An Integrated Robust Parsing of Spontaneous Utterances

Yuji MATSUMOTO and Osamu IMAICHI
Graduate School of Information Science,
Nara Institute of Science and Technology
Takayama, Ikoma, Nara 630-01, JAPAN
e-mail: matsu@is.aist-nara.ac.jp

1. Introduction

Spontaneous speech has various kinds of ill-formed phenomena such as omissions, inversions, word fragments, interjections, speech repairs. In particular, word fragments, interjections, and speech repairs, which are called disfluencies, are peculiar to spoken language and do not occur in written language. Speech repair is a repair of an utterance in the middle by the speaker, and occurs when some of the words uttered by the speaker need to be removed in order to understand his/her intended meaning. Word fragments, which is a kind of speech repairs, occur in case where speech repair occurs in the middle of a word. Of spontaneous Japanese utterances, approximately 26speech repairs and 64construct spoken language systems, we cannot ignore such disfluencies.

Most of previous approaches to ill-formed inputs distinguish the process of speech repairs from the other linguistic processing and place it as the preprocessing of parsing. However, any processing should not be a pre-process of the other without clear reasons. This report proposes an integrated method for dealing with grammatical errors and disfluencies. The framework is based on cost-reward tradeoff. A cost is defined according to the degree of ill-formedness and a reward is defined by the effect of a phase to work for the eventual analysis of the input.

2. Integrated method based on cost and reward trade-off

2.1 Outline

The proposed method comprises a two-stage model consisting of the normal parsing process and the recovery process. When the normal parsing process has become stuck in finding a complete parse for the input, the recovery process is invoked. The parsing process employs the bottom up Chart parsing algorithm, and the recovery process can use partial parses for recovering from the parsing failure. The main mechanism of the recovery process is the selection of partial parses appropriate for error recovery. In order to select the most appropriate one, cost and reward are assigned to each partial parse. Cost indicates the degree of inconsistency included in a partial parse, and is calculated by the cost-based unification, while reward indicates the degree of their contribution to error recovery.

Ill-formed inputs are handled by three modules depending on the types of ill-formedness.

1. Module A handles local ill-formedness such as constraint violations.
2. Module B handles non-local ill-formedness such as word order violations.
3. Module C handles non-local ill-formedness such as incomplete sentential fragments and contextual ellipses.

Modules A and B try to find a phrase that would cover the whole input, whereas Module C tries to find an appropriate interpretation using contextual information. This report summarizes mainly the former two modules.

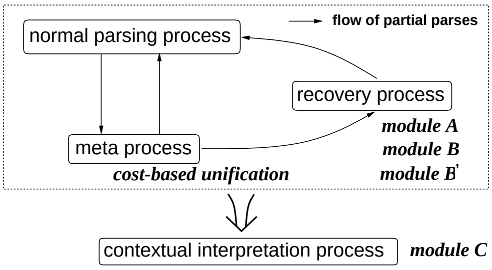

Fig.1:

(Fig. 1) shows the architecture and flow of partial parses employed in our system. There are three components: the normal parsing process, the meta-process, and the recovery process. The normal parsing process performs syntactic parsing. This process only constructs context-free skeletons and passes the intermediate results to the meta-process, which evaluates the inconsistency in partial parses using cost-based unification. The meta-process returns the results to the normal parsing process only when inconsistent information is not detected under cost-based unification. The meta-process passes all the partial parses to the recovery process.

If the normal parsing process succeeds in finding a complete parse, the recovery process is never invoked. Otherwise, the recovery process is invoked by the meta-process. The recovery process starts finding an appropriate partial parse according to the criteria for Modules A and B to be described below. It selects a partial parse, and passes it to the normal parsing process. Then the normal parsing process, which has been suspended, resumes its process taking into account the selected partial parse. This process terminates when the normal parsing process finds a phrase covering the whole input.

2.2 Recovery process

If the normal parsing process fails in finding a complete parse for the input, the recovery process is invoked and selects a partial parse that gives the most plausible explanation of the ill-formedness. The selection is based on the following criteria.

1. Selection criterion:

Select a partial parse with the maximal reward and the minimal cost.

The criteria of the reward are defined as follows:

Reward criterion A: Verb phrases are preferred to noun phrases and noun phrases are preferred to other phrases.

Reward criterion B: Phrases covering the widest range of the input are preferred. And, the ones at the righter position are preferred to others (in the case of Japanese).

The Reward criteria specify the ranking-order of the reward. If the partial parse with the maximal reward is not decided by the Reward criterion A, the Reward criterion B is applied. For example, if two verb phrases exist, the partial parse with the maximal reward is selected using the Reward criterion B. The verb phrase that covers the wider range of the input string and exists at the righter portion in the input string is selected.

Module A

When the selected partial parse is not used in the normal parsing process, Module A is invoked. It simply passes the partial parse to the normal parsing process, and then the normal parsing process resumes the suspended process. This corresponds to the relaxation of constraints violated in the partial parse.

Module B

When the selected partial parse has already been used in the normal parsing process, Module B is invoked. Module B uses the lexical expectation mechanism. That is, it uses the syntactic and semantic information described in the feature structure of the selected phrase. The selected phrase is called the seed phrase in Module B. Module B is defined by the following procedure.

Procedure of Module B

1. Lexical expectation:

   Looks at the syntactic and semantic information described in the seed phrase for finding the candidates that satisfy the requirement imposed by the seed phrase.

2. Searches for the candidates according to the search criterion.

   Search criterion: Search for phrases that the seed phrase requires.

3. There are two cases depending on whether the found phrase is complete or not.

   If complete: check if the phrase is suitable for the argument of the seed phrase, combines this phrase and the seed phrase, and passes the result to the normal parsing process.

   If incomplete: the phrase becomes the next seed phrase, and call this procedure recursively.

It is important to notice that Modules A and B are defined independently but work in an interleaving manner. That is, they are controlled by the principle of cost and reward trade-off that is defined in the search criterion.

3. Module for Disfluencies

We introduce a new module called Module B' for dealing with disfluencies such as interjections, repetition, speech repairs and so on. In order to detect and correct speech repairs in an utterance, it is necessary to identify the reparandum and repair parts in the utterance.

In a repaired utterance, a fragment in the reparandum part often matches some words in the repair part.

In example 1, the word "koto" appears in both the reparandum and the repair part. Thus the repetition of a word/phrase is an important clue for detecting speech repairs.

Example: kaigi no koto, eto, touroku no koto de ukagai tai.

We use another clues for detecting speech repairs. In transcribed texts from spoken language, the information on pause is often indicated by using a comma. When the length of the duration of pause between two words has some length, a comma is inserted between the two words by the transcriber. In the example, a comma is inserted between the reparandum and disfluency parts, and between the disfluency part and the repair part. Thus we use it as a clue for detecting the interruption point, because a comma might indicate the interruption point of speech repairs. In addition, we use interjections as a clue that might indicate the interruption point.

In order to process these types of speech repairs, it is necessary to delete the disfluency part and the reparandum part. The main mechanism of Module B' is a skipping operation. By posing a cost in the skipping operation, this module is naturally integrated into the framework based on cost and reward trade-off.

The cost is defined by the easiness of skipping. Similar words are repeated in speech repairs, both in morphologically and in semantically. Accordingly, we define the similarity measure of the two phrases; one is the skipped phrase and the other is the selected phrase. The similarity is evaluated from morphological, syntactic, and semantic viewpoints. A cost is assigned to the new phrase that is constructed by skipping a word/phrase. When the degree of the similarity between the skipped phrase and the selected phrase is high, the phrase constructed by the skipping operation receives a small cost. Moreover, the easiness of skipping is also defined by the length of the skipped words.

Module B' is invoked when the selected partial parse is already used and is preceded by a comma or an interjection. The procedure of Module B' is defined as follows.
Procedure of Module B'

1. Interjection:

   The preceding word or fragment is skipped at a low cost defined for each interjection.

2. Repetition:

   The morphological similarity is evaluated and the word is skipped and is given a cost on the basis of the similarity.

3. Speech repairs:

   The morphological, syntactic, and semantic similarities between the reparandum and repair parts are evaluated and the reparandum part is skipped on the basis of the similarity. The cost is calculated according to the similarity.

4. Apposition:

   The semantic similarity between the apposition part and the selected phrase is evaluated and is given a cost based on the similarity.

New partial parses with a cost that are constructed in Module B' are used when Module A is invoked. Like other modules, Module B' is controlled by the principle of cost and reward trade-off.

Consider the above example. The normal parsing process leaves the phrase 'kaigi no koto', the word 'eto,' and the phrase 'touroku no koto de ukagai tai'. Using the selection criterion, the recovery process selects the phrase 'touroku no koto de ukagai tai". Since this phrase is preceded by a comma, Module B' is invoked. Module B' skips the word 'eto' as an interjection and builds the new partial parse. This new partial parse, which receives a new cost, is passed to the normal parsing process. Since this phrase has a cost, it is not used in the normal parsing process. When the recovery process is invoked again, the recovery process selects the phrase '(eto), touroku no koto de ukagai tai' using the selection criterion. The Module B' is invoked with this phrase because it is also preceded by a comma. Module B',then, skips the noun phrase "kaigi no koto" as speech repairs, and the new phrase is passed to the normal parsing process. Since this phrase receives another cost, it is not used in normal parsing. Then, the recovery process selects the phrase that covers the whole input, and then Module A passes this phrase to the normal parsing process. Since the phrase that covers the whole input is found, the recovery process terminates successfully.

4. Conclusion

This paper proposed a new integrated framework for handling grammatically ill-formed inputs. It deals with wide phenomena of ill-formedness on a single principle of cost-reward trade-off. The following modules are introduced for handling various types of ill-formedness.

Module A handles local ill-formedness such as constraint violations.

Module B handles non-local ill-formedness such as word order violations.

Module B' handles local ill-formedness such as disfluencies.

Module C, which is not described in this report, handles non-local ill-formedness such as incomplete sentential fragments and contextual ellipses.

We have integrated these modules into a uniform framework on the basis of the notion of cost and reward. Cost represents the degree of inconsistency included in a partial parse, and reward represents the goodness of a partial parse. Though each module alone is not enough to deal with various types of ill-formedness, the integration is done in a straightforward way.

Keywords: grammatically ill-formed sentence, robust parsing, integration, spontaneous speech