EAI Endorsed Transactions
on Context-Aware Systems and Applications
Research Article ICST.ORG

An approach for summarization of two-sentences Vietnamese paragraph

Trung Tran[1,*] and Dang Tuan Nguyen[1]

[1]Faculty of Computer Science, University of Information Technology, VNU-HCM, Ho Chi Minh City, Vietnam

Abstract

The purpose of this paper is to introduce a general approach for summarizing the meaning of Vietnamese paragraphs based on simple two-sentences. The studied objects are paragraphs having the common characteristics: the first sentence has one or two nouns indicating human objects; the second sentence has one or two anaphoric pronouns. We only consider two types of Vietnamese human pronouns in this research: the pronouns standing alone in the sentence; the pronouns standing with demonstrative adjective in the sentence. At the first phase, depending on the context of pronouns in the second sentence, we propose appropriate strategies to find the exact human object at the first sentence which is referred to by each pronoun. A discourse structure is also built to represent the meaning of each paragraph. At the second phase, each discourse representation will be transformed to a syntactic structure of meaning-summarizing sentence. The final phase complete the new sentence of summarization.

Keywords: inter-sentential anaphoric pronoun, referent resolution, discourse representation, meaning summarization, sentence generation.

Received on 09 January, 2016, accepted on 29, March, 2016, published on 02 May, 2016

doi: 10.4108/eai.2-5-2016.151210

1. Introduction

In the state of the art of the field document summarization ([7, 9, 10, 12, 13, 16, 17, 19, 20, 23]), according to the type of the summary, there are two different directions in which many research groups followed and tried to create a good summary: (i) the first direction called "extractive summarization" in which the researchers applied methods and techniques in machine learning to extract the most important sentences in the original text and combined these ones to build the output paragraphs; (ii) the second direction called "abstractive summarization" in which the researchers had to propose the methods to understand and represent the semantic of the source text as well as generate the new summary. The main purpose of this article is to present our new approach in this field, follow the "abstractive summarization" direction with the combination of anaphora resolution, text understanding, meaning representation and sentence summarization. To transform the original paragraph to the complete new summary, we propose a new solution containing two main phases: (i) phase one is to understand and represent the semantic of the source text; (ii) phase two is to generate the new sentence from the representation in phase one. These two phases help for answering in general three important parts which K. S. Jones ([12, 13]) set out for every summarization system: (a) create the first representation from the original text; (b) transform the first representation to the second representation of the summary; (c) generate the summary from the second representation.

In this article, base one the idea of T. Tran and D. T. Nguyen ([27, 28]), we apply the solution to Vietnamese paragraphs which compose two simple sentences. These

This paper is extended version of two preliminary papers [27, 28] which were presented at the 4th EAI International Conference on Context-Aware Systems and Applications (ICCASA 2015), November 26 – 27, 2015, Vung Tau, Vietnam.

*Corresponding author. Email: ttrung@nlke-group.net

two-sentences paragraphs have in common the characteristic: the first sentence has one or two nouns indicating human objects; the second sentence has one or two anaphoric human pronouns. We only consider two types of Vietnamese human pronouns in this research: type one includes pronouns standing alone in the sentence; type two includes pronouns standing with one demonstrative adjective [“ta” / “ấy” / “này”] in the sentence.

With the above characteristic, based on the ideas of T. Tran and D. T Nguyen ([25, 27]), at the first phase of our approach, we propose the appropriate finding strategies to determine the exact antecedent which is the noun indicating human object in the first sentence for each pronoun appearing in the second sentence. To implement these strategies, with the foundation is framework Graph Unification Logic Programming (GULP [4]), the proposed method contains following steps:

- Step one: we split the input Vietnamese paragraph into two separate sentences. Then we set a syntactic characteristic about the position of each sentence in the original text.
- Step two: with each sentence from step one, we analyse the syntactic structure. In each node of the syntactic tree, we apply the method of Unification-based Grammar (UBG [24]) about transferring the information up and down to define the appropriate grammatical characteristics. These characteristics help for resolving the anaphoric pronouns as well as building the meaning representation of the original paragraph.
- Step three: at the level of lexicons in the syntactic tree from step two, there are two actions which will be taken depending on the kind of lexicon: (a) if the lexicon is noun, verb or adjective, then we define the appropriate syntactic and semantic characteristics which will be used to find the antecedent for each pronoun and build the meaning representation of the original paragraph; (b) if the lexicon is pronoun, then we apply the algorithm using the characteristics of lexicon and larger constituents to resolve.

The result of phase one after three above steps is a meaning representation of the original paragraph. In T. Tran and D. T Nguyen ([25, 26, 27, 28]) and this article, we represent the semantic of Vietnamese paragraph by a structure call Discourse Representation Structure (DRS). According to Discourse Representation Theory ([1, 5, 6, 15]), we have to build two ordered list in order to complete a DRS: (a) the first list called U in which we in turn put unique numbers which indicate different human objects; (b) the second list called Con in which we in turn put suitable predicates. Each predicate in list Con has one or two arguments which are the numbers in list U, indicates that this is the condition which each human object has to satisfy.

At the second phase of our approach, based on the ideas in T. Tran and D. T. Nguyen ([26, 28]), we propose methods to transform the DRS to a complete new Vietnamese sentence called meaning-summarizing sentence. These methods are included in three following steps:

- Step one: we analyse the DRS, determine the information which express the semantic: all numbers in list U, semantic predicates in list Con which are predicates express the semantic of nouns, verbs and adjectives.
- Step two: with the information from step one we determine two types of relationships: the inner relationship between object and verb or adjective inside one sentence; the inter relationship between two objects or object with verb or adjective at different sentence. Depending on the relationships, we apply the algorithm to generate the syntactic structure of the new meaning-summarizing sentence.
- Step three: we combine the syntactic structure from step two with the appropriate lexicons in the built set to complete the new sentence.

Although following the idea in the field natural language generation (Reiter and Dale [21, 22]), our approach for generating new sentence has some different points in comparison with the traditional ways: (i) the input and output of the system are complete paragraph and sentence; (ii) reuse the set of lexicons of the input text to generate the output summary. With this approach, the output sentence is more natural for the native speakers.

The rest of this article is classified into following sections. Section 2 describes in details groups of two-sentences Vietnamese paragraphs which are considered in this research. In Section 3, we present the strategies and implementations of three steps of phase one. The methods for implementing phase two will be presented in Section 4. Next, we describe in Section 5 the experiment as well as indicate some remaining points of our approach. Finally, we conclude this paper and point out some future improvements in Section 6.

2. Classify two-sentences Vietnamese paragraphs

The main content of this section is to express in detail each group of pairs of simple Vietnamese sentences which are considered in this research. As mentioned in Section 1, all these groups have the similar common characteristics: there are one or two nouns in the first sentence and one or two pronouns in the second sentence. Analyzing further, we classify the considered paragraphs into four groups depending on: the number of noun in the first sentence; the number of pronouns in the second sentence; two human objects are identical or not.

2.1. Group A

Base on the idea of T. Tran and D. T. Nguyen ([25]), we consider in this research some two-sentences paragraphs which are included in group A. In each member, the first and second sentence has the characteristics as follows:

Characteristics of the first sentence

In the first sentence of each paragraph, there is only one noun indicating human object. As a simple sentence, this object can take the subject role of one intransitive verb or adjective.

- The human object takes the subject role of an intransitive verb.

Example 1: "*David mỉm cười.*"
(English: David smiles.)

- The human object takes the subject role of an adjective.

Example 2: "*Peter vui vẻ.*"
(English: Peter is happy.)

Characteristics of the second sentence

We consider simple Vietnamese sentences in which there is the appearance of only one anaphoric human pronoun. This pronoun can belong to one of two types: type one includes pronouns which stand alone; type two includes pronouns which stand with a demonstrative adjective ["ta" / "ấy" / "này"]. The sentences having these characteristics are also the second sentence of each Vietnamese paragraph which we consider in group B (Section 2.2) and group C (Section 2.3).

- The pronoun takes the subject role of copula "là" which means is identical with a human object.

Example 3: "*Anh / Anh ta là John.*"
(English: He is John.)

- The pronoun takes the object role of copula "là" which means is identical with a human object.

Example 4: "*Hiệu trưởng là ông / ông ấy.*"
(English: The principle is him.)

- The pronoun takes the subject role of an adjective.

Example 5: "*Cô / Cô ấy xinh đẹp.*"
(English: She is beautiful.)

- The pronoun takes the subject role of an intransitive verb.

Example 6: "*Anh / Anh ta đứng dậy.*"
(English: He stands up.)

- The pronoun takes the subject role of a transitive verb.

Example 7: "*Ông / Ông ấy dạy môn toán.*"
(English: He teaches maths.)

- The pronoun takes the object role of a transitive verb.

Example 8: "*Johan nói chuyện với cô / cô ấy.*"
(English: Johan talks to her.)

2.2. Group B

In T. Tran and D. T. Nguyen ([28]) and this research, we consider some special paragraphs in which: the first sentence is formed by identical relationship with copula "là"; the second sentence belongs to one of the types mentioned in Section 2.1.2.

With the first sentence, the general characteristic is that there are two nouns indicating human objects are identical and connected by a copula "là". Analyze further, we categorize into four types depending on the context in which each noun is the common or proper noun. Each unique context helps us propose the appropriate finding strategy to resolve the anaphoric ambiguity.

- The proper noun takes the subject role and the common noun takes the object role of copula "là".

Example 9: "*Susan là y tá.*"
(English: Susan is a nurse.)

- The proper noun takes the object role and the common noun takes the subject role of copula "là".

Example 10: "*Giám đốc là ông Kim.*"
(English: The chief officer is Mr Kim.)

- There are two proper nouns.

Example 11: "*Mary là bà Lopez.*"
(English: Mary is Mrs Lopez.)

- There are two common nouns.

Example 12: "*Cậu bé là đội trưởng.*"
(English: The kid is the captain.)

2.3. Group C

As mentioned in Section 2.1.2, the second sentence of each member in group C has the characteristic belong to one of the types which similar to group A.

With the first sentence, the common considered characteristic in T. Tran and D. T. Nguyen ([25, 26]) and this research is: there are two nouns indicating two different human objects. These two objects are not identical and connected by a transitive verb.

Example 13: "*Jim đấu với Bill.*"
(English: Jim fights Bill.)

2.4. Group D

As the special group, based on the idea in T. Tran and D. T. Nguyen ([28]), each paragraph has the general characteristic: the first sentence has two human objects which are not identical; the second sentence has two pronouns belonging to different types. In the second sentence: the pronoun standing with a demonstrative adjective takes the subject role of the transitive verb; the pronoun standing alone takes the object role of the transitive verb.

To determine the exact antecedent for each pronoun in the second sentence, we firstly establish a presupposition about the property of each transitive verb appearing at the first and second sentence. Based on the reality experience in using Vietnamese, we see that depending on the context in which the event happens, one transitive verb can have the property "affect" or "communication". With this presupposition, we categorize the considered paragraphs into four smaller groups:

- Group D1: the pair of transitive verbs has the pair of properties "affect" – "affect".

Example 14: "*Paul đánh nhau với Joe. Anh ta đánh anh.*"
(English: Paul fight Joe. He hits him.)

- Group D2: the pair of transitive verbs has the pair of properties "affect" – "communication".

Example 15: "*Cô y tá tiêm cho Mary. Cô ấy cảm ơn cô.*"
(English: The nurse injects Mary. She thanks her.)

- Group D3: the pair of transitive verbs has the pair of properties "communication" – "affect".

Example 16: "*Ông Smith nói chuyện với bác sĩ. Ông ta khám cho ông.*"
(English: Mr Smith talks to the doctor. He examines him.)

- Group D4: the pair of transitive verbs has the pair of properties "communication" – "communication".

Example 17: "*Luật sư thảo luận với bị cáo. Ông ấy bào chữa cho ông.*"
(English: The lawyer discusses with the accused. He defends him.)

3. Understand and represent the semantic of two-sentences Vietnamese paragraphs

At the first main part of this section, we present the general strategies for determining the exact antecedent for each pronoun appearing in the second sentence. With each group, depending on the context of the paragraph, we have the suitable strategy. Then, we present in detail the implementation of these strategies through three main steps of phase one of our approach as mentioned in Section 1. The main content of this section is also based on the idea in T. Tran and D. T. Nguyen ([25, 26, 27, 28])

3.1. The general finding strategies

The proposed strategies here are commonly based on the idea: depending on the different context, use the constraints which are grammatical characteristics of nouns, transitive verbs, intransitive verbs, adjectives, pronouns and the position of each sentence.

The strategy for group A

With the paragraphs in which there is only one noun indicating human object at the first sentence and only one human pronoun at the second sentence, we establish the finding strategy as follows:

- Firstly identify the position of each sentence.
- Then determine the only human object at the first sentence is the antecedent for the only human pronoun at the second sentence.

The strategy for group B

As the special context in which there are two identical human objects in the first sentence, we acknowledge that there is only one object is actually referred to by the only human pronoun at the second sentence. To resolve this ambiguity, we propose the finding strategy according to the reality experiences: the main object in the paragraph is often mentioned first or indicated by a proper noun.

With the context in which the first sentence has one or two proper noun, the general strategy is:

- If there is only one proper noun, then we determine that this one is the antecedent of the only human pronoun at the second sentence.
- If there are two proper nouns, either one can be the antecedent of the only human pronoun at the second sentence. We determine the first object which takes the subject role of copula "là".

With the context in which the first sentence does not have any proper noun, the general strategy is: we determine the antecedent is the common noun taking the subject role of copula "là".

The strategy for group C

Because there are two human objects which are not identical in the first sentence but there is only one human pronoun in the second sentence, the finding strategy here is based on the type of the pronoun. According to the common using in Vietnamese, with the only one pronoun then the referred object should be the one as follows:

- If the pronoun stands alone, then it refers to the human object which takes the subject role of the transitive verb in the first sentence.
- If the pronoun stands with a demonstrative adjective, then it refers to the human object which takes the object role of the transitive verb in the first sentence.

The strategy for group D

Differ from group C, the second sentence of each member in group D has two human pronouns belonging to two different types. To determine the exact antecedent for each pronoun, we base on the presupposition in Section 2.4 about the property of each transitive verb. With this presupposition, we have two comments about the relationship between object and transitive verb: (i) commonly, one object does two consecutive actions which have the same properties which are "affect" or "communication"; (ii) if two transitive verbs have different properties, then commonly the object taking the subject role of the transitive verb in the first sentence is referred to by the pronoun standing alone, and the other object is referred to by the pronoun standing with a demonstrative adjective.

With two above comments, we establish the finding strategy for each smaller group as follows:

(i) The finding strategy for group D1: because both of transitive verbs have the property "affect" then according to comment (i) above, one object does both these actions. Therefore, we determine that:

- The human object taking the subject role of the transitive verb at the first sentence is the antecedent of the human pronoun standing with a demonstrative adjective.
- The human object taking the object role of the transitive verb at the first sentence is the antecedent of the human pronoun standing alone.

(ii) The finding strategy for group D2: because two transitive verbs have different properties then according to comment (ii) above, these actions are performed by different objects. Therefore, we determine that:

- The human object taking the subject role of the transitive verb at the first sentence is the antecedent of the human pronoun standing alone.
- The human object taking the object role of the transitive verb at the first sentence is the antecedent of the human pronoun standing with a demonstrative adjective.

(iii) The finding strategy for group D3: because two transitive verbs have different properties then according to comment (ii) above, these actions are performed by different objects. Therefore, we determine that:

- The human object taking the subject role of the transitive verb at the first sentence is the antecedent of the human pronoun standing alone.
- The human object taking the object role of the transitive verb at the first sentence is the antecedent of the human pronoun standing with a demonstrative adjective.

(iv) The finding strategy for group D4: because both of transitive verbs have the property "communication" then according to comment (i) above, an object does both these actions. Therefore, we determine that:

- The human object taking the subject role of the transitive verb at the first sentence is the antecedent of the human pronoun standing with a demonstrative adjective.
- The human object taking the object role of the transitive verb at the first sentence is the antecedent of the human pronoun standing alone.

3.2. Methods for implementing the strategies and creating the meaning representations

As mentioned in Section 1, we propose the methods for implementing the strategies in Section 3.1 with the foundation of framework GULP ([4]). These implementations is for phase one of our approach (see Section 1).

The method for implementing the step one of phase one

At the first step of phase one, we split the original paragraph into two separate sentences and set the position with value `[first]` or `[second]` for each one.

```
discourse(D) -->   {
        S1 = syn~flag_position~[first],
        D = sem~in~A,
        S1 = sem~in~A,
        S1 = sem~out~B,
        S2 = syn~flag_position~[second],
        S2 = sem~in~B,
        S2 = sem~out~C,
        D = sem~out~C
},
statement(S1),
endpunct,
statement(S2),
endpunct.
```

Figure 1. Split the original paragraph into two separate sentences based on framework GULP (Source: T. Tran and D. T. Nguyen [25])

The method for implementing the step two of phase one

To implement the second step of phase one, we firstly analyse each sentence into smaller constituents depending

on the syntactic structure. According to the characteristics of the first and second sentence of all groups which are mentioned in Section 2, there are two structures of considered sentences:

- The first structure: Sentence → Noun Phrase + Verb Phrase / Adjective Phrase. With this structure, we define Pphrase (denoted by PP) for both Verb Phrase and Adjective Phrase in the implementation.

```
s(S,H1,H3) -->  {
        NP = sem~A,
        S = sem~A,
        PP = sem~C,
        NP = sem~scope~C,
        NP = syn~flag_index~D,
        PP = syn~flag_arg1~D,
        NP = syn~flag_role~[subject],
        S = syn~flag_position~E,
        NP = syn~flag_position~E,
        PP = syn~flag_position~E
},
np(NP,H1,H2), pp(PP,H2,H3).
```

Figure 2. Analyse the sentence with the first structure based on framework GULP

As mentioned in Section 1 about applying methods of UBG ([24]) to transfer the information up and down in the syntactic tree, when analyze the sentence in Fig. 2, we describe some special grammatical characteristics as follows. The variable `flag_index` take the value which is the unique index of the human object taking the subject role of the verb or adjective. This value will be transferred up from the description of lexicon and then set to noun phrase. This value is also set to the variable `flag_arg1` which is the first argument of the verb or adjective. The next variable `flag_role` takes value `[subject]` which shows that this noun phrase and then the noun takes the subject role of the verb or adjective. The value of variable `flag_position` is transferred down from the description of paragraph.

- The second structure: Sentence → Noun Phrase + "là" + Noun Phrase.

```
s(S,H1,H3) -->  {
        S = syn~flag_position~FP,
        NP1 = syn~flag_position~FP,
        NP2 = syn~flag_position~FP,
        NP1 = syn~flag_role~[subject],
        NP2 = syn~flag_role~[object],
        S = sem~A,
        NP1 = sem~A,
        NP2 = sem~B,
        NP1 = sem~scope~B,
        NP1 = syn~flag_index~I1,
        NP2 = syn~flag_index~I2,
        NP2 = sem~scope~(in~[drs(U,Con)|Super] ..
        out~ [drs(U,[(I1=I2)|Con])|Super])
},
np(NP1,H1,H2), [là], np(NP2,H2,H3).
```

Figure 3. Analyze the sentence with the second structure based on framework GULP (Source: T. Tran and D. T. Nguyen [28])

Similar to the description in Fig.2, we set the suitable value for each grammatical characteristic when describing sentence in Fig. 3.

With Pphrase, there are two descriptions for this kind of phrase depending on the characteristic of the original paragraph and finding strategy.

- With the first three groups, because the finding strategies (see Section 3.1) do not depend on the grammatical structure of Pphrase, then we have the description as in Fig. 4.

```
pp(PP,H1,H2) -->        {
        P = syn~D,
        PP = syn~D,
        NP = sem~A,
        PP = sem~A,
        NP = syn~flag_index~C,
        PP = syn~flag_arg2~C,
        NP = syn~flag_role~[object],
        PP = syn~flag_position~E,
        NP = syn~flag_position~E,
        P = sem~B,
        NP = sem~scope~B
},
p(P), np(NP,H1,H2).
```

Figure 4. Analyse the Pphrase in first three group based on framework GULP

In the description in Fig. 4, we set the suitable value for each grammatical variable as follows. The variable `flag_index` takes the value is the unique index of the noun taking the object role of the transitive verb. This value will be also set to the variable `flag_arg2` which is the second argument of the transitive verb. The variable `flag_role` takes value `[object]` which shows that this noun phrase and then the noun takes the object role of the transitive verb.

- For implementing the finding strategy of group D (see Section 3.1), we add two more grammatical variables in the description of Pphrase in Fig. 5. The variable `flag_property_of_verb` takes value which is transferred up from the description of verb and then set to noun phrase. The variable `flag_index_other` takes value when determine the antecedent of two pronouns, then set to the variable `flag_arg1` of verb. Two variable `flag_index` and `flag_index_other` in turn takes the indexes of the antecedents of the second and first pronoun.

```
pp(PP,H1,H2) -->        {
        PP = syn~flag_position~E,
        NP = syn~flag_position~E,
        P = syn~flag_position~E,
        P = syn~D,
        PP = syn~D,
```

```
        NP = sem~A,
        PP = sem~A,
        NP = syn~flag_index~C,
        PP = syn~flag_arg2~C,
        NP = syn~flag_index_other~G,
        PP = syn~flag_arg1~G,
        NP = syn~flag_role~[goal],
        PP = syn~flag_property_of_verb~F,
        NP = syn~flag_property_of_verb~F,
        P = sem~B,
        NP = sem~scope~B
},
p(P), np(NP,H1,H2).
```

Figure 5. Analyze the Pphrase in group D based on framework GULP (Source: T. Tran and D. T. Nguyen [27])

With noun phrase, we have the same description for all finding strategies of four groups (see Section 3.1) in Fig. 6.

```
np(NP,H,H) -->  {
        NP = syn~flag_position~F,
        N = syn~flag_position~F,
        NP = syn~flag_role~G,
        N = syn~flag_role~G,
        N = syn~A,
        NP = syn~A,
        N = sem~B,
        NP = sem~res~B,
        NP = sem~in~C,
        NP = sem~res~in~C,
        NP = sem~res~out~D,
        NP = sem~scope~in~D,
        NP = sem~scope~out~E,
        NP = sem~out~E
},
n(N).
```

Figure 6. Analyze the noun phrase based on framework GULP (Source: T. Tran and D. T. Nguyen [25])

The methods for implementing the step three of phase one

At step three of phase one, we firstly consider the type of lexicon and determine which action should be performed. If the lexicon is noun, verb or adjective, then we describe the grammatical characteristics with appropriate variables. If the lexicon is pronoun, then we apply the suitable algorithm to find the antecedent. In both cases, the implementation also help for building the meaning representation DRS of the source text: the unique index of each noun will be added to list U; the predicates of each lexicon will be added to list Con.

With the noun indicating human object, there are two descriptions according to the finding strategies in Section 3.1.

- With the strategies in the first three groups, we focus on the constraints including: position in the paragraph; role in the relationships with verb or adjective; common or proper noun. Therefore we define the corresponding variables in the description in Fig. 7.

```
n(N) --> [david],{
        append([position(I,FP),role(I,FR),
                f_proper(I,FPR),
                named(I,[david],noun,proper)],
                Con,NewCon),
        unique_integer(I),
        FPR = [proper],
        N = syn~(flag_index~I ..
                flag_position~FP ..
                flag_role~FR ..
                flag_proper~FPR) ..
        sem~    (in~DRSList ..
                out~ NewDRSList)
}.
```

Figure 7. Describe proper noun “David” (example 1) in first three groups based on framework GULP

- With the strategy in group D, we focus on the constraints including: position in the paragraph; role in the relationship with the transitive verb; the property of each transitive verb. Therefore we define the corresponding variables in the description in Fig. 8. We add two more variables: `flag_property_of_verb` to indicate the property of the related transitive verb; `flag_index_other` with the notice that when determine the antecedent for one pronoun, then the other object is the antecedent of the remaining pronoun. The value of `flag_index` and `flag_index_other` are different.

```
n(N) --> [paul],{
        append([position(I,FP),role(I,FR),
                paul(I,CO,CAT)],Con,NewCon),
        unique_integer(I),
        CO = [paul],
        CAT = [object],
        N = syn~(flag_index~I ..
                flag_position~FP ..
                flag_property_of_verb~FPOV ..
                flag_index_other~FIO ..
                flag_role~FR) ..
        sem~     (in~[drs(U,Con)|Super] ..
                out~[drs([I|U],NewCon)|Super])
}.
```

Figure 8. Describe proper noun “Paul” (example 14) in group D based on framework GULP

With the intransitive verb, this type of lexicon appears in the first three groups (see Section 2). We have the same description in Fig. 9 for each member of this type because the constraints in finding strategies (see Section 3.1) do not include these grammatical characteristics.

```
p(P) --> [đứng,dậy],{
        append([đứng_dậy(Arg,[đứng,dậy],
                verb,intransitive)],
                Con,NewCon),
        P = syn~(flag_arg1~Arg) ..
        sem~(in~ [drs(U,Con)|Super] ..
                out~ [drs(U,NewCon)|Super])
```

```
}.
```

Figure 9. Describe intransitive verb "đứng dậy" (example 6) in the first three groups based on framework GULP

With the adjective, this type of lexicon appears in the first three groups (see Section 2). We have the same description in Fig. 10 for each member of this type because the constraints in finding strategies (see Section 3.1) do not include these grammatical characteristics.

```
p(P) --> [xinh,đẹp],{
        append([xinh_đẹp(Arg,
                [xinh,đẹp],adjective)],
                Con,NewCon),
        P = syn~(flag_arg1~Arg) ..
        sem~(in~ [drs(U,Con)|Super] ..
                out~ [drs(U,NewCon)|Super])
}.
```

Figure 10. Describe adjective "xinh đẹp" (example 5) based on framework GULP

With the transitive verb, we propose two descriptions depending on the different context of groups in Section 2 and finding strategies in Section 3.1.

- In Fig. 11 we describe a transitive verb in the first three groups because the constraints which are used in the finding strategies (see Section 3.1) do not depend on these grammatical characteristics.

```
p(P) --> [dạy], {
        append([dạy(Arg1,Arg2,[dạy],
                verb,transitive)],
                Con,NewCon),
        P = syn~(flag_arg1~Arg1 ..
                flag_arg2~Arg2) ..
        sem~(in~ [drs(U,Con)|Super] ..
                out~ [drs(U,NewCon)|Super])
}.
```

Figure 11. Describe transitive verb "dạy" (example 5) in the first three groups based on framework GULP

- We present in Fig. 12 the description of a transitive verb in group D. Because of the finding strategy (see Section 3.1), we add one variable and one predicate. Variable `flag_property_of_verb` indicates the property of the transitive verb. This variable takes value `[affect]` or `[communication]` depending on the corresponding property. Predicate `property_of_verb` takes two arguments: the first argument indicates the position of transitive verb in the paragraph; the second argument indicates the property of the transitive verb.

```
p(P) --> [đánh], {
        append([property_of_verb(FP,FPOV),
        đánh(Arg1,Arg2,CO,CAT,FCLASS,FPOV)],
                Con,NewCon),
        unique_integer(I),
        CO = [đánh],
        CAT = [verb],
        FCLASS = [transitive],
        FPOV = [affect],
        P = syn~(flag_arg1~Arg1 ..
                flag_arg2~Arg2 ..
                flag_position~FP ..
                flag_property_of_verb~FPOV) ..
        sem~(in~ [drs(U,Con)|Super] ..
                out~ [drs(U,NewCon)|Super])
}.
```

Figure 12. Describe transitive verb "đánh" (example 14) in group D based on framework GULP

With the pronouns, according to the constraints in each finding strategy, we propose the suitable algorithm and corresponding implementation for determine the antecedent.

- According to the strategy for group A (see Section 3.1), we have the algorithm and implementation to find the antecedent for the only one pronoun at the second sentence:

Algorithm 1: Determine the antecedent for the only pronoun with finding strategy in group A.

Consider DRS structure at current time;
Step 1: Check object having index `I` and value of predicate `position(I)`
- Check the value is `[first]`

Step 2: The antecedence of the only pronoun is object having index `I`
- Set feature `flag_index` of second pronoun value `I`;

```
np(NP,H,H) -->  ([anh];[anh,ấy]),{
        NP=sem~in~DrsList,
        member(drs(U,Con),DrsList),
        member(Index,U),
        member(position(Index2,[first]),Con),
        member(f_proper(Index2,[proper]),Con),
                Index == Index2,
        NP=syn~flag_index~Index,
        NP=sem~scope~in~DrsList,
        NP=sem~scope~out~DrsOut,
        NP=sem~out~DrsOut
}.
```

Figure 14. Implement Algorithm 1 based on framework GULP

- According to the strategy for group B (see Section 3.1), we have two algorithms as well as corresponding implementations. The first algorithm (Alg. 2) determine the antecedent is the proper noun at the first sentence. The second algorithm (Alg. 3) determine the antecedent is the noun taking the subject role of copula "là" at the first sentence.

Algorithm 2: Determine the antecedent which is the proper noun for the only pronoun with finding strategy in group B.

Consider DRS structure at current time;
Step 1: Check object having index `I` and value of predicate `position(I), f_proper(I)`
- Check these values are `[first], [proper]` respectively.

Step 2: The antecedence of the only pronoun is object having index `I`
- Set feature `flag_index` of this pronoun value `I`;

```
np(NP,H,H) -->  ([anh];[anh,ấy]),{
        NP=sem~in~DrsList,
        member(drs(U,Con),DrsList),
        member(Index,U),
        member(position(Index2,[first]),Con),
        member(f_proper(Index2,[proper]),Con),
                Index == Index2,
        NP=syn~flag_index~Index,
        NP=sem~scope~in~DrsList,
        NP=sem~scope~out~DrsOut,
        NP=sem~out~DrsOut
}.
```

Figure 15. Implement Algorithm 2 based on framework GULP

Algorithm 3: Determine the antecedent which is the noun taking the subject role of copula "là" for the only pronoun with finding strategy in group B.

Consider DRS structure at current time;
Step 1: Check object having index `I` and value of predicate `position(I), role(I)`
- Check these values are `[first], [subject]` respectively.

Step 2: The antecedence of the only pronoun is object having index `I`
- Set feature `flag_index` of this pronoun value `I`;

```
np(NP,H,H) -->  ([anh];[anh,ấy]),{
        NP=sem~in~DrsList,
        member(drs(U,Con),DrsList),
        member(Index,U),
        member(position(Index2,[first]),Con),

        member(role(Index2,[subject]),Con),
                Index == Index2,
        NP=syn~flag_index~Index,
        NP=sem~scope~in~DrsList,
        NP=sem~scope~out~DrsOut,
        NP=sem~out~DrsOut
}.
```

Figure 16. Implement Algorithm 3 based on framework GULP

- According to the strategy for group C (see Section 3.1), we have two algorithms as well as corresponding implementations. The first algorithm (Alg. 4) determine the antecedent is the noun taking the subject role of the transitive verb at the first sentence. The second algorithm (Alg. 5) determine the antecedent is the noun taking the object role of the transitive verb at the first sentence.

Algorithm 4: Determine the antecedent which is the noun taking the subject role of the transitive verb for the only pronoun standing alone with finding strategy in group C.

Consider DRS structure at current time;
Step 1: Check object having index `I` and value of predicate `position(I), role(I)`
- Check these values are `[first], [subject]` respectively.

Step 2: The antecedence of the only pronoun is object having index `I`
- Set feature `flag_index` of the only pronoun value `I`;

```
np(NP,H,H) -->  ([anh]),{
        NP=sem~in~DrsList,
        member(drs(U,Con),DrsList),
        member(Index,U),
        member(position(Index2,[first]),Con),

        member(role(Index2,[subject]),Con),
                Index == Index2,
        NP=syn~flag_index~Index,
        NP=sem~scope~in~DrsList,
        NP=sem~scope~out~DrsOut,
        NP=sem~out~DrsOut
}.
```

Figure 17. Implement Algorithm 4 based on framework GULP

Algorithm 5: Determine the antecedent which is the noun taking the object role of the transitive verb for the only pronoun standing with a demonstrative adjective with finding strategy in group C.

Consider DRS structure at current time;
Step 1: Check object having index `I` and value of predicate `position(I), role(I)`
- If group A then these values are `[first], [object]` respectively.

Step 2: The antecedence of the only pronoun is object having index `I`
- Set feature `flag_index` of the only pronoun value `I`;

```
np(NP,H,H) -->  ([anh,ấy]),{
        NP=sem~in~DrsList,
        member(drs(U,Con),DrsList),
        member(Index,U),
        member(position(Index2,[first]),Con),

        member(role(Index2,[object]),Con),
                Index == Index2,
        NP=syn~flag_index~Index,
        NP=sem~scope~in~DrsList,
```

```
        NP=sem~scope~out~DrsOut,
        NP=sem~out~DrsOut
}.
```

Figure 18. Implement Algorithm 5 based on framework GULP

- According to the strategy for group D (see Section 3.1), we have the algorithm (Alg. 6) with the main idea: when determine the antecedent for the pronoun standing alone and set the index of this object for variable `flag_index`, we simultaneously determine the antecedent for the remaining pronoun and its index becomes the value of variable `flag_index_other`.

Algorithm 6: Determine the antecedents for pronouns with finding strategy in group D.

Consider DRS structure at current time;

Step 1: Check property of the second transitive verb

- Variable `flag_property_of_verb` of second pronoun takes this value.

Step 2: Check property of the first transitive verb

- The second argument of predicate `property_of_verb()` of the first transitive verb takes this value.

Step 3: Check object having index `I` and value of the second argument of predicate `position(I), role(I)`

- If group D1 then these values are `[first]`, `[object]` respectively.
- If group D2 then these values are `[first]`, `[subject]` respectively.
- If group D3 then these values are `[first]`, `[subject]` respectively.
- If group D4 then these values are `[first]`, `[object]` respectively.

Step 4: The antecedence of the second pronoun is object having index `I`

- Set feature `flag_index` of second pronoun value `I`;
- Transfer this value to feature `flag_arg2` of the second verb;

Step 5: Check object having index `I_other` and

- `I_other` is different from `I`

Step 6: The antecedence of the first pronoun is object having index `I_other`

- Set feature `flag_index_other` of second pronoun value `I_other`;
- Transfer this value to feature `flag_arg1` of the second verb;

```
np(NP,H,H) -->  ([anh]),{
        NP=sem~in~DrsList,
        NP=syn~flag_property_of_verb~FPOV,
        FPOV == [affect],
        member(drs(U,Con),DrsList),
        member(property_of_verb([first],[affect])
,Con),
        member(Index,U),
        member(position(Index2,[first]),Con),

        member(role(Index2,[object]),Con),
                Index == Index2,
        NP=syn~flag_index~Index,
        member(Index_Other,U),
                Index_Other \= Index2,
        NP=syn~flag_index_other~Index_Other,
        NP=sem~scope~in~DrsList,
        NP=sem~scope~out~DrsOut,
        NP=sem~out~DrsOut
}.
```

Figure 19. Implement Algorithm 6 based on framework GULP

After resolving the ambiguity for all the pronouns at the second sentences, the final result of phase one is a meaning representation DRS of the original paragraph. This DRS will becomes the input of phase two when we select the main predicates of list Con to generate the syntactic structure of the new meaning-summarizing sentence.

4. Generate Vietnamese sentence for meaning summarization of two-sentences paragraph

The main content of this section is based on the ideas in T. Tran and D. T. Nguyen ([26, 27, 28]). We present steps in phase two of our approach from analyzing the DRS of each original text to generating the new complete Vietnamese sentence.

As mentioned in Section 1, at the first steps of phase two, we analyse the output DRS from phase one and determine the main information which actually represent the main content of the source paragraph. In our approach, the main information includes: all unique numbers in list U; the predicates indicate semantic of noun, verb, and adjective. With the selected information, we can determine the inner and inter relationship in order to generate the syntactic structure of the new Vietnamese sentence.

4.1. Generate the new sentences from the representations in group A

As an example, we consider the DRS with main information in Fig. 20 of paragraph in example 18.

Example 18: “David hạnh phúc. Anh tìm thấy mỏ vàng.”

(English: David is happy. He found the gold-mine.)

```
[david,hạnh,phúc,.,anh,tìm,thấy,mỏ,vàng,.]
[1,2]
named(1,[david],noun,proper)
hạnh_phúc(1,[hạnh,phúc],adjective)
mỏ_vàng(2,[mỏ,vàng],noun,common)
tìm_thấy(1,2,[tìm,thấy],verb,transitive)
```

Figure 20. The DRS with main predicates of two-sentences Vietnamese paragraph "David hạnh phúc. Anh tìm thấy mỏ vàng."

We analyze the relationships when considering the semantic predicates of adjective and transitive verb in Fig. 20:

- With predicate `hạnh_phúc()`, we have the inner relationship: `1` → `[hạnh,phúc]`.
- With predicate `tìm_thấy()`, we have the inter relationship: `1` → `[tìm,thấy]` → `2`.

To generate the syntactic structure of the new Vietnamese sentence, we propose the algorithm in Alg. 7.

Algorithm 7: Generate the syntactic structure of new sentence for group A.

Step 1: Consider the first relationship
- Put the semantic predicate of the first object into the syntactic structure.
- Put the semantic predicate of the first adjective or intransitive verb into the syntactic structure.

Step 2: Consider the second relationship
- Put "và" (and) into the syntactic structure.
- If the second predicate is the semantic predicate of adjective or intransitive verb: put this predicate into the syntactic structure.
- If the second predicate is the identical predicate: (i) put "là" (is) into the syntactic structure; (ii) put the semantic predicate of the second object into the syntactic structure.
- If the second predicate is the semantic predicate of transitive verb:
  - If the first object takes the subject role: (i) put this predicate into the syntactic structure; (ii) put the semantic predicate of the second object into the syntactic structure.
  - If the second object takes the object role: (i) put "được" (is – passive voice) into the syntactic structure; (ii) put the second predicate into the syntactic structure; (iii) put "bởi" (by) into the syntactic structure; (iv) put the semantic predicate of the second object into the syntactic structure.

Applying Alg. 7 to the above relationships, we have the syntactic structure:

```
named(1) + hạnh_phúc(1) + "và" + tìm_thấy(1,2) +
                    mỏ_vàng(2)
```

To complete the new sentence, we in turn replace each predicate in the syntactic structure by the corresponding Vietnamese lexicon. With the above syntactic structure, we have the complete sentence:

"David hạnh phúc và tìm thấy mỏ vàng."
(English: David is happy and found the gold-mine.)

4.2. Generate the new sentences from the representations in group B

As an example, we consider the DRS with main information in Fig. 21 of paragraph in example 19.

Example 19: "Ông Kim là giám đốc. Ông ấy tài năng."
(English: Mr Kim is the chief officer. He is talented.)

```
[ông,kim,là,giám,đốc,.,ông,ấy,tài,năng,.]
[1,2]
named(1,[ông,kim],noun,proper)
giám_đốc(2,[giám,đốc],noun,common)
1=2
tài_năng(1,[tài,năng],adjective)
```

Figure 21. The DRS with main predicates of two-sentences Vietnamese paragraph "Ông Kim là giám đốc. Ông ấy tài năng."

We analyze the relationships when considering the semantic predicates of adjective and identical in Fig. 21:

- With predicate `1=2`, we have the inner relationship: `1 = 2`.
- With predicate `tài_năng()`, we have the inner relationship: `1` → `[tài,năng]`.

To generate the syntactic structure of the new Vietnamese sentence, we propose the algorithm in Alg. 8.

Algorithm 8: Generate the syntactic structure of the new sentence for group B.

Step 1: Consider the index appearing in both relationships.
- Put the semantic predicate of noun which have this index into the syntactic structure.
- Put "là" (is) into the syntactic structure.
- Put the semantic predicate of the other noun into the syntactic structure.

Step 2: Consider the second relationship
- Put "và" (and) into the syntactic structure.
- If the second predicate is the semantic predicate of adjective or intransitive verb: put this predicate into the syntactic structure.
- If the second predicate is the identical predicate: (i) put "là" (is) into the syntactic structure; (ii) put the semantic predicate of the third object into the syntactic structure.
- If the second predicate is the semantic predicate of transitive verb:
  - If the first or second object takes the subject role: (i) put this predicate into the syntactic structure; (ii) put the semantic predicate of the third object into the syntactic structure.
  - If the first or second object takes the object role: (i) put "được" (is – passive voice) into

the syntactic structure; (ii) put the second predicate into the syntactic structure; (iii) put "bởi" (by) into the syntactic structure; (iv) put the semantic predicate of the third object into the syntactic structure.

Applying Alg. 8 to the above relationships, we have the syntactic structure:

```
named(1) + "là" + giám_đốc(2) + "và" +
                tài_năng(1)
```

To complete the new sentence, we in turn replace each predicate in the syntactic structure by the corresponding Vietnamese lexicon. With the above syntactic structure, we have the complete sentence:

"Ông Kim là giám đốc và tài năng."
(English: Mr Kim is the chief officer and talented.)

4.3. Generate the new sentences from the representations in group C

As an example, we consider the DRS with main information in Fig. 22 of paragraph in example 20.

Example 20: "Jim biết bà Susan. Bà ta là nông dân."
(English: Jim knows Mrs Susan. She is a farmer.)

```
[jim,biết,bà,susan,.,bà,ta,là,nông,dân,.]
[1,2,3]
named(1,[jim],noun,proper)
named(2,[bà,susan],noun,proper)
biết(1,2,[biết],verb,transitive)
nông_dân(3,[nông,dân],noun,common)
2=3
```

Figure 22. The DRS with main predicates of two-sentences Vietnamese paragraph "Jim biết bà Susan. Bà ta là nông dân."

We analyze the relationships when considering the semantic predicates of transitive verb and identical in Fig. 22:

- With predicate `biết()`, we have the inner relationship: `1` → `[biết]` → `2`.
- With predicate `2=3`, we have the inter relationship: `2 = 3`.

To generate the syntactic structure of the new Vietnamese sentence, we propose the algorithm in Alg. 9.

Algorithm 9: Generate the syntactic structure of the new sentence for group C.

Step 1: Consider the first relationship
- Put the semantic predicate of the first noun into the syntactic structure.
- Put this predicate into the syntactic structure.
- Put the semantic predicate of the second noun into the syntactic structure.

Step 2: Consider the second relationship
- If the second predicate is the semantic predicate of adjective or intransitive verb:
  - If the first noun takes the subject role: (i) put "và" (and) into the syntactic structure; (ii) put this predicate into the syntactic structure.
  - If the second noun takes the subject role: put this predicate into the syntactic structure.
- If the second predicate is the identical predicate:
  - If the first noun takes the subject role: (i) put "và" (and) into the syntactic structure; (ii) put "là" (is) into the syntactic structure; (iii) put the semantic predicate of the third noun into the syntactic structure.
  - If the second noun takes the subject role: (i) put "là" (is) into the syntactic structure; (ii) put the semantic predicate of the third noun into the syntactic structure.
- If the second predicate is the semantic predicate of transitive verb:
  - If the first noun takes the subject role: (i) put "và" (and) into the syntactic structure; (ii) put this predicate into the syntactic structure; (iii) put the semantic predicate of the third noun into the syntactic structure.
  - If the first noun takes the object role: (i) put "và" (and) into the syntactic structure; (ii) put "được" (is – passive voice) into the syntactic structure; (iii) put the second predicate into the syntactic structure; (iv) put "bởi" (by) into the syntactic structure; (v) put the semantic predicate of the third object into the syntactic structure.
  - If the second noun takes the subject role: (i) put this predicate into the syntactic structure; (ii) put the semantic predicate of the third noun into the syntactic structure.
  - If the second noun takes the object role: (i) put "được" (is – passive voice) into the syntactic structure; (ii) put the second predicate into the syntactic structure; (iii) put "bởi" (by) into the syntactic structure; (iv) put the semantic predicate of the third object into the syntactic structure.

Applying Alg. 9 to the above relationships, we have the syntactic structure:

```
named(1) + biết(1,2) + named(2) + "là" +
                nông_dân(3)
```

To complete the new sentence, we in turn replace each predicate in the syntactic structure by the corresponding Vietnamese lexicon. With the above syntactic structure, we have the complete sentence:

"Jim biết bà Susan là nông dân."
(English: Jim knows Mrs Susan is the farmer.)

4.4. Generate the new sentences from the representations in group D

As an example, we consider the DRS with main information in Fig. 23 of paragraph in example 17.

```
[luật,sư,thảo,luận,với,bị,cáo,.,ông,ấy,bào,chữa,
cho,ông,.]
[1,2]
luật_sư(1,[luật,sư],[object],[human])
bị_cáo(2,[bị,cáo],[object],[human])
thảo_luận_với(1,2,[thảo,luận,với],[action],[tran
sitive],[communication])
bào_chữa_cho(1,2,[bào,chữa,cho],[action],[transi
tive],[communication])
```

Figure 23. The DRS with main predicates of two-sentences Vietnamese paragraph "Luật sư thảo luận với bị cáo. Ông ấy bào chữa cho ông."

We analyze the relationships when considering the semantic predicates of transitive verb and identical in Fig. 23:

- With predicate `thảo_luận_với()`, we have the inner relationship: `1` → `[thảo,luận,với]` → `2`.
- With predicate `bào_chữa_cho()`, we have the inter relationship: `1` → `[bào,chữa,cho]` → `2`.

To generate the syntactic structure of the new Vietnamese sentence, we propose the algorithm in Alg. 10.

Algorithm 10: Generate the syntactic structure of the new sentence for group D.

Step 1: Consider the first relationship
- Put the semantic predicate of the first noun into the syntactic structure.
- Put this predicate into the syntactic structure.

Step 2: Consider the second relationship
- Put "và" (and) into the syntactic structure.
- If the first object takes the subject role: (i) put this predicate into the syntactic structure; (ii) put the semantic predicate of the second noun into the syntactic structure.
- If the first object takes the object role: (i) put "được" (is – passive voice) into the syntactic structure; (ii) put the second predicate into the syntactic structure; (iv) put "bởi" (by) into the syntactic structure; (v) put the semantic predicate of the second object into the syntactic structure.

Applying Alg. 10 to the above relationships, we have the syntactic structure:

```
luật_sư(1) + thảo_luận_với(1,2) + "và" +
    bào_chữa_cho(1,2) + bị_cáo(2)
```

To complete the new sentence, we in turn replace each predicate in the syntactic structure by the corresponding Vietnamese lexicon. With the above syntactic structure, we have the complete sentence:

"Luật sư thảo luận với và bào chữa cho bị cáo."
(English: The lawyer discusses with and defends the accused)

5. Experiment and Discussion

We randomly collected the paragraphs for the experiment in T. Tran and D. T. Nguyen ([25, 26, 27, 28]) and this research as follow: group A includes 15 paragraphs; group B includes 120 paragraphs; group C includes 18 paragraphs; group D includes 200 paragraphs. We also test these paragraphs through two phases of our approach.

At the first experiment phase, the results of anaphoric pronoun disambiguation as follows:

- With group A: the system finds correctly the antecedent and create the DRS for 14 paragraphs. The successful rate is 0.933.
- With group B: the system finds correctly the antecedent and create the DRS for all 120 paragraphs. The successful rate is 1.
- With group C: the system finds correctly the antecedent and create the DRS for all 18 paragraphs. The successful rate is 1.
- With group D:
  - Group D1: the system finds correctly the antecedent and create the DRS for 29 paragraphs in 45 paragraphs. The successful rate is 0.644.
  - Group D2: the system finds correctly the antecedent and create the DRS for 41 paragraphs in 47 paragraphs. The successful rate is 0.872.
  - Group D3: the system finds correctly the antecedent and create the DRS for 37 paragraphs in 48. The successful rate is 0.771.
  - Group D4: the system finds correctly the antecedent and create the DRS for 51 paragraphs in 60 paragraphs. The successful rate is 0.933.

At the second experiment phase, the system generates the new meaning-summarizing sentences with all the DRSs which are created from the first experiment phase. Consider these sentences, we see that each sentence is grammatical correct and has the content which summarizes the semantic of the original paragraph.

Further analysing, there are some shortcomings which should be overcome in the future research:

- Our finding strategies in Section 3.1 are not suitable for all forms of two-sentences Vietnamese paragraphs. This lead to another point is that the DRS may be not exactly or moreover cannot be

created, and the generated sentence does not summarize the meaning of the source text.
- With the algorithms for generating the syntactic structure of the new sentence, there are some generated ones are not commonly used in the reality of the native speakers.
- The current approach has been applied for paragraph composing only two sentences.

The above limitations will become the main objectives in our future researches.

6. Conclusion

In this paper we presented in detail two phases of our new approach for summarizing two-sentences Vietnamese paragraphs. Through three steps in phase one, we proposed the methods for performing the disambiguation resolution and representing the semantic by a discourse representation structure. This structure becomes the input of phase two, in which we proposed the methods for transforming to the syntactic structure and completing the new meaning-summarizing Vietnamese sentence.

Further analysing the experiments indicates that the new sentences satisfy the requirements about grammatical correctness and summarize the meaning of the source texts. Besides, in the future research, we still follow the current approach, establish new methods to overcome the discussed limitations.

References

[1] BLACKBURN, P. and BOS, J., (1999) *Representation and Inference for Natural Language – Volume II: Working with Discourse Representation Structures*. (Department of Computational Linguistics, University of Saarland).

[2] CARBONELL, J. G. and BROWN, R. D., (1988) Anaphora Resolution: A Multi-Strategy Approach. In *Proceedings of the 12th International Conference on Computational Linguistics*, 96–101.

[3] CORNISH, F., (2009) Inter-sentential anaphora and coherence relations in discourse: a perfectmatch. *Language Sciences* 31(5): 572–592.

[4] COVINGTON, M. A., (2007) *GULP 4: An Extension of Prolog for Unification Based Grammar*. Research Report number: AI-1994-06. (USA: Artificial Intelligence Center, The University of Georgia).

[5] COVINGTON, M. A. and SCHMITZ, N., (1989) *An Implementation of Discourse Representation Theory*. ACMC Research Report number: 01-0023. (Advanced Computational Methods Center, The University of Georgia).

[6] COVINGTON, M. A., NUTE, D., SCHMITZ, N. and GOODMAN, D., (1988) *From English to Prolog via Discourse Representation Theory*. ACMC Research Report number: 01-0024. (Advanced Computational Methods Center, University of Georgia).

[7] DAS, D. and MARTINS, A. F. T., (2007) *A survey on automatic text summarization*. (Language Technologies Institute, Carnegie Mellon University).

[8] GROSZ, B. J., JOSHI, A. K. and WEINSTEIN, S., (1995) Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics* **21**(2): 203–225.

[9] GUPTA, V., LEHAL, G. S., (2010) A survey of text summarization extractive techniques. *Journal of Emerging Technology in Web Intelligence* **2**(3): 258–268.

[10] JEZEK, K. and STEINBERGER, J., (2008) Automatic Text Summarization. In: SNASEL, V. (ed.), *Znalosti*, ISBN 978-80-227-2827-0 (FIIT STU Brarislava, Ustav Informatiky a softveroveho inzinierstva), pp. 1–12.

[11] JOHNSON, M. and KLEIN, E., (1986) *Discourse, anaphora and parsing*. Report number: CSLI-86-63. (USA: Center for the Study of Language and Information, Stanford University).

[12] JONES, K. S., (1999) Automatic summarizing: factors and directions. In: MANI, I., MARBURY, M. (eds.), *Advances in Automatic Text Summarization* (MIT Press, Cambridge).

[13] JONES, K. S., (2007) *Automatic summarising: a review and discussion of the state of the art*. Technical Report 679. (Computer Laboratory, University of Cambridge).

[14] JOSHI, A. K. and WEINSTEIN, S., (1981) Control of inference: Role of some aspects of discourse structure centering. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'1981)*, 385–387.

[15] KAMP, H., (1981) A theory of truth and semantic representation. In GROENENDIJK, JEROEN, JANSSEN, THEO M. V and STOKHOF, M. (eds.), *Formal Methods in the Study of Language* (Mathematical Centre Tracts) Part 1, 277–322.

[16] KASTURE, N. R., YARGAL, N., SINGH, N. N., KULKARNI, N. and MATHUR, V., (2014) A Survey on Methods of Abstractive Text Summarization. *International Journal for Research in Merging Science and Technology* **1**(6): 53–57.

[17] KHAN, A. and SALIM, N., (2014) A Review on Abstractive Summarization Methods. *Journal of Theoretical and Applied Information Technology* **59**(1): 64–72.

[18] LIANG, T. and WU, D.S., (2004) Automatic Pronominal Anaphora Resolution in English Texts. *Computational Linguistics and Chinese Language Processing* **9**(1): 21-40.

[19] LLORET, E., (2008) *Text summarization: an overview*. Paper supported by the Spanish Government under the project TEXT-MESS (TIN2006-15265-C06-01).

[20] MANI, I. and MAYBURY, M. T., (1999) *Advances in Automatic Text Summarization*. (MIT Press).

[21] REITER, E. and DALE, R., (1997) *Building Natural Language Generation System*. (Cambridge University Press).

[22] REITER, E., DALE, R., (1997) Building applied natural language generation systems. *Natural Language Engineering* **3**(1): 57–87.

[23] SARANYAMOL, C. S. and SINDHU, L., (2014) A Survey on Automatic Text Summarization. *International Journal of Computer Science and Information Technologies* **5**(6): 7889–7893.

[24] SHIEBER, S. M., (2003) *An introduction to unification-based approaches to grammar*. (Massachusetts: Microtome Publishing Brookline).

[25] TRAN, T. and NGUYEN, D. T., (2013) A Solution for Resolving Inter-sentential Anaphoric Pronouns for Vietnamese Paragraphs Composing Two Single Sentences. In *Proceedings of the 5th International Conference of Soft*

Computing and Pattern Recognition (SoCPaR 2013), Hanoi, Vietnam, 172–177.

[26] TRAN, T. and NGUYEN, D. T., (2014) Merging Two Vietnamese Sentences Related by Inter-sentential Anaphoric Pronouns for Summarizing. In *Proceedings of the 1st NAFOSTED Conference on Information and Computer Science (NCIS 2014)*, Hanoi, Vietnam, 371–381.

[27] TRAN, T. and NGUYEN, D. T., (2015) Referent Disambiguation for Anaphoric Human Pronouns in Contexts of Pairs of Vietnamese Sentences Using Affecting and Communicating Verbs. In *Proceedings of the 4th International Conference on Context-Aware Systems and Applications (ICCASA 2015)*, Vung Tau, Vietnam. In press.

[28] TRAN, T. and NGUYEN, D. T., (2015) Ambiguous Antecedents Resolution in Summarization of Vietnamese Sentences Formed by Identical Relationship with Copula "LÀ". In *Proceedings of the 4th International Conference on Context-Aware Systems and Applications (ICCASA 2015)*, Vung Tau, Vietnam. In press.