Towards an Error Correction Memory to Enhance Technical Texts Authoring in LELIE

Juyeon Kang, Patrick Saint-Dizier

IRIT-CNRS, Prometil, Toulouse, France

Motivations

- Technical documents are designed to be easy to read and as efficient and unambiguous as possible for their users and readers.
- They tend to follow relatively strict controlled natural language principles concerning both their form and contents.
- However, these principles are not always followed for various reasons, e.g. temporal constraints, technical level of writers, lack of understanding of CNL importance, etc.

→Aim: develop and test several facets of an **error correction memory** system that would, after a period of observation of technical writers making corrections, automatically propose corrections from the LELIE alerts:

(1) memorize errors which are not or almost never corrected so that they are no longer displayed in texts and

(2) memorize corrections and propose correction recommendations via generalizations and mediation.

Secondary aims

→Contributes to controlled natural language authoring and its natural evolution, whatever the application (e.g. learning from texts)

→Improves safety in procedures and requirements

→Allows or facilitates further controls on procedures (coherence, feasibility, etc.).

Our approach allows to revise texts a posteriori writen without any constraints or when using e.g. boilerplates / templates / chunks.

Situation

→ Starting point: LELIE: a system to check the quality of procedures (Barcellini, Saint-Dizier 2012), Implemented on <TextCoop> our NLP platform for processing discourse.

- CNL: (many refs) general principles, minimalism, guidelines (general or domain related), etc.
- Error correction memory originates principles from memory-based NLP (Daelemans et al. 2005): TiMBL, (Buchholz 2002) devoted to grammatical memory and generalizations. Memory-based systems are also used to resolve ambiguities, using notions such as analogies (Schriever et al. 1989).

- Finally, memory-based techniques are used in programming languages support systems to help programmers to resolve frequent errors.

→ Not yet much devoted to authoring systems.

Implementation in Dislog:
the TextCoop platform deszigned for discourse processing

- (1) **Dislog**, which is a logic-based language designed to describe in a declarative way discourse structures and the way they can be bound via selective binding rules,
- (2) **an engine associated with a set of processing strategies**. This engine offers several mechanisms to deal with ambiguity and concurrency
- (3) **a set of active constraints**, that state well-formedness typical language and of discourse
- (4) input-output facilities (XML, MS Word), and interfaces with other environments
- (5) a set of of lexical resources which are frequently used in discourse analysis (e.g. connectors),
- (6) a set of about 180 generic discourse analysis rules

The situation in LELIE

Lelie is rule-based with constraints anf filters. It produces alerts on lexical, grammatical style and business errors which do not follow recommendations of CNL or of a company.

However:

- (1) Lelie displays numerous false positives (about 25% of the alerts) which must be filtered out (e.g.: fuzzy terms, modals, passives, negation cannot be avoided in certain contexts) and
- (2) help must be provided to technical writers under the form of generic correction patterns paired with recommendations (domain and practice dependent) whenever possible since this is a difficult task.

→ Our approach is designed to be more flexible and adapted to the user needs and company context, compared e.g. to Rat-Rqa, Attempto, Rubric or Rabbit.

Example: Alert distribution in LELIE

error type	frequency / 1000 lines	A	B	C
fuzzy lexical items	66	44	89	49
modals in instructions	5	0	12	1
pronouns with unclear reference	22	4	48	2
negation	52	8	109	9
complex discourse structures	43	12	65	50
complex coordinations	19	30	10	17
heavy N+N or noun complements	46	58	62	15
passives	34	16	72	4
sentences too complex	108	16	221	24
incorrect references	13	33	22	2

Develop a 2-level method that shows how to construct:

- (1) relatively generic correction patterns paired with
- (2) accurate contextual correction recommendations, based on previously memorized and analyzed corrections.

→ Experiments in this paper on fuzzy lexical items

Exploring the case of fuzzy lexical items

A fuzzy lexical item denotes a concept whose meaning, interpretation, or boundaries can vary considerably according to context, readers or conditions, instead of being fixed once and for all.

(1) it is difficult to precisely define and identify what a fuzzy lexical item is, must be contrasted with:
- vague and
- underspecified expressions,
which involve different forms of corrections.

(2) there are several categories of fuzzy lexical items. These categories include:
- adverbs (manner, temporal, location, and modal adverbs),
- adjectives (adapted, appropriate)
- determiners (some, a few),
- prepositions (near, around),
- a few verbs (minimize, increase) and
- some nouns.

Categories are not homogeneous in terms of fuzziness:

- e.g. determiners and prepositions are always fuzzy in most context.
- the degree of fuzziness is also quite different from one term to another in a category.

Contrast definition of fuzziness with:

A verb such as *damaged* in *the mother card risks to be damaged*

is not fuzzy but **vague** because the importance and the nature of the damage is unknown;

heat the probe to reach 500 degrees

is not fuzzy but **underspecified** because the means to heat the probe are not given an adjunct is missing in this instruction.

→Correction strategies are different for vague and underspecified situations.

→The context in which a fuzzy lexical item is uttered may also have an influence on its severity level.

'progressively' used in a short action (*progressively close the water pipe*) or used in an action that has a substantial length (*progressively heat the probe till 300 degrees Celsius are reached*) may entail different severity levels.

→This motivates the need to memorize the context of the error to establish an accurate error diagnosis.

Observing technical writers at work

– What are the strategies deployed by technical writers when they see the alerts? what do they think of the relevance of each alert?

- How do they feel about making a correction? How much do they interact with each other ?

– Over large documents, how do they produce stable and homogeneous corrections?

– How much of the sentence is modified, besides the fuzzy lexical item? Does the modification affect the sentence content?

– How difficult is a modification and what resources does this requires (e.g. they spend about 50% of their time looking for external documentation, asking someone else for help, looking for similar situations (Barcellini et al. 2012))

- How many corrections have effectively been done? How many are left pending and why?

Some principles for a correction memory

- Corrections must take into account their utterance context,

– Corrections must result from a consensus among technical writers via mediation or an administrator.

- These corrections are then proposed in future correction tasks in similar situations.

– Corrections are directly accessible to technical writers: as a result, a lot of time is saved; furthermore, corrections become more homogeneous over the various documents of the company,

– Corrections reflect a certain know-how of the authoring habits and guidelines of a company, therefore they can be used to train novices.

The system: (1) Construction of a lexicon of fuzzy terms

category	number of entries	a priori severity level
manner adverbs	130	2 to 3
temporal and location adverbs	107	in general 2
determiners	24	3
prepositions	31	2 to 3
verbs and modals	73	1 to 2
adjectives	87	in general 1

(2) Memorizing corrections: database example

```
fuzzyitem([term], [category], [severity],
[[text fragment with alert, text after correction with tags,
ID of writer], ....] ).
```

For example:

```
fuzzyitem([progressively], [adverb], [3],
[[[<fuzzy>, progressively, </fuzzy>, heat, the, probe],
   [[heat, the, probe, <revised>, progressively,
              in, 5, seconds, </revised>]],  [John] ] .... ]
```

(3) Error correction memory scenarios

- (1) A fuzzy lexical item not corrected over several similar cases, within a certain word context or in general, no longer originates an alert.

- (2a) A fuzzy lexical item replaced or complemented by a value, a set of values or an interval, may originate, via generalizations, the development of correction patterns:
- *Progressively heat the probe → heat the probe progressively over a 2 to 4 mns period.*

→*Generic pattern (interval) + contextual recommendation (values)*

- (2b) In parallel with generalizing over corrections, the above item can be complemented by the observation of correctly realized utterances in the same context.

- (3) A fuzzy lexical item simply erased in a certain context (probably because it is judged to be useless, of little relevance or redundant): *proc. 690 used as a basic reference applicable to airborne → proc. 690 used as a reference....*
- (4) A fuzzy lexical item replaced by another term or expression in context that is not fuzzy, e.g. *aircraft used in normal operation → aircraft used with side winds below 35 kts and outside air temperature below 50 Celsius,*
- (5) A fuzzy lexical item may involve a complete rewriting of the sentence in which it occurs.

case nb.	number of cases	rate (%)
1	60	18
2	154	46
3	44	13
4	46	14
5	28	9

Taking into account the context of a correction: evaluating the size of the context

- Contexts are composed of nouns, verbs, adjectives that appear to the left or to the right of the term to be corrected.
- Important to consider to have a correct contextual analysis and correction recommendation.
- Experiments made on 332 situations, with contexts of various sizes, to evaluate stability of correction recommendations w.r.t. corrections:

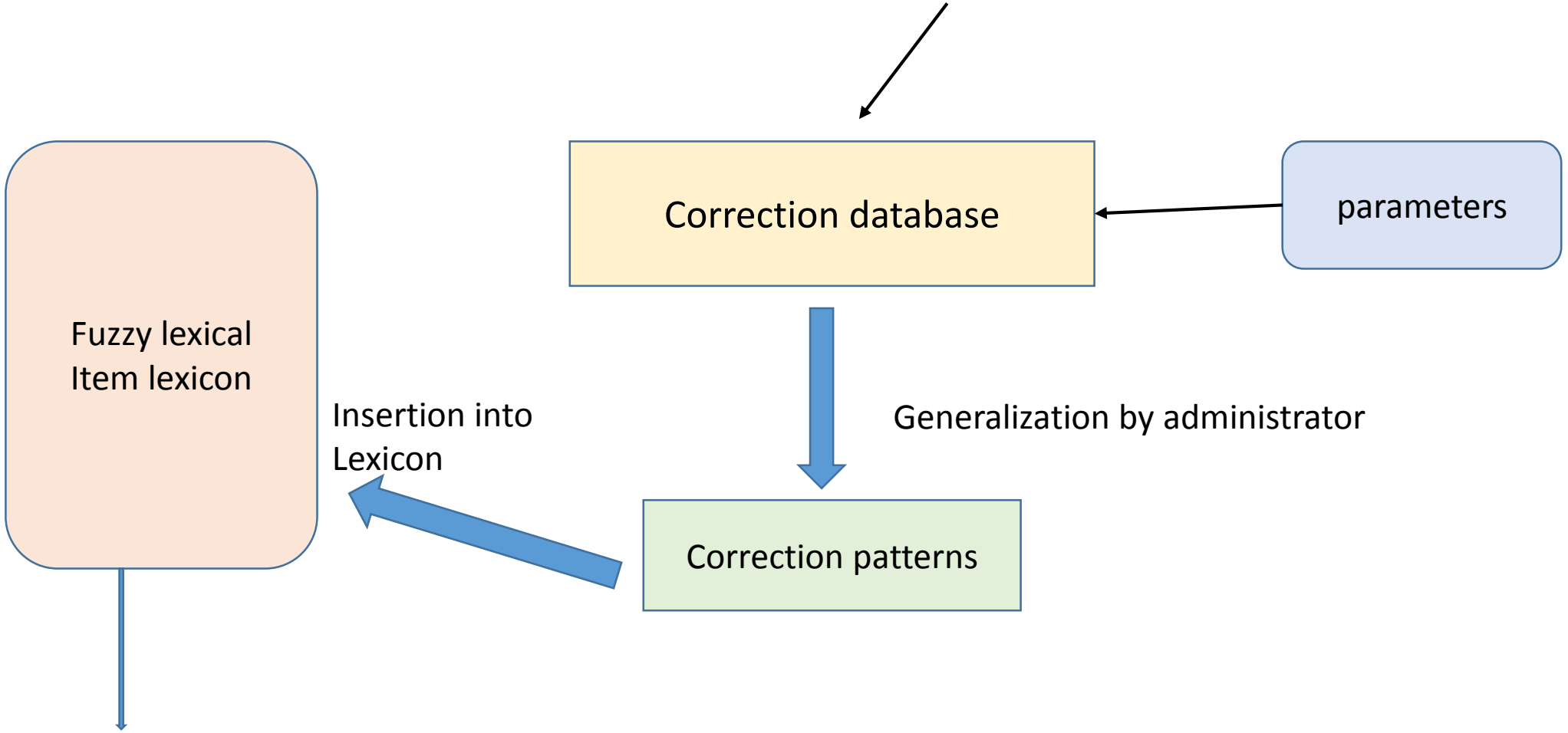
number of additional words	stability from previous set
3	85%
4	92%
5	94%

Observing the database of corrections

- For each entry (a fuzzy lexical item), define one or more patterns depending on context, then generalize over entries if possible,
- In a first stage, to experiment, patterns are defined manually to identify their nature, linguistic and conceptual structure and scope.
- This is confirmed by a technical writer – administrator, possibly via mediation with other writers
- Very much time consuming and error prone → needs at some stage to be partly automated from a set of preliminary patterns
- Patterns are included in the fuzzy lexical item lexicon together with their context and recommendations

Overall process

Technical texts/ alerts from LELIE / corrections by technical writers



Error correction patterns, simple samples: fuzzy determiners

(1) fuzzy determiners: specification of an upper or a lower boundary (N) or an interval, e.g. pattern:

[a few X] → [less than N X], [most X] → [more than N X].

- Besides patterns, which are generic, the context may induce a **correction recommendation** for the value of X: depending on X and its usage (context) a value for X can be suggested,

e.g. ’12’ *in: take-off a few knots above V1* → *take-off less than 12 knots above V1*,

with Context = main term: *a few knots*, additional: *take-off, above V1*.

Adverbs (temporal and manner)

- **temporal adverbs**, combined with an action verb, such as frequently, regularly: specification of a temporal value with an adequate quantifier, e.g.:

[regularly Action] → [every Time Action],

where Time is a variable that is instantiated on the basis of the context or the Action.

[progressively verb(durative)] → [progressively verb(durative) in Time],

e.g. progressively close the pipe → progressively close the pipe in 10 seconds.

Time is suggested by the correction recommendation level.

- **manner adverbs,** such as carefully which do not have any direct measurable interpretation, recommendation is:

(1) to produce a warning that describes the reasons of the care if there is a risk, or

(2) to explain how to make the action in more detail, via a kind of ’zoom in’, or

(3) to simply skip the adverb in case it is not crucial.

- For example, case (1):

[carefully Action] → [carefully Action Warning],

e.g. carefully plug-in the mother card → carefully plug-in the mother card otherwise you may damage the connectors.

Prepositions and adjectives

- **prepositions** such as near, next to, around, about require the specification of a value or an interval of values that depends on the context. A pattern is for example:

[near noun(location)] → [less than Distance from noun(location)], where Distance depends on the context, e.g. park

near the gate → park less than 100 meters from the gate.

- **adjectives** such as acceptable, convenient, specific as in a specific procedure,

e.g. *a convenient programming language*

can only be corrected via a short paraphrase of what the fuzzy adjective means.

Some challenging cases

- Whenever possible, when necessary the system shall operate…: rewrite the whole clause ?
- If the card is installed incorrectly then a message must be produced
- General corrosion should be detected …: temporal dimension ?
- Potential multiple states may occur and must be …
- Specific pulse spacing are defined on the basis of…
- When equivalent proofs can be defined, then….
- etc.

Perspectives

- Settled a framework for an error correction memory, tested on fuzzy terms, 27 non-overlapping patterns have been defined,
- Evaluate complexity of patterns in real cases
- Elaborate protocol method for evaluation with users: feasibility, usability, etc.
- Full implementation on top of LELIE, in TEXTCOOP ongoing, will be freely available in CC BY NC.
- Investigate other types of errors which can be treated similarly (e.g. negation, sentences too complex, etc.)
- Investigate other uses of this method for other applications (language simplification,etc.).