Evaluation of the Role Analysis Task

NTCIR Workshop TMREC Group
(Teruo Koyama, Masaharu Yoshioka, Koichi Takeuchi, & Kyo Kageura)
Research and Development, NACSIS
{koyama,yoshioka,koichi,kyo}@rd.nacsis.ac.jp

Abstract: In this paper we summarise the results of the evaluation of the role analysis task. Two groups submitted one result file each. The evaluation was carried out for randomly selected samples, based on manually extracted answer sets.

1 Introduction

The evaluation of the role analysis task was carried out by matching the extracted descriptive structure with answer sets that were manually constructed by the TMREC group. In this process, we randomly selected 90 documents for evaluation.

During the evaluation process, we discovered that our original task description had an ambiguity which caused different understandings of the task. Because of that, the results of each group reflect their own understanding, which differs from that of the other group, making it difficult to compare them.

Therefore, we first clarify the task description. Then, we describe the criteria used for making the answer sets and evaluation results. Finally we discuss issues relating to the role analysis task and possible future work.

2 Clarification of the Task Description

Figure 1 shows the original description of the role analysis task. While evaluating the results, we found that the task description was ambiguous and thus the following two different understandings were possible.

1. Extract main descriptive structure and represent it in the form of triplets (This is what we intended in this task.)
2. Extract the description that matches the form of triplets.

In order to make the discussion of this paper clearer, we revised the task description shown in Figure 2 to reflect our intended understanding. In the following, we use this description as the basis for the evaluation. For this reason, the results of the 'evaluation' does not show the superiority or inferiority of the submitted results.

3 Making the Answer Sets

In order to make the answer sets from the documents, it was necessary to give a guideline to extract

> The main descriptive structure of academic documents consists of the following elements:
> 1. what they treat (main subject)
> 2. by what means (means or method)
> 3. what is done (procedure or action applied to the main subject)
>
> For instance, a document may treat the topic that can be described as "the construction of educational system by means of learners' profile analysis", whose constituents are [{educational system (main subject)}, {learners' profile analysis (means)}, and {construction (procedure)}]
> The role analysis task aims at extracting terms with these three roles from the title and the documents. The five possible triplets may be extracted.

Figure 1: Task description of the role analysis task

the "main description of an academic document." So, we made the following guidelines for defining the criteria.

1. Classify the description into the following three types.

   Background description Description of the problem definition, previous work and so on.

   Main description Description of the topics that are most emphasised in the paper.

   Derivative description Description that is derived from the main description or other derivative descriptions.

2. Consideration of the limitation of modern NLP systems
   Because of the limitation, we decided not to use the following information to make the answer sets.

   - Do not use contextual information such as the resolution of demonstrative pronouns.

The role analysis task aims to extract the main descriptive structure from academic documents (title and abstract). The five possible structures may be extracted in the following representational form. Representational form of the descriptive structure: The main descriptive structure of academic documents consists of the following elements:

1. what they treat (main subject)
2. by what means (means or method)
3. what is done (procedure or action applied to the main subject)

For instance, a document may treat a topic that can be described as "the construction of an educational system by means of learner profile analysis", whose constituents are [{educational system (main subject)}, {learner profile analysis (means)}, and {construction (procedure)}]

Figure 2: Revised task description of the role analysis task

- Do not use dependency information within the combined words.
- Since the extraction of combined words is another research issue, we will not consider how each element in the triplet form is represented.

Based on this guideline, we defined the following criteria for making the answer sets.

- We excluded background descriptions from the answer sets.
- We set two description levels for the answer sets; "main" and "derivative." For the "derivative" answer, we only dealt with "derivatives" from the "main" description.
- We only extracted the root element of the combined word for each element in the triplet form. However, for words that were frequently used in the documents such as "こと (that)", we extracted longer elements to clarify the reference.
- For some verbs such as "述べる (describe)", and "提案する (propose)", it is better to extract the previous part of these verbs as answer sets in some cases. Therefore, we extracted both of them in such cases and represented the dependency by using the "-" symbol.

Example :

「システムの構築について述べる (we describe the construction of the system)」
1. 構築 - 述べる
(construction) - (describe)
-1. システム - 構築する
(system) - (construct)

- When means or methods were described in different sentences, we did not extract them for the descriptive structure.
- We did not resolve demonstrative pronouns.
- When some topics were described in parallel, we expanded each topic with triplet forms and treated all the descriptions in the same description level.
- We normalised verb phrases into standard forms.
- We did not split combined words (especially 'sahen' verb) to extract the descriptions.
- We did not limit the number of manually generated answers to 5. Each answer was placed in appearance order.

Table 1 shows the extracted result from 90 documents by the TMREC group.

Table 1: Manually extracted answer sets

Main description	397
Derivative description	126

4 Evaluation

We defined the following criteria for evaluating the results.

- The results of the evaluation are categorised into three categories by taking into account the description level and the correspondence of each element in the triplet form (Table 2).
- The triplet form was checked by observing whether each element in the answer set was included or not. In the case of parallel description, if we found the correspondence between the element in the answer set and the result easily, we regarded them as correspondent.
- When means or methods were described in different sentences, we only evaluated the description by itself. We did not consider the relation with the original description.

Table 3 shows the results of the evaluation. Since some descriptions were extracted more than once in the result files, recall and precision were not calculated.

Table 2: Evaluation results category

	Correspondent	Partly correspondent	Not correspondent
Main description	○	△	×
Derived description	△	×	×
Other description	×	×	×

Table 3: Evaluation result for each group

	A	B
○	163	78
△	92	83
×	92	97

5 Discussion and future work

Since this was the first attempt to organise a workshop for this task, we came up against several problems. The ambiguity of task description was a serious problem, which made it difficult to compare the results. In addition, even though we proposed criteria for the role analysis task, we did not have good consensus on several concepts relating to this task; e.g. definition of "main description".

In this paper, we proposed our definition and extracted example results based on it. We consider our definition to be just a starting point in discussing the possible directions of the role analysis task. So, in future we would like to discuss not only the methodology for extracting the role description but also the framework of the role analysis task itself.