Towards a computational generation of récit: evaluating the perception of the récit plan

Belén A. Baez Miranda **Sybille Caffiau** **Catherine Garbay** **François Portet**
Univ. Grenoble Alpes, LIG, F-38000 Grenoble, France
FirstName.LastName@imag.fr

1 Introduction

The aim for this PhD project is to generate stories about human activities from real ambient data. Computational Narratology (CN) is the study of narratives from the point of view of computation and information processing (Mani, 2012). Most of the current researches in CN are related to creativity, where the stories emerge from a set of predefined parameters, trying to imitate literary genres like fairy tales (Riedl and Young, 2010). In this paper we are interested in stories emerging from real ambient data for which we have no control. We name these stories ”activity récits”.

One challenge addressed is the coherence between a story plan (i.e. a sequence of events) rendered as text read by humans and real activity performed by actors and gathered through sensors. This paper aims to present an evaluation of the perception of the story plan from a tool employed to express human activities: a task model sequence (Baez et al., 2014).

2 The activity récit and its plan

According to Adam (2011), a *récit* is a succession of events related to facts that have been effectively experienced, observed or captured. His definition is suited to fictional as well as actually lived events. Our problem statement lays onto the récit plan. On this subject, we take the Adam's (2011) definition of a récit plan, where a récit is a set of chained scenes told from an actor's perspective and linked in a temporal and causal way. The récit plan used in our approach is the narrative structure.

To represent the récit plan we propose to use the task model (Caffiau et al., 2011), which has been used in fictional but interactive stories (Cavazza et al, 2002). We use ski touring as a use case. During their ski journey, skiers, alone or in groups, use special devices such as portable GPS. Once finished, some of them share on websites such as `www.skitour.fr` the experience, observations, and other evaluative elements (weather conditions, terrain and key places to visit) that could be useful or interesting for others skiers.

In Figure 1 we present the different layers of our approach for story generation. First, the raw data is captured and interpreted. Second, those interpretations of the data are structured and organized in order to link them together according to a task model. Third, one of the sequences of the task model identified as the récit plan is selected and used as a document plan during the generation stage. The result is an activity récit that emerges directly from the sensor data but is organized according to the task model expressing the human activity. The work presented in this paper focuses on the use of a sequence extracted from a task model to generate an activity récit.

3 Evaluation

To evaluate the temporal perception of the récit plan, we asked 18 French speakers (12 men and 6 women) aged between 19 and 38 to rebuild the ski touring sortie in chronological order after reading three texts separately. The text selection was performed based on the text size, the complexity of the ski touring sortie; the clarity of the description of the sortie, the linguistic quality and finally the number of protagonists of the sortie and the level of expertise shown in the narration of the sortie. The duration of the experiment was 25 min on average. Each text was presented in either two versions, (i) the original human written text from the collected corpus of ski touring récits and (ii) the generated text based on the task model. The experiment consisted in sorting cards of basic activities into the sequence of the actual sortie using adhesive tape and a paper-made timeline. Once the reading was finished, the reader choose the cards corresponding to the events encountered in the text. Then, all the cards were arranged on the timeline according to the chronological order perceived during the lecture. The participants did not know whether the text presented was generated automatically or not.

The distance between the participant's answers and the reference récit plan was computed using the Word Error Rate (WER), the standard evaluation metric in Speech recognition. Figure 2 shows

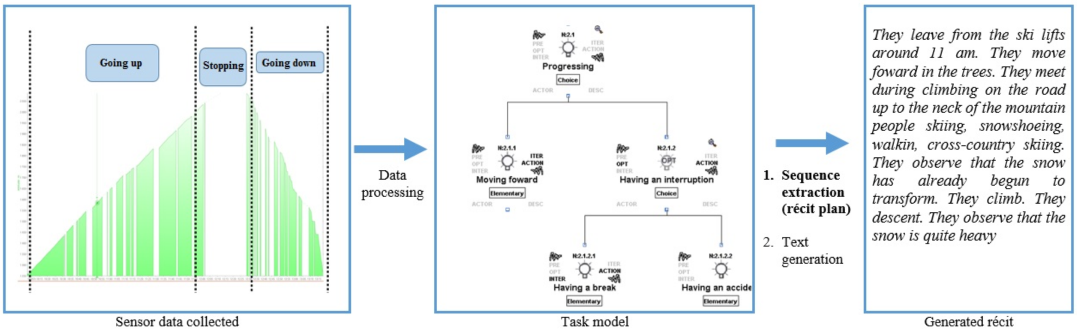

Figure 1: Representation layers for story generation.

the box plot of the "distance" from participant to the reference according to the "mode" of presentation as well as the number of cards used in each case ("size"). An ANOVA performed on the distance value showed a significant effect of text ($F(1,18) = 7.583$, $p=0.0131$). A participant effect was also found ($F(17,18)=2.281$, $p=0.0457$). Regarding the size of the participant's sequences a difference between the human text and the generated one was found ($F(1,48) = 5.604$, $p = 0.022$) and a text effect ($F(1,18)=3.666$, $p=0.033$), that appears significant when the text is taken as a factor. It thus seems that the generated texts induce significantly less errors during the activity identification than the original ones ($F(1,18)=8.993$, $p=0.00771$).

4 Discussion

In our evaluation, the identification and sequence reconstruction on the timeline was slightly better for the generated texts than for the human written ones, even if the number of activities presented in both was the same. We can observe differences in the number of activities perceived. The results could suggest that the generated texts can more explicitly show the sequence of events because of the lack of detail describing them (one sentence is associated to each event and each task in the task model is linked to a verb for making sentences during the generation stage). Human texts have a higher amount of information, making believe to the participant that there is more activities.

Regarding the distance, the generated texts presents a more explicit chronological order and that may explain why participants were able to perceive easily the structure of the events sequence and to reconstruct the path. In human texts, the chronological order is more implicit because of the text configuration, which can include much detail or events omissions (e.g ellipses).

However, it could be possible to find that some activities were identified in the human texts but not in the generated ones. This could be explained by many reasons, such as a possible lack of coverage during the task model construction; activities not identified during the corpus analysis or the fact that, due to the presence of ambiguity in human texts, the participants did not distinguish the activities correctly. Improvements in the task model and in the collection and analysis of the corpus would thus be needed in order to make the approach more robust.

In order to improve and to produce a more natural text, we need to explore other aspects such temporality. Currently, the récit plan from the task model can produce a sequence of events linked in a causal way by establishing preconditions and effects during the task model construction. However, this is not reflected in the generated texts. So, we need to add discourse connectors that indicate this causal links. Rendering simultaneous tasks is also an important feature to add to the model. The task model can express this, but it is not yet reflected in the generated text.

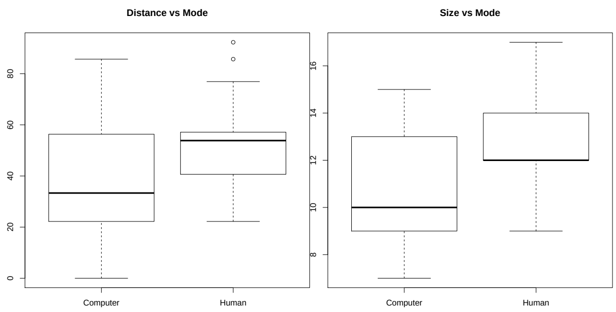

Figure 2: Distance score (left) and size of sequences (right) wrt the writing mode.

Finally, generating an activity récit from sensor data raises have specific issues, for example a loss of information during collection process. If an event is not recorded, it will not be in the generated récit. Inferencing and reasoning processes are then needed to cope with this lack of information and keep the récit consistent.

References

Jean-Michel Adam 1997. *Genre de récits. Narrativité et généricité des textes* , Academia, 2011

Belén Baez Miranda, Sybille Caffiau, Catherine Garbay, François Portet. 2014 Task based model for récit generation from sensor data: an early experiment., In 5th International Workshop on Computational Models of Narrative (pp. 1-10).

Sybille Caffiau, Dominique Scapin, Patrick Girard, Mickaël Baron and Francis Jambon 2011. *Increasing the expressive power of task analysis: Systematic comparision and empirical assessment of tool-supported task models.*, Interacting with computers, 22(6), 569-593. doi:10.1016/j.intcom.2010.06.003, 2010

Marc Cavazza, Fred Charles, Steven J. Mead. 2002 Planning characters' behaviour in interactive storytelling., The Journal of Visualization and Computer Animation, 13(2), 121-131.

Mani, Inderjeet 2013. *Computational Narratology*, Synthesis Lectures on Human Language Technologies, December 2012, Vol. 5, No. 3, Pages 1-142

Mark Owen Rield and R. Michel Young 2010. *Narrative planning: Balancing plot and character*, Journal of Artificial Intelligence Research, 39:217-268