

Knowledge Engineering for NLP

May 1, 2005

Clausal semantics

Overview

- Why clausal semantics?
- What's a clause?
- Messages in G&S, MRS, the Matrix
- Messages and the syntax-semantics interface
- Details about this week's implementation

Why clausal semantics?

- Illocutionary force correlates with syntactic form
- MRS representations should include all semantic information that is syntactically marked
- Illocutionary force is ‘predicated of’ situations — i.e., the rest of the MRS

Aside: Perlocutionary, Locutionary, Illocutionary

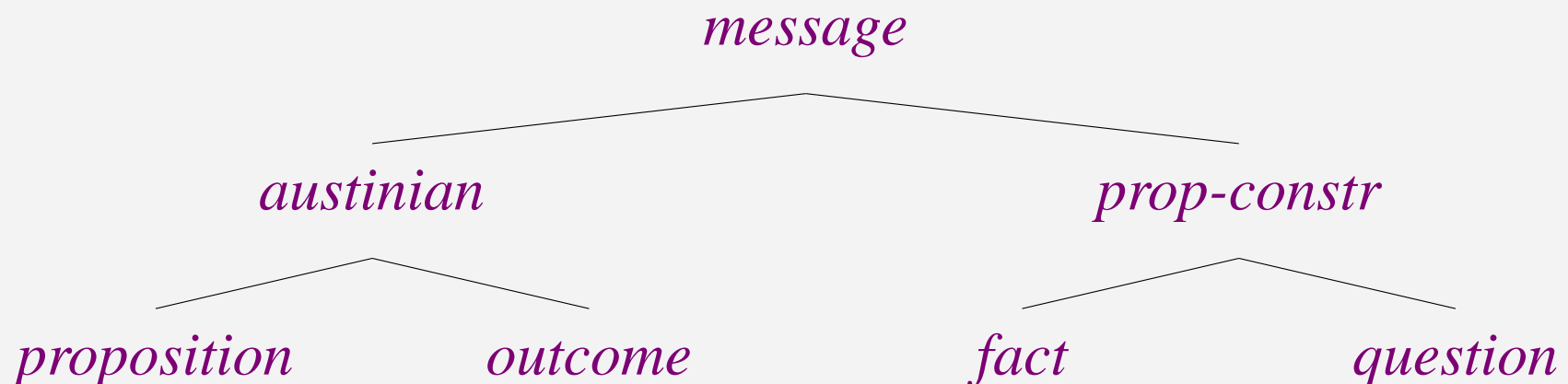
- locutionary act: The act of saying something
- illocutionary act: The act of asking, asserting, commanding, etc. by saying something
- perlocutionary act: The act of getting someone to do or believe something by asking, asserting, etc. something

What's a clause?

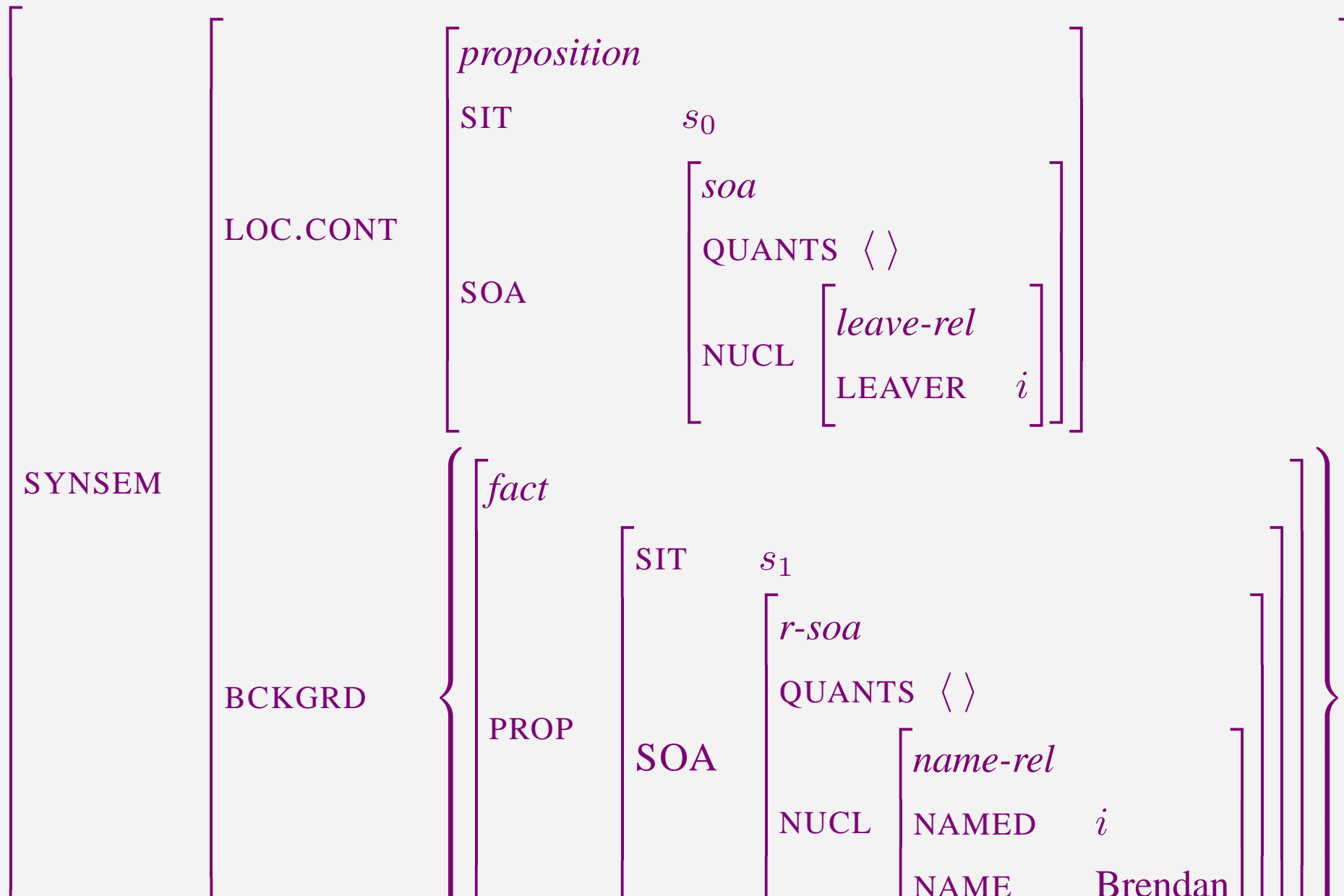
- Syntactically complete
- Expresses some illocutionary force
- Contrasts with fragments, some of which can also carry illocutionary force.
- Marking of illocutionary force is often associated with the clause as a whole or with its head verb.
- Clauses can be matrix or embedded.
- Embedded clauses carry messages, too.

Clausal semantics: Messages

- “*Message* is the semantic type that is the most basic to communication—its (maximal) subtypes constitute the descriptive contents of basic illocutionary acts such as assertion, querying, commanding, exclaiming and the like.” (Ginzburg & Sag 2000:121)
- Partial hierarchy under *message*:



Clausal semantics in recursive representation (1/2)



Clausal semantics in recursive representation (2/2)

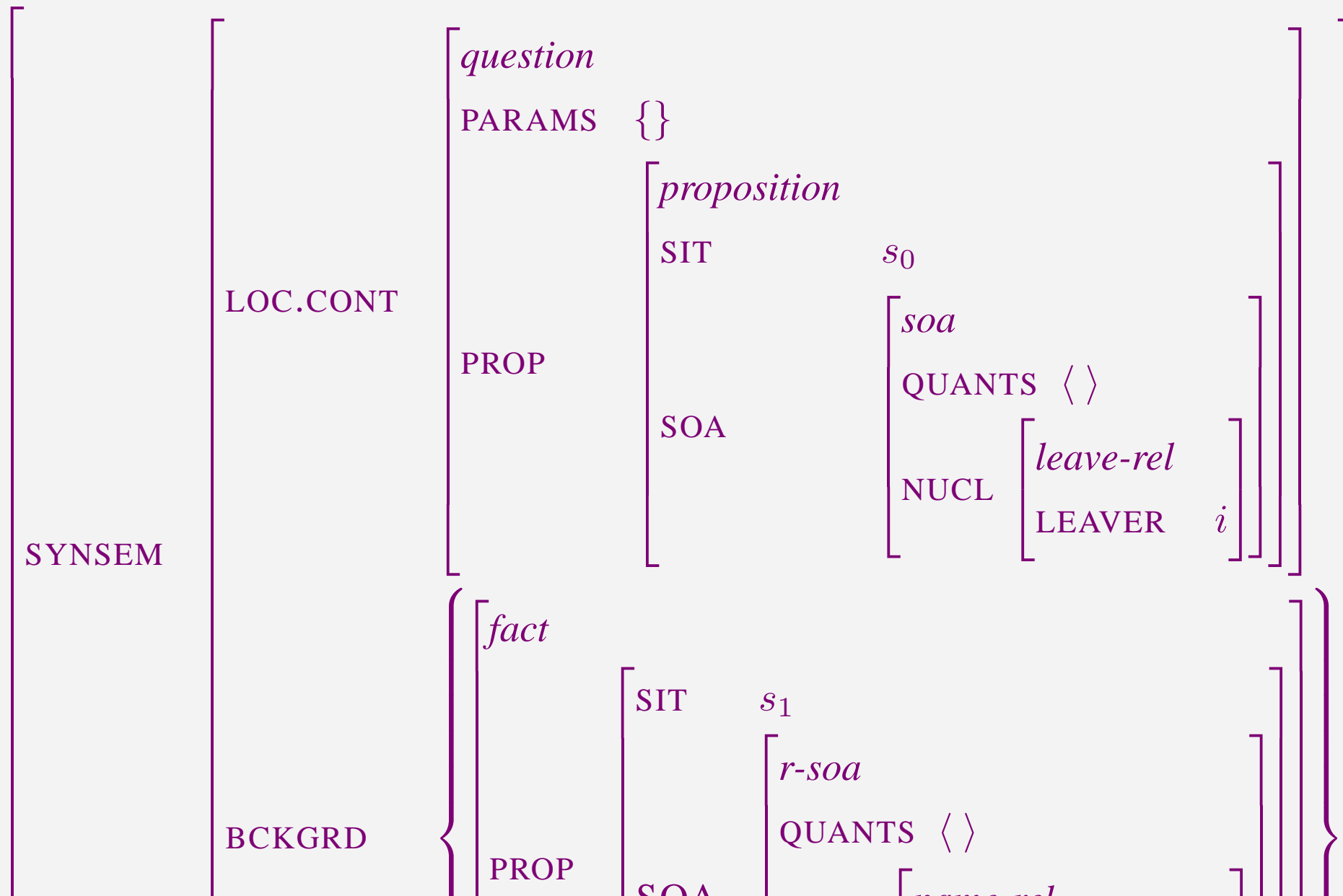


Diagram illustrating the structure of the sentence "Kim left the room" using a hierarchical tree structure. The root node is **LTOP**, which branches into **INDEX** and **RELS**.

- INDEX** branches into **HCONS** and **RELS**.
 - HCONS** branches into **h6** and **h14**.
 - RELS** branches into **int_m_rel**, **prpstn_m_rel**, and **proper_q_rel**.
- int_m_rel** branches into **LBL** (*h1*) and **MARG** (*h5*).
- prpstn_m_rel** branches into **LBL** (*h5*) and **MARG** (*h6*).
- proper_q_rel** branches into **LBL** (*h9*), **ARG0** (*x11*), **RSTR** (*h10*), and **BODY** (*h12*).
- BODY** (*h12*) branches into **named_rel** and **_leave_v_1_rel**.
 - named_rel** branches into **LBL** (*h13*), **ARG0** (*x11*), and **CARG** (*"kim"*).
 - _leave_v_1_rel** branches into **LBL** (*h14*), **ARG0** (*e2*), **ARG1** (*x11*), and **ARG2** (*i15*).
- HCONS** also branches into **h10** and **h13**.

The final structure is a list of **h6**, **h14**, **h10**, and **h13**.

Messages in the Matrix (1/3)

```
mrs := mrs-min &  
  [ HOOK hook,  
    RELS diff-list,  
    HCONS diff-list,  
    MSG basic_message ].
```

- Messages appear on the RELS list, but also have a dedicated pointer in CONT.MSG.
- We can use CONT.MSG to ensure that only clauses are accepted as stand alone utterances.

Messages in the Matrix (2/3)

```
basic_message := relation.  
message := basic_message &  
    [ PRED message_m_rel,  
      MARG handle ].  
  
no-msg := basic_message.
```

Messages in the Matrix (3/3)

```
message_m_rel := predsor.  
command_m_rel := message_m_rel.  
prop-or-ques_m_rel := message_m_rel.  
    ;for COMPS of e.g. 'know'  
proposition_m_rel := prop-or-ques_m_rel.  
abstr-ques_m_rel := prop-or-ques_m_rel.  
question_m_rel := abstr-ques_m_rel.  
ne_m_rel := abstr-ques_m_rel.
```

Messages in compositional semantics

- Ginzburg & Sag and the ERG cross-classify phrase structure rules along dimensions of ‘clausality’ and ‘headedness’.
- For English, one can identify certain constructions as licensing clauses.

E.g. *decl-hd-subj-ph* pairs a VP head and its subject, while introducing the *proposition*.

- What about languages with freer word order?
- (Still see some of the effects of this in matrix.tdl.)

Our general strategy

- Unary-branching clausal constructions
- Daughter is [MSG no-msg]
- C-CONT has a *message-relation* on its RELS list and associated *qeq* on its HCONS
- Mother's CONT.HOOK.MSG points to the message on the RELS list
- Daughter is constrained to have appropriate syntactic properties

Why do we need MSG?

- Constrain the initial symbol to only accept complete clauses
- Allow clause embedding verbs to select for the right semantic type of complement

Overview

- Why clausal semantics?
- What's a clause?
- Messages in G&S, MRS, the Matrix
- Messages and the syntax-semantics interface
- Details about this week's implementation