The LIMSI participation to the QAst track

Sophie Rosset, Olivier Galibert, Guillaume Bernard,
Eric Bilinski, Gilles Adda

Spoken Language Processing Group
LIMSI-CNRS
France

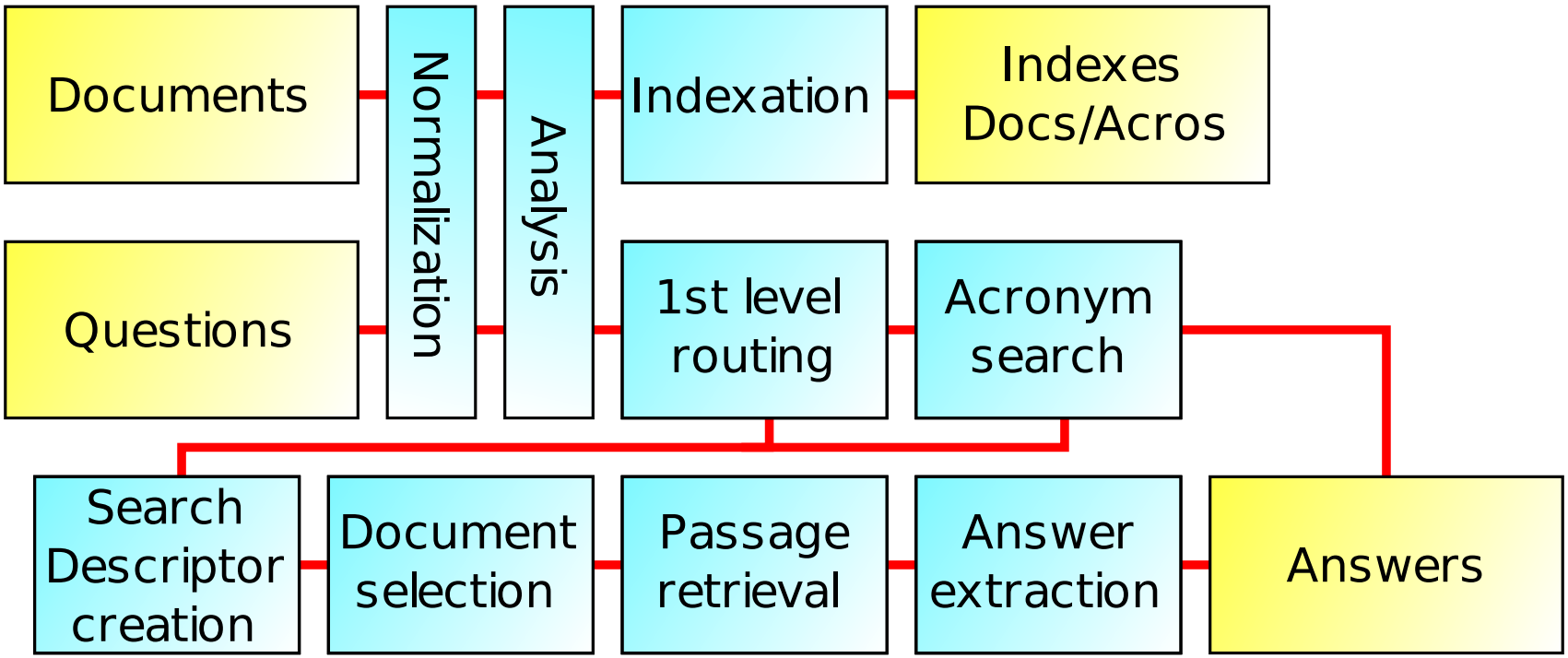
The QAst tasks

Monolingual Question Answering on manual and automatic speech transcriptions. 5 tasks, 16 sub-tasks:

1. English lectures (CHIL): manual + 1 ASR
2. English meetings (AMI): manual + 1 ASR
3. French broadcast news (ESTER): manual + 3 ASR
4. English european parliament (TCSTAR/EPPS): manual + 3 ASR
5. Spanish european parliament (TCSTAR/EPPS): manual + 3 ASR

Factoid (75%) and definitional (25%) questions, 100 questions per task.

System structure



Normalization

- Punctuation seperation
- Case reconstruction
- Punctuation addition
- Sentence splitting

Before

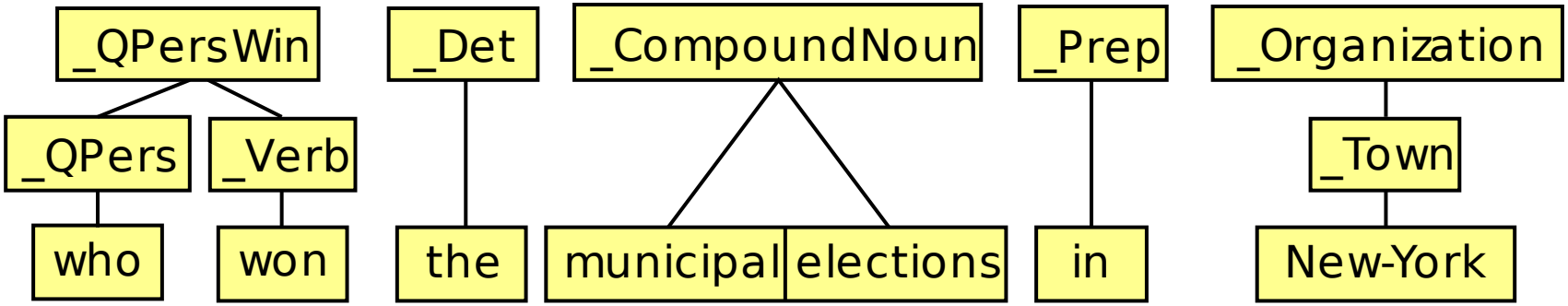
UM ACCORDING TO ROLLING STONES IT'S...
Yeah I dunno. But they're they're okay. Uh I think.

After

{fw} according to Rolling Stones it's ...
yeah I don't know . but they're they're okay . {fw} I think .

Documents and questions analysis

- Extract all the pertinent information we can detect.
  - Extended, hierarchical, standard and specific Named Entities
  - Linguistic chunks
  - Question words
- French system has 70+ passes of analyses with word-based regular expressions (≈1500 rules) plus POS.
- Spanish and English are partial *translations* of the French system, keeping the tag names.



Question answering

Question Routing

- Given an analysed question selects:
  - the search methodology to use
  - a rough categorisation of the expected answer type
- Only two methodologies for QAst: simple factoid search and predetected acronym lookup.
- ≈170 simple rules, 19 categories.

Factoid search

- Starts by building a *Search Descriptor*.
- The search descriptor describes what the search is about:
  - Key elements to find in the text
  - Possible variations on these elements (*query expansion*)
  - Possible answer types

Question answering

Search descriptor example

Question: *when was Hans Krasa killed?*

- Critical element
  - 1,0 *pers* identity(Hans Krasa)
  - 0,2 *pers* expand(Hans Krasa)
- Secondary element
  - 1,0 *verb* identity(killed)
  - 0,7 *verb* lemma(killed)
  - 0,5 *verb* synonym(killed)
  - 0,5 *subs* verb_subs(killed)
- Answer types
  - 1,0 *full_date*
  - 0,9 *month_year*, *day_month*, *hour*
  - 0,7 *year*

Question answering

Information Retrieval

- Score the documents using the SD element counts in them and the weights
- Extract snippets of the documents by extracting blocks of lines around the SD elements
- Score the snippets by their counts and the document score

Answer extraction and scoring

- All elements of the appropriate types are candidates
- The candidates are scored using their distances to the SD elements, the snippet scores, their occurence counts
- Uses a set of tuning constants optimized by trials using the development data

Results

Task		Acc.	Best
T1	manual	41%	-
	ASR	27%	31% UPC
T2	manual	33%	-
	ASR	16%	18% UPC
T3	manual	45%	-
	ASR A	41%	-
	ASR B	25%	-
	ASR C	21%	-

Task		Acc.	Best
T4	manual	33%	34% UPC
	ASR A	21%	30% INAOE
	ASR B	20%	-
	ASR C	19%	-
T5	manual	33%	-
	ASR A	24%	-
	ASR B	19%	-
	ASR C	23%	-

Conclusions and perspectives

- Decent results
- Separation of language-dependant and language-independant parts
- Multi-level integrated language analysis
- Specialized IR with integrated query expansion
- Fast (but QAst is small enough that everything is fast)

- Work being done for better answer scoring through tree editing distance
- Common tags set should simplify cross-lingualism