Assembling Narratives with Associative Threads

Pierre-Luc Vaudry and **Guy Lapalme**
RALI-DIRO – Université de Montréal
C.P. 6128, succ. Centre-Ville
Montréal, QC, Canada H3C 3J8
{vaudrypl,lapalme}@iro.umontreal.ca

Abstract

A model is proposed showing how automatically extracted and manually written association rules can be used to build the structure of a narrative from real-life temporal data. The generated text's communicative goal is to help the reader construct a causal representation of the events. A connecting associative thread allows the reader to follow associations from the beginning to the end of the text. It is created using a spanning tree over a selected associative sub-network. The results of a text quality evaluation show that the texts were understandable, but that flow between sentences, although not bad, could still be improved.

1 Introduction

A narrative is a text presenting with a certain angle a series of logically and chronologically related events caused or experienced by actors (Bal, 2009, p. 5). A data-to-text system summarizing temporal data including actions or activities should aim at generating such a text, if that corresponds to its users' needs. Some have pointed at causal relations as a means of improving the narrative aspect of temporal data-to-text (Hunter et al., 2012; Gervás, 2014).

The concepts of causal network and causal chain have been used to explain the process of narrative comprehension in humans (Trabasso and van Den Broek, 1985; Trabasso et al., 1989). Those causal networks are essentially composed of physical and mental events and states (of which goals and actions) connected by causal relations. Restrictions apply on which types of causal relation can connect which types of event or state. The causal chain comprises the events that are on a path traversing the causal network from the introduction of the protagonists and setting to either goal attainment or the consequences of failure. Being on a causal chain and having more causal connections have both been found to increase chances of an event being recalled, included in a summary or judged important by the reader.

Swartjes and Theune (2006) and Theune et al. (2007) applied causal networks to the automatic creation of fairy tales. Several narrative data-to-text systems already identify and make use of some causal relations (Hallett, 2008; Hunter et al., 2012; Wanner et al., 2010; Bouayad-Agha et al., 2012). Going further, in Vaudry and Lapalme (2015) we have tried to extract a form of causal network from temporal data and use it to build the structure of the generated narrative. We used data mining techniques to extract sequential association rules and interpreted them as indicating potential, approximate causal relations. The resulting causal network was used to express locally some rhetorical relations in the sense of the Rhetorical Structure Theory (RST) (Mann and Thompson, 1987). However we did not succeed at exploiting it to build a complete rhetorical structure that would give the text a global coherence.

Building on what was begun, this paper proposes a model showing how automatically extracted and manually written association rules can be used to build the entire structure of a narrative from real-life temporal data.

In the course of our research, we found that it was very difficult to infer even the direction of a

Proceedings of the INLG 2016 Workshop on Computational Creativity and Natural Language Generation, pages 1–10,
Edinburgh, September 2016.

potential causal relation from an extracted association (see Section 4). If even that could not be determined, how could we claim to identify causal relations? We prefer to simply name *associations* the relations found during data interpretation. The task of inferring causal relations is left to the human reader of the generated text.

By association, we mean a connection between events or states without specifying the nature of the underlying relation. For example, an association can be based on a frequent sequence or a formal similarity. For the purpose of narrative comprehension, we assume that interesting associations are those that can help formulate causal hypotheses.

Note that although this is not a model for creating fictional narratives, its function is to suggest new associations between previously unassociated events. In this sense and to the extent that it accomplishes this, it can be considered to produce original, creative text (Jordanous, 2012, p. 257).

The proposed model assumes that the human reader can follow an associative thread from the beginning to the end of the text. The associations expressed between some of the events can give him hints toward building a mental representation of the events. His world and domain knowledge can enable him to sort through the expressed associations to retain and enrich the relevant ones. This can lead him to fill the gaps left by the text towards a causal interpretation of the events.

Section 2 presents our model of assisted temporal data interpretation. Section 3 presents the results of our efforts so far to evaluate this model. Related work is discussed in Section 4.

2 Model

This section presents our model of assisted temporal data interpretation using narrative generation. Figure 1 gives an overview of this model. We will refer to its components by using numbers for steps and letters for representation levels. Association rules come from two sources: data mining (1) for sequential association rules (B) from training data (A) and world and domain knowledge (C) formalized as rules (D). The data about a specific period (E) is interpreted (2) using the association rules to create an associative network (F). Then a sub-network containing the most unusual facts (G) is selected (3) using the probabilities of the corresponding sequential association rules (B). The following step of document structuring (4) involves determining the connecting associative thread going from the beginning to the end of the narrative (H). Microplanning (5) produces from this the lexico-syntactic specification (I). This specification is then realized (6) as a text (J) read by a human (7). The human reader uses his knowledge (C) to reason about the associations expressed in the text. From this he forms a mental representation which hypothetically includes a form of causal network (K). The following subsections detail each of these steps.

The communicative goal of the generated text in the context of this model is to communicate effectively the facts necessary to facilitate the construction of a causal network by the reader. By necessary facts, we mean the least easily predictable facts. Those facts are the most unusual (or least usual) of the summarized period compared to a typical period of the same kind of data. They are what makes this period unique. The associations expressed in the generated text should give valuable hints to the reader in constructing a causal mental representation of the events. Moreover, they should generally help see the events of the period as a coherent whole if such coherence can be found. This should help the reader assimilate effectively the text's content.

The facts not mentioned in the text should be implicitly understood as "same as usual" and the reader should be able to infer them approximately from the text's content if needed. According to Niehaus and Young (2014), the reader will make such an inference if it is necessary to the comprehension of the text (because of a break in a causal chain, for example) and not too difficult to make. The knowledge that the reader has of what usually happens, if the sequential association rules model that correctly enough, should enable the reader to make such inferences. In the case of the inferences that could be triggered in the reader by the expressed associations, it is much more difficult to use the criteria of necessity and enabledness, as exactly what should be inferred or not is not known by the computer.

To illustrate the various representation levels of the model, an example in the Activity of Daily Living (ADL) domain is provided in Figures 2, 3, 4, and

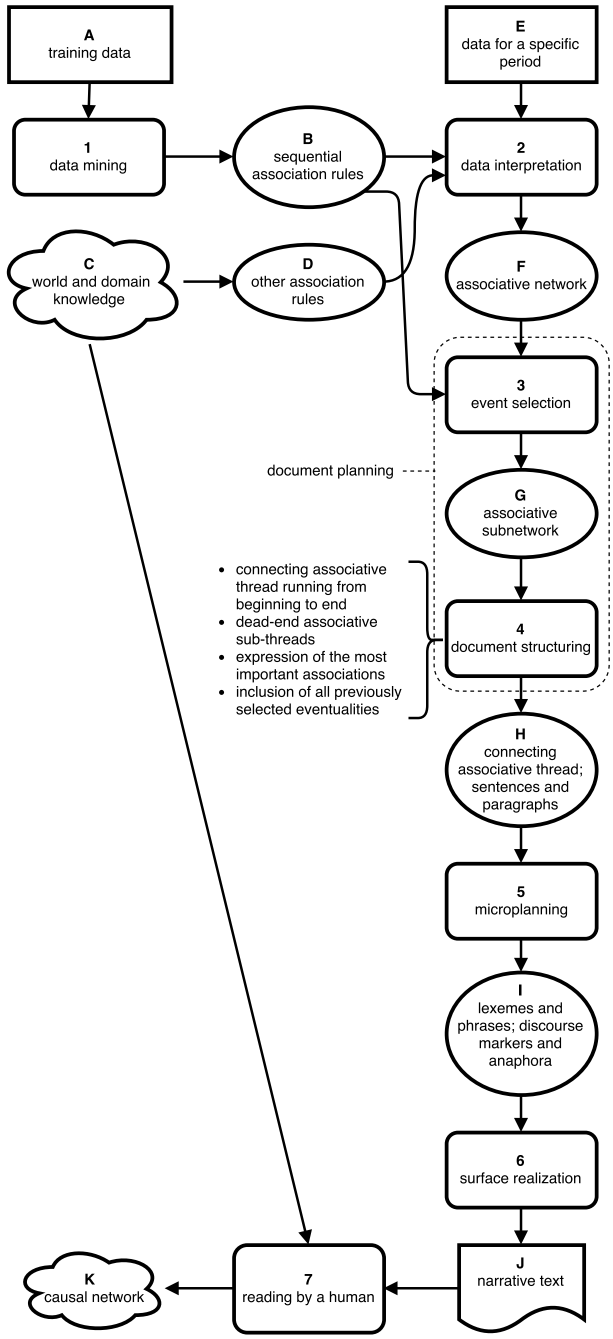

Figure 1: Assisted temporal data interpretation model. Rectangles represent input data; rounded rectangles: computational representations; ellipses: steps; clouds: hypothesized mental representations; rectangle with S-shaped bottom side: natural language document. For ease of reference, steps are identified by a number and representations by a letter.

5. The data it is based on is taken from the publicly available UCI ADL Binary Dataset (Ordóñez et al., 2013). This dataset contains 14 and 21 consecutive days of ADL data for users A and B, respectively. The data for each ADL occurrence consists of: start time, end time and activity label. The ADL label set is: *Sleeping*, *Toileting*, *Grooming*, *Showering*, *Breakfast*, *Lunch*, *Dinner*, *Snack*, *Spare_Time/TV*, *Leaving*. The input for this example consists of the data for user B as training data (A) and the portion covering the day of November 24, 2012 as the data to summarize (E).

2.1 Association Rules

Sequential Association Rule Mining: In step 1 on Figure 1, data mining techniques are used to select candidate sequential association rules based on confidence and significance. Confidence (cf in Figure 2) is computed as the conditional probability of encountering an instance of the rule given that the left side has been encountered. Depending on the confidence, associations are considered expected or unexpected. In the example, expected association rule candidates had $cf > 0.2$ and unexpected association rule candidates had $cf < 0.07$. This is roughly justified by the fact that since there are 10 activity types, the prior probability of one happening at any place in the sequence is 0.1. Significance measures the chances of the left and right sides of the rule of actually being independent according to the binomial distribution. In Figure 2, the p-values according to this distribution are called $p_{expected}$ and $p_{unexpected}$ for expected and unexpected association rules, respectively. In the example, a p-value lower than 0.05 was considered significant.

A rule can express a backward prediction. The chronological direction of each association rule is determined by computing the confidence for the two possible directions (chronological and reverse chronological) and retaining the direction with the highest one. That means that for candidate association AB, we checked which we could predict with more confidence: that B follows A or that A precedes B. This enabled us to better estimate the unusualness of each fact and thus improve content selection.

Rules 1 to 5 of Figure 2 are examples of mined sequential association rules.

Mined sequential association rules:

1. $H_{11,p} \rightarrow A_{Grooming,p}$
$cf = 0.67, p_{expected} = 0.000005$

2. $A_{Sleeping,p-1} \leftarrow A_{Breakfast,p}$
$cf = 0.23, p_{expected} = 0.01$

3. $A_{Showering,p-2} \rightarrow A_{Grooming,p}$
$cf = 0.64, p_{expected} = 0.01$

4. $A_{Grooming,p-2} \wedge A_{Toileting,p-1} \rightarrow A_{Grooming,p}$
$cf = 0.58, p_{expected} = 0.001$

5. $A_{Toileting,p-1} \wedge H_{10,p} \not\rightarrow A_{Spare\ time/TV,p}$
$cf = 0.04, p_{unexpected} = 0.03$

World and domain knowledge association rule:

6. $A_{i,p} \xleftrightarrow{Same\ category} A_{j,q} \iff category(i) = category(j)$

Figure 2: Association rule examples. A and H are categorical variables and stand respectively for activity and hour of the day (hours 0-23, not considering minutes). A_i, p stands for a particular type of activity i at position p in the event sequence. cf stands for confidence. $p_{expected}$ and $p_{unexpected}$ are p-values that measure the significance of expected and unexpected association rules, respectively (lower is better).

World and Domain Knowledge Rules: World and domain knowledge can be formalized as rules (C and D in Figure 1). Those rules can be manually entered or come from an existing ontology, for example. The associations they create have the advantage of linking events regardless of their place in the sequence. That means that we can use them to create long-distance links in the text while keeping temporally close events also close in the text.

Rule 6 of Figure 2 is a simple but effective example of a manually entered association rule. It defines a *Same category* association. For the purpose of the ADL example, we arbitrarily grouped the ADL types into categories in the following manner. *Toileting*, *Grooming*, and *Showering* were placed in the category of personal hygiene activities. *Breakfast*, *Lunch*, *Dinner*, and *Snack* were grouped as eating activities. *Spare_Time/TV*, *Leaving*, and *Sleeping* were kept in separate categories.

2.2 Data Interpretation

Step 2 of Figure 1 consists of searching the data to summarize for instances where an association rule applies. Sequential associations are derived from rules such as Rules 2 to 5 from Figure 2. They are shown as arrows going from one row to another at the left of Figure 3. The arrow labels indicate the confidence of the corresponding association rule. Temporal associations are derived from rules such as Rules 1 and 5 from Figure 2. They are indicated by the *Time prob.* and *Temporal association* columns in Figure 3. *Usual* means that an expected association was found and *Unusual* indicates an unexpected association. No indication means that time was not considered significantly useful in predicting those occurrences (no association rule). The probability conditional on time (the confidence of the corresponding association rule candidate) is in any case indicated as it will be used for content selection.

From there, some extra associations are derived and added to the network. The *Repetition* association is generated whenever the type of activity that appears on the right side of the association rule also appears on the left side. *Conjunction* is added when two sequential associations start or end at the same activity. Their other ends are then linked by a *Conjunction* association. The *Instead* association appears when an unexpected association is found. It indicates what would have been the most probable alternate activity according to the sequential association rule model. Derived associations are shown on the right of the first column of Figure 3.

2.3 Event Selection

As can be seen on Figure 1, event selection (step 3) takes as input the associative network and outputs a sub-network of its input. Note that final association selection takes place later, during document structuring, as they are used to build the document structure.

In Figure 3, the output of event selection is shown in bold type. Event selection has one parameter: a maximum probability threshold. Events that have either a probability conditioned on time or an association with a confidence lower or equal to the threshold are selected. In this example, the maximum probability threshold was set to 0.3. Generally the ideal value of the threshold varies in function of

	Start time	Activity	Time prob.	Temporal association
0.23	**00:33**	**Sleeping**	0.33	Usual
	10:04	**Breakfast**	0.33	Usual
	10:17	**Toileting**	0.37	Usual
0.04	**10:19**	**Spare time/TV**	**0.04**	Unusual
0.45	***10:19***	***Grooming***	–	–
	11:16	Snack	0.36	–
	11:30	**Showering**	**0.17**	–
0.91 0.64	11:39	Grooming	0.67	Usual
	11:59	Grooming	0.67	Usual
	12:01	**Toileting**	**0.30**	–
	12:09	**Snack**	**0.28**	–
0.51	12:31	Spare time/TV	0.40	Usual
	13:50	Spare time/TV	0.57	Usual
	14:32	Grooming	0.42	Usual
	14:36	**Leaving**	**0.29**	–
	16:00	Toileting	0.52	Usual
0.37	16:01	Grooming	0.35	–
0.58	16:02	Toileting	0.52	Usual
0.58	16:03	Grooming	0.35	–
	16:04	Spare time/TV	0.65	–
0.45	19:58	Snack	0.44	–
0.51	20:08	Spare time/TV	0.83	–
	22:01	Toileting	0.14	–
0.37	22:02	Spare time/TV	0.62	Usual
0.37	22:17	Dinner	0.55	Usual
0.64	22:19	Spare time/TV	0.62	Usual
0.45	**23:21**	**Snack**	**0.27**	–
0.51	23:23	Spare time/TV	0.87	–
	00:45	Grooming	0.74	Usual
	00:48	Spare time/TV	0.44	–
	01:50	Sleeping	0.45	Usual

Figure 3: Associative network for user B on November 24, 2012. The events selected with maximum probability 0.3 are shown in bold type. Sequential associations are on the left. The X-headed arrow represents an unexpected association. On the right are *Instead* (dotted), *Conjunction* (dashed), and *Repetition* (double). *Same category* associations are not shown.

how well the sequential rule model captures what usually happens and the desired average length of the generated text.

2.4 Document Structuring

Connecting Associative Thread: The main goal of document structuring (step 4 in Figure 1) is to give the text a simple narrative structure including a beginning, a middle section, and an end. The importance of this structure for narrative generation was highlighted by a comparison with human written texts (McKinlay et al., 2009). The first event of the period (chronologically) is selected to be the beginning of the text and is called the initial situation (*Sleeping 00:33* in the example of Figure 3). The last event of the period is correspondingly called the final situation (*Sleeping 01:50* in the example). The (rest of the) selected associative sub-network will form the middle section (in bold type in Figure 3). The best event pairs are then chosen to link the selected events with each other. In the example, event pairs with sequential associations are preferred over those with only *Same category* associations. Manually set parameters, called association preferences, define which association types are preferred. They take a value between 0.0 and 1.0. A smaller value gives an event pair with this association type more chances to be chosen. When no other association is present, the default association of temporal proximity is used with association preference 1.0. The association preference is combined (by averaging) with the relative temporal distance in order to favor temporally close event pairs. The resulting score is then used as a distance to compute a minimum spanning tree on the selected associative sub-network.

This minimum spanning tree is converted into a directed rooted tree by designating the initial situation as its root. This tree is hereafter called the connecting associative thread. The path from the initial situation to the final situation is the main associative thread. The other branches of the spanning tree are said to be dead-end threads because once the text has reached their end, it must go back to the connection point with the main thread before continuing toward the final situation. The connecting associative thread connects every event together through the main thread and the dead-end threads. This is illustrated in Figure 4.

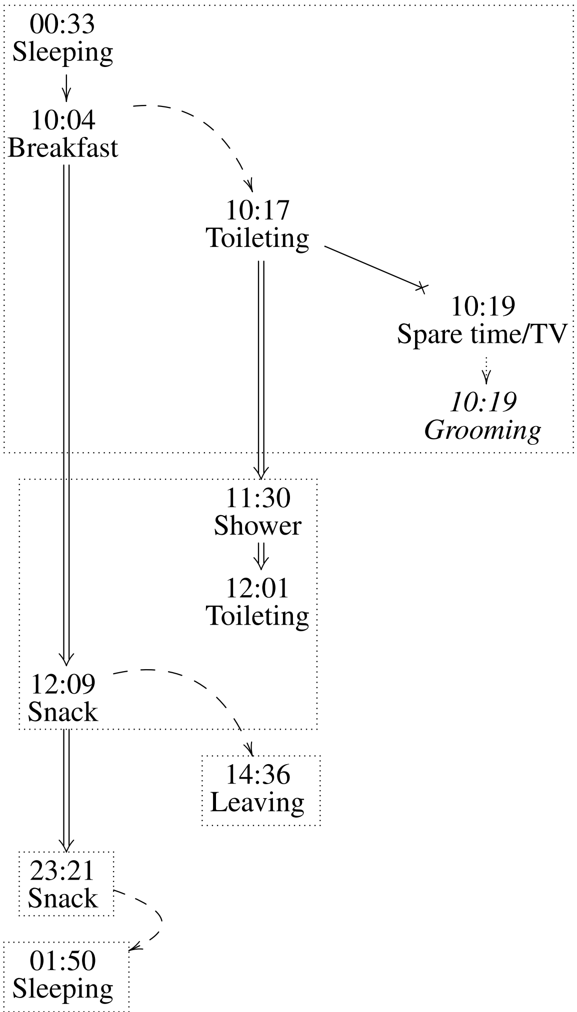

Figure 4: Connecting associative thread for user B on November 24, 2012. Arrows represent associations: simple: expected sequence; X-headed: unexpected sequence; double: *Same category*; dotted: *Instead*; curved and dashed: temporal proximity. Paragraphs are boxed. The vertical order of presentation is the order of mention in the generated text (Figure 5). For event selection, the maximum probability threshold was set to 0.3.

Research on causality in narrative comprehension has uncovered that events on the causal chain going from the beginning to the end of the story are more often recalled than those on dead-end parts of the causal network (Trabasso and van Den Broek, 1985). In the future, it may be interesting to verify if eventualities on the associative sub-threads are less remembered than those on the main associative thread. If this is the case, the content structuring algorithm should be modified to optimize the importance of the expressed associations together with the proportion and importance of the eventualities included in the main associative thread. However taking into account the relative temporal distance in the computation of the minimum spanning tree already tends to avoid a too short main associative thread.

Paragraph and Sentence Segmentation: The document content is then segmented into sentences and paragraphs. The style can be varied by adjusting two parameters: the average number of events introduced in one sentence and the average number of sentences in one paragraph. Those parameters are used to calculate the number of breaks needed between sentences and paragraphs. The candidate break points are between consecutive event pairs in the document plan. The actual break points are selected according to the distance computed previously for the determination of the minimum spanning tree. The greatest distances correspond to paragraph breaks, then sentence breaks, and lastly phrase boundaries. Paragraphs are boxed in Figure 4.

At this point, a mapping is made between the selected associations and the rhetorical relations that will be expressed in the text. In the example, sequential associations are expressed by a Temporal Sequence relation and *Same category* associations are expressed by a Conjunction relation.

2.5 Microplanning

Microplanning (step 5 of Figure 1) translates the rhetorical structure into a lexico-syntactic specification. Each sentence plan tree is traversed depth-first. When a leaf is visited, a specification of the corresponding eventuality's description is produced from lexico-syntactic templates. When an internal node is visited, the rhetorical relations linking the two children nodes are expressed with appropriate discourse markers. Those markers are then used to assemble the lexico-syntactic specifications obtained from the children nodes.

However the marking of rhetorical relations between sentences is handled differently. Each sentence has a main event, which is the one expressed by its first independent clause. The main event of a paragraph is the main event of its first sentence. A rhetorical relation marker is placed at the front of a sentence to indicate its parent relation in the connecting associative thread. If its parent is the main event of the preceding sentence, or the main event of the preceding paragraph in the case of the first sentence of a paragraph, the marker appears alone. If not, an anaphoric expression is added that restates the parent event. For example, the parent of *Toileting 12:01* in Figure 4 is *Shower 11:40*. Since it is the main event of the preceding sentence, no anaphor is added and we have just the marker *also* in the generated text (Figure 5). On the contrary, the parent of *Snack 12:09* is *Breakfast 10:04*. It is located in another paragraph. Consequently, the marker becomes *beside his 10:04 PM breakfast*.

2.6 Surface Realization

Surface realization (step 6 of Figure 1) was performed using the SimpleNLG-EnFr Java library (Vaudry and Lapalme, 2013). During surface realization, the syntactic and lexical specifications are combined with the output language grammar and lexicon to generate formatted natural language text. The lexico-syntactic templates used in microplanning were written for both English and French output languages. In combination with SimpleNLG-EnFr, this enabled bilingual generation.

An example of English generated text corresponding to the preceding figures is given in Figure 5.

```
OrdonezB Saturday, 24 November
2012 12:33 AM - Sunday,
25 November 2012 09:24 AM
------------------------------------

OrdonezB got up at 10:02 AM and then
he ate his breakfast.  As usual at
10:17 AM he went to the toilet but
then he unexpectedly spent 1 hour in
the living room instead of grooming.

In addition to having gone to the
toilet at 10:17 AM, he took a shower
at 11:30 AM. Also at 12:01 PM he went
to the toilet.  Beside his 10:04 AM
breakfast, he had a snack at 12:09
PM.

At 2:36 PM he left for 1 hour.

In addition to his 12:09 PM snack, he
had a snack at 11:21 PM.

As usual at 1:50 AM he went to bed.
```

Figure 5: Generated text example for user B on November 24, 2012. The maximum probability threshold was set to 0.3.

2.7 Human Reading

Finally, in step 7 of Figure 1 a human reader combines his world and domain knowledge with the generated text to construct a causal mental representation of the events. For that the reader can follow the connecting associative thread through the text while trying to infer possible causal relations.

We hypothesize that statistically identifying sequential associations is a useful pre-processing of the data for the purpose of determining causal relations. Association rules based on type could also be helpful because events of the same type sometimes have the same cause or the same type of cause. Other association rules based on such causal reasoning could also give useful hints. In any case, the reader can choose to ignore irrelevant associations.

For example, the fact that the clauses expressing *Sleeping 00:33* and *Breakfast 10:04* are coordinated in the same sentence and linked by the temporal marker *then* could lead the reader to different conclusions depending of his knowledge. On one hand, he could think that maybe the user was particularly hungry when he woke up that morning; he could ponder why. On the other hand, he could also ignore this sequence as just a random happening. Another example: the fact that *Snack 23:21* references *Snack 12:09* could make the reader conclude that maybe the user was often hungry on that day and maybe there was a common cause for that. Or the reader may ignore this, reasoning that *Snack 12:09* was probably in reality a *Lunch* activity. The point is that some of the associations can help the reader in forming causal hypotheses. The reader can later verify those, for example by asking the user. Moreover, those causal hypotheses can help the reader remember the content of the text.

3 Evaluation

We asked judges to evaluate the textual quality of the reports. To assemble the evaluation corpus, a report was generated for the 32 complete days (starting and ending with a long *Sleeping* activity) of the dataset. The selection parameter was adjusted in order that texts for both users have comparable average length. The maximum probability threshold was thus set to

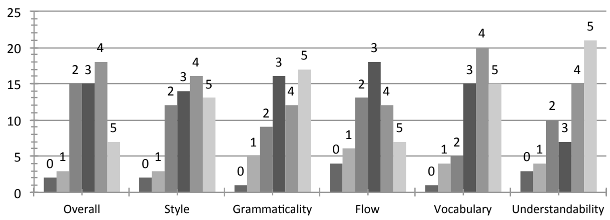

Figure 6: Results of the text quality evaluation.

0.4 for user A and 0.3 for user B. User A's routine seems to be easier to capture by the sequential rule model than user B's. Hence the probability for user A's activities is generally higher than for user B's.

Because no human-written equivalent of the generated ADL reports exists, it would have been meaningless to try to write some to make the comparison. Therefore only the generated texts are evaluated.

13 judges evaluated four to five generated texts each, so that 28 texts were evaluated by two judges each and 4 texts by one judge. The judges had to evaluate the texts on a 0 to 5 scale for six criteria: Overall, Style, Grammaticality, Flow (between sentences), Vocabulary, and Understandability. They could also leave comments. The evaluation forms, generated texts and answers are publicly available[1].

If we view all the evaluations taken together as evaluating the data-to-text system as a whole, as opposed to individual texts, we get the results shown in Figure 6. The best ratings are for Understandability and Vocabulary with peaks at 5 and 4, respectively. The worst ratings are for Flow with a peak at 3. This could indicate some deficiencies in document planning and/or microplanning. However, according to the good Understandability ratings, the texts do not seem as badly planned as to be confusing. The results for Grammaticality are hard to interpret, since there are two peaks: one at 3 and one at 5. By looking at the evaluations, we think it could be because this criterion was not defined clearly enough. Overall and Style have most ratings ranging from 2 to 5, with peaks at 4.

4 Related Work

Chambers and Jurafsky (2008) learn narrative event chains (partially ordered sets of events with a common protagonist) from a news stories corpus. For this they use pointwise mutual information (PMI) to measure the relation between two events, instead of the probability of independence according to the binomial distribution. They then use a temporal classifier to determine a partial order. Finally, they cluster events using the PMI scores to form in effect undirected n-ary associations. Those could be converted to directed associations if confidence was also computed.

With the help of focus and inferencing models, Niehaus and Young (2014) generate narratives in which some events need to be causally inferred by the reader. Those inferences are precisely defined as part of the input, whereas in our model only hints are available about the causal relations to be found by the reader.

León and Gervás (2010) also use causality-related relations to structure narratives. Their algorithm learns preconditional rules between events of a fictional story with the help of human feedback. An assumption is made that every event must be directly or indirectly a precondition to the last event of the story. Although this may make sense for a fictional story, it could involve selecting out important information when starting from real-life data.

In the context of generating a narrative from data with multiple actors, Gervás (2014) associates actions having the same actor. This makes sense, because actions by the same actor can certainly be directly or indirectly causally related. However, our prototype having been tested only on data with a single actor, this tactic would not have been adequate here.

Farrell et al. (2015) use regular expressions to define explanation specifications for error trace data. Regular expressions could also be used to manually define association rules in the context of our model.

Baez Miranda et al. (2014) use a task model to provide top-down constraints on the sequence of scenes that can be identified in the data to form the structure of the narrative. In contrast, our model can be said to be more bottom-up in the importance it gives to automatically extracted associations.

In Vaudry and Lapalme (2015), we tried to structure the narrative using hierarchical clustering. This did not achieve a structure fully labeled with rhetorical relations as our current spanning tree algorithm. Paragraph and sentence segmentation was

[1] http://www-etud.iro.umontreal.ca/%7Evaudrypl/ADL/eval/

less transparent since we did not use dedicated parameters. Furthermore we extracted only chronological sequential association rules. We interpreted this chronological direction as the direction of causality. This partly justified our claim of identifying approximate causal relations. By looking at the results of data interpretation, we now come to the conclusion that there is no clear link between the direction of the rule and the direction of a potential causal relation.

In addition, we selected which events to include in the text in what we called the summarization step, which we placed after document planning. This has the disadvantage of undoing some of the document planner's work. We selected events using the probability of this event type happening at any point in the event sequence. We now find that using the probability conditioned on time results in a greater proportion of the associative sub-network being connected. This leads to a better text structure. Note that we use for our current example the same data as before, with a maximum probability threshold of 0.3 instead of 0.4.

5 Conclusion

We presented a data-to-text model demonstrating that it is possible to structure a narrative around a mix of automatically mined and manually defined associations. The model also relies on sequential associations for event selection. The generated text's communicative goal is to help the reader assimilate the facts necessary to construct a causal representation of the events. According to the model, the connecting associative thread allows the reader to follow associations from the beginning to the end of the text. This structure takes the form of a spanning tree over a selected associative sub-network.

The textual quality of the generated texts was rated by judges. The results show that the texts were understandable, but that flow between sentences, although not bad, could still be improved. A possible solution would be to modify document structuring such as to minimize discontinuities. According to the event-indexing model (Zwaan et al., 1995), sentence-reading times increase with the number of discontinuities in temporality, spatiality, protagonist, causality, or intentionality.

We are currently designing a memorization experiment to test if the generated texts help the reader assimilate unusual facts independently of the domain. Apart from that, a task-oriented evaluation with domain experts could be organized. Furthermore texts could be generated from bigger datasets or datasets belonging to other domains. It would be interesting to fine-tune all parameters for each of those to see if ideal values vary from domain to domain.

References

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

Mieke Bal. 2009. *Narratology : introduction to the theory of narrative*. University of Toronto Press, Toronto, 3rd ed. edition.

Nadjet Bouayad-Agha, Gerard Casamayor, Simon Mille, and Leo Wanner. 2012. Perspective-oriented Generation of Football Match Summaries: Old Tasks, New Challenges. *ACM Trans. Speech Lang. Process.*, 9(2):3:1–3:31, August.

Nathanael Chambers and Daniel Jurafsky. 2008. Unsupervised Learning of Narrative Event Chains. In *ACL*, volume 94305, pages 789–797. Citeseer.

Rachel Farrell, Gordon Pace, and M Rosner. 2015. A Framework for the Generation of Computer System Diagnostics in Natural Language using Finite State Methods. In *Proceedings of the 15th European Workshop on Natural Language Generation (ENLG)*, pages 52–56, Brighton, UK, September. Association for Computational Linguistics.

Pablo Gervás. 2014. Composing narrative discourse for stories of many characters: A case study over a chess game. *Literary and Linguistic Computing*, August.

Catalina Hallett. 2008. Multi-modal presentation of medical histories. In *Proceedings of the 13th international conference on Intelligent user interfaces*, pages 80–89.

James Hunter, Yvonne Freer, Albert Gatt, Ehud Reiter, Somayajulu Sripada, and Cindy Sykes. 2012. Automatic generation of natural language nursing shift summaries in neonatal intensive care: BT-Nurse. *Artificial intelligence in medicine*.

Anna Jordanous. 2012. A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation*, 4(3):246–279.

Carlos León and Pablo Gervás. 2010. Towards a Black Box Approximation to Human Processing of Narratives Based on Heuristics over Surface Form. In *2010 AAAI Fall Symposium Series*, November.

William C. Mann and Sandra A. Thompson. 1987. *Rhetorical structure theory: A theory of text organization*. University of Southern California, Information Sciences Institute.

A. McKinlay, C. McVittie, E. Reiter, Y. Freer, C. Sykes, and R. Logie. 2009. Design Issues for Socially Intelligent User Interfaces: A Discourse Analysis of a Data-to-text System for Summarizing Clinical Data. *Methods of Information in Medicine*, 49(4):379–387, December.

James Niehaus and R. Michael Young. 2014. Cognitive models of discourse comprehension for narrative generation. *Literary and Linguistic Computing*, 29(4):561–582, December.

Fco Javier Ordóñez, Paula de Toledo, and Araceli Sanchis. 2013. Activity Recognition Using Hybrid Generative/Discriminative Models on Home Environments Using Binary Sensors. *Sensors*, 13(5):5460–5477, April.

Ivo Swartjes and Marit Theune. 2006. A fabula model for emergent narrative. In *Technologies for Interactive Digital Storytelling and Entertainment*, pages 49–60. Springer.

Marit Theune, Nanda Slabbers, and Feikje Hielkema. 2007. The Narrator: NLG for digital storytelling. In *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pages 109–112. Association for Computational Linguistics.

Tom Trabasso and Paul van Den Broek. 1985. Causal Thinking and the Representation of Narrative Events. *Journal of Memory and Language*, 24(5):612–630, October.

Tom Trabasso, Paul Van den Broek, and So Young Suh. 1989. Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, 12(1):1–25.

Pierre-Luc Vaudry and Guy Lapalme. 2013. Adapting SimpleNLG for bilingual English-French realisation. In *Proceedings of the 14th European Workshop on Natural Language Generation*, pages 183–187, Sofia, Bulgaria, August. Association for Computational Linguistics.

Pierre-Luc Vaudry and Guy Lapalme. 2015. Narrative Generation from Extracted Associations. In *Proceedings of the 15th European Workshop on Natural Language Generation*, pages 136–145, Brighton, United Kingdom, September. Association for Computational Linguistics.

Leo Wanner, Bernd Bohnet, Nadjet Bouayad-Agha, Franois Lareau, and Daniel Nickla. 2010. Marquis: Generation of User-Tailored Multilingual Air Quality Bulletins. *Applied Artificial Intelligence*, 24(10):914–952.

Rolf A. Zwaan, Mark C. Langston, and Arthur C. Graesser. 1995. The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, 6(5):292–297, September.