Emerald Insight

International Journal of Pervasive Computing and Communications

Modelling consequence relationships between two action, state or process Vietnamese sentences for improving the quality of new meaning-summarizing sentence

Trung Tran Dang Tuan Nguyen

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

*Related content and download information correct at time of download.

Modelling consequence relationships between two action, state or process Vietnamese sentences for improving the quality of new meaning-summarizing sentence

Trung Tran and Dang Tuan Nguyen

Faculty of Computer Science, University of Information Technology,
Vietnam National University, Ho Chi Minh City, Vietnam

Abstract

Purpose – The purpose of this paper is to enhance the quality of new reducing sentence in sentence-generation-based summarizing method by establishing consequence relationship between two action, state or process Vietnamese sentences.

Design/methodology/approach – First, types of pairs of Vietnamese sentences based on presupposition about the consequence relationship is classified: the verb indicating action or state at the first sentence is considered as the consequence of the verb indicating action, state or process at the second sentence. Then main predicates in Discourse Representation Structure – a logical form which represents the semantic of a given pair of sentences – is analyzed and inner- and inter-sentential relationships are determined. The next step is to generate the syntactic structure of the new reducing sentence. Finally, a combination with the built set of lexicons is done to complete the new meaning-summarizing Vietnamese sentence.

Findings – This method makes the new meaning-summarizing Vietnamese sentence satisfy two requirements: summarize the semantic of the given pair of Vietnamese sentences and have naturalism in common Vietnamese communication. In addition, it is possible to extend the method and apply for the purpose of summarizing the more complex Vietnamese paragraphs as well as paragraphs in other languages.

Research limitations/implications – At the first step, only inter-sentential consequence relationship is considered and this is applied to the limit types of pairs of Vietnamese sentences which have a simple structure.

Originality/value – This study presents improvements in sentence-generation-based summarization method to enhance the quality of new meaning-summarizing Vietnamese sentences. This method proves effective in summarizing the considered pairs of sentences.

Keywords Anaphoric pronoun, Discourse representation, Inter-sentential anaphora, Meaning summarization, Sentence generation

Paper type Research paper

The preliminary version of this paper was presented at the 16th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2014), December 4-6, 2014, Hanoi, Viet Nam (Tran and Nguyen, 2014c).

1. Introduction

From the 1950s, the idea about summarizing the meaning of the given paragraph has attracted much attention of computer scientists. Until recent years, the study in text summarization field also has a boom with many different solutions and techniques which have been applied for several types of paragraphs (Das and Martins, 2007; Fattah and Ren, 2008; Hovy and Lin, 1999; Jezek and Steinberger, 2008; Jones, 2007, 1999; Le *et al.*, 2010; Lloret, 2008; Mani and Maybury, 1999; Mani, 2001b; Nguyen and Nguyen, 2013; Nguyen *et al.*, 2013; Nguyen and Le, 2008; Radev *et al.*, 2002). Although there are many different approaches, to make a transition from a given paragraph into a more concise text which contains the main source content, a common summarizing system has to perform three main phases (Jones, 2007, 1999):

(1) the first phase is to transform the input text into the first representation form;

(2) the second phase is to generate the second representation form from the first representation form; and

(3) the third phase is to generate the concise text from the second representation form.

Based on the classification of the results of performing the third phase, there are two main direction approaches to summarize the given paragraph (Jezek and Steinberger, 2008; Lloret, 2008):

(1) "extraction" in which every sentence in the source text is determined the degree of importance based on different factors and methods, from which the sentence which is to be the most important will be chosen to form the summarizing paragraph; and

(2) "abstraction" in which the system determines the main content of the input text so that it generates a new, meaningful, concise paragraph.

Starting from the elementary research (Tran and Nguyen, 2014b), we proposed the new summarizing solution based on the "abstraction" approach including the combination of ideas and techniques in text-generation field (Dale, 1992; Dale and Haddock, 1991; Reiter and Dale, 1997a, 1997b, 1995; Reiter *et al.*, 2000). The main idea in the solution presented by Tran and Nguyen (2014b) is to generate a new complete Vietnamese sentence having content that summarizes the meaning of a given pair of Vietnamese sentences. We followed the above three summarizing phases into the following three phases:

(1) the first phase is to transform the original pair of Vietnamese sentences into a Discourse Representation Structure (DRS) (Blackburn and Bos, 1999; Covington and Schmitz, 1989; Covington *et al.*, 1988; Kamp, 1981) representing the semantic;

(2) the second phase is to analyze this DRS for determining the predicates containing the main content and generating the syntactic structure of the new Vietnamese sentence; and

(3) the third phase is to combine with the built set of lexicons for completing the new Vietnamese sentence.

According to Discourse Representation Theory (Blackburn and Bos, 1999; Covington and Schmitz, 1989; Covington *et al.*, 1988; Kamp, 1981), a DRS represents the semantic of a paragraph through two ordered lists:

(1) the first list called U contains indexes denoting objects (expressed by nouns) in the paragraph; and

(2) the second list called Con contains predicates (or can be called conditions – a semantic representation form) describing the semantic information of objects, actions or states (respectively expressed by nouns, verbs or adjectives) which associates with indexes in list U.

Continuing with sentence-generation-based summarization solution, we set out the requirement that is to improve the quality of the generated Vietnamese sentence for satisfying the naturalism for the cognition of native Vietnamese speakers. To satisfy this requirement, in Tran and Nguyen (2014c), we made some improvements, in comparison with Tran and Nguyen (2014b), with the following main points:

- Re-classified verbal categories base on the Functional Grammar theory (Cao, 2006; Halliday and Matthiessen, 2004) about the classification of action and state sentences. These verbal categories also established the priority orders based on the sustainable level for the context: the verb denoting status state had priority 1; the verb denoting intransitive action had priority 2; the verb denoting transitive action had priority 3; and the verb denoting property state had priority 4. The studied pairs of Vietnamese sentences in Tran and Nguyen (2014c) had the general characteristic: the first sentence had the verb denoting action or state which had the higher priority and the second sentence had the verb denoting action or state which had the lower priority.

- Considered consequence relationships in the original pairs of sentences: the verb at the first sentence having the higher priority took the consequent role of the verb at the second sentence.

- Re-defined the information in semantic predicates of lexicons in list Con of DRS to fit the classification and defined only one lexical class when building original lexical set.

- Adjusted the algorithm for generating the syntactic structure of the new reducing Vietnamese sentence.

- Proposed the new evaluation method suitable for the sentence-generation approach with the main idea is to compare the semantic similarity between automatically generated Vietnamese sentence by the proposed solution with some manually generated Vietnamese sentences.

Based on the idea of Tran and Nguyen (2014c), we consider additional process sentences besides action and state sentences. According to the Functional Grammar theory (Cao, 2006; Halliday and Matthiessen, 2004), in the process sentence, a non-animated object experiences an event involuntarily. There are three kinds of process sentence (Cao, 2006; Halliday and Matthiessen, 2004):

(1) State-changed process: the occurrence that changes the state of a non-animated object.
    Example 1: *Bức tượng bị vỡ.*
    (English: The statue is broken.)

(2) Position-changed process: the occurrence that changes the position of a non-animated object.

Example 2: *Chiếc bình bị rơi.*

(English: The vase falls.)

(3) Impact process: the occurrence when a natural or artificial phenomenon affects the state of position of a non-animated object.

Example 3: *Sét đánh chiếc thuyền.*

(English: Lightning hits the boat.)

The additional pairs of sentences have main contents: a non-animated object is affected by an action at the first sentence and experiences a process at the second sentence. To summarize these types of pair of sentences, the presupposition about the consequence relationship is used as foundation: the verb denoting the transitive action at the first sentence takes the consequence role of the verb denoting the process at the second sentence. The reason of this presupposition is that, in our experience, in most reality contexts, a process that will happen first will then lead to an action being performed. Besides, we adjust the information in semantic predicates of lexicons for two purposes:

(1) to determine more exactly the antecedence for pronoun "nó" – a special pronoun in Vietnamese, indicates human, animate or non-animated object depending on the context of the paragraph; and

(2) to determine which verb in the context denotes the transitive action and which verb in the context denotes the impact process.

To evaluate the quality of generated Vietnamese sentences in the experiment, we apply the new human-based evaluation method introduced by Tran and Nguyen (2014c).

2. The process of performing summarization

2.1 Classifying types of pairs of sentences based on the presupposition about the consequence relationship

The main objective of Tran and Nguyen (2014c) and this research is to summarize some types of pairs of simple Vietnamese sentences based on considering the presupposition about the consequence relationship for enhancing the quality of new reducing Vietnamese sentences. The main studied objects are pairs of Vietnamese sentences in which each sentence belongs to one of three forms which are action, state and process. We establish the inter-sentential relationship <Consequence Relationship> between the verb in the first sentence and the verb in the second sentence for limiting considered sentence pairs:

• In the study by Tran and Nguyen (2014c), based on establishing priorities for verbs denoting action or state (presented in Introduction), we considered pairs of sentences with consequence relationships: the verb denoting action or state in the first sentence has the higher priority and the verb denoting action or state in the second sentence has the lower priority.

- In this research, with the consideration in the common context, an action that is taken against an object will follow a process which this object experiences, we consider pairs of sentences with consequence relationships: the verb denoting transitive action in the first sentence takes the consequence role of the verb denoting process in the second sentence.

Based on categorizing verbs denoting action, state and process, we synthesize, in Table I, the considered types of pairs of Vietnamese sentences in the study by Tran and Nguyen (2014c) and this research. We use the notations A, B, C, D, E and F that indicate the verb denoting intransitive action, transitive action, status state, property state, state-changed process and position-changed process, respectively. For the impact process sentence, a verb denoting transitive action, which takes the role, is the impact process.

Type	First verb	Second verb	Example
1	C	D	Example 4: *Mai hãnh diện. Cô ấy xinh đẹp.* (English: Mai is proud. She is beautiful.)
2	A	D	Example 5: *Trí tính toán. Anh ta gian giảo.* (English: Trí calculates. He is shifty.)
3	B	D	Type 3.1: The pronoun in the second sentence stands alone, indicates the object taking the subject role in the first sentence. Example 6: *Lan học võ. Cô ấy mạnh mẽ.* (English: Lan learns martial arts. She is strong.) Type 3.2: The pronoun in the second sentence stands with a demonstrative adjective ["ta" / "ấy" / "này"], indicates the object taking the object role in the first sentence. Example 7: *Nghĩa ghét Tín. Anh ta keo kiệt.* (English: Nghĩa hates Tín. He is stingy.)
4	C	A	Example 8: *Lễ thư thái. Anh khiêu vũ.* (English: Lễ is relaxed. He dances.)
5	C	B	Type 5.1: The pronoun in the second sentence takes the subject role. Example 9: *Nhân thoải mai. Anh mặc đồ thể thao.* (English: Nhân is comfortable. He dresses sport clothes.) Type 5.2: The pronoun in the second sentence takes the object role. Example 10: *Trúc hạnh phúc. Tín cầu hôn cô.* (English: Trúc is happy. Tín proposes to her.)
6	B	E	Example 11: *Lễ vá lốp xe. Nó bị thủng.* (English: Lễ patches the tyre. It is punctured.)
7	B	F	Example 12: *Cúc nhặt chiếc bình. Nó bị rơi.* (English: Cúc picks up the vase. It falls.)
8	B	B	Example 13: *Nghĩa sửa angten. Sét đánh nó.* (English: Nghĩa fixes the antenna. Lightning hits it.)

Table I.
Synthesize
considered pairs of
Vietnamese
sentences based on
consequence
relationships

2.2 Building the DRS to represent the semantic of the source pair of Vietnamese sentences and determine inner- and inter-sentential relationships

As presented in the Introduction, at the phase of transforming the original pair of Vietnamese sentences into a DRS representing the semantic of this pair, we perform the following steps as in the study by Tran and Nguyen (2014c): apply methods and techniques as in the studies by Tran and Nguyen (2013b) and (2014b) to:

- analyze the syntactic structure of the source pair of sentences into two separated sentences and each sentence into smaller constituents;
- describe appropriate characteristic information for each lexicon; and
- determine the antecedent for the anaphoric pronoun appearing at the second sentence.

However, to consider additional types of pairs of sentences in which the second sentence has the verb denoting process and a pronoun "nó", indicating the object which experiences the process, we also add some adjustment in comparison with Tran and Nguyen (2014c) as follows:

(1) Add the factor identifying the context in the sentence indicating process or action in the lexical semantic predicate form in DRS. This factor helps for determining:

- a non-animated object taking the subject role in the context of the sentence indicating process; and
- a verb denoting transitive action which when in the context of action sentence and when in the context of process sentence.

```
semantic(index_1, index_2, content, category, sub_category, sentence_class
```

Figure 1. The lexical semantic predicate form with information: (i) `semantic` takes the value is a morphology form of lexicon, expresses the lexical name and `semantic`; (ii) `index_1` takes the value is the index of the object taking the subject role of the verb denoting action, state or process; (iii) `index_2` takes the value is the index of the object taking the object role of the verb denoting transitive action; (iv) `content` takes the value is the lexical content, includes lexical entries; (v) `category` takes the value is category [`object` / `action` / `state` / `process`]; (vi) `sub_category` takes the value is the subclass of category – [`human` / `nonanimated` / `phenomenon`] for category [`object`], [`intransitive` / `transitive`] for category [`action`], [`status` / `property`] for category [`state`], [`position_changed` / `state_changed`] for category [`process`]; (vii) `sentence_class` takes the value is the context of action or process sentence [`action_sentence` / `process_sentence`].

(2) Resolve pronoun "nó" appearing in pairs of sentences belonging to Type 6, 7 and 8 in Table I. The strategy in this research is to determine the non-animated object appearing at the first sentence as the antecedent of anaphoric pronoun "nó".

The result of this phase is the DRS representing the semantic of the original pair of Vietnamese sentences. As an example, consider a pair of sentences in Example 13 in Table I, the DRS representing the semantic of this pair is:

```
[1,2,3]
nghĩa(1,[nghĩa],[object],[human],[action_sentence])
subject_role_at(1,[action_sentence])
age_gen(1,[young_age])
species(1,[human])
role(1,[agent])
position(1,[first])
angten(2,[angten],[object],[nonanimated],[process_sentence])
subject_role_at(2,[process_sentence])
species(2,[nonanimated])
role(2,[goal])
position(2,[first])
sửa(1,2,[sửa],[action],[transitive],[action_sentence])
sét(3,[sét],[object],[phenomenon],[process_sentence])
subject_role_at(3,[process_sentence])
species(3,[phenomenon])
role(3,[agent])
position(3,[second])
đánh(3,2,[đánh],[action],[transitive],[process_sentence])
```

Figure 2. The DRS represents the semantic of pair "*Nghĩa sửa angten. Sét đánh nở.*" with two lists: (i) list U contains indexes "1" / "2" / "3", respectively, denotes objects "nghĩa" / "angten" / "sét"; and (ii) list Con contains predicates associating to these indexes. The lexical semantic predicates (are highlighted) are determined as main predicates in the DRS according to Tran and Nguyen (2014b). The verb denoting transitive action "sửa" (fix) is in the context of the action sentence, so information sentence_class takes value [action_sentence]. The verb denoting transitive action "đánh" (hit), which is in the context of process sentence so information sentence_class takes value [process_sentence].

Based on the idea and techniques in the study by Tran and Nguyen (2014c), we analyze main predicates in the DRS of each pair of sentences to determine:

- the inner relationship is determined when considering information index_1 and index_2 in the semantic predicate of verb; and
- based on the presupposition about relationship <Consequence Relationship> to determine the inter-sentential relationship when considering information category, sub_category and sentence_class in the semantic predicate of the first and second verb.

In Table II, we synthesize and analyze the DRS with main predicates of each pair type in Table I.

Type	The DRS structure with main predicates

1	`[1]`
	`named(1,[mai],[object],[human],[action_sentence])`
	`hãnh_diện(1,[hãnh,diện],[state],[status])`
	`xinh_đẹp(1,[xinh,đẹp],[state],[property])`

Figure 3. The DRS structure with main predicates of the pair of sentences in Example 4 "*Mai hãnh diện. Cô ấy xinh đẹp.*".

Analyze:
- The inner relationship:
  - The first sentence: [1] → `hãnh_diện(1)`
  - The second sentence: [1] → `xinh_đẹp(1)`
- The inter-sentential relationship:
  - Information `category` in the semantic predicate of verb "hãnh diện" (proud) and "xinh đẹp" (beautiful) take value `[state]`.
  - Information `sub_category` in the semantic predicate of verb "hãnh diện" (proud) takes value `[status]`.
  - Information `sub_category` in the semantic predicate of verb "xinh đẹp" (beautiful) takes value `[property]`.
- → Relationship: `hãnh_diện(1)` → <Consequence Relationship> → `xinh_đẹp(1)`

Type	

2	`[1]`
	`trí(1,[trí],[object],[human],[action_sentence])`
	`tính_toán(1,[tính,toán],[action],[intransitive])`
	`gian_giảo(1,[gian,giảo],[state],[property])`

Figure 4. The DRS structure with main predicates of the pair of sentences in Example 5 "*Trí tính toán. Anh ta gian giảo.*".

Analyze:
- The inner relationship:
  - The first sentence: [1] → `tính_toán(1)`
  - The second sentence: [1] → `gian_giảo(1)`
- The inter-sentential relationship:
  - Information `category` in the semantic predicate of verb "tính toán" (calculate) takes value `[action]`.
  - Information `category` in the semantic predicate of verb "gian giảo" (shifty) takes value `[state]`.
  - Information `sub_category` in the semantic predicate of verb "tính toán" (calculates) takes value `[intransitive]`.
  - Information `sub_category` in the semantic predicate of verb "gian giảo" (shifty) takes value `[property]`.
- → Relationship: `tính_toán(1)` → <Consequence Relationship> → `gian_giảo(1)`

Table II.
The DRS structure with main predicates and determine the inner and inter-sentential relationship

(*continued*)

3 | Type 3.1:

```
[1,2]
lan(1,[lan],[object],[human],[action_sentence])
võ(2,[võ],[object],[nonanimated],[process_sentence])
học(1,2,[học],[action],[transitive],[action_sentence])
mạnh_mẽ(1,[mạnh,mẽ],[state],[property])
```

Figure 5. The DRS structure with main predicates of the pair of sentences
in Example 6 "*Lan học võ. Cô ấy mạnh mẽ.*".

Analyze:
- The inner relationship:
  - The first sentence: [1] → học(1,2) → [2]
  - The second sentence: [1] → mạnh_mẽ(1)
- The inter-sentential relationship:
  - Information category in the semantic predicate of verb "học"
    (learn) takes value [action].
  - Information category in the semantic predicate of verb "mạnh mẽ"
    (strong) takes value [state].
  - Information sub_category in the semantic predicate of verb "học"
    (learn) takes value [transitive].
  - Information sub_category in the semantic predicate of verb
    "mạnh mẽ" (strong) takes value [property].
→ Relationship: học(1,2) → <Consequence Relationship> →
  mạnh_mẽ(1)

Type 3.2:

```
[1,2]
nghĩa(1,[nghĩa],[object],[human],[action_sentence])
tín(1,[tín],[object],[human],[action_sentence])
ghét(1,2,[ghét],[action],[transitive],[action_sentence])
keo_kiệt(2,[keo,kiệt],[state],[property])
```

Figure 6. The DRS structure with main predicates of the pair of sentences
in Example 7 "*Nghĩa ghét Tín. Anh ta keo kiệt.*".

Analyze:
- The inner relationship:
  - The first sentence: [1] → ghét(1,2) → [2]
  - The second sentence: [2] → keo_kiệt(2)
- The inter-sentential relationship:
  - Information category in the semantic predicate of verb "ghét"
    (hate) takes value [action].
  - Information category in the semantic predicate of verb "keo kiệt"
    (stingy) takes value [state].
  - Information sub_category in the semantic predicate of verb
    "ghét" (hate) takes value [transitive].

(*continued*)

Table II.

		o Information `sub_category` in the semantic predicate of verb "keo kiệt" (stingy) takes value `[property]`. ➔ Relationship: `ghét(1,2)` ➔ `<Consequence Relationship>` ➔ `keo_kiệt(2)`

4		```[1]```
		```lễ(1,[lễ],[object],[human],[action_sentence])```
		```thư_thái(1,[thư,thái],[state],[status])```
		```khiêu_vũ(1,[khiêu,vũ],[action],[intransitive])```
		Figure 7. The DRS structure with main predicates of the pair of sentences in Example 8 "*Lễ thư thái. Anh khiêu vũ.*".
		
		Analyze:
		● The inner relationship:
		  o The first sentence: `[1]` ➔ `thư_thái(1)`
		  o The second sentence: `[1]` ➔ `khiêu_vũ(1)`
		● The inter-sentential relationship:
		  o Information `category` in the semantic predicate of verb "thư thái" (relax) takes value `[state]`.
		  o Information `category` in the semantic predicate of verb "khiêu vũ" (dance) takes value `[action]`.
		  o Information `sub_category` in the semantic predicate of verb "thư thái" (relax) takes value `[status]`.
		  o Information `sub_category` in the semantic predicate of verb "khiêu vũ" (dance) takes value `[intransitive]`.
		➔ Relationship: `thư_thái(1)` ➔ `<Consequence Relationship>` ➔ `khiêu_vũ(1)`

5		Type 5.1:
		
		```[1,2]```
		```nhân(1,[nhân],[object],[human],[action_sentence])```
		```thoải_mái(1,[thoải,mái],[state],[status])```
		```đồ_thể_thao(2,[đồ,thể,thao],[object],[nonanimated],[process_sentence])```
		```mặc(1,2,[mặc],[action],[transitive],[action_sentence])```
		Figure 8. The DRS structure with main predicates of the pair of sentences in Example 9 "*Nhân thoải mái. Anh mặc đồ thể thao.*".
		
		Analyze:
		● The inner relationship:
		  o The first sentence: `[1]` ➔ `thoải_mái(1)`
		  o The second sentence: `[1]` ➔ `mặc(1,2)` ➔ `[2]`
		● The inter-sentential relationship:
		  o Information `category` in the semantic predicate of verb "thoải mái" (comfortable) takes value `[state]`.

(*continued*)

Table II.

- o Information `category` in the semantic predicate of verb "mặc" (dress) takes value `[action]`.
- o Information `sub_category` in the semantic predicate of verb "thoải mái" (comfortable) takes value `[status]`.
- o Information `sub_category` in the semantic predicate of verb "mặc" (dress) takes value `[transitive]`.
- ➔ Relationship: `thoải_mái(1)` ➔ `<Consequence Relationship>` ➔ `mặc(1,2)`

Type 5.2:

```
[1,2]
trúc(1,[trúc],[object],[human],[action_sentence])
hạnh_phúc(1,[hạnh,phúc],[state],[status])
tín(2,[tín],[object],[human],[action_sentence])
cầu_hôn(2,1,[cầu,hôn],[action],[transitive],[action_sentence])
```

Figure 9. The DRS structure with main predicates of the pair of sentences in Example 10 "*Trúc hạnh phúc. Tín cầu hôn cô.*".

Analyze:
- The inner relationship:
  - o The first sentence: `[1]` ➔ `hạnh_phúc(1)`
  - o The second sentence: `[2]` ➔ `cầu_hôn(2,1)` ➔ `[1]`
- The inter-sentential relationship:
  - o Information `category` in the semantic predicate of verb "hạnh phúc" (happy) takes value `[state]`.
  - o Information `category` in the semantic predicate of verb "cầu hôn" (propose) takes value `[action]`.
  - o Information `sub_category` in the semantic predicate of verb "hạnh phúc" (happy) takes value `[status]`.
  - o Information `sub_category` in the semantic predicate of verb "cầu hôn" (propose) takes value `[transitive]`.
- ➔ Relationship: `hạnh_phúc(1)` ➔ `<Consequence Relationship>` ➔ `cầu_hôn(2,1)`

6

```
[1,2]
lễ(1,[lễ],[object],[human],[action_sentence])
lốp_xe(2,[lốp,xe],[object],[nonanimated],[process_sentence])
vá(1,2,[vá],[action],[transitive],[action_sentence])
thủng(2,[thủng],[process],[state_changed])
```

Figure 10. The DRS structure with main predicates of the pair of sentences in Example 11 "*Lễ vá lốp xe. Nó bị thủng.*".

Analyze:
- The inner relationship:
  - o The first sentence: `[1]` ➔ `vá(1,2)` ➔ `[2]`

(*continued*)

Table II.

o The second sentence: [2] → thủng(2)
- The inter-sentential relationship:
  o Information category in the semantic predicate of verb "vá" (patch) takes value [action].
  o Information category in the semantic predicate of verb "thủng" (punctured) takes value [process].
→ Relationship: vá(1,2) → <Consequence Relationship> → thủng(2)

7	`[1,2]`
	`cúc(1,[cúc],[object],[human],[action_sentence])`
	`chiếc_bình(2,[chiếc,bình],[object],[nonanimated],[process_sentence])`
	`nhặt(1,2,[nhặt],[action],[transitive],[action_sentence])`
	`rơi(2,[rơi],[process],[position_changed])`

Figure 11. The DRS structure with main predicates of the pair of sentences in Example 12 "*Cúc nhặt chiếc bình. Nó bị rơi.*".

Analyze:
- The inner relationship:
  o The first sentence: [1] → nhặt(1,2) → [2]
  o The second sentence: [2] → rơi(2)
- The inter-sentential relationship:
  o Information category in the semantic predicate of verb "nhặt" (pick up) takes value [action].
  o Information category in the semantic predicate of verb "rơi" (fall) takes value [process].
→ Relationship: nhặt(1,2) → <Consequence Relationship> → rơi(2)

8	`[1,2,3]`
	`nghĩa(1,[nghĩa],[object],[human],[action_sentence])`
	`angten(2,[angten],[object],[nonanimated],[process_sentence])`
	`sửa(1,2,[sửa],[action],[transitive],[action_sentence])`
	`sét(3,[sét],[object],[phenomenon],[process_sentence])`
	`đánh(3,2,[đánh],[action],[transitive],[process_sentence])`

Figure 12. The DRS structure with main predicates of the pair of sentences in Example 13 "*Nghĩa sửa angten. Sét đánh nó.*".

Analyze:
- The inner relationship:
  o The first sentence: [1] → sửa(1,2) → [2]
  o The second sentence: [3] → đánh(3,2) → [2]
- The inter-sentential relationship:
  o Information category in the semantic predicate of verb "sửa" (fix) takes value [action].

(*continued*)

Table II.

o Information `category` in the semantic predicate of verb "đánh" (hit) takes value `[action]`.   o Information `sub_category` in the semantic predicate of verb "sửa" (fix) takes value `[transitive]`.   o Information `sub_category` in the semantic predicate of verb "đánh" (hit) takes value `[transitive]`.   o Information `sentence_class` in the semantic predicate of verb "sửa" (fix) takes value `[action_sentence]`.   o Information `sentence_class` in the semantic predicate of verb "đánh" (hit) takes value `[process_sentence]`.   → Relationship: sửa (1,2) → `<Consequence Relationship>` → đánh (3,2)

Table II.

2.3 Generate the syntactic structure of the new reducing Vietnamese sentence
After determining inner- and inter-sentential relationships, at the next performing phase, we generate the syntactic structure of the new meaning-summarizing Vietnamese sentence. The main idea for generating the syntactic structure is to combine elements in the inner- and inter-sentential relationship structures. The general algorithm is performed through the following steps:

```
(i)   Step 1: With first inner relationship
   If (Type 1 / Type 2 / Type 4 / Type 5.1 / Type 5.2) Then
      Put: [1] → first_predicate(1);
   Else If (Type 3.1 / Type 3.2 / Type 6 / Type 7 / Type 8) Then
      Put: [1] → first_predicate(1,2);
(ii)  Step 2: With inter relationship
      Put: <Consequence Relationship>;
(iii) Step 3: With second inner relationship
   If (Type 1 / Type 2 / Type 3.1 / Type 4) Then
      Put: second_predicate(1);
   Else If (Type 3.2) Then
      Put: [2] → second_predicate(2);
   Else If (Type 5.1) Then
      Put: second_predicate(1,2) → [2];
   Else If (Type 5.2) Then
      ⇨ Passive voice
      Put: second_predicate(2,1) ← [2];
   Else If (Type 6 / Type 7) Then
      Put: [2] → second_predicate(2);
   Else If (Type 8) Then
      ⇨ Passive voice
      Put: [2] ← second_predicate(3,2) ← [3];
```

Figure 13. The general algorithm for generating the syntactic structure of the new meaning-summarizing Vietnamese sentence.

We perform steps of this algorithm for each pair type in Table I as follows:

- Type 1:
  - o Step 1: With the first inner relationship
    ⇨ Add:[1] → hãnh_diện(1)
  - o Step 2: With the inter-sentential relationship
    ⇨ Add: <Consequence Relationship>
  - o Step 3: With the second inner relationship
    ⇨ Add:xinh_đẹp(1)

➔ The syntactic structure:

[1]➔ hãnh_diện(1) + <Consequence Relationship> + xinh_đẹp(1)

- Type 2:
  - o Step 1: With the first inner relationship
    ⇨ Add:[1] → tính_toán(1)
  - o Step 2: With the inter-sentential relationship
    ⇨ Add: <Consequence Relationship>
  - o Step 3: With the second inner relationship
    ⇨ Add: gian_giảo(1)

➔ The syntactic structure:

[1]➔ tính_toán(1) + <Consequence Relationship> + gian_giảo(1)

- Type 3.1:
  - o Step 1: With the first inner relationship
    ⇨ Add: [1] → học(1,2) → [2]
  - o Step 2: With the inter-sentential relationship
    ⇨ Add: <Consequence Relationship>
  - o Step 3: With the second inner relationship
    ⇨ Add: mạnh_mẽ(1)

➔ The syntactic structure:

[1] → học(1,2) → [2] + <Consequence Relationship>
    + mạnh_mẽ(1)

- Type 3.2:
  - o Step 1: With the first inner relationship
    ⇨ Add: [1] → ghét(1,2)→ [2]
  - o Step 2: With the inter-sentential relationship
    ⇨ Add: <Consequence Relationship>
  - o Step 3: With the second inner relationship
    ⇨ Add: [2] → keo_kiệt(2)

➔ The syntactic structure:

[1] → ghét(1,2) → [2] + <Consequence Relationship> + [2]
    → keo_kiệt(2)

- Type 4:
  - o Step 1: With the first inner relationship
    - ⇨ Add: `[1]` → `thu_thái(1)`
  - o Step 2: With the inter-sentential relationship
    - ⇨ Add: `<Consequence Relationship>`
  - o Step 3: With the second inner relationship
    - ⇨ Add: `khiêu_vũ(1)`

→ The syntactic structure:

```
[1] → thu_thái(1) + <Consequence Relationship>
    + khiêu_vũ(1)
```

- Type 5.1:
  - o Step 1: With the first inner relationship
    - ⇨ Add: `[1]` → `thoải_mái(1)`
  - o Step 2: With the inter-sentential relationship
    - ⇨ Add: `<Consequence Relationship>`
  - o Step 3: With the second inner relationship
    - ⇨ Add: `mặc(1,2)` → `[2]`

→ The syntactic structure:

```
[1] → thoải_mái(1) + <Consequence Relationship>
    + mặc(1,2) → [2]
```

- Type 5.2:
  - o Step 1: With the first inner relationship
    - ⇨ Add: `[1]` → `hạnh_phúc(1)`
  - o Step 2: With the inter-sentential relationship
    - ⇨ Add: `<Consequence Relationship>`
  - o Step 3: With the second inner relationship
    - ⇨ Add: `cầu_hôn(2,1)` ← `[2]`

→ The syntactic structure:

```
[1] → hạnh_phúc(1) + <Consequence Relationship>
    + cầu_hôn(2,1) ← [2]
```

- Type 6:
  - o Step 1: With the first inner relationship
    - ⇨ Add: `[1]` → `vá(1,2)` → `[2]`
  - o Step 2: With the inter-sentential relationship
    - ⇨ Add: `<Consequence Relationship>`
  - o Step 3: With the second inner relationship
    - ⇨ Add: `[2]` → `thủng(2)`

→ The syntactic structure:

```
[1] → vá(1,2) → [2] + <Consequence Relationship> + [2]
    → thủng(2)
```

- Type 7:
  o Step 1: With the first inner relationship
    ⇨ Add: `[1] → nhặt(1,2) → [2]`
  o Step 2: With the inter-sentential relationship
    ⇨ Add: `<Consequence Relationship>`
  o Step 3: With the second inner relationship
    ⇨ Add: `[2] → rơi(2)`

→ The syntactic structure:

```
[1] → nhặt(1,2) → [2] + <Consequence Relationship> + [2]
    → rơi(2)
```

- Type 8:
  o Step 1: With the first inner relationship
    ⇨ Add: `[1] → sửa(1,2) → [2]`
  o Step 2: With the inter-sentential relationship
    ⇨ Add: `<Consequence Relationship>`
  o Step 3: With the second inner relationship
    ⇨ Add: `[2] ←đánh(3,2) ← [3]`

→ The syntactic structure:

```
[1] → sửa(1,2) → [2] + <Consequence Relationship> + [2]
    ← đánh(3,2) ← [3]
```

2.4 Complete the new reducing Vietnamese sentence
To perform the last phase which is to complete the new reducing Vietnamese sentence, we proceed in two steps:

(1) Step 1 – build the lexical set; and

(2) Step 2 – replace elements in the syntactic structure by appropriate lexicons.

Based on the idea of Tran and Nguyen (2014c) when proceeding Step 1, we define only one object class Lexicon (is an improvement compared with the study by Tran and Nguyen, 2014b when defining many object classes for different lexical categories) in which each property corresponding to one information in lexical semantic predicate: Lexicon{flagSemantic – corresponding to information semantic; flagContent – corresponding to information content; flagCategory – corresponding to information category; flagSubCategory – corresponding to information sub_category; flagMorphology – is a form of practical use of lexicon}.

Consider examples are lexicons belonging to different categories:

- Consider lexicon "Nghĩa" in the pair of sentences at Example 13 in Table I, which has the semantic predicate presented in the DRS in Figure 12, the defined lexical object: ohNghia{flagSemantic – nghĩa; flagContent – [nghĩa]; flagCategory – [object]; flagSubCategory – [human]; flagMorphology – Nghĩa}.

- Consider lexicon "đánh" () in the pair of sentences at Example 13 in Table I, which has the semantic predicate presented in the DRS in Figure 12, the defined lexical object: atDanh{flagSemantic – đánh; flagContent – [đánh]; flagCategory – [action]; flagSubCategory – [transitive]; flagMorphology – đánh }.

- Consider lexicon "rơi" () in the pair of sentences at Example 12 in Table I, which has the semantic predicate presented in the DRS in Figure 11, the defined lexical object: ppcNghia{flagSemantic – rơi; flagContent – [rơi]; flagCategory – [process]; flagSubCategory – [position_changed]; flagMorphology – rơi}.

- Consider lexicon "hạnh_phúc" () in the pair of sentences at Example 10 in Table I, which has the semantic predicate presented in the DRS in Figure 9, the defined lexical object: ssHanhPhuc{flagSemantic – hạnh_phúc; flagContent – [hạnh,phúc]; flagCategory – [state]; flagSubCategory – [status]; flagMorphology – hạnh_phúc}.

The general algorithm for performing Step 2:

```
Consider syntactic structure
  With element ∈ first inner relationship
    Replace semantic predicate by flagMorphology of corresponding lexicon
    object;
  With inter relationship
    Replace <Consequence Relationship> by "vì" (English: Because of);
  With element ∈ second inner relationship
    If (Type 1 / Type 2 / Type 3.1 / Type 3.2 / Type 4 / Type 5.1)
        Replace semantic predicate by flagMorphology of corresponding
        lexicon object;
    Else If (Type 5.2)
        Add "được" (English: is – passive voice);
        Replace semantic predicate of second transitive action by
        flagMorphology of corresponding lexicon object;
        Add "bởi" (English: by);
        Replace semantic predicate of second object by flagMorphology of
        corresponding lexicon object;
    Else If (Type 6 / Type 7)
        Replace semantic predicate of second object by flagMorphology of
        corresponding lexicon object;
        Add "bị" (English: is);
        Replace semantic predicate of process by flagMorphology of
        corresponding lexicon object;
    Else If (Type 8)
        Replace semantic predicate of second object by flagMorphology of
        corresponding lexicon object;
        Add "bị" (English: is);
        Replace semantic predicate of third object by flagMorphology of
        corresponding lexicon object;
        Replace semantic predicate of second transitive action by
        flagMorphology of corresponding lexicon object;
```

Figure 14. The algorithm for completing the new reducing Vietnamese sentence.

The result of performing the algorithm in Figure 14 for each example sentence pair in Table I is presented in Table III.

3. Experiment and evaluation
The main content of this section is based on the idea of Tran and Nguyen (2014c).

With the given requirement that the generated Vietnamese sentence through proposed method has to satisfy naturalism in common Vietnamese communication, we test and evaluate the quality of new reducing Vietnamese sentence based on these criteria. To be consistent with the sentence-generation approach (in Tran and Nguyen, 2014c) and to extend this research, we propose the method for evaluating the sentential quality with the main idea: for each tested pair of Vietnamese sentences, compare the semantic similarity between new reducing Vietnamese sentence which is automatically generated through proposed method in Section 2 with some Vietnamese sentences which are manually generated in different real contexts. These semantic similarities are used to calculate a quality index. We compare the quality index with an established threshold to identify: the generated

Type	The syntactic structure	The complete Vietnamese sentence
1	"Mai" + "hãnh diện" + "vì" + "xinh đẹp"	"*Mai hãnh diện vì xinh đẹp*" (English: Mai is proud because of being beautiful.)
2	"Trí" + "tính toán" + "vì" + "gian giảo"	"*Trí tính toán vì gian giảo*" (English: Trí calculates because of being shifty.)
3	"Lan" + "học" + "võ" + "vì" + "mạnh mẽ"	"*Lan học võ vì mạnh mẽ*" (English: Lan learns martial arts because of being strong.)
	"Nghĩa" + "ghét" + "Tín" + "vì" + "Tín" + "keo kiệt"	"*Nghĩa ghét Tín vì Tín keo kiệt*" (English: Nghĩa hates Tín because Tín is stingy.)
4	"Lễ" + "thư thái" + "vì" + "khiêu vũ"	"*Lễ thư thái vì khiêu vũ*" (English: Lễ is relaxed because of dancing.)
5	"Nhân" + "thoải mái" + "vì" + "mặc" + "đồ thể thao"	"*Nhân thoải mái vì mặc đồ thể thao*" (English: Nhân is comfortable because of dressing sport clothes.)
	"Trúc" + "hạnh phúc" + "vì" + "được" + "cầu hôn" + "bởi" + "Tín"	"*Trúc hạnh phúc vì được cầu hôn bởi Tín*" (English: Trúc is happy because of being proposed by Tín.)
6	"Lễ" + "vá" + "lốp xe" + "vì" + "lốp xe" + "bị" + "thủng"	"*Lễ vá lốp xe vì lốp xe bị thủng*" (English: Lễ patches the tyre because the tyre is punctured.)
7	"Cúc" + "nhặt" + "chiếc bình" + "vì" + "chiếc bình" + "bị" + "rơi"	"*Cúc nhặt chiếc bình vì chiếc bình bị rơi*" (English: Cúc picks up the vase because the vase falls.)
8	"Nghĩa" + "sửa" + "angten" + "vì" + "angten" + "bị" + "sét" + "đánh"	"*Nghĩa sửa angten vì angten bị sét đánh*" (English: Nghĩa fixes the antenna because the antenna is hit by lightning.)

Table III.
The complete
Vietnamese sentence

Vietnamese sentence has the quality which is accepted if the quality index has the value which is higher than the threshold.

At the first step, we build a dataset which is used in the experiment composing pairs of Vietnamese sentences which satisfy following requirements:

(1) Have the characteristic belonging to one of the types in Table I.

- Types 1, 2, 3, 4 and 5: the first sentence has the verb having the higher priority than the verb in the second sentence.
- Types 6, 7 and 8: The first sentence denotes transitive action, and the second sentence denotes process.

(2) Can exactly determine the antecedence at the first sentence for the anaphoric pronoun at the second sentence and build the DRS. We only focus on evaluating the quality of the new reducing Vietnamese sentence which is generated through the presented method in Section 2.

With the above requirements, we collected 292 pairs of Vietnamese sentences for testing: following Tran and Nguyen (2014c), there are 29 sentence pairs belonging to Type 1; 47 sentence pairs belonging to Type 2; 83 sentence pairs belonging to Type 3; 11 sentence pairs belonging to Type 4; and 22 sentence pairs belonging to Type 5; added in this research are 30 sentence pairs belonging to Type 6, 11 sentence pairs belonging to Type 7 and 59 sentence pairs belonging to Type 8.

Next, we test each pair of sentence in the dataset and generate the new meaning-summarizing Vietnamese sentence by the system which is developed according to the presented method in Section 2. We also manually generate some Vietnamese sentence according to different real context. The semantic similarity between the automatically generated Vietnamese sentence and each manually generated Vietnamese sentence is calculated using the Jaccard formula (Manning and Schutze, 1999; Nivattanakul *et al.*, 2013; Tanimoto, 1958). In this formula, set AuS as set of lexicons in the automatically generated Vietnamese sentence and MaS as the set of lexicons in the manually generated Vietnamese sentence. The Jaccard formula for calculating the semantic similarity is:

$$J(AuS, MaS) \, = \, \frac{|AuS \cap MaS|}{|AuS \cup MaS|}$$

After calculating the semantic similarity by the Jaccard formula between the automatically generated Vietnamese sentence and each manually generated Vietnamese sentence, we calculate the sum of these semantic similarities and divide by the number of manually generated Vietnamese sentences. The obtained result is an index called quality index. The formula for calculating this quality index is as follows (in which n is the number of manually generated Vietnamese sentences and can be different for different original pairs of Vietnamese sentences):

$$Quality\ Index \, = \, \frac{\Sigma_n^i j(AuS, MaS_i)}{n}$$

Finally, we compare the quality index of each automatically generated Vietnamese sentence with an established threshold to evaluate: if the quality index a has the higher value than the threshold, then the quality of the automatically generated Vietnamese sentence is accepted. We establish the threshold as follows: for pairs of sentences belonging to Types 1, 2, 3, 4 and 5, the threshold is 0.7 according to Tran and Nguyen (2014c) and for pairs of sentences belonging to Types 6, 7 and 8, the threshold is 0.75. To evaluate the degree of accuracy, we calculate indexes Precision, Recall and F-measure (Nivattanakul et al., 2013; Powers, 2011).
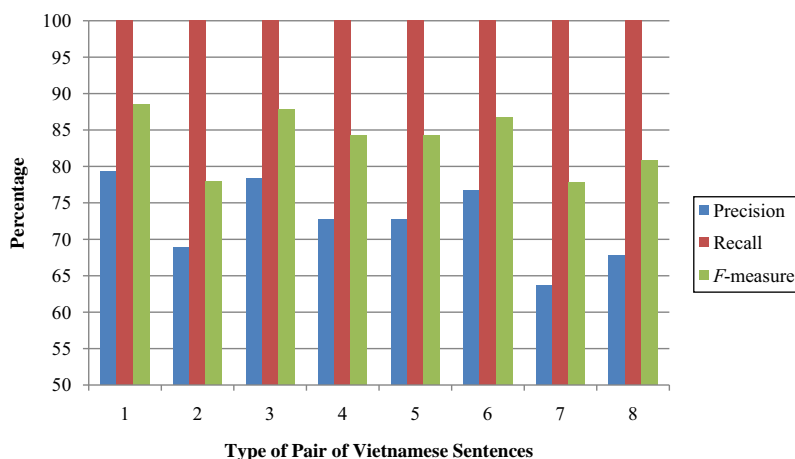
Experimental and evaluation result is presented in Table IV and Figure 15.

4. Conclusion
To summarize the given pair of Vietnamese sentences by the method of generating a new reducing Vietnamese sentence is a new approach in the text summarization field with the combination of ideas in the text-generation field. As different traditional text summarization systems, but with the new approach, we perform phases according to the definition of Jones (2007), (1999):

Type	The no. of tested pairs of sentences	The no. of new reducing Vietnamese sentences having accepted quality	Precision	Recall	F-measure
1	29	23	79.31	100	88.46
2	47	30	68.83	100	77.92
3	83	65	78.31	100	87.84
4	11	8	72.73	100	84.21
5	22	16	72.73	100	84.21
6	30	23	76.67	100	86.79
7	11	7	63.64	100	77.78
8	59	40	67.80	100	80.81

Table IV.
The experiment result and degree of accuracy with Precision, Recall and F-measure index

Figure 15.
The result of calculating Precision, Recall and F-measure index for 292 tested pairs of Vietnamese sentences

- the first phase is to understand and represent the semantic of the original pair of Vietnamese sentences by a logical form is DRS (Blackburn and Bos, 1999; Covington and Schmitz, 1989; Covington *et al.*, 1988; Kamp, 1981);

- the second phase is to transform this meaning representation form into the syntactic structure of the new meaning-summarizing Vietnamese sentence; and

- the third phase is to combine the syntactic structure with the lexical set for completing the new meaning-summarizing Vietnamese sentence.

With this new approach, the question is how to guarantee the quality of the generated Vietnamese sentence so that it can satisfy the naturalism in the common Vietnamese communication. To handle this issue, we based our study on the appropriate linguistic theory which is the Functional Grammar theory (Cao, 2006; Halliday and Matthiessen, 2004) to establish the inter-sentential relationship <Consequence Relationship> in universal contexts and classify considered pairs of sentences. The proposed experiment and evaluation method shows that the percentage of new reducing Vietnamese sentences having accepted quality is high (from 70-80 per cent for Precision index and from 80-90 per cent for *F*-measure index).

To analyze the performing process and tested result, we acknowledge some remaining points:

- The sentence-generation method was applied for pairs of Vietnamese sentences having simple structures. In some situations, the new generated Vietnamese sentence has the quality that was not accepted; the reason is that there is no factor expressing the context about time or space in the source pair of sentences. This leads to the real context in which these pairs appear; the inter-sentential relationship is not identical with <Consequence Relationship> which we established at the beginning.

- For paragraphs composed of more than two sentences, we realize that we can apply the new summarizing method. However, an important issue that needs to be set out is to establish some presuppositions about relationships and considering priorities.

- The Jaccard index (Manning and Schutze, 1999; Nivattanakul *et al.*, 2013; Tanimoto, 1958) which is applied in the new experiment and evaluation method may be not totally correct in some situations. This index only calculates the semantic similarity between two sentences based on the number of resembling lexicons but not on other factors.

With obtained results, we will still apply the sentence-generation approach to summarize the meaning for pairs of sentences having complex structures and paragraphs composing more than two sentences. In next studies, we also consider and establish some more relationships to enhance the quality of generated Vietnamese sentences.

References

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, Germany.

Cao, H.X. (2006), "Tiếng Việt: Sơ thảo ngữ pháp chức năng [Vietnamese: Brief of Functional Grammar], "Nhà xuất bản giáo dục [Education Publisher].

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, Athens, Georgia.

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, Athens, Georgia.

Dale, R. (1992), *Generating Referring Expressions: Constructing Descriptions in a Domain of Objects and Processes*, MIT Press, Cambridge.

Dale, R. and Haddock, N. (1991), "Generating referring expressions involving relations", *Proceedings of The 5th Conference on European chapter of the Association for Computational Linguistics (EACL '91), Berlin, Germany, 9-11 April*, pp. 161-166.

Das, D. and Martins, A.F.T. (2007), *A Survey on Automatic Text Summarization*, Language Technologies Institute, Carnegie Mellon University, 5000 Forbes Avenue Pittsburgh, PA.

Fattah, M.A. and Ren, F. (2008), "Automatic text summarization", *World Academy of Science, Engineering and Technology*, Vol. 27, pp. 192-195, ISSN 13076884.

Halliday, M.A.K. and Matthiessen, C.M.I.M. (2004), *An Introduction to Functional Grammar*, 3rd ed., Hodder Arnold, London.

Hovy, E. and Lin, C.Y. (1999), "Automated text summarization in summarist", in Mani, I. and Maybury, M.T. (Eds), *Advances in Automatic Text Summarization*, MIT Press, Cambridge, MA, pp. 81-94.

Jezek, K. and Steinberger, J. (2008), "Automatic text summarization", in Snasel, V. (Ed.), *Znalosti 2008*, ISBN 978-80-227-2827-0. FIIT STU Brarislava, Ustav Informatiky a softveroveho inzinierstva, pp. 1-12.

Jones, K.S. (1999), "Automatic summarizing: factors and directions", in Mani, I. and Marbury, M. (Eds), *Advances in Automatic Text Summarization*, MIT Press, Cambridge, MA.

Jones, K.S. (2007), "Automatic summarising: a review and discussion of the state of the art", Technical Report number: 67, Computer Laboratory, University of Cambridge, Cambridge.

Kamp, H. (1981), "A theory of truth and semantic representation", in Groenendijk, J., Janssen, T.M.V and Stokhof, M. (Eds), *Formal Methods in the Study of Language, Part 1*, Mathematical Centre Tracts, pp. 277-322.

Le, H.T., Sam, R.C. and Nguyen, P.T. (2010), "Extracting phrases in Vietnamese document for summary generation", *Proceedings International Conference on Asian Language Processing (IALP), Harbin, China*, pp. 207-210.

Lloret, E. (2008), "Text summarization: an overview", paper supported by the Spanish Government under the Project TEXT-MESS (TIN2006-15265-C06-01).

Mani, I. (2001b), *Automatic Summarization*, John Benjamins Publishing, Amsterdam.

Mani, I. and Maybury, M.T. (1999), *Advances in Automatic Text Summarization*, MIT Press, Cambridge, MA.

Manning, C.D. and Schutze, H. (1999), *Foundations of Statistical Natural Language Processing*, MIT Press, Cambridge, MA.

Nguyen, H.T.T. and Nguyen, Q.H. (2013), "A semi-supervised learning method combined with dimensionality reduction in vietnamese text summarization", *International Journal of Innovative Computing, Information and Control*, Vol. 9 No. 12, pp. 4903-4915.

Nguyen, H.T.T., Nguyen, Q.H. and Nguyen, T.N.T. (2013), "A supervised learning method combine with dimensionality reduction in vietnamese text summarization", *Proceedings*

2013 Computing, Communications and IT Applications Conference (ComComAp), Hong Kong, 1-4 April, pp. 69-73.

Nguyen, P.T. and Le, H.T. (2008), "Vietnamese text summarisation using discourse structures", *ICT.rda Conference*, Hanoi, Vietnam.

Nivattanakul, S., Singthongchai, J., Naenudorn, E. and Wanapu, S. (2013), "Using of Jaccard coefficient for keywords similarity", *Proceedings of the International Multi Conference Engineers and Computer Scientists*, Hong Kong, pp. 380-384.

Powers, D.M.W. (2011), "Evaluation: from precision, recall and *F*-measure to ROC, Informedness, markedness & correlation", *Journal of Machine Learning Technologies*, Vol. 2 No. 1, pp. 37-63.

Radev, D.R., Hovy, E. and McKeown, K. (2002), "Introduction to the special issue on summarization", *Computational Linguistics*, Vol. 28 No. 4, pp. 399-408.

Reiter, E. and Dale, R. (1995), "Computational interpretations of the Gricean maxims in the generation of referring expressions", *Cognitive Science*, Vol. 19 No. 2, pp. 233-263.

Reiter, E. and Dale, R. (1997a), *Building Natural Language Generation System*, Cambridge University Press, Cambridge.

Reiter, E. and Dale, R. (1997b), "Building applied natural language generation systems", *Natural Language Engineering*, Vol. 3 No. 1, pp. 57-87.

Reiter, E., Robertson, R. and Osman, L. (2000), "Knowledge acquisition for natural language generation", *Proceedings of the 1st International Conference on Natural Language Generation*, Vol. 14, pp. 217-224.

Tanimoto, T.T. (1958), "An element mathematical theory of classification", Technical Report (internal report), I.B.M. Research, New York, NY.

Tran, T. and Nguyen, D.T. (2013b), "A solution for resolving inter-sentential anaphoric pronouns for Vietnamese paragraphs composing two single sentences", *Proceedings of the 5th International Conference of Soft Computing and Pattern Recognition (SoCPaR 2013)*, Hanoi, Vietnam, pp. 172-177.

Tran, T. and Nguyen, D.T. (2014b), "Merging two Vietnamese sentences related by inter-sentential anaphoric pronouns for summarizing", *Proceedings of the 1st NAFOSTED Conference on Information and Computer Science*, Hanoi, Vietnam, pp. 371-381.

Tran, T. and Nguyen, D.T. (2014c), "Enhancement of sentence-generation based summarization method by modelling inter-sentential consequent-relationships", *Proceedings of the 16th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2014)*, Hanoi, Vietnam, pp. 302-309.

Further reading

Amigó, E., Gonzalo, J., Peñas, A. and Verdejo, F. (2005), "QARLA: a framework for the evaluation of text summarization systems", *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, Ann Arbor, MI, pp. 280-289.

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

Jing, H., Barzilay, R., McKeown, K. and Elhadad, M. (1998), "Summarization evaluation methods experiments and analysis", in *AAAI Intelligent Text Summarization Workshop*, Stanford, CA, pp. 60-68.

Johnson, M. and Klein, E. (1986), *Discourse, Anaphora and Parsing*, Report number: CSLI-86-63, Center for the Study of Language and Information, Stanford University, Germany.

Lin, C.Y. (2004), "ROUGE: a package for automatic evaluation of summaries", *Proceedings of ACL Text Summarization Workshop*, Barcelona, pp. 74-81.

Lin, C.Y. and Hovy, E. (2002), "Manual and automatic evaluation of summaries", *Proceedings of the Workshop on Automatic Summarization (including DUC 2002), Association for Computational Linguistics*, Philadelphia, PA, pp. 45-51.

Lloret, E. and Palomar, M. (2010), "Challenging issues of automatic summarization: relevance detection and quality-based evaluation", *Informatica*, Vol. 34 No. 1, pp. 29-35.

Mani, I. (2001a), "Summarization evaluation: an overview", *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL) 2001 Workshop on Automatic Summarization*, National Institute of Informatics, Tokyo.

Mani, I., House, D., Klein, G., Hirschman, L., Obrst, L., Firmin, T., Chrzanowski, M. and Sundheim, B. (1998), "The TIPSTER SUMMAC text summarization evaluation: final report", Technical Report, MITRE Corp, Virginia.

Tran, T. and Nguyen, D.T. (2013a), "Improve effectiveness resolving some inter-sentential anaphoric pronouns indicating human objects in Vietnamese paragraphs using finding heuristics with priority", *Proceedings of the 10th IEEE RIVF International Conference on Computing and Communication Technologies – Research, Innovation, and Vision for the Future (RIVF 2013)*, Hanoi, Vietnam, pp. 109-114.

Tran, T. and Nguyen, D.T. (2014a), "The solution for resolving inter-sentential anaphoric pronoun 'nó' in Vietnamese paragraphs composing 3 to 5 simple sentences", *International Journal of Advanced Science and Technology*, Vol. 65, pp. 95-112.

Zhou, L., Lin, C.Y., Munteanu, D.S. and Hovy, E. (2006), "PARAEVAL: using paraphrases to evaluate summaries automatically", *Proceedings of the HLT Conference of NAACL*, New York, NY, pp. 447-454.

Corresponding author
Trung Tran can be contacted at: ttrung@nlke-group.net