Natural Language Generation of Museum Object Descriptions based on User Model

Hsiao Wei Chen, Mary Grace Lim, Patricia Bea Perez,
Joanna Patricia Reyes, and Nathalie Rose Lim

De La Salle University- Manila
{hsiao_wei_chen, mary_graces, lil_nawty_gurl13, snowdrop_106}@yahoo.com,
and limn@dlsu.edu.ph

Abstract. Natural Language Generation (NLG) techniques can be applied in generating virtual documents dynamically using information from a database (Dale et al, 1999). One of the applications of NLG techniques to generate documents dynamically is the web-based interactive virtual museum, VIGAN. NLG is used to generate the descriptions of the objects in a virtual museum dynamically based on the profile and interests of the visitor. The focus of the research is on incorporating user's interest, age group, and visit history in the generation of museum object descriptions. The descriptions do not vary only on user's profile but also in lexicalization. Facts are not only in describing the objects, but also in describing Ilocano personalities. User Acceptance Testing proved that object descriptions do vary based on age groups, category of interest, and lexicalization. They commented that the descriptions are easy to understand, the user interface is user friendly and the suggested objects are appropriate.

Keywords: Natural Language Generation, Dynamic User Modeling, Knowledge Representation, Virtual Museums

1. Introduction

With the emergence of the Internet, lack of information is no longer a major problem most users face. The vast amount of information available is overwhelming and it is important to hide information that is irrelevant to the users. One example that provides vast information to the people is museums. A museum is an institution providing services—acquiring, conserving, researching, communicating, and exhibiting objects for education and enjoyment (ICOM, 2001). Some museums are made available online through virtual museums so that more people can have access to the information provided by the museum.In order to improve the services that virtual museums provide, they should be able to determine which information is most relevant to a particular user and object information should be presented in such a way that the user can understand (Bandelli, 1999).

Similar to internet users, visitors of museums have different expectations, needs and behaviors, and it is important to address these differences. Ambeth Ocampo, a columnist at a notable Philippine newspaper, says that, "While I appreciate the educational task of museums, I would like to think that all our best efforts are still not enough to get the youth into museums and keep them returning… The problem lies not with a museum but a child's first encounter with it" (Ocampo, 2007). Ocampo also says that for many college students who had to endure a grade school trip to the museum, going there a second or third time is considered a cruel and unusual punishment. This mind-set is not the fault of the museum; it is the fault of the teacher or museum guide who did not infect the students with a sense of discovery and appreciation of our past. Thus, improving the encounter with a museum through a web-based interactive virtual museum that generates descriptions based on user profile is the motivation for this research.

The next section introduces the virtual museum we have created. Section 3 expounds more on the components of the virtual museum. In Section 4, we discuss how we modeled the user preference and how these are updated. Section 5 presents the rules created to generate the different types of description. Samples of generated descriptions based on user profile is also discussed in this section. In Section 6, we discuss the testing done for the system. Lastly, we give our conclusion and indicate some future work in Section 7.

2. Our Virtual Museum

We have chosen to create a virtual museum for three (3) well-known musuems in Vigan, Philippines, namely: Padre Jose Burgos Museum, Crisologo Museum and Syquia Museum. These musuems are quite small, but contain a substantial collection of objects that represent the culture of the Southern Ilocos Region. We call this interactive virtual museum VIGAN.

As VIGAN controls the amount of information generated depending on age group, the user is first asked to register upon first use of the virtual museum. Refer to Section 3 for a more detailed discussion on the information requested in the registration page. On succeeding sessions, the user is tasked to log in his user name and to choose which of the three-mentioned museums he wants to visit first. Once the user has chosen a particular museum, he has to choose a room from a list that is provided to him. After selecting a room, a 2D panoramic image of the actual museum is displayed. Hotspots in green represent another room to enter, while those in red represent objects. Clicking on the green hotspot will introduce the user to the next room with another 2D panoramic view consisting of one or more objects. On the other hand, clicking on the red hotspots indicates that the user wants to view the description for that particular object. Aside from generating the description of that object, each click on objects updates the user model. Updates in the user model involve increasing the weight of preference for that particular type of object, as well as storing the information that was already produced. Storing the information would allow the program to determine which new information can be given to the user, should he request for it. Moreover, storing the actual statements used will provide consistency in the generated text should the user choose to revisit the same object. The preference weight and the items that were visited is also used in producing a list of related objects that the user might be interested in viewing.

It should be noted that the 2D panoramic view is done via the Panorama Tool Viewer (PTViewer 2.8 of Senore, 2006). The pictures were personally taken by the researchers and the information stored in the database where from consultations with locals or the curators of the museums. The information stored in the database is called the Knowledge Base. The objects from the knowledge base are divided into four categories: customs, economic life, history and household items. Customs refer to rituals, practices, traditions and way of life of the Ilocanos, such as musical instruments used in traditional dances and way of life of the Syquia, Crisologo and Burgos families during the Spanish era. Economic life refers to how the Ilocanos made a living, such as cottage industries, farming, fishing, hunting. History refers to events and personalities that had happened in Ilocos that has historical importance. Household items refers to items that are seen at home, such as cooking and gardening equipments, paintings and antiques. The knowledge base is hand-constructed and contains keywords that will be used by the NLG to generate descriptions. The type and amount of information given to the user is analyzed from the information given by the curators and from the consultation with a psychologist. The generation of the descriptions in natural language is through SimpleNLG 3.5 (Reiter, 2007), but the various sentence structures for each type of description is defined by the researchers. The types of descriptions for the objects are in the form of messages. Aside from the name of the object, descriptions generated can be a combination of the following:

- basicMsg(<noun>basic_name,<noun>basic_what-is, <adjective>basic_description)

  The basicMsg includes the name of the object, what the object is and the description of the object. It is applicable to all object categories and subcategories.

- purposeMsg(<verb>purpose_action,<noun>purpose_object)

  The purposeMsg includes the name of the object, what the object is and the description of the object. It is applicable to all object categories and subcategories.

- make-upMsg(<noun>make-up,<adjective>make-up_description, <noun>make-up_location,<noun>make-up_alternative-make-up)

  The make-upMsg includes what the object is made up of, the description of the makeup of the object, where the makeup of the object is found, and alternative makeup of the object. It is applicable to the household item and economic item category. It is applicable to the some subcategories.

- people-involvedMsg(<noun>people-involved_relation,<noun>people-involved)

  The people-involvedMsg includes the people involved with the object and the relation of the people involved with the object. It is applicable to all categories and subcategories. For example, for the object bolo, the people-involvedMsg will be people-involvedMsg(<noun>owner, <noun>Elpidio Quirino).

- akaMsg(<noun>aka)

  The akaMsg includes what the object is also known as. It is applicable to all categories and subcategories.

- used-inMsg(<noun>used-in)

  The used-inMsg includes where the object is used in. It is applicable to household item, customs and economic item categories. It is also applicable to some subcategories.

- consist-of(<noun>consist-of)

  The consist-ofMsg includes what the object consists of, because some objects are made up of more than one parts. It is applicable to household item, customs and economic item categories. It is applicable to some subcategories:

- personMsg(<noun>person_name,<noun>person_bdate,<noun>person_bplace ,<noun>person_ddate,<noun>person_gender)

  The personMsg includes the name of the person, his/her birth date birth place, death date and gender. It is applicable to the history category and the personality subcategory.

- paintingMsg(<noun>painting_name,<noun>painting_theme,<adjective> painting_description, <noun>painting_age)

  The paintingMsg includes the name, the theme, description and age of the painting. It is applicable to the household item category and art subcategory.

- parentsMsg(<noun>parents_fathersName,<noun>parents_mothersName)

  The parentsMsg includes the father's name and the mother's name. It is applicable to the history category and personality subcategory.

- educationMsg(<noun>education_degree,<noun>education_year, <noun>education_school)

  The educationMsg includes the degree, year and school. It is applicable to the history category and personality subcategory.

- `siblingMsg(<noun>sibling_siblingsName)`

  The siblingMsg includes the sibling's name. It is applicable to the history category and personality subcategory.

- `original-paintingMsg(<adjective>original-painting_description, <noun>original-painting_artist,<noun>original-painting_location)`

  The original-paintingMsg includes the description, artist and location of the original painting. It is applicable to the household item category and art subcategory.

- `sizeMsg(<noun>size_unit-of-measure, <noun>size_height, <noun>size_width)`

  The sizeMsg includes the unit of measure, height and width of the object. It is applicable to all categories and all subcategories.

- `award(<noun>award_ prize,<noun>award_ sponsor,<noun>award_ year)`

  The awardMsg includes the prize, sponsor and year of the award. It is applicable to the household item category and art and personality subcategory.

3. VIGAN's Architectural Design

The architecture of VIGAN can be seen in Figure 1. There are basically two (2) main processes in VIGAN: the Natural Language Generator (NLG) and the Virtual Environment (VE). These components access and update data from the databases User Model, Plan Library, Knowledge Base, Discourse History, Lexicon, and the text file Grammar.

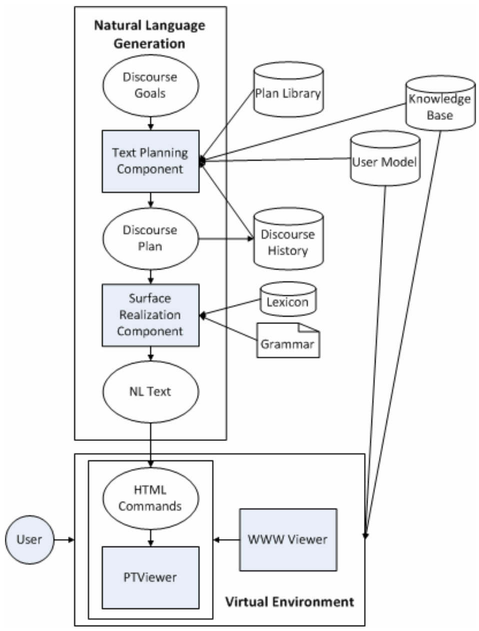

Figure 1: The Architecture of the VIGAN System

The user chooses an object from the 2D panoramic view of the museum in the Virtual Environment. The chosen object's identification number is then passed on to the NLG component where the object description will be generated.

The NLG component starts with the Discourse Goal, in order to provide a description for the selected object. The Discourse Goal is represented as the ID number of the object selected. Based on the Discourse Goal, the Text Planning component will decide which information will be included and how it will be ordered in text that will be generated. The Text Planning component will get the User Model (UM), which is used to tailor the description according to the user, Discourse History, a database that contains the facts about the objects that have been presented to the user, a message (1), which is a pre-configuration of domain elements grouped

```
basicMsg(name, what-is, description) (1)
```

in categories and arranged according to level of importance, from the Plan Library, and facts about the object stored in the database. The Text Planning component will decide based on the age group of the user (as obtained from the UM) how much information will be included. The UM is also used to make suggestion of objects more appropriate to the user. UM is a database which contains non-decision and decision properties. Non-decision properties about the user which includes username, password, name and gender are gathered from the user explicitly during registration. The user model is dynamic or always changing because new information about the user is gathered each time the user visits an object in the museum. The new information gathered as the session progresses are decision properties that will aid the system in generating descriptions for the objects in the museum. In VIGAN, decision properties are interest rating on each category of objects, importance rating on each available fact and age, because the amount of facts to be presented to the user for each description will depend on it.

The Text Planning component will fill in messages with information stored in the Knowledge Base, such as that message in (1) will be (2).

```
basicMsg(burnay, Ilocano earthen jar, null) (2)
```

Once the messages have been filled, it will include it in the Discourse Plan, which is an ordered and structured set of messages that will be included in the generated text. The Discourse Plan is passed to the Surface Realization component, and the facts used in the Discourse Plan are stored in the user's Discourse History so that the system will know which information has already been presented to the user to avoid redundancy for future presentation of object description.

VIGAN uses SimpleNLG (Reiter, 2007), a Java class library developed by the University of Aberdeen, which performs NLG lexicalisation and realisation, as its Surface Realization component. The Surface Realization component makes use of the Grammar and Lexicon to produce Natural Language expressions based on the Discourse Plan. For the Grammar, VIGAN uses a set of grammar rules used to form English sentences (see Example below).

Example of grammar rules for English language:

```
<Simple Sentence> = <Declarative Sentence>
<Declarative Sentence> = <subject> <predicate>
<subject> = <simple subject> | <compound subject>
<simple subject> = <noun phrase> | <nominative personal pronoun>
<noun phrase> = "the" <specific proper noun>
              | <proper noun>
              | <non-personal pronoun>
              | <article> [<adverb>* <adjective>] <noun>
              | [<adverb>* <adjective>] <noun-plural>
<noun> = <noun> [<prep phr>*]
```

The system reads the Grammar file using the bottom up approach, where it will start with filling the grammar rules at the bottom with facts, before working its way up. Grammar rules that have not been filled up will be removed, while grammar rules that have been filled will be used to generate the description. A message can be applied to more than one grammar rule. A message can result in more than one type of sentences. The selection of sentences to use for the description is done randomly.

VIGAN uses WordNet 2.0 (Miller, 2005) as its Lexicon, along with Java WordNet Library (JWNL), a Java API that is used to access WordNet. It will be used to map the messages in the Discourse Plan into words and phrases. Words and phrases that are used frequently in the sentence generation are stored in a Word Bank along with its senses, will be replaced with different lexicalizations, including incorporating synonyms, based on the senses automatically. The building of the Word Bank is done manually before runtime.

Natural Language (NL) Text will be the output of the Surface Realization component. The NL Text will be the description of the museum object selected by the user. The NL text will be passed to the WWW Viewer.

The WWW Viewer is the website which also links the user to the VE. It is in the VE that user can explore the virtual museum, select objects from the rooms in the virtual museum, and view descriptions of an object. Every time the user selects a uniqueobject, the user model is updated by increasing the interest rating of the user based on which category the object belongs to, refer to Section 4 for a more detailed discussion. For example, the user selects burnay which belongs to Household Items category, therefore the system will increase the percentage of the user's interest on household items. While exploring the virtual museum, WWW Viewer also suggest objects that might be of interest to the user by determining which category the user is most interested with and getting random objects with that category from the database.

4. User Modeling

User's preferences are modeled in the UM. Each user will have a rating of interest per category of object and rating of importance per facts. Interest rating is the rate of how interested the user is to each category of object (i.e., history, economic life, customs, household items). Interest rating of each category of object will initially be 0.0. During the sign-up phase, the users will be asked to select the categories of object that they are interested in. The rating of the categories is computed using (3). So when the user selects Household Items and Customs as his interests, then his initial interest rating will be 50.0/50.0/0.0/0.0 where the rating corresponds to Household Items, Customs, History, and Economics, respectively.

`100/number_of_selected_category` (3)

UM will also be updated whenever the user clicks on a new object and whenever the user clicks on "more information". The interest rating of each category will be updated using (4)

`(number_of_clicks_per_category/total_clicks) x 100` (4)

where `total_clicks` is non-unique because it represent the total number of objects that the user clicks to view.

Importance rating per fact is the assumed rate of how important for the user to know the facts of objects. Facts of objects are the properties of objectives like description of the object, creator of the object, and the like. The text planning component will decide the information that will be included in the generated description through the rate of importance. The rating will be from 1 (lowest) to 5 (highest). The rating of a description will depend on the age group of the user where in a child is 6-16 years old and an adult is 17 years old and above. The facts that have higher rating will be presented to the user. For example, it is assumed that a twelve-year old user is not so interested in the creator of the painting therefore the fact creator will have a lower rating compared to other facts so that the generated description will not include the creator of the painting.

The database design of the user model is divided into two types, namely, user profile and user model. User profile stores the non-decision properties (i.e., name, password, etc.) while user model stores the decision properties (interest rating, importance rating, and age group) for each user.

5. Generation of Natural Language Descriptions based on Message Type

The Natural Language Generation will generate descriptions based on messages. Messages, as discussed in Section 3, are pre-configuration of domain elements grouped in categories and arranged according to level of importance. Messages are classified per category, similar to the category of objects, can have more than one category per message. For example, basic message, which include the name, brief description and what the object is, has 4 categories, because it is applicable to all objects, while painting message, which includes the title, artist, theme and description of the painting will only have one category, because it can only be applied to objects in the household item category.

The Text Planning component will take all messages that are applicable to a selected object, including messages that are applicable to the category of the object.

For example, for the object Burnay, the messages would incude: (1).

```
basicMsg(name, what-is, description)                        (1)
purposeMsg(action, object)
make-up(name, description, location, alternative)
akaMsg(aka)
related-itemMsg(name, description)
```

The messages will then be filled with facts about the object from the database. Some facts that are not available in the database will be set as null in the message. Some facts have multiple entries in the database, such as there are more than one purpose for the object Burnay, in this case, only the first entry will be filled in the message. The other entries will only be used if the user asks for more information.

After the messages for the object Burnay have been filled, it would be (2).

```
basicMsg(burnay, Ilocano earthen jar, null)            (2)
purposeMsg(purify, bagoong)
make-up(red clay, seashores of Vigan, null)
akaMsg(tapayan)
related-itemMsg(null, null)
```

Some messages, although applicable to the category of the object, may not have facts available in the database. These messages, such as the related-itemMsg in the example, will be removed. The other messages will be ordered according to level of importance, a static value that is based on how the facts included in the messages are presented in the descriptions obtained from curator and captions, and passed on to the Surface Realizer.

VIGAN will tag each element in the message with its part of speech, for example, in basicMsg, burnay will be tagged as noun, while Ilocano earthen jar will be tagged as noun phrase. After tagging all the elements of a message, it will be passed to the grammar rules. Using the grammar rules, the noun 'burnay' will be tagged as subject, while the noun phrase Ilocano earthen jar will be tagged as predicate with the verb 'is'. The filled up grammar rule will then be passed to SimpleNLG, which will perform surface realization and then generate the description for the selected object. For the example, the basicMsg will produce the sentence: "Burnay is an Ilocano earthen jar."

The process is similar for the rest of the messages.

Words and phrases in the sentences can be changed if it has been used often, using the Lexicon, as discussed in Section 4. Also "burnay" can be alternated with its pronoun form.

The description for the object burnay would be: (3)

```
 Burnay is an Ilocano earthen jar. It is used for purifying bagoong.
It is made up of red clay form the seashores of Vigan. It is also
known as tapayan.
```

6. Testing Results

The system was tested using component testing and user acceptance testing.

For the Component testing, the functionality of the system was tested, particularly the internal processing, such as how data was passed from one module to another.

For the User Acceptance Testing, three approaches were used, namely the Reader-focused approach, Text-focused approach and Expert-Judgment approach.

The reader-focused approach was used to properly assess if the user interface is user friendly, the suggested objects are appropriate and descriptions are adequate and formed properly on the perspective of the users. The reader-focused approach was done by asking 20 users, 10 for each age group, to explore and rate the system. From a range of 1 to 5 (5 being the highest score), the average score for the user friendliness of the user interface is 4.7 for the child age group, while 4.3 for the adult age group. The average score for the grammar of the descriptions is 3.4 for the child age group, while 4.4 for the adult age group. This is because the child users explored the museum more and they tend to click on all the objects, even though some of the objects have limited descriptions. The average score for the structure of the descriptions is 4.3 for the child age group, while 3.8 for the adult age group, because some adult users think the sentences should be ordered differently. The average score for the appropriateness of the suggested objects is 5 for both age groups, because the system only suggests objects that are interesting to the users based on the user model. The average score for the descriptions being easy to understand is 4.8 for the child age group, while 4.9 for the adult age group. The average score for the descriptions being adequate is 4.1 for the child age group, while 4.8 for the adult age group. This is because the child users viewed some objects that had limited descriptions.

The text-focused approach was used to check if there is a variation in the descriptions generated for the different types of user age groups and interest. We forced different user types and generated descriptions for objects for each of the user types. We then compared the descriptions generated. Based on the test, we observed that the descriptions vary on number of facts and type of facts based on age group, type of facts based on the interest and words in the sentences based on lexicalization.

The expert-judgment-focused approach was used to assess the quality of the description generated. Linguist evaluation was conducted to assess the grammatical errors and paragraph structures. The feedback is that the sentences were too simple as to become repetitive. As the study does not yet cover difference in sentence structure depending on age group, the researchers made use of simple sentence structures so child users would also be able to understand the generated descriptions. Curator evaluation was also conducted to assess if the content of the description generated is sufficient. On this criterion, the curator commented that descriptions generated should also include some background information (e.g., significance of a particular artifact based on the owner or the era, etc.) to provide context. Since these details were not initially provided, the information were not encoded in the database. However, this can easily be remedied by populating the knowledge base.

7. Conclusion and Future Work

The system is able to generate different descriptions depending on user's age group and on user's preference.

Currently, the system stores the descriptions generated as part of the user model, thus making the database quite large, but it facilitates non-redundant generation of object descriptions if the object was viewed by the same user before. The researchers have considered storing the descriptions in a central database that is accessed depending on the user preference category to make the database smaller. However, this study aims to show that the system is able to generate different sentences for each type of object, at the same time keeping the generated description consistent per user.

The system is able to generate the object descriptions through the use of SimpleNLG. Although SimpleNLG is an existing system, grammar structure rules based on the English language and addtional rules had to be created to automatically feed the facts from the messages to the SimpleNLG.

The system is also able to add variations to the sentences produced by replacing commonly used words with its synonyms. The proponents created a Word Bank that contains the commonly used words along with their sense based on WordNet. The Word Bank was created manually. The system would access WordNet during runtime to replace commonly used words with their synonyms based on the sense of the word stored in Word Bank.

Currently, the system produces based on the sentences based on message types that are based on fact types. Another format can be recommended for the messages, such that it is not based on fact types, but generalized instead based on the structure of the sentence to be produced. For example, in the message, the first content is always the subject of the sentence, the second content is always the object, the third indirect object, and so on. Using this format, the facts can be mapped to the message based on its part of sentence. However, each type of facts should also include the part of sentence, such as subject and object; it will be used to map the facts to the message.

In the present system, it is only capable of producing simple sentences. In future, compound and complex sentences will also be included in the object descriptions so that sentences that are related can be combined and better structured. Also, to improve on the quality of the text in terms of coherence, Rhetorical Structure Theory, which is appropriate for multi-sentential and monologic texts, may be implemented. The sentence structures and lexical choices will also vary depending on age group.

The system built a Word Bank of frequently used words manually and accesses the synonyms of the words from the WordNet during runtime. In the improved system, the researchers may opt not to use a Word Bank, and instead will count the occurrences of the words included in the descriptions during run time to determine whether a word is to be replaced with its synonym or not. In doing so, word sense disambiguation would have to be done so that the words replaced would be appropriate.

In the future, it might be more feasible (in actual deployment) to do the generation of object descriptions offline, with an administrator verifying the generated descriptions, based on grammar, structure, lexicalization and context. This entails a minor adjustment in the current work and an incorporation of an interface for the administrator.

References

Bandelli, A. (1999). Virtual spaces and museums. *Journal of Museum Education*, 24 (1&2), pp. 20-26.

Dale, R., Green, S.J., Milosavljevic, M., and Paris, C. (1999) When virtual documents meet the real world. Proceedings of the Workshop on Virtual Documents, Hypertext Functionality and the Web, held in conjunction with the Eighth International World Wide Web Conference.

Dersch, H. (2001). *PTViewer Documentation*. Technical University Furtwangen. Retrieved March 22. 2008 from http://webuser.hs-furtwangen.de/~dersch/PTVJ/doc.html.

Finin, T. and R. Kass. (1988). Modeling the user in natural language systems. *Computational Linguistics* 4(3), pp. 5-22.

Fulvio, S. (2006). *PTViewer 2.8*. Retrieved March 22, 2008 from http://www.fsoft.it/panorama/ptviewer.htm.

International Council of Museums (ICOM). (2001). Development of the museum definition according to ICOM statues. Retrieved June 20, 2007, from http://icom.museum/hist_def_eng.html.

Miller, G. (2005). WordNet 2.0. Princeton University. Retrieved April 19, 2008 from http://wordnet.princeton.edu.

Mine, M. R. (1995). Virtual environment interaction techniques. UNC Chapel Hill Computer Science Technical Report, TR95-018.

Ocampo, A. (2007, April 13). Looking back museums. *Philippine Daily Inquirer*. Retrieved June 20, 2007, from http://opinion.inquirer.net/inquireropinion/columns/view_article.php?article_id=60116.

Reiter, E. (2007). SimpleNLG v3.5. University of Aberdeen. Retrieved April 19, 2008 from http://www.csd.abdn.ac.uk/~ereiter/simplenlg.