



DELPHIN

Blocks, Bitse and pieces: Update on developments in Linköping

Fefor, June 16th 2006

Lars Ahrenberg
NLPLab
Department of Computer and Information
Science, Linköpings universitet

1

Overview

- Three projects 2005-2006
 - The NL1 project (*Martin Magnusson*)
 - Martin Magnusson (2006). [Natural Language Understanding using Temporal Action Logic](#). Proceedings of the workshop on *Knowledge and Reasoning for Language Processing (KRAQ'06)*, Trento, April 2006.
 - The BiTSE En-Se Grammar (*Sara Stymne*)
 - Sara Stymne and Lars Ahrenberg (2006). A bilingual grammar for translation of English-Swedish verb frame divergencies. Proceedings of EAMT 2006.
 - SweCore (*self*)

2

The PHLAI and NL1 project



Martin Magnusson.

“The aim of the NL1 project is the use of logic to integrate several artificial intelligence technologies in a question answering system and the long term goal of PHLAI* is the construction of a general artificial intelligence.”

*Project Hyper-Logistic Artificial Intelligence

3

Temporal Action Logic (TAL)

- Non-monotonic logic for time, action and change using an explicit time line
- Features concurrency and methods to deal with the frame, ramification, and qualification problems
- Fluents, formulas, actions, and narratives
- Circumscription in TAL is computable
- Translation to first-order logic

4

Dialogue with NL1

U1: B1 is on the table.

S1: OK

U2: Does the table support B2?

S2: I don't know.

U3: B2 is on the table and B3 is on B1.

S3: OK

U4: Does the table support B2?

S4: Yes.

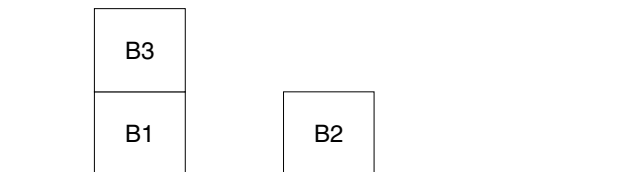
U5: What blocks are clear?

S5: B2 is clear and B3 is clear.

5

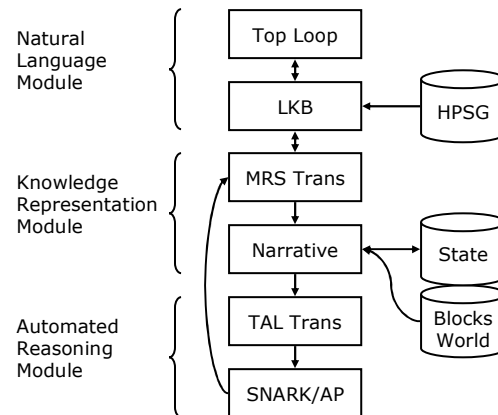
NL1 blocks world

State after S3



6

NL1



7

Knowledge Representation

- Translation between MRS and TAL
- World and state knowledge bases
- Imperative commands are added to the narrative
- Propositions and questions form reasoning problems

8

Automated Reasoning

- Compilation into first-order logic enables the application of regular theorem provers
 - SNARK by Mark Stickel
 - Allegro Prolog

9

Pros and Cons

- Advantages
 - Simplicity means easily modified
 - Generality of LKB, TAL, and theorem proving
- Disadvantages
 - Similar approaches have scalability problems

10

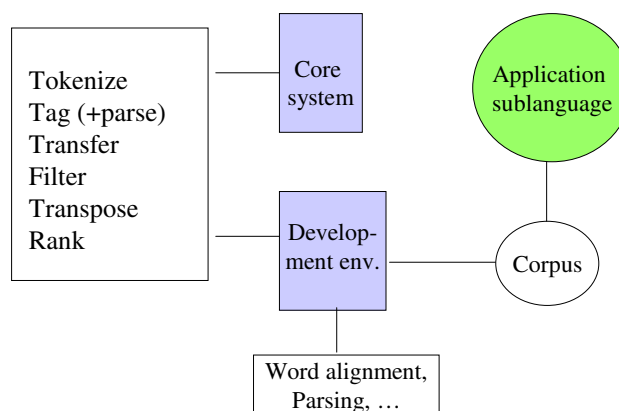
BiTSE: A bilingual English-Swedish grammar

■ Goals

- Explore potential of bilingual grammars using an interlingua as a methodology for MT
 - compare with semantic transfer a la LOGON
 - compare with lexical transfer a la T⁴F
- Focus on (and initial restriction to) verb frame divergencies
- Explore potential of building bilingual grammars for similar languages

11

The T⁴F framework



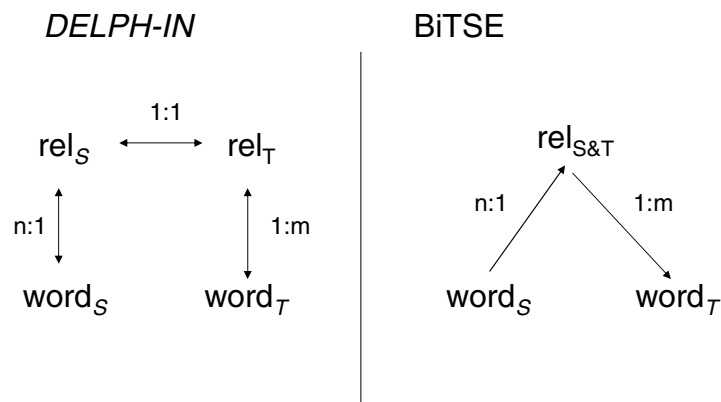
12

VFD examples

- Particles, prepositions and reflexives
 - *endure something / stå ut med något* (stand out with something)
 - *turn out / visa sig* (show itself)
- Support verbs
 - *gave no answer / svarade inte* (make decision)
 - *want something / vill ha något* (want have something)
- Head inversions
 - *finish packing / packa färdigt* (pack ready)
- Verb vs. copula + adjective
 - *owe / vara skyldig*
 - *be able to / kunna*

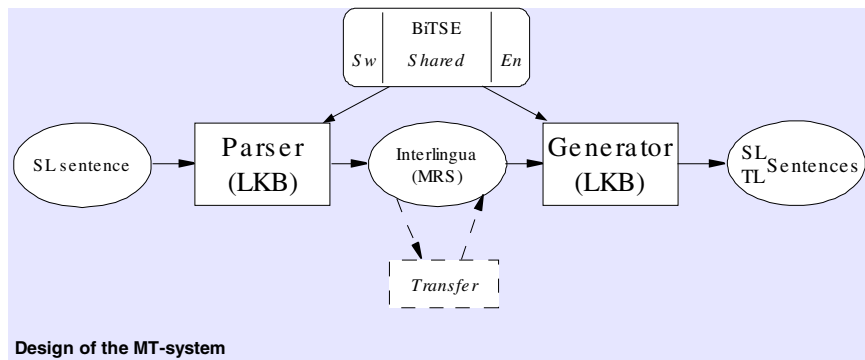
13

Non-compositional interpretation



14

BiTSE architecture



15

Implemented grammar

- Small parallel grammars for English and Swedish using Matrix v 0.8
 - verb frame differences as above
 - declarative main clauses and yes-no questions
 - main and subordinate clause word order
 - lexicon includes representative items only (~ 70 entries)
 - LANG-feature to separate the languages

swedish-only-rule := headed-phrase &
[SYNSEM.LANG #sw & sw,
HEAD-DTR.SYNSEM.LANG #sw].

swedish-only-lex := basic-lex &
[SYNSEM.LANG sw].

16

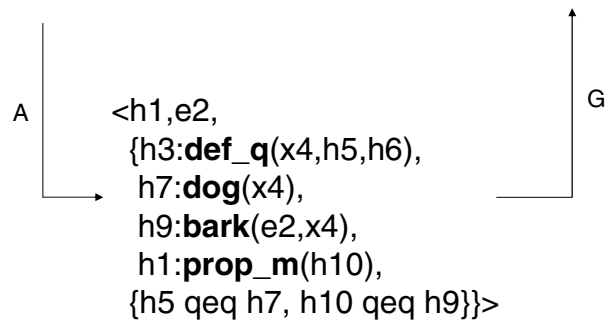
Size of the grammar

	Types
Shared	188
Swedish (most for inflections)	76
English	32

17

Simple example

The dog barks { The dog barks, Hunden skäller }



18

Sample solutions

- Structural divergencies
 - Employs empty prepositions, particles and reflexives
 - Preps and particles have a feature PFORM, used by other constituents to choose a correct prep./part.
 - PFORM-values are organised in an hierarchy
 - Empty particles and reflexives are on the comps list. Thus many new base types facilitating empty comps have to be introduced, e.g.
 - Basic-four-arg, basic-five-arg
 - Trans-empty2arg-lex-item, intrans-empty2-3ndarg-lex-item

19

Empty vs. non-empty items

	Empty	Non-empty
Reflexive pronoun	perjure <i>oneself</i>	kill <i>oneself</i>
Particle	throw <i>up</i> (vomit)	get <i>up</i>
Pronoun	depend <i>on</i> something	sit <i>on</i> something

20

Sample solutions

- Syntactic divergencies, e.g.

She told me a story

Hon berättade en historia för mig

requires no extra machinery

- Conflational divergencies, e.g.

He shaved [himself].

Han rakade sig.

The feature OPTTYPE is given a value **refl-opt** that licenses a reflexive relation being added when the optional object is removed and assures agreement with the subject.

21

SweCore

- Background in the Scandinavian Matrix Network (2003-2005)

- Create a Matrix for Scandinavian languages

- Create moderately large grammars for Danish, Norwegian and Swedish

- NorSource

- SweCore

- Danish (?)

- ...

22

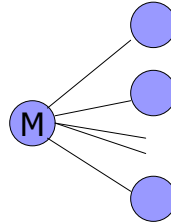
Some observations and conclusions

Matrix methodology

To model n languages,

- model what is common to them and then
- model what is specific for each of them

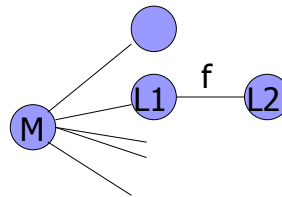
NOTE: The 'modules' project



Complementary methodology

To model a language,

- find a model of a similar language
- define a mapping from that language



23

The SweCore Grammar

- Based on NorSource
 - NorSource type and rule files with only very small modifications
 - Swedish inflectional morphology and lexicon
 - Very little development since last time

24

Some observations and conclusions

- NorSource rules and types generally worked very well for Swedish
 - Norwegian more varied than Swedish
 - delete and restrict rather than add new types and features
 - NorSource treatment of word order for verb complements found suboptimal

25

Current work

- Towards a first public release
- Core dictionary (~ 500 entries + dummies)
- Adopt *DELPH-IN* naming conventions
- Swedish MRS test suite (done!)

26