Syncretic Text Composition in Artificial Museum Guides

Antonio Sorgente, Antonio Calabrese, Gianluca Coda, Paolo Vanacore, and Francesco Mele

Institute of Applied Sciences and Intelligent Systems "Eduardo Caianiello" of the National Research Council
Via Campi Flegrei 34, 80078 Pozzuoli (Naples) Italy
{a.sorgente, a.calabrese, g.coda, p.vanacore, f.mele}@isasi.cnr.it

Abstract. In this paper, we present our ongoing research about the composition of syncretic text for artificial museum guides. During a museum visit, the visitors receive information about the cultural assets and responses to their questions. The aim is to reuse existing texts(for example those already published on the web) to compose responses for visitors that take into account the time at their disposal, and are balanced with respect to possible insights. Finally, system responses will result from a composition process that coherently synchronises media elements with a synthetic voice related to selected text.

Keywords: syncretic text, multimedia composition, cultural heritage

1 Introduction

Nowadays, the diffusion of new technologies (such as mobile and wearable devices) has allowed the practitioners and organisations operating in the area of Cultural Heritage to propose new approaches for the fruition of cultural assets. This approaches allow us to access to museum collections in multiple ways, both in and off site. Also, the amount of information related to the domain of Cultural Heritage built by experts, and published on the web, is growing day by day. In this scenario, an important aspect is related the extraction of information that must be coherent with the query submitted to the system.

The goal of a generation system is to produce text in response to a given stimulus. It must be able to choose what to include in such text and how to organise this information so that it can be easily understood, and increase the knowledge of the user. The most common information available on cultural heritage is related to the story of the asset or what it depicts. The best way to represent stories consists of using natural language. Our aim in this work is to propose an approach that allows us to dynamically generate information that can be close to the user request, re-using textual information provided by experts and/or already published on the web, integrating these lastly with media resources (photos and video) to generate a unique multimedia response.

In this paper, we present an ongoing research about the production of syncretic text for artificial museum guides. The texts have to take into account the time available to visitors, and to be balanced with respect to the possible insights. The latter means that the response provides an explanation of equal length for each topic involved in the dialogue. The main characteristic of this methodology is to propose an approach based on the thematic structure of the text, selecting appropriate contents related to a cultural item and then aggregating them with media resources (photo and video).

The construction of the thematic structure is based on CSWL formalism[1]. Also, we want to use the thematic progression as pattern for the selection and composition of the text to be proposed to the user. The application of the thematic progression permits us to improve the cohesion and coherence of the composed text provided to visitors.

This activity has been developed within the SIMArt project[1]. The aim of the project is to design interactive multimedia systems for the use of the cultural heritage based on the augmented knowledge paradigm[2].

We will briefly introduce the concept of thematic organisation of text and how we define it using CSWL annotation. We will present the approach adopted to compose text and how to synchronise it with media resources. Finally, some conclusions and future work will be presented.

2 A dialog model based on Theme-Rheme structure

In the construction of text, the speakers/writers construct their messages gradually introducing concepts in such a way that the message is clear, like a touristic guide. A way to achieve this aim is to organise the text through a thematic structure[3]. This structure is based on two elements: *theme* and *rheme*. The *theme* (called also topic) is related to 'what' the text is talking about, and *rheme* (called also comment or focus) is related to 'what' is said about the topic. This structure is known as thematic organisation of text. For the automatic text composition, our model is based on such structure.

As reported in [4], the theme typically contains information which has been previously mentioned or refers to the context of discussion, for example in a museum it can be a cultural asset. It is followed by the rheme that is the part of text that explains the theme introducing new information. An example of theme and rheme in a sentence is the following: *(The Basilica of Saint Clara)*$_{theme}$ *(in Naples was built between 1310 and 1340)*$_{rheme}$. In the sentence, the goal is to talk about the Saint Clara church (theme) and to say something about the story of its construction (rheme).

In order to build the response text in a dialog, some principles, based on the thematic organisation (called thematic progression[5]) have been defined. The thematic progression defines how the theme and rheme are introduced in order to have cohesion and coherence in the text. It can be seen as the skeleton of the plot.

[1] SIMArt (Interactive Multimedia Systems for the use of Art objects) is a project of National Research Council (CNR)

The main types of thematic progression, called patterns, are: linear progression, constant (or parallel) progression, split-theme, and split-rheme progression[5, 6].

In the Constant Progression (CP) the same theme appears in sequence in a series of sentences, in some cases using words having equivalent meaning (see Fig. 1.a). With Linear Progression (LP), the rheme of a sentence will be the theme of the next sentence (see Fig. 1.b). Instead, in Derived-Theme Progression (DTP), the theme of a sentence is linked to the theme of the next sentences. This means there exists a relation (hypertheme) between the themes. An example is the meronomic relation as shown in Fig. 1.c. Finally, the Split-Rheme Progression (SRP), that can be considered a general case of LP, is applied when a sentence introduces more rhemes. In this case, each rheme becomes the theme of a sentence (see Fig. 1.d). Analysing some cultural texts, it has been necessary to introduce another pattern that we called: Inverse Progression (IP). With respect to this pattern, the theme of the sentence becomes a rheme of the new sentence (see Fig. 1.e).

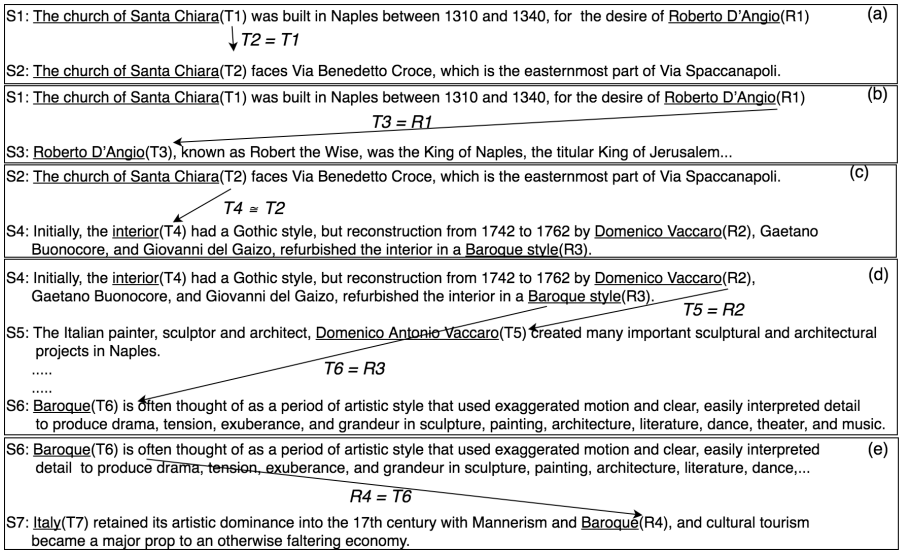

Fig. 1. Examples of Thematic Progression Pattern: (a) Constant; (b) Linear; (c) Derived-Theme; (d) Split-Rheme; (e) Inverse.

2.1 Correspondence between event components and theme-rheme elements

One of the key ideas of this work is the use of sentences annotated by the CSWL formalism to represent the thematic structure. Generally, a theme corresponds to

the first noun phrase of a sentence that is participant, circumstance, or process, while the last part of sentence contains the rhemes[7].

Our starting point is the annotation in CSWL[1]. CSWL is an event-based formalism and defines three types of entities changing over time: *simple events*, *complex events* and *fluents*. **Simple Events** are represented by four components: *When* - the time interval in which the event happens; *What* - the action happening in the event; *Where* - the location where an event takes place; and *Who* - the participants in the event. In CSWL, stories are represented through complex events. A **Complex Event** is constituted by a set of events, causal and temporal relationships between them, and all properties holding true over the time in which the story unfolds. Through the **Fluents** one can represent properties, mental events, spatial relations, and meronomic relations, which are all entities that can change over the time.

Using event-based annotation, for each sentence S we have one or more events annotated, and for each event e_i there are four components:

$$\begin{aligned} Event(S) &= e_1, e_2, ...e_n \\ Components(e_i) &=< what_i, who_i, where_i, when_i > \end{aligned}$$

In our approach, we select as theme the first event component that appears in the sentence and the remaining components as rhemes. If we consider the following sentence with related CSWL annotation:

S_1: La basilica di Santa Chiara$_{po1}$ in Napoli$_{loc1}$ fu costruita$_{act1}$ tra 1310 e 1340$_{int1}$.
- Event$(S_1) = e_1$
- Compoments(e_1)= $< act_1, po_1, loc_1, int_1 >$
  with $act1 : Action, po1 : PhysicalObject, loc1 : Location, int1 : Interval$

In accordance with this representation, the theme is po_1, that is "la basilica di Santa Chiara" and the rhemes are act_1, loc_1, int_1. Each thematic schema defined starting from the CSWL annotation, can be enriched with other semantic relations such as: meronomic, hyperonomic, synonymous. Then, using this representation, we implement the thematic progression presented in the previous section. Through this approach we choose the next sentences according to both thematic progression and semantic relations.

3 Response as regular expansion

The construction of natural and self-explanatory responses for the visitor needs to be built with respect to some criteria. In a museum, a key constraint of the visit is time. Each visitor has a limited amount of time to spend for museum visits. So, the systems to generate responses have to take in to account such value. Also, during a presentation, the text need to be clear, so it is necessary to present and explain all new terms introduced for the first time. Of course, in such process, the text does not explain obvious things. The choice whether to explain a concept, or not, can be made in accordance to the visit context and/or user

profile (background). For example, if we consider the sentence *The Basilica of Saint Clara in Naples was built between 1310 and 1340*, and the visitor is located in Naples, it's pointless to explain something about Naples, vice versa, it can be useful if the visitor is listening to the story in another country. Finally, the response built for the user has to be balanced, ensuring that the text does not present insights too large with respect to specific subjects and thus, deflecting attention from the main topic.

3.1 Dialogue responses using thematic progression

In this section we present, through some examples, how we can expand a sentence using a thematic propagation that takes into account the semantic annotation of sentences. An important step is to define the procedures for browsing the thematic structure of sentences, to research the text that composes the response.

Listing 1.1. Pseudocode 1

```
1  def spQuery_Expansion(DialogueState,Sx, Tr)
2      Tr = Tr - time(Sx) # response time
3      Expansion = Sx     # expanded response
4      DS = DialogueState # sentences already uttered
5
6      if Tr>0:
7          for each S in (Text -(DS U Expansion)):
8              if  (thematic_progression(S,SX) and
9                   semantic_relation(S, Sx) and time(S) <= Tr+Dt):
10                     Expansion = Expansion + S
11                     Tr = Tr - time(S)
12     if Tr>0:
13         for each S in (Text - (DS U Expansion))
14             if  (linear_expansion(S,Sx) and time(S)<= Tr+Dt):
15                     Expansion = Expansion + S
16                     Tr = Tr - time(S)
17     if Tr>0:
18         for each S in (Text- (DS U Expansion))
19             if  (constant_progression(S,Sx) and Time(S)<= Tr+Dt):
20                     Expansion = Expansion + S
21                     Tr = Tr - time(S)
22     return Expansion
```

Generally, the main theme is related to a cultural asset, but in an interactive system based on dialogue, the starting theme for the search depends on the user query (Q). As a first step, the system identifies the event, and the corresponding sentence (Sx)[8], that answers the user query Q. So in the expansion phase we will have as the starting point the sentence Sx. The listing 1.1 shows the pseudocode of algorithm for expansion of sentence Sx with respect to a specific query Q (e.g. *"When did Vaccaro renovated the Basilica?"* or *"When was it bombed?"*.

As a first step (lines 6..10), the algorithm finds all sentences strongly connected with Sx. We assume that two sentences S_1 and S_2 are strongly connected

if there is a thematic progression and a semantic relation between them. For semantic relations, we have used causal relations and hyperonomy relations. In the first case, this means that there is a cause and effect relation between two sentences, and in the composition process it is more useful to choose both sentences for the explanation. While in the second case, there exists a specialisation of some sentence component, which is an insight.

After this selection, the algorithm, for the next text, searches (lines 11..16) the sentences S which have a Linear Progression ($Theme(S) \in Rheme(Sx)$), which are sentences that have as theme one of the rhemes belonging to Sx.

Listing 1.2. Peudocode 2

```
1  def genQuery_Expansion(DialogueState,Sx, Tr)
2      Tr = Tr - time(Sx) # response time
3      Expansion = Sx     # expanded response
4      DS = DialogueState # contains the sentences already uttered
5
6      if Tr>0:
7          for each S in (Text- (DS U Expansion)):
8              if (constant_progression(S,Sx) and time(S)<= Tr+Dt):
9                      Expansion = Expansion + S
10                     Tr = Tr - time(S)
11     if Tr>0:
12         for each S in (Text- (DS U Expansion)):
13             if (inverse_progression(S,Sx) and time(S)<= Tr+Dt):
14                     Expansion = Expansion + S
15                     Tr = Tr - time(S)
16     if Tr>0:
17         for each S in (Text- (DS U Expansion)):
18             if (derived_progression(S,Sx) and time(S)<= Tr+DX):
19                     Expansion = Expansion + S
20                     Tr = Tr - time(S)
21     if Tr>0:
22         for each S in (Text-(DS U Expansion)) and St in Expansion:
23             if (linear_progression(S,St) and time(S)<=Tr+DX):
24                     Expansion = Expansion + S
25                     Tr = Tr - time(S)
26     return Expansion
```

If the user's query contains a generic request (e.g. *"can you tell me something about the basilica?"* or *"can you give me some information about the style"*, this means that he/she does not ask specific information about the topic expressed in the query. The listing 1.2 shows the pseudo-code of the algorithm for the expansion of sentences with respect to a generic query Q. As in the previous case, the starting point is the first sentence Sx that the system provides as a response. In this case, the request is generic, and it asks information about a topic. So, the algorithm finds all sentences that have a Constant progression with respect to Sx (lines 6..10) and if there is more narration time, it searches (lines 11..15) sentences that have inverse progression ($Theme(Sx) \in Rhemes(S)$) with

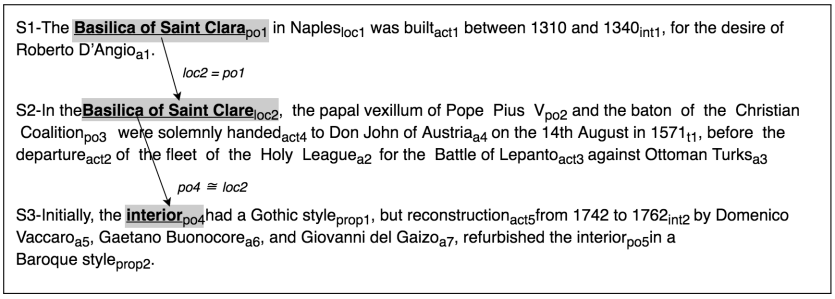

Fig. 2. Answer of the query *"Can you tell me about the Basilica?"*

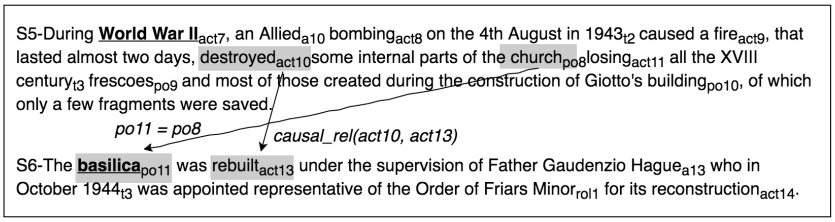

Fig. 3. Answer of the query *"When was it bombed?"*

respect to the theme of the query, or try to find sentences with derived theme progression (lines 16..20). If there is more available time for the narration, the algorithm finds some deepening of the sentences that were already selected in the previous steps (lines 21..24). Using this approach, the composed text will present a generic description of the topic required and, in accordance with the available time for the response, some deeper insights.

As experimentation of the adopted approach, we have considered a cultural text about the Basilica of Santa Chiara in Naples. Asking the query *"can you tell me about the basilica?"* we can obtain the answer shown in Fig. 2. In this figure, the first sentence is selected as the response by the system. To build an expansion, the algorithm in listing 1.2 is applied. It adds two sentences to the response: the first with constant progression and the second with derived theme progression. In the latter, the system can detect the theme through a meronomic relation. In fact, the interior (po_4) is *part of* of the basilica (loc_2).

If we consider the query *"When was it bombed?"* we can obtain an answer as shown in Fig. 3. In this case, starting from the sentence that contains the response, for the expansion the algorithm presented in listing 1.1 is used. We can observe that between the two sentences exist a linear progression and a causal relation, so they are strongly connected.

3.2 Syncretic approach for multimedia responses

To build the composed text for responding to the user question, we create a multimedia response temporally synchronising text and media according to semantic

annotations. This approach is called **syncretic text**[9], namely a text composed of heterogeneous languages within a unitary communications model[2], having features of cohesion and coherence, respect to a same enunciation instance. For these goals, the system selects, and carries out a ranking, using available multimedia objects, that can be associated to the composed text. The selection is based on annotated entities using the semantic of CSWL[8] formalism. Then, multimedia objects selected are synchronised with synthesised text, so that media items are coherently visualised with the time intervals in which a synthetic voice talks about the content represented in the media.

4 Conclusions and Future Work

In this work, we presented an ongoing research activity about the composition of balanced texts that uses a thematic progression structure built through an event based formalism. What we have presented here represents just a first application that composes texts using information coming from a single document, but we believe that this approach can be adopted for building texts integrating more documents. Future work will consist of analysing more texts to validate the patterns of thematic progression and discover new ones. We also believe that presented algorithms can be improved taking into account some characteristics of the user profile. In addition, because the approach is based on CSWL annotation, to reduce the time in such phase, we are working on an assisted tool that helps the users in the annotation.

References

1. Mele, F., Sorgente, A.: CSWL - Cultural Stories Web Language. Technical Report 180/15, Institute of Cybernetics "E. Caianiello" of CNR (06 March 2015)
2. Sorgente, A., Calabrese, A., Coda, G., Vanacore, P., Mele, F.: Building multimedial dialogues annotating heterogeneous resources. In Bordoni, L., Mele, F., Sorgente, A., eds.: Artificial Intelligence for Cultural Heritage. Cambridge Scholars Publishing (2016) 49–82
3. Ebrahimi, S.F., Ebrahimi, S.: Information development in efl students composition writing. Advances in Asian Social Science **1**(2) (2012) 212–217
4. Halliday, M.A.K., Matthiessen, C.M.: Halliday's introduction to functional grammar. Routledge (2013)
5. Danes, F.: Functional sentence perspective and the organization of the text. In Papers on Functional Sentence Perspective (1974) 106–128
6. Soepriatmadji, L.: Thematic progression in thesis abstacts written by english students of fbib unisbank semarang. Jurnal Ilmiah Dinamika Bahasa dan Budaya **3**(1) (2009) 28–40
7. Jing, W.: Theme and thematic progression in english writing teaching. Journal of Education and Practice **6**(21) (2015) 178–187
8. Sorgente, A., Vanacore, P., Origlia, A., Leone, E., Cutugno, F., Mele, F.: Multimedia responses in natural language dialogues. In: Proceedings of AVI*CH 2016. Volume 1621 of CEUR Workshop Proceedings., CEUR-WS.org (2016) 15–18
9. Cosenza, G.: Semiotica dei nuovi media. Laterza (2010)