traditional coordinated clauses)
  - Peel oranges, and slice crosswise.

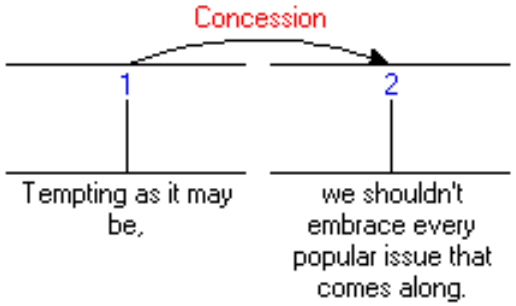


- But also across sentences
  - 1. Peel oranges, 2. and slice crosswise. 3. Arrange in a bowl 4. and sprinkle with rum and coconut. 5. Chill until ready to serve.

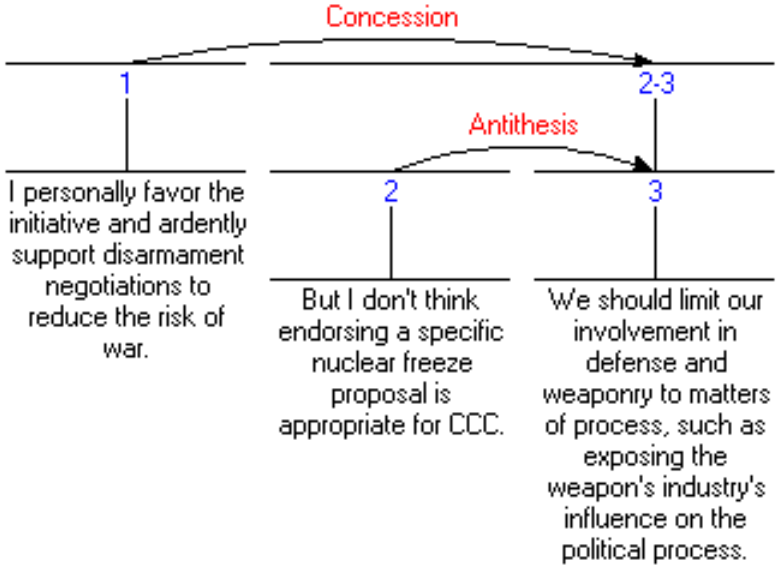


Hypotactic (subordinate)

- Sub-sentential Concession relation



- Concession across sentences
  - Nucleus (spans 2-3) made up of two spans in an Antithesis relation



Relation Types

- Relations are of different types:
  - Subject matter: they relate the content of the text spans
    - Cause, Purpose, Condition, Summary
  - Presentational: more rhetorical in nature. They are meant to achieve some effect on the reader
    - Motivation, Antithesis, Background, Evidence

Other possible classifications

- Relations that hold outside the text
  - Condition, Cause, Result

  vs. those that are only internal to the text
  - Summary, Elaboration

- Relations frequently marked by a discourse marker
  - Concession (*although, however*); Condition (*if, in case*)

  vs. relations that are rarely, or never, marked
  - Background, Restatement, Interpretation

- Preferred order of spans: nucleus before satellite
  - Elaboration – usually first the nucleus (material being elaborated on) and then satellite (extra information)

  vs. satellite-nucleus
  - Concession – usually the satellite (the *although*-type clause or span) before the nucleus

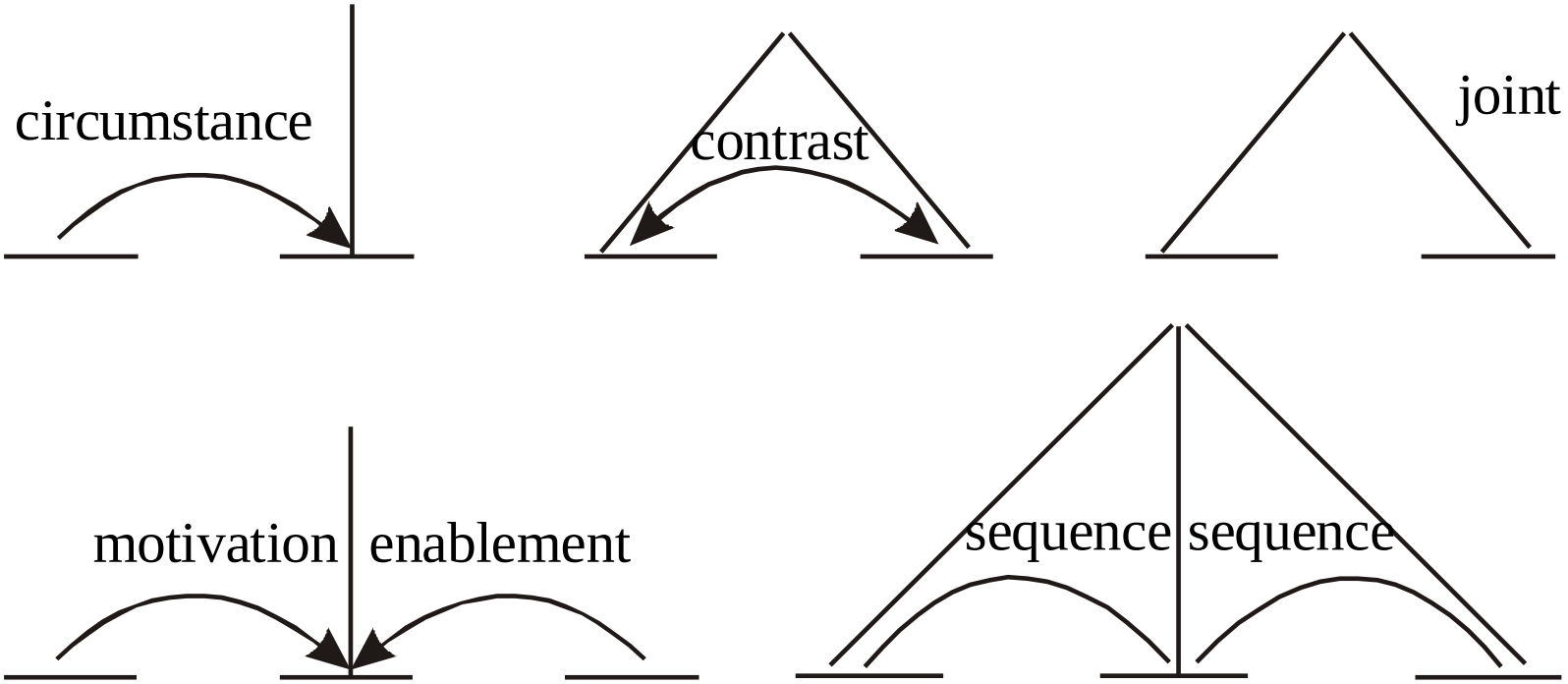
Relation names (in M&T 1988)

Circumstance	**Antithesis and Concession**
Solutionhood	**Antithesis**
Elaboration	**Concession**
Background	**Condition and Otherwise**
Enablement and Motivation	**Condition**
Enablement	**Otherwise**
Motivation	**Interpretation and Evaluation**
Evidence and Justify	**Interpretation**
Evidence	**Evaluation**
Justify	**Restatement and Summary**
Relations of Cause	**Restatement**
Volitional Cause	**Summary**
Non-Volitional Cause	**Other Relations**
Volitional Result	**Sequence**
Non-Volitional Result	**Contrast**
Purpose	

Other classifications are possible, and longer and shorter lists have been proposed

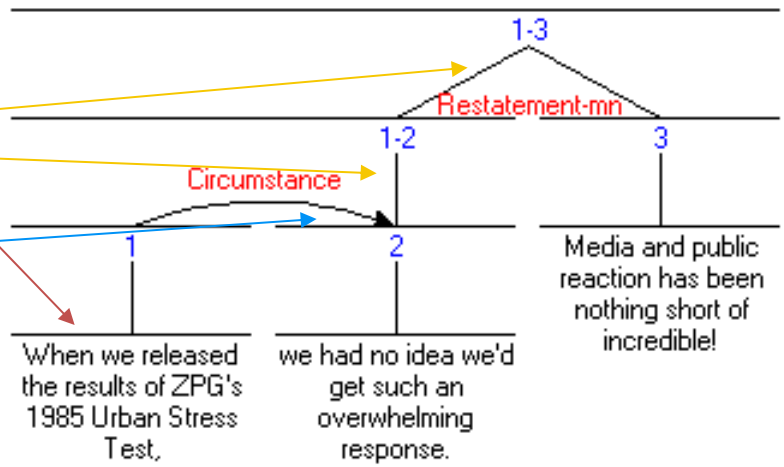
Schemas

- They specify how spans of text can co-occur, determining possible RST text structures



Graphical representation

- A horizontal line covers a span of text (possibly made up of further spans
- A vertical line signals the nucleus or nuclei
- A curve represents a relation, and the direction of the arrow, the direction of satellite towards nucleus



Taboada

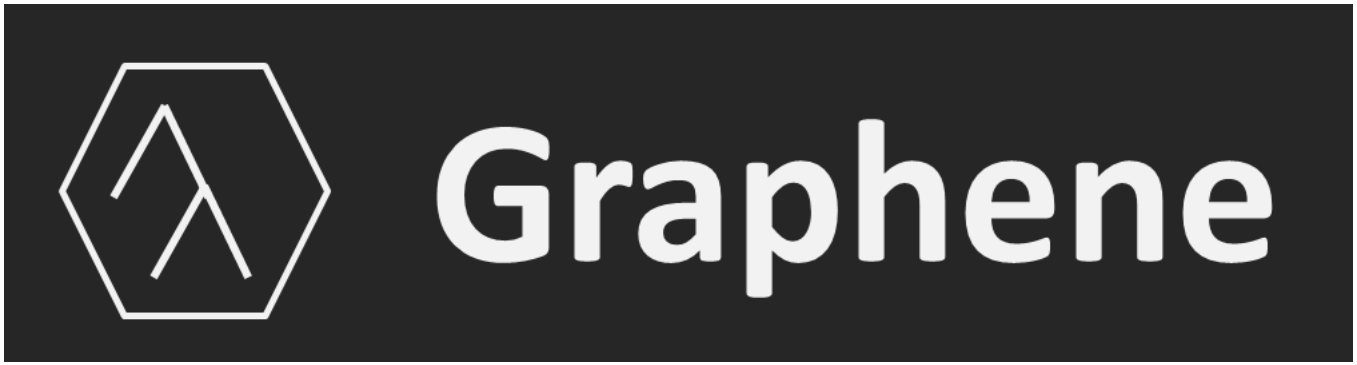
How to do an RST analysis

1. Divide the text into units
   - Unit size may vary, depending on the goals of the analysis
   - Typically, units are clauses (but not complement clauses)
2. Examine each unit, and its neighbours. Is there a clear relation holding between them?
3. If yes, then mark that relation (e.g., Condition)
4. If not, the unit might be at the boundary of a higher-level relation. Look at relations holding between larger units (spans)
5. Continue until all the units in the text are accounted for
6. Remember, marking a relation involves satisfying all 4 fields (especially the Effect). The Effect is the plausible intention that the text creator had.

Taboada

Putting All Together

Software: Extracting Knowledge Graphs from Text



https://github.com/Lambda-3/Graphene

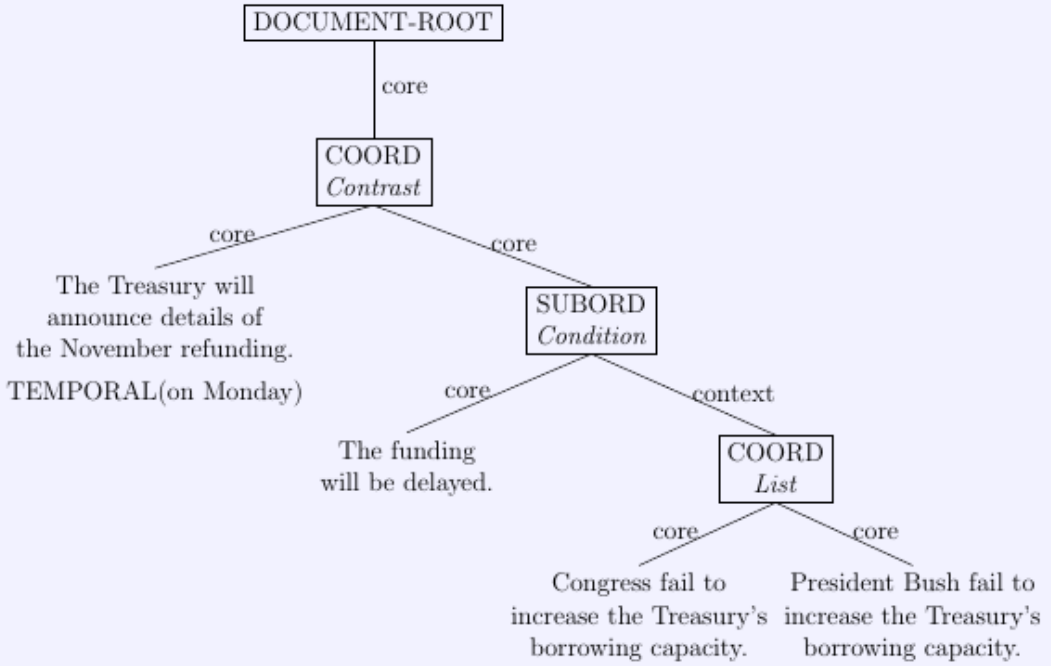
Niklaus et al., A Sentence Simplification System for Improving Relation Extraction, COLING (2017)

Coreference-Resolution

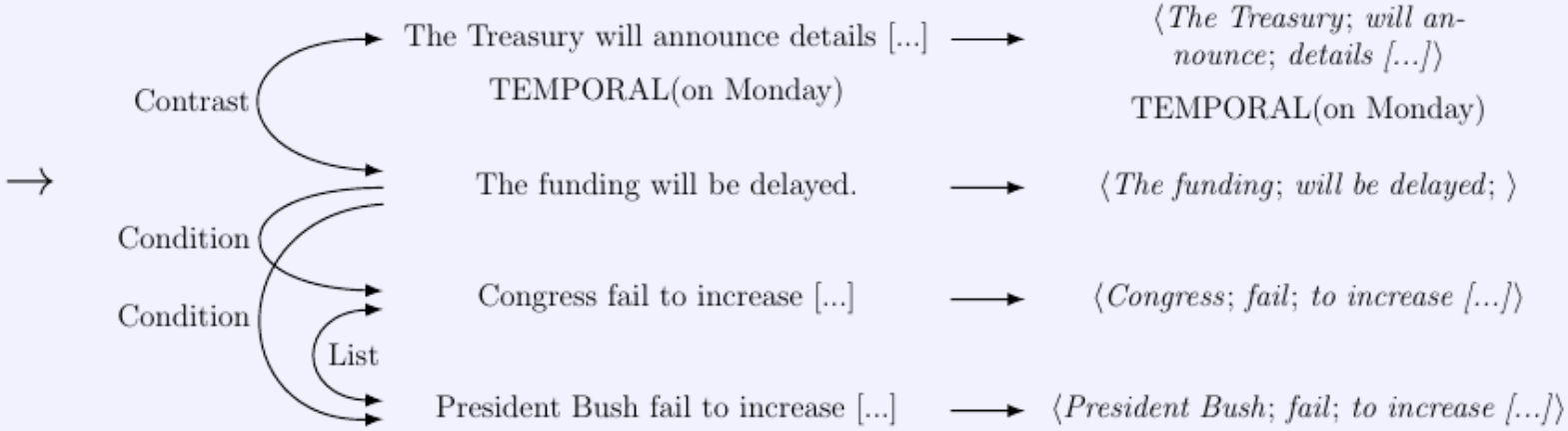
[...] Although the Treasury will announce details of the November refunding on Monday, the funding will be delayed if Congress and President Bush fail to increase the Treasury's borrowing capacity. [...]

→

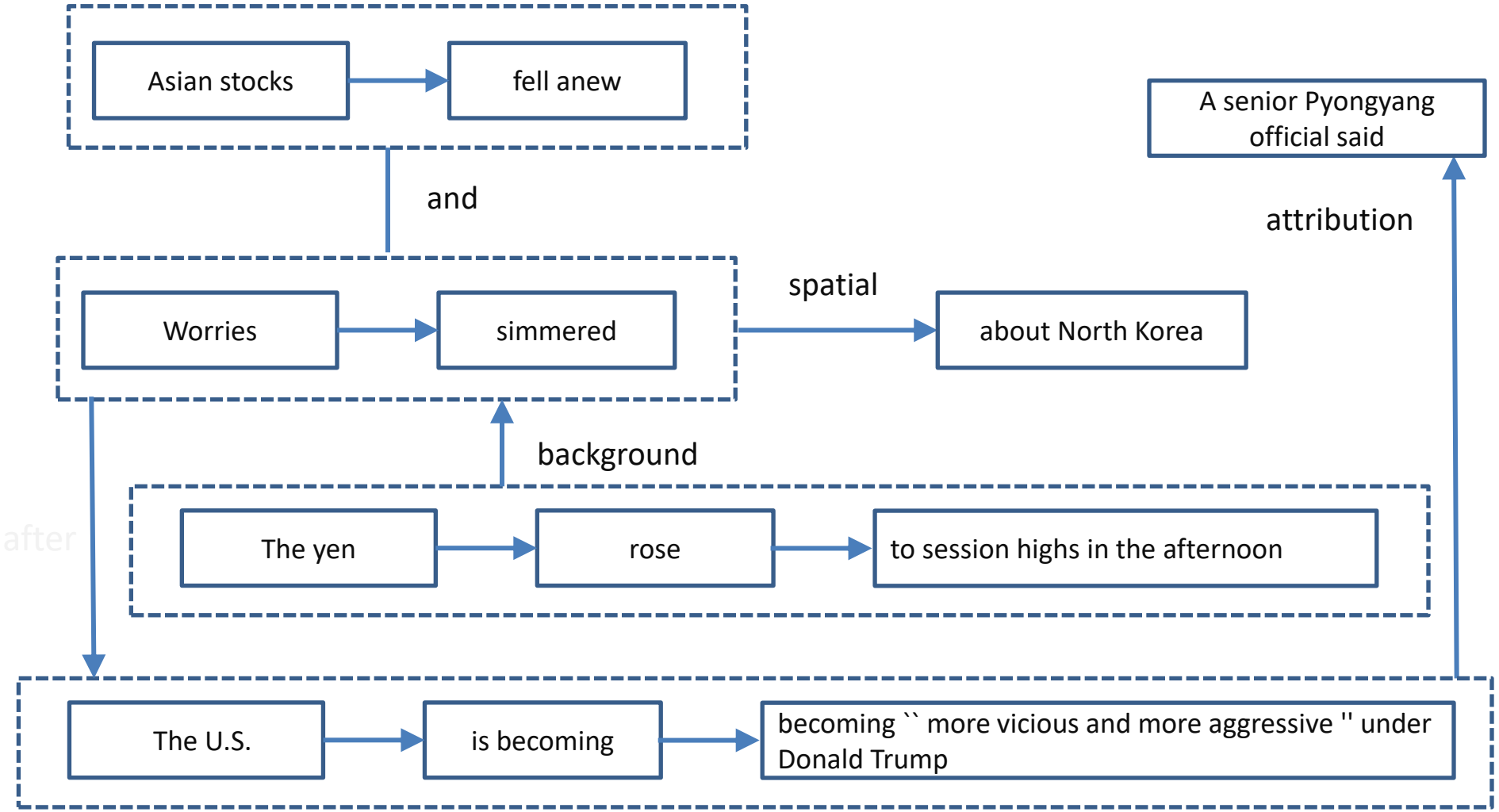
Transformation Stage



Relation Extraction



Asian stocks fell anew and the yen rose to session highs in the afternoon as worries about North Korea simmered, after a senior Pyongyang official said the U.S. is becoming ``more vicious and more aggressive'' under President Donald Trump .



The RDF-NL Format

Although the Treasury will announce details of the November refunding on Monday , the funding will be delayed if Congress and President Bush fail to increase the Treasury 's borrowing capacity .

bacf06771e0f4fc5a8e68c30fc77c9c4 the Treasury will announce details of the November refunding
S:TEMPORAL on Monday .
L:CONTRAST 948eeebd73564adab7dee5c6f177b3b9

948eeebd73564adab7dee5c6f177b3b9 the funding will be delayed
L:CONDITION 006a71e51295440fab7a8e8c697d2ba6
L:CONDITION e4d86228cff443b7a8e9f6d8a5c5987b
L:CONTRAST bacf06771e0f4fc5a8e68c30fc77c9c4

006a71e51295440fab7a8e8c697d2ba6 Congress fail to increase the Treasury 's borrowing capacity

e4d86228cff443b7a8e9f6d8a5c5987b president Bush fail to increase the Treasury 's borrowing capacity

The RDF-NL Format

Although the Treasury will announce details of the November refunding on Monday , the funding will be delayed if Congress and President Bush fail to increase the Treasury 's borrowing capacity .

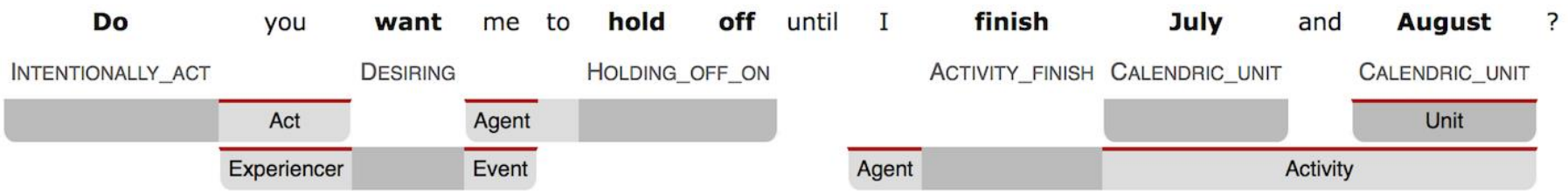
bacf06771e0f4fc5a8e68c30fc77c9c4 the Treasury will announce details of the November refunding
S:TEMPORAL on Monday .
L:CONTRAST 948eeebd73564adab7dee5c6f177b3b9

948eeebd73564adab7dee5c6f177b3b9 the funding will be delayed
L:CONDITION 006a71e51295440fab7a8e8c697d2ba6
L:CONDITION e4d86228cff443b7a8e9f6d8a5c5987b
L:CONTRAST bacf06771e0f4fc5a8e68c30fc77c9c4

006a71e51295440fab7a8e8c697d2ba6 Congress fail to increase the Treasury 's borrowing capacity

e4d86228cff443b7a8e9f6d8a5c5987b president Bush fail to increase the Treasury 's borrowing capacity

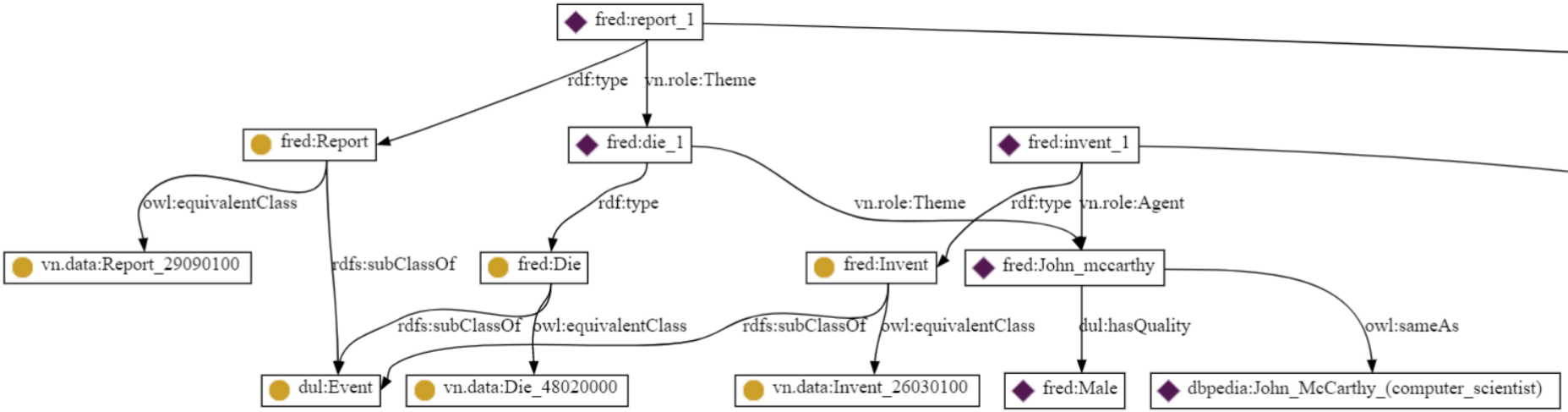
Semafor



http://www.cs.cmu.edu/~ark/SEMAFOR/

Frame-based Extraction

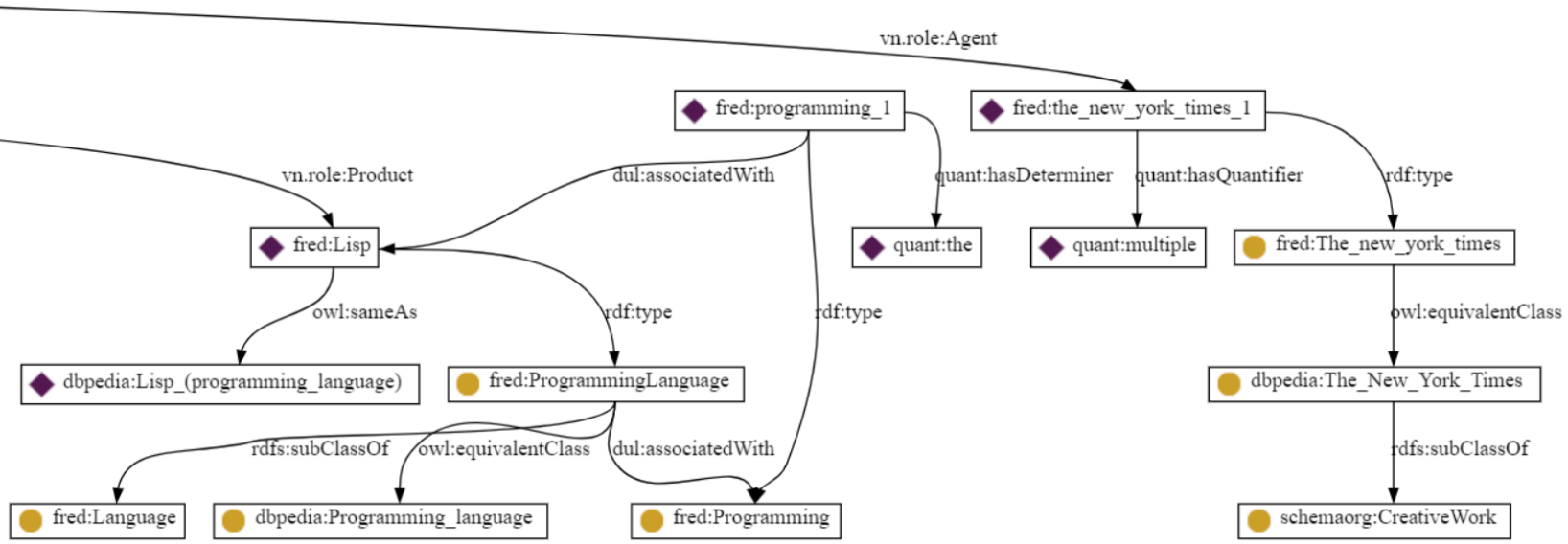
The New York Times reported that John McCarthy died. He invented the programming language LISP.



Gangemini et al., Semantic Web Machine Reading with FRED, Semantic Web Journal, 2017

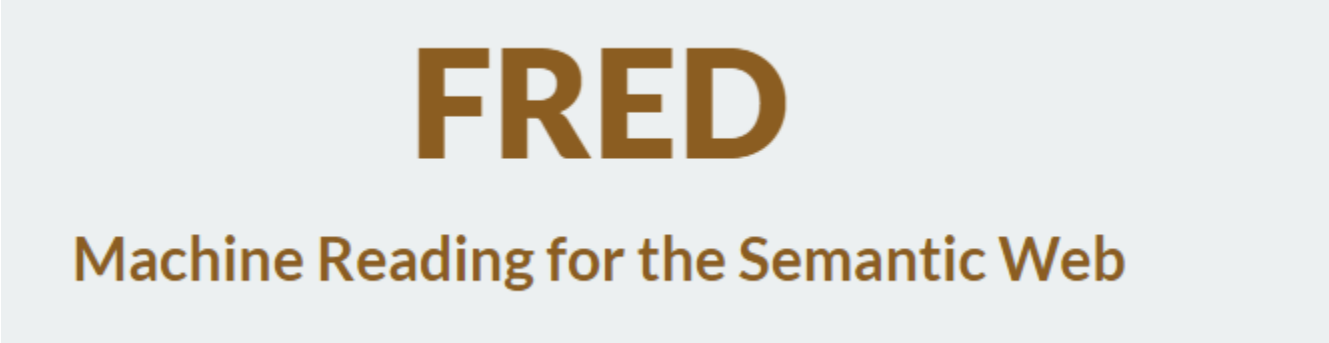
Frame-based Extraction

The New York Times reported that John McCarthy died. He invented the programming language LISP.



Gangemini et al., Semantic Web Machine Reading with FRED, Semantic Web Journal, 2017

Software: FRED



http://wit.istc.cnr.it/stlab-tools/fred/

Gangemini et al., Semantic Web Machine Reading with FRED, Semantic Web Journal, 2017

Recommended Reading

Bundy, Alan, & Fiona McNeill (2006) "Representation as a Fluent: An AI Challenge for the Next Half Century," AAAI Fellows Symposium, Boston, MA.