Frame Semantics

semantic characterisation of
situations or states of affairs

---

structure of the presentation

1. introduction (partially taken from a presentation of Markus Egg):
    i. what is a frame supposed to be?
    ii. basic features of the approach
2. the FrameNet project:
    i. aims of the project
    ii. FrameNet annotation
3. the SALSA project

1. Introduction

---

knowledge-representation in larger-scale structures

- declarative representations:
  - frames (Marvin Minsky)
  - schemata (David Rumelhart)
  - scripts (Roger Schank)
- procedural representations (productions):
  - conditionals that specify actions to be performed if certain conditions are satisfied
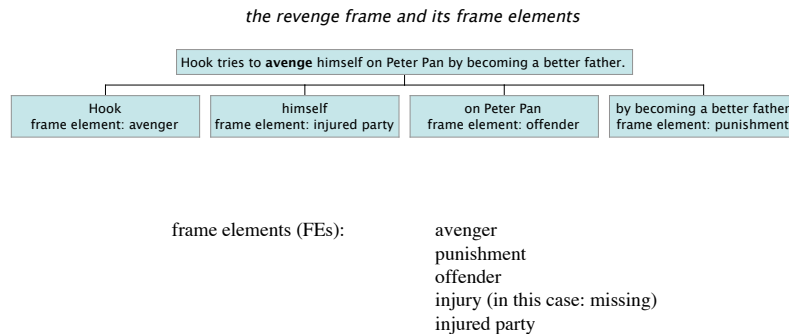
1. Introduction

---

the concept of frame

- Minsky's initial notion of frame:
  - a data structure representing a stereotyped situation
- Fillmore's notion of frame:
  - a frame is to be understood as a cognitive structuring device
  - frames charakterise situations or states of affairs
  - they are in principle independent of their linguistic verbalisation
  - parts of frames are connected to specific words
  - since verbs refer to whole situations, they are most closely associated with frames

1. Introduction

an example

the revenge frame and its frame elements

Hook tries to **avenge** himself on Peter Pan by becoming a better father.

Hook frame element: avenger	himself frame element: injured party	on Peter Pan frame element: offender	by becoming a better father frame element: punishment

frame elements (FEs):   avenger
punishment
offender
injury (in this case: missing)
injured party

definition of frame

- for the purposes of this context:

  A frame is an intuitive construct that allows us to formalize the links between semantics and syntax in the results of lexical analysis. Semantic frames are schematic representations of situations involving various participants, props, and other conceptual roles, each of which is a frame element. The semantic arguments of a predicating word correspond to the frame elements of the frame (or frames) associated with that word.

frames introduce a perspective on situations

- verbs refer to whole situations (i. e. frames); they can focus and background different aspects of the same frame:
  - e.g. : buy, sell, pay, spend; all these verbs evoke the same frame
- to know the meaning of each of these verbs, one has to know the frame as a whole: this brings together these verbs in one semantic group
- thus, frames introduce a perspective on situations, even if the same arguments are realised syntactically:
  - buy X from Y for Z: perspective of buyer
  - sell X to Y for Z: perspective of seller

frames as interface

- frames are supposed to serve as a kind of interface between purely linguistic and conceptual knowledge
- thus, the description of a verb comprises
  - the frame it evokes
  - linking information: the grammatical realisation of the frame elements (which case, obligatory vs optional)

the organisation of frames

- frames are related in several ways
  - they constitute a taxonomy or hierarchy parallel to the hierarchical organisation of concepts (e.g.: the lease of an appartment is a specific commercial transaction which in turn is a change-of-state)
  - sub-frame relations: wholes and parts (e.g.: commercial transaction breaks down in two subparts: transfer of goods and transfer of money)

frames as paradigmatic descriptions of situations

- frames have prototypes, e.g.: commercial transaction:
  - prototypical buyer/seller: human
  - prototypical goods: concrete objects or substances
  - prototypical payment: cash
- problematic cases:
  - what kind of goods is transferred in fine and alimony?
  - a county might obtain public services from a company in exchange for tax liabilities

frames and knowledge of the world

- frames often involve a rather large slice of the surrounding culture:
  - cp. the concept of breakfast: we (typically) eat three meals a day, the first of which is breakfast. it takes place after getting up in the morning and involves special kinds of food, e.g., jam
  - but, one can have breakfast at noon, too, or, after not having slept all night

frames as device for text understanding

- frames are used for theories of text understanding:
  - first, the general topic of the text (which means, for our purposes, sentences) is recognised and encoded in an appropriate frame
  - then the derivation of the text meaning is modelled as filling in the elements of a frame
  - this is actually the way in which some computational approaches to information extraction work

the Berkeley FrameNet II project

- successor of FrameNet I which in turn based on earlier research done in terms of frame semantics (the DELIS project: frame analyses of lexical units in the domains of communication and perception; http://www.ims.uni-stuttgart.de/projekte/delis/)
- FRAME NET II: http://www.icsi.berkeley.edu/~framenet/

the Berkeley FrameNet II project

- creation of an on-line lexical resource for English
- based on frame semantics
- supported by corpus evidence

the FrameNet corpus

- the main FrameNet corpus is the 100-million-word British National Corpus
- subprojects are using U. S. newswire texts
- semantic and syntactic annotation carried out by in-house software, respectively

the FrameNet database

- the project's deliverables will consist of the FrameNet database itself:
  - lexical entries for individual word senses
  - descriptions of frames and frame elements
  - annotated subcorpora

intended purpose

- the database is to be used by NLP researchers working on a variety of NLP applications, including:
  - word sense disambiguation: FrameNet gives the syntactic and collocational information
  - machine translation: FrameNets for other languages will reveal cross-linguistic differences in the meanings and grammatical behavior of words belonging to shared frames
  - information extraction: FrameNet annotations provide a high precision seed for building information extraction patterns

intended purpose

- the database is to be used by NLP researchers working on a variety of NLP applications, including
  - information extraction: FrameNet annotations provide a high precision seed for building information extraction patterns
  - question answering: FrameNet data will facilitate the recognition of semantic frame relations between the language of user questions and that of passages in the documents that answer the question

the FrameNet database

- the database serves as:
  - a dictionary:
    - definition (from the Concise Oxford Dictionary, 10th Edition, Oxford University Press, or a definition written by a FrameNet staff member)
    - tables showing how frame elements are syntactically expressed in sentences containing each word. this includes a complete characterisation of the headword's grammar and combinatorial properties
    - annotated examples from the corpus
    - alphabetical index

the FrameNet database

- the database serves as:
  - a thesaurus:
    - words are linked to the semantic frames in which they participate and hence to the other words which evoke this frames
    - frames are related to other frames

semantic and syntactic annotation

- there are three layers of annotation on a tagged constituent:
  - the frame element realisation consists of a frame Element (say, patient), a grammatical function (say, object) and a phrase type (say, NP)
  - valence descriptions of predicating words are generalisations over such structures

organisation of the database

The annotated sentences are the building blocks of the database: marked up in XML, they form the data from which the lexical entry descriptions are derived. This format supports searching by lexical unit, frame, frame element, and combinations of these.

general principles of annotation

- a lexical unit (LU) is a pairing of a word with a meaning. Typically, each sense of a polysemous word belongs to a different semantic frame
- FrameNet annotation is always done relative to one particular linguistic unit, the target, which is most often a single word but can also be a multi-word expression such as a phrasal verb (e.g. give in) or an idiom (say, take into account)
- core FEs vs non-core FEs

general principles of annotation

- whole constituents are annotated
- each dependent is annotated for frame element identity, phrase type and grammatical function relative to the target LU
- annotation of single sentences instead of annotating running text.
- the sentences used in describing a single LU are annotated only in respect to that LU
- there's no information about frequency of occurrence

the FrameNet desktop: "Annotate"

Sentence with target lexical unit avenge.v

The annotators actually only need to apply a frame element label; grammatical function and phrase type are derived algorithmically based on patterns of POS labels but may require manual correction.

problems of annotation

- missing frame elements:
  - Constructional Null Instantiation:
    Tie together loosely (CNI items)
  - Definite Null Instantiation:
    The monkey avenged himself (DNI offender)
  - Indefinite Null Instantiation
    He drinks a lot (INI ingestibles)
- frame element conflation:
    He avenged Pedro's death (injury AND injured party) by taking out the poker-faced Guards Officer

SALSA - The SAarbrücken Lexical Semantics Annotation and Analysis Project

the SALSA project

- creating a semantic annotation schema which allows for the integration of a wide range of phenomena
- annotating the German 1,5 million word TIGER corpus by hand with FrameNet semantic roles
- concurrent development of a German "FrameNet light"
- development of machine learning tools for supervised and unsupervised annotation of larger corpora