The LOGON MT infrastructure

Ling 567
February 27, 2024

Overview

- MRS and MT: Some history (reprise)
- The LOGON architecture
  - Processing steps
  - Transfer rules
  - VPM
- Stepping through our MMT sentences

But first: non-obligatory null affix rules

- Why do you create null affix rules?
- What happens if they are in non-obligatory PCs?
- How do you make PCs obligatory?

MRS and MT: Some history

- Copestake et al 1995: Original motivation for MRS included MT applications
- Resolving scope ambiguities is hard, and usually not necessary
  - Logical form equivalence is undecidable even in FOPL (Shieber 1993)
- Mimicking syntactic structure in semantics makes transfer harder
  - *fierce black cat <> gato negro y feroz* (Spanish)
  - *young black bull <> novillo negro*
- MRS gives logical forms with less syntactic complexity and underspecification wherever possible.

MRS and MT: Some history

- MRS originally developed in the context of Verb*Mobil* but not fully deployed for transfer-based MT in that project.

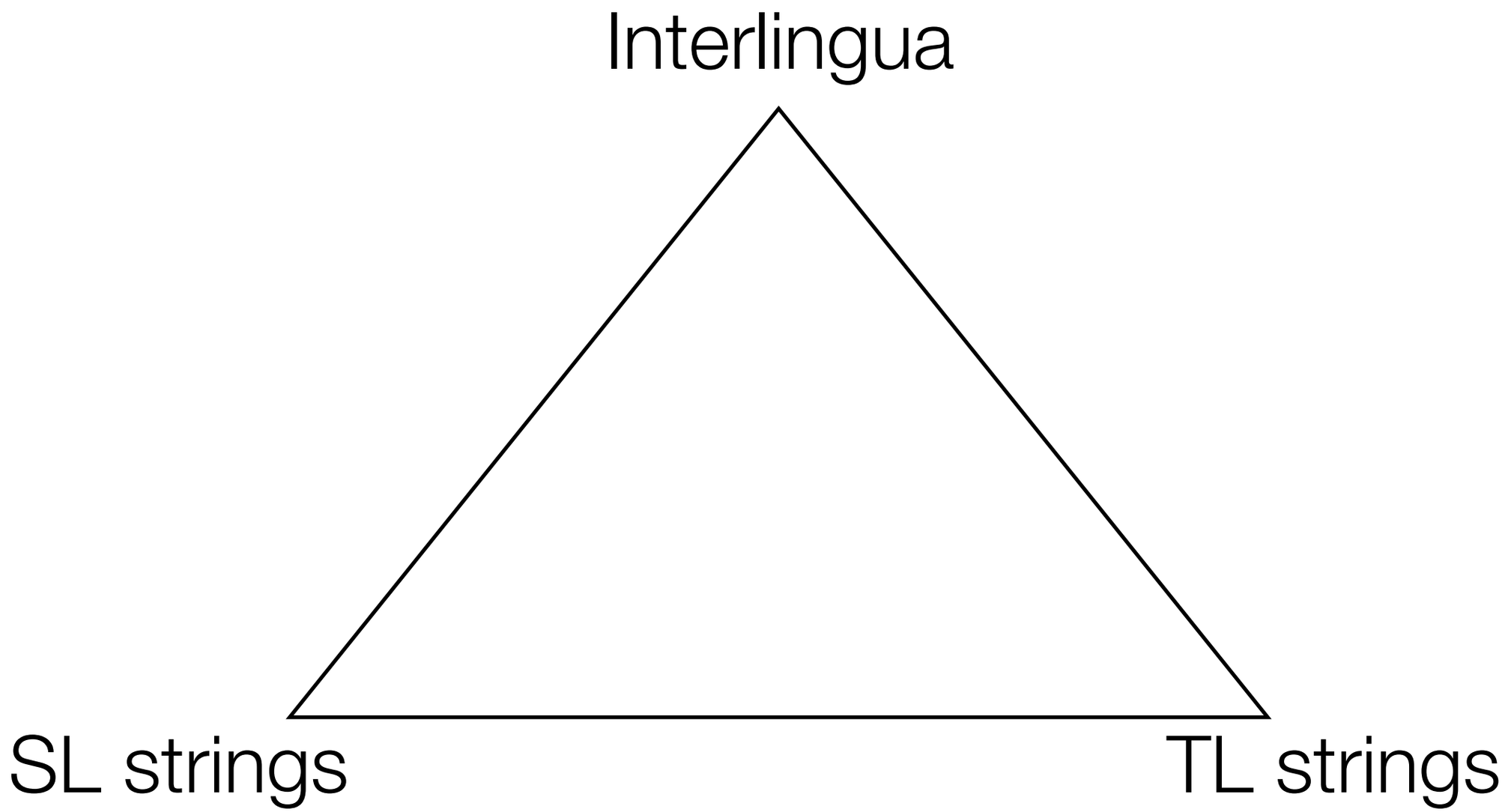
- In 2003, LOGON (Oepen et al 2004, 2007) up the thread and builds the first MRS-based MT system. (Norwegian -> English; tourism brochures)

- Input is LFG, with MRSs projected from f-structure.

- Output is generated by the English Resource Grammar (HPSG; Flickinger 2000)

- Further projects have built LOGON-based MT systems for other language pairs, notably JaEn (Bond et al 2011)

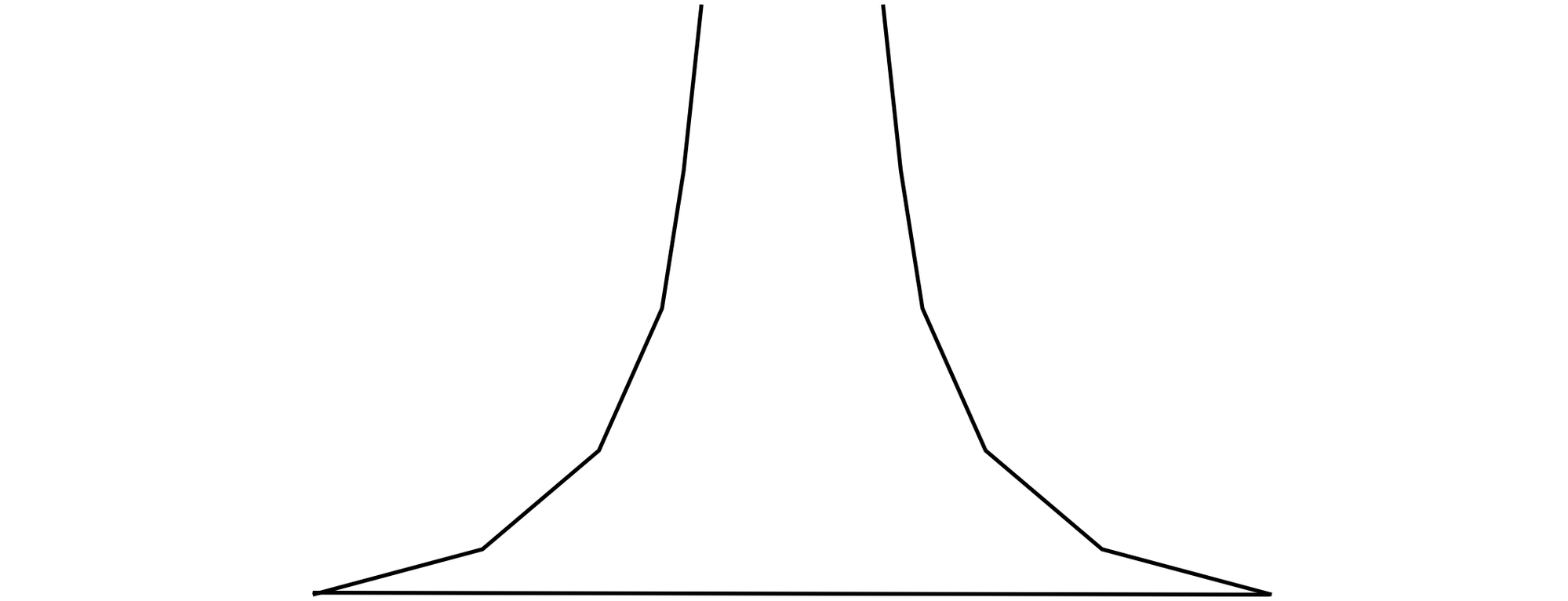
Vauquois Pyramid (ObMT Triangle)



Is MRS an interlingua?

- Could MRS be used to encode an interlingua?
- Could our grammars produce such an MRS-encoded interlingua?

Copestake Volcano



Massively Multilingual MT

- Problem of combinatory explosion (*n x n*):
- 2 languages: 2 sets of transfer rules
- 4 languages: 9 sets of transfer rules
- 24 languages: 552 sets of transfer rules
- 6000 languages: 35,994,000 sets of transfer rules

What are the alternatives?

- Design an interlingua (or select a pivot language), and create two grammars for each language
  - strings <> ordinary MRS
  - ordinary MRS <> interlingua (transfer grammar)
- Hybrid interlingual/transfer-based model
  - partial lexical interlingua or PanLex-derived rules
  - TL-side “accommodation” transfer grammars: O(n)
  - transfer matrix to capture generalizations
- How far will approach 2 scale?
- How much mismatch is there?

Mismatch: Translation divergences (Dorr 1994)

- Categorial divergence: Translation of words in one language into words that have a different part of speech in another language.

- Conflational divergence: The translation of two or more words in one language into one word in another language

- Structural divergence: The realization of verb arguments in different syntactic configurations in different languages.

- Head swapping divergence: The inversion of the structural dominance relation between two semantically equivalent words when translating from one language to another.

- Thematic divergence: The realization of verb arguments in different configurations that reflect different thematic to syntactic mapping orders.

MRS ‘harmonization’ helps

- Just because it’s not an interlingua doesn’t mean the grammars can’t be brought closer together.
- Example 1: Demonstratives (adjectives v. determiners)
- Example 2: COG-ST et al, reduction in quantifier-rel inventory
- Further potential for harmonization: pronouns v. pro-drop (but cf. information structure marking on overt pronouns)
- Other examples?

LOGON processing steps

- Parse in source language
    - visualization tools for parses and MRSs
- Apply transfer grammar to produce new MRS
    - visualization tools for transfer outputs
- Generate in target language from new MRSs
    - visualization tools for input MRSs
    - compare to MRS produced by parsing expected output
    - generator chart

Anatomy of a transfer rule

- Quadruple: [CONTEXT:] INPUT [!FILTER] -> OUTPUT
- Each item above is a (partial) MRS
- Rules apply to complete MRSs to produce partially rewritten MRSs.
- Resource sensitive: INPUT is consumed in producing OUTPUT.
- CONTEXT: Additional properties beyond the INPUT that must be satisfied. (Not consumed.)
- FILTER: Negative constraints; contexts in which the rule should not apply.

Anatomy of a transfer rule

- Rules can be obligatory or optional.
- Optional rules produce non-determinism in the transfer process.
- Pairing each optional rule with one obligatory rule cuts down the transfer search space.
- Rules can also be grouped into sets for ‘extrinsic’ ordering (which we probably won’t need).
- Handled with chart-based processing.

Example type

```
monotonic_mtr := mrs_transfer_rule &
 [ CONTEXT.HOOK.LTOP #h,
   INPUT.HOOK.LTOP #h,
   OUTPUT.HOOK.LTOP #h ].
```

Example rule instance

```
pro-insert-arg1-mtr := monotonic_mtr &
[ INPUT.RELS <! !>,
  CONTEXT.RELS <! [ ARG0.SF prop-or-ques,
                    ARG1 #x & x ] !>,
  FILTER.RELS <! [ ARG0 #x ] !>,
  OUTPUT [ RELS <! [ PRED "_pronoun_n_rel",
                     ARG0 #x,
                     LBL #larg ],
                   [ PRED "exist_q_rel",
                     ARG0 #x,
                     RSTR #harg ] !>,
        HCONS <! qeq &
                 [ HARG #harg,
                   LARG #larg ] !> ],
FLAGS.EQUAL < #x > ].
```

Other transfer rules

- ==> Demo acm.tdl

What about features of indices?

- Can’t change value from input to output while maintaining identity of index with other positions.

- Person and number can be harmonized (in principle at least) by extending hierarchies on both sides, but we can’t harmonize between PERNUM and separate PER and NUM features.

- Tense and aspect (and others) can likewise be harmonized at least somewhat, but inventories vary greatly.

- Variable property mapping allows grammar-internal variable properties to differ from grammar-external universe.

    - We’ll use this for harmonization (e.g., of PERNUM) and setting of defaults.

MMT sentences: Preview

- Step through sentences one by one
- For each grammar, working or no?
- If not working, how do you tell where the point of failure is?
  - Intended output doesn’t parse
  - Search space too large
  - Need a trigger rule
  - MRS doesn’t quite match
  - Might we need a transfer rule?