Symbolic AI

Andre Freitas
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.

\[ \exists w \text{ ISA}(w, Arriving) \land Arriver(w, Speaker) \land Destination(w, NewYork) \]

• The temporal information provided by the tense of the verbs can be exploited by predicing additional information about the \textbf{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.

\[ \exists i, e, w, t \text{ ISA}(w, Arriving) \]
\[ \land Arriver(w, Speaker) \land Destination(w, NewYork) \]
\[ IntervalOf(w, i) \land EndPoint(i, e) \land Precedes(e, Now) \]

\[ \exists i, e, w, t \text{ ISA}(w, Arriving) \]
\[ \land Arriver(w, Speaker) \land Destination(w, NewYork) \]
\[ IntervalOf(w, i) \land MemberOf(i, Now) \]
Semantic Roles: Frame!

Who did what to whom at where?

The police officer detained the suspect at the scene of the crime

Agent Predicate Theme Location
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

**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

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

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
agree.01
Arg0: Agreer
Arg1: Proposition
Arg2: Other entity agreeing

Ex1: [Arg0 The group] agreed [Arg1 it wouldn’t make an offer].
Ex2: [ArgM-TMP Usually] [Arg0 John] agrees [Arg2 with Mary] [Arg1 on everything].
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

[Arg0 Big Fruit Co.] increased [Arg1 the price of bananas].
[Arg1 The price of bananas] was increased again [Arg0 by Big Fruit Co.]
[Arg1 The price of bananas] increased [Arg2 5%].
Frege’s Begriffsschrift for the Same Sentence

• Peirce’s algebraic notation (1885):

\[ \Sigma_x \Sigma_y (\text{Go}(x) \cdot \text{Person}(\text{John}) \cdot \text{City}(\text{Boston}) \cdot \text{Bus}(y) \cdot \text{Agnt}(x,\text{John}) \cdot \text{Dest}(x,\text{Boston}) \cdot \text{Inst}(x,y)) \]

• With Peano’s choice of symbols:

\[ (\exists x)(\exists y)(\text{Go}(x) \land \text{Person}(\text{John}) \land \text{City}(\text{Boston}) \land \text{Bus}(y) \land \text{Agnt}(x,\text{John}) \land \text{Dest}(x,\text{Boston}) \land \text{Inst}(x,y)) \]
Frege’s Begriffsschrift for the Same Sentence

- Translation to Peirce-Peano notation:

\[ \neg(\forall x)(\forall y)(Go(x) \Rightarrow (Person(John) \Rightarrow (City(Boston) \Rightarrow (Bus(y) \Rightarrow (Agnt(x,John) \Rightarrow (Dest(x,Boston) \Rightarrow \neg Inst(x,y)))))\]
Frege’s Begriffsschrift for the Same Sentence

• Translation to Peirce-Peano notation:

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

• 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.

\[(\exists x)(\exists y)(\text{Person}(Sue) \land \text{Child}(x) \land \text{Book}(y) \land \text{Gives}(Sue, x, y)).\]
• 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:

\[(\exists x)(\exists y)(\exists z) (\text{Person}(\text{Sue}) \land \text{Child}(x) \land \text{Book}(y) \land \text{Give}(z) \land \text{Agnt}(z,\text{Sue}) \land \text{Rcpt}(z,x) \land \text{Thme}(z,y))\]
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.

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

(∃s)(Situation(s) ∧ Time(2pm) ∧ PTim(s,2pm) ∧ descr(s, (∃y)(∃z)(∃u)(∃v)(∃w)(Person(Sue) ∧ Child(Bob) ∧ Book(y) ∧ Give(z) ∧ Gracious(u) ∧ Poor(v) ∧ New(w) ∧ Manr(z,u) ∧ Attr(Bob,v) ∧ Attr(y,w) ∧ Agnt(z,Sue) ∧ Rcpt(z,Bob) ∧ Thme(z,y))))
• The concept of type Situation with a nested CG represents a situation described by that CG.

• In the algebraic formula, the relation \(\text{dscr}(s,p)\) 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.
• 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):

\[
\exists x \exists y \text{Cat}(x) \land \text{Mat}(y) \land \text{On}(x, y)
\]

Charles Sanders Peirce (1885):

\[
\exists x \exists y \text{Cat}(x) \land \text{Mat}(y) \land \text{On}(x, y)
\]

Giuseppe Peano (1895):

\[
\exists x \exists y \text{Cat}(x) \land \text{Mat}(y) \land \text{On}(x, y)
\]

Charles Sanders Peirce (1897):

\[
\text{Cat} \rightarrow \text{On} \rightarrow \text{Mat}
\]

J. Sowa
First-Order Logic

• 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

Cat—Black  
*Some cat is black.*

Cat—Black  
*Some cat is not black.*

Cat—Black  
*No cat is black.*

Cat—Black  
*It is false that some cat is not black.*

Cat—Black  
*If there is a cat, then it is black.*

Cat—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._
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)]
Example: *Tom believes Mary wants to marry a sailor.*
Example: There is a sailor that Tom believes Mary wants to marry.
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.

\[ \exists x (\text{woman}(x) \land \text{walk}(x)) \quad \text{smoke}(x) \]

\[ \exists x (\text{woman}(x) \land \text{walk}(x) \land \text{smoke}(x)) \]

Need to Expand the scope of the existential quantifier.
A woman walks. She smokes.

- A woman walks. She smokes.

<table>
<thead>
<tr>
<th>x, y</th>
</tr>
</thead>
<tbody>
<tr>
<td>woman(x)</td>
</tr>
<tr>
<td>walk(x)</td>
</tr>
<tr>
<td>y = x</td>
</tr>
<tr>
<td>smoke(y)</td>
</tr>
</tbody>
</table>

discourse referent x, y, in the top part of the box.

conditions upon these discourse referents in the lower part of the box.
Discourse Structure and Accessibility

- Mary ordered a milk shake, John tasted it.

| $x, y$ | $x = Mary$
|-------|-----------------
|       | order$(x, y)$
|       | shake$(y)$

| $x, y, u, v$
|-----------------|
| $x = Mary$
| order$(x, y)$
| shake$(y)$
| $u = John$
| taste$(u, v)$
| $v = y$

- 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.

<table>
<thead>
<tr>
<th>$x, y$</th>
<th>$f\text{armer}(x)$</th>
<th>$d\text{onkey}(y)$</th>
<th>$\text{own}(x, y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$z$</td>
<td>$\text{beat}(x, z)$</td>
<td>$z = y$</td>
</tr>
</tbody>
</table>

• 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
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.
    - 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)

<table>
<thead>
<tr>
<th>Circumstance</th>
<th>Antithesis and Concession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solutionhood</td>
<td>Antithesis</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Concession</td>
</tr>
<tr>
<td>Background</td>
<td>Condition and Otherwise</td>
</tr>
<tr>
<td>Enablement and Motivation</td>
<td>Condition</td>
</tr>
<tr>
<td>Enablement</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Motivation</td>
<td>Interpretation and Evaluation</td>
</tr>
<tr>
<td>Evidence and Justify</td>
<td>Interpretation</td>
</tr>
<tr>
<td>Evidence</td>
<td>Evaluation</td>
</tr>
<tr>
<td>Justify</td>
<td>Restatement and Summary</td>
</tr>
<tr>
<td>Relations of Cause</td>
<td>Restatement</td>
</tr>
<tr>
<td>Volitional Cause</td>
<td>Summary</td>
</tr>
<tr>
<td>Non-Volitional Cause</td>
<td>Other Relations</td>
</tr>
<tr>
<td>Volitional Result</td>
<td>Sequence</td>
</tr>
<tr>
<td>Non-Volitional Result</td>
<td>Contrast</td>
</tr>
</tbody>
</table>

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

---

**When we released the results of ZPG's 1985 Urban Stress Test,**

**we had no idea we'd get such an overwhelming response.**

**Restatement-mn**  

**Circumstance**  

**Media and public reaction has been nothing short of incredible!**
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.
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)
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. [...]

The Treasury will announce details of the November refunding. (The Treasury; will announce; details [...])

The funding will be delayed. (The funding; will be delayed; )

Congress fail to increase the Treasury’s borrowing capacity. (Congress; fail; to increase [...])

President Bush fail to increase the Treasury’s borrowing capacity. (President Bush; fail; to increase [...])
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.
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.
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

Do you want me to hold off until I finish July and August?

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
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

FRED
Machine Reading for the Semantic Web

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

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