Open Research Online

The Open University's repository of research publications and other research outputs

Harvesting re-usable high-level rules for expository dialogue generation

Conference or Workshop Item

How to cite:

Stoyanchev, Svetlana and Piwek, Paul (2010). Harvesting re-usable high-level rules for expository dialogue generation. In: 6th International Natural Language Generation Conference (INLG 2010), 7-8 Jul 2010, Dublin, Ireland.

For guidance on citations see FAQs.

© 2010 The Authors

Version: [not recorded]

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.

oro.open.ac.uk

Harvesting Re-usable High-level Rules for Expository Dialogue Generation

Svetlana Stoyanchev
Centre for Research in Computing
The Open University
Walton Hall, Milton Keynes, UK
s.stoyanchev@open.ac.uk

Paul Piwek
Centre for Research in Computing
The Open University
Walton Hall, Milton Keynes, UK
p.piwek@open.ac.uk

Abstract

This paper proposes a method for extracting high-level rules for expository dialogue generation. The rules are extracted from dialogues that have been authored by expert dialogue writers. We examine the rules that can be extracted by this method, focusing on whether different dialogues and authors exhibit different dialogue styles.

1 Introduction

In the past decade, a new area of Natural Language Generation (NLG) has emerged: the automated generation of expository dialogue, also often referred to as scripted, authored or fictive dialogue. Research in this area began with the seminal study by André et al. (2000), which explored generation of dialogues between a virtual car buyer and seller from technical data on a car. This strand of work was developed further in the NECA project (van Deemter et al., 2008) and has since been extended to other domains, including explanation of medical histories (Williams et al., 2007), patient information leaflets (Piwek et al., 2007) and Wall Street Journal articles (Hernault et al., 2008).

Systems for generating expository dialogue have explored different inputs (databases, knowledge representations and text), generation methods (e.g., rule versus constraint-based approaches) and outputs (from dialogue scripts in text form to audio and computer-animated dialogue). A common trait of all these systems is, however, that at some point in the generation process, they produce a dialogue script, a text file which specifies what the interlocutors say, possibly enriched with mark-up for dialogue acts, speech and gestures – see, e.g., Piwek et al. (2002). These systems are different from conventional dialogue systems in that the system does not engage in a dialogue with the user; rather, the system generates a dialogue between two or more fictitious characters for the user/audience to view and learn from. In other words, the dialogue is used to deliver information to the user or audience, rather than between the interlocutors. Piwek (2008) discusses several empirical studies that identify benefits of the use of expository dialogue for education and persuasion.

In this paper, we take a step towards addressing two shortcomings of the work so far. Firstly, all the work cited has relied on hand-crafted resources (typically rules) for creating the dialogue. With the resources being created by non-expert dialogue authors (e.g., academic researchers), generated dialogues based on these resources may not be optimal; for instance, Williams et al. (2007) found that generated dialogues can be too information-dense, requiring conversational padding. Secondly, the resources for creating dialogue are tied to a specific domain, making it hard to redeploy a system in new domains.

We propose to address the first issue by automatically creating dialogue generation resources from a corpus of dialogues written by known effective dialogue authors. This fits in with a trend in dialogue modelling and generation to create resources from empirical data (Oh and Rudnicky, 2002; DeVault et al., 2008; Henderson et al., 2008; Belz and Kow, 2009).

The second issue is addressed by specifying dialogue generation rules at a level of detail that abstracts over the particulars of the domain and fits in with existing NLG architectures. The reference architecture of Reiter and Dale (2000) identifies three principal NLG tasks: Document Planning (DP), Microplanning and Realisation. DP is primarily non-linguistic: it concerns selection of information and organization of this information into a coherent whole. The latter is achieved by making sure that the information is tied together by Rhetorical Relations such as Contrast, Elaboration and Explanation, in other words, it is part of a Rhetorical Structure. We propose that dialogue generation rules interface with Rhetorical Structure and map to a Sequence of Dialogue Acts.

Interestingly, the interface between DP and Microplanning has also been identified as a place where decisions and preferences regarding style take an effect (McDonald and Pustejovsky, 1985). A question that we explore in this paper is whether dialogue styles exist at the highly abstract level we focus on in this paper. We concentrate on style in the sense of '[t]he manner of expression characteristic of a particular writer'[1].

The remainder of this paper is set up as follows. In Section 2, we introduce the corpus that we use to extract dialogue generation resources. Section 3 examines the dialogues in the corpus for prima facie evidence for stylistic differences between authors at the dialogue level. In Section 4, we describe our approach to extracting high-level dialogue generation rules from the corpus. Next, in Section 5 we analyse the resulting rules, looking for further evidence of different dialogue styles. We also compare the rules that were harvested from our corpus with handcrafted rules in terms of content and variety. Finally, Section 6 contains our conclusions and a discussion of avenues for further research.

2 A Parallel Monologue-Dialogue Corpus

The current work makes use of a corpus of human-authored dialogues, the CODA corpus.[2] In total, this corpus consist of about 800 dialogue turns. This paper is based on three dialogues from the corpus: George Berkeley's 'Dialogues between Hylas and Philonous' (extract of 172 turns), Mark Twain's 'What is man?' (extract of 445 turns) and Yuri Gurevich's 'Evolving Algebras' (extract of 89 turns). Berkeley's dialogue is one of the classics of philosophy, arguing for the, at first sight, extravagant claim that 'there is no such thing as *material substance* in the world'. Twain, according to the Encyclopaedia Britannica 'one of America's best and most beloved writers', takes on the concept of free will. Gurevich's dialogue deals with the mathematical concept of evolving algebras. Of these dialogues, Twain is by a large margin the longest (over 800 turns in total) and the only one which is aimed specifically at the general public, rather than an academic/specialist audience.

For each of the dialogues, the corpus also contains human-authored monologue which expresses the same content as the dialogue. Monologue and dialogue are aligned through mappings from monologue snippets to dialogue spans. As a result, the CODA corpus is a parallel monologue-dialogue corpus. Both the monologue and dialogue come with annotations: the monologue with Rhetorical Structure Theory (RST) relations (Mann and Thompson, 1988; Carlson and Marcu, 2001) and the dialogue side with an adaptation of existing Dialogue Act annotation schemes (Carletta et al., 1997; Core and Allen, 1997). Table 2 contains an overview of these RST relations and Dialogue Act labels.

3 Dialogue Analysis

In this section we examine whether there is prima facie evidence for differences in style between the three dialogues. Whereas existing work in NLG on style has focused on lexical and syntactic choice, see Reiter and Williams (2008), here we focus on higher-level characteristics of the dialogues, in particular, proportion of turns with multiple dialogue acts, frequencies of dialogue act bigrams, and relation between dialogue acts and speaker roles.

An important reason for determining whether there are different styles involved, is that this has implications for how we use the corpus to create expository dialogue generation resources. If different dialogues employ different styles, we need to be

[1] From definition 13.a. of the Oxford English Dictionary at http://dictionary.oed.com

[2] Further information on the construction of this corpus can be found in the annotation manual at computing.open.ac.uk/coda/AnnotationManual.pdf.

RST relations	Dialogue Acts
Enablement, Cause, Evaluation *(Subjective, Inferred)*, Comment, Attribution, Condition-Hypothetical, Contrast, Comparison, Summary, Manner-means, Topic-Comment *(Problem-Solution, Statement-Response, Question-Answer, Rhetorical Question)* Background, Temporal, Elaboration/Explanation, *(Additional, General-Specific, Example, Object-attribute, Definition, Evidence, Reason)*, Same-unit	Explain, Info-Request *(Init-Factoid-InfoReq, Init-YN-InfoReq, Init-Complex-InfReq)*, Init-Request-Clarify, Response-Answer *(Resp-Answer-Yes/No, and Resp-Answer-Factoid)*, Resp-Agree, Resp-Contradict

Table 1: RST relations and Dialogue Acts used in the CODA corpus. Annotators used the fine-grained categories in italics that are listed in brackets. For the current study, we rely, however, on the higher-level categories that preceed the fine-grained categories and which combine several of them.

careful with creating resources which combine data from different dialogues. Merging such data, if anything, may lead to the generation of dialogues which exhibit features from several possibly incompatible styles. Since our aim is specifically to generate dialogues that emulate the masters of dialogue authoring, it is then probably better to create resources based on data from a single master or dialogue.

3.1 Multi-act Turns

One of the characteristics of dialogue is the pace and the amount of information presented in each of the speaker's turns. In a fast-paced dialogue turns are concise containing a single dialogue act. Such dialogues of the form *A:Init B:Response A:Init B:Response ...* are known as 'pingpong' dialogue. Twain's 'What is man?' dialogue starts in this fashion (O.M. = Old Man; Y.M = Young Man):

O.M. What are the materials of which a steam-engine is made?
Y.M. Iron, steel, brass, white-metal, and so on.
O.M. Where are these found?
Y.M In the rocks.
O.M. In a pure state?
Y.M. No–in ores.
. . .

One character serves as the initiator and the other replies with a response. With turns that contain more than one dialogue, henceforth multi-act turns, this pattern can be broken:

O.M. . . .
And you not only did not make that machinery yourself, but you have NOT EVEN ANY COMMAND OVER IT.
Y.M. This is too much.
You think I could have formed no opinion but that one?
O.M. Spontaneously? No. And . . .

Author	Twain	Gurevich	Berkeley
Multi-act	34%	43%	24%
Layman/Expert	45%/55%	36%/64%	51%/49%

Table 2: Proportion of multi-act utterances and their distribution between Layman and Expert

Multi-act turns are turns comprised of multiple dialogue acts, such as the Young Man's in the example above, where a Resp-Contradict ('This is too much.') is followed by an Init-YN-Request ('You think I could have formed no opinion but that one?').

The dialogue pace may vary throughout a dialogue. We, however, find that overall proportions of multi-act turns and their distribution between expert and layman vary between the authors (see Table 2). Gurevich's dialogue has the highest proportion (43%) of multi-act turns and majority of them are attributed to the *expert*. Only 24% of Berkeley's dialogue turns consist of multiple dialogue acts and they are evenly split between the expert and the layman. Gurevich's dialogue is the type of dialogue where an expert gives a lesson to a layman while in Berkeley's dialogue one character often complements ideas of the other character making it difficult to determine which of the characters is an *expert*. The amount of multi-act turns seems to be one of the stylistic choices made by a dialogue author.

3.2 Dialogue Diversity

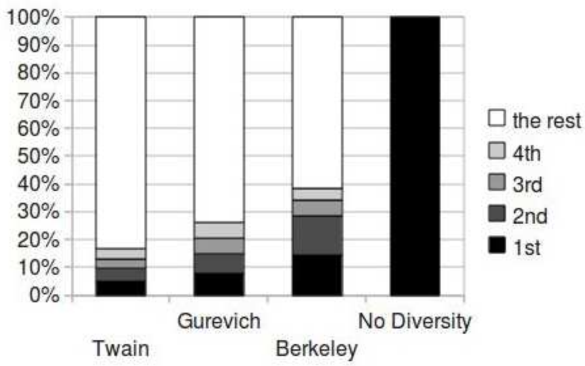

Figure 1: *Bigram coverage* for the 1-st to 4th most frequent bigrams.

Dialogues are essentially a sequence of turns, where each turn consists of one or more dialogue acts. For our measure of dialogue diversity we focus on two-turn sequences (i.e., turn bigrams), where a turn is identified by the sequence of dialogue acts it contains.

We define *bigram coverage for i* as the percentage that the top *i* most frequent bigrams contribute to all bigrams in the corpus. Diversity of the dialogue is inversely related to the *dialogue coverage*. In a dialogue with minimal diversity, the same turn, consisting of one or more dialogue acts, is repeated throughout the dialogue. The turn bigram consisting of two such turns has 100% bigram coverage.

Figure 1 shows the coverage for $1 \leq i \leq 4$ for each author in the corpus.[3] Out of the three authors, Twain's dialogues are the most diverse where the top 4 bigrams constitute only 15% of all bigrams. In Gurevich's dialogues the four most frequent bigrams constitute 25% and in Berkeley 40%.

Note that for all three authors the dialogue coverage for the 4 most frequent bigrams is quite low indicating high variability in bigrams used. To achieve such variability in automatically generated dialogues we need a large number of distinct generation rules.

[3]This range was chosen for illustration purposes. *Bigram coverage* can be compared for any $i \leq$total number of distinct bigrams.

3.3 Dialogue Acts and Speaker Roles

One of the most frequent bigrams for all three authors was, not unexpectedly, the sequence:

A: InfoRequest
B: Response-Answer

There is, however, a difference in the roles of speakers A and B. In all dialogues, one of the speakers took on the expert role and the other the layman role. For the aforementioned bigram, both in Berkeley's and Gurevich's dialogues the layman typically initiates the request for information and the expert responds (and often goes on to explain the response in Gurevich's dialogue):

Q: Is it difficult to define basic transition rules in full generality?
A: No. Here is the definition.
– Any local function update is a rule.
. . .
(From Gurevich's dialogue)

In contrast, in Twain's dialogues the roles are typically reversed: the expert asks and the layman responds:

O.M. Then the impulse which moves you to submit to the tax is not ALL compassion, charity, benevolence?
Y.M. Well–perhaps not.

Both techniques allow the author to convey a particular piece of information, but each giving rise its very own dialogue style.

4 Approach to Rule Extraction

Comparing statistics for individual dialogues gives us some idea about whether different styles are involved. The true test for whether different styles are involved is, however, whether for the *same* content different realizations are generated. Unfortunately, for our three dialogues the content is different to begin with. The parallel corpus allows us, however, to get around this problem. From the parallel corpus we can extract rules which map RST structures to dialogue act sequences. The Lefthand Side (LHS) of a rule represents a particular rhetorical structure found in the monologue side, whereas the Righthand Side (RHS) of the rule represents the dialogue act sequence with which it is aligned in the corpus.

Such rules can be compared between the different dialogues: in particular, we can examine whether the *same* LHS gives rise to similar or different RHSs.

4.1 Comparison with previous work

Hernault et al. (2008) manually construct surface-level rules mapping monologue to dialogue. Surface-level rules execute text-to-text conversion operating directly on the input string. In our approach, we separate the conversion into two stages. A first stage converts RST structures to Dialogue Act sequences. A second stage, which is beyond the scope of this paper, converts Dialogue Act sequences to text.

A further difference between the current approach and Hernault et al.'s is that the LHS of our rules can match nested RST structures. This covers, what we call, *simple rules* (involving a single RST relation, e.g., Contrast(X,Y)) and *complex rules* (involving 2 or more nested RST relations, e.g., Contrast(Condition(X,Y),Z)). Hernault et al. only allow for simple rules. A detailed comparison between our approach and that of Hernault et al., using the attribution rule as an example, can be found in Section 5.3.

id	DA	turns
0	Init-YN-InfoReq	Is your mind a part of your PHYSICAL equipment ?
0	Resp-Answer-No	No.
1	Explain	It is independent of it ; it is spiritual
2	Init-YN-InfoReq	Being spiritual, it cannot be affected by physical influences?
2	Resp-Answer-No	No.
3	Init-YN-InfoReq	Does the mind remain sober with the body is drunk ?
-	decorative	Well–
3	Resp-Answer-No	No.

Table 3: Example of annotated dialogue (from Mark Twain's 'What is man?').

4.2 Rule Extraction Algorithm

Table 3 and Figure 2 show annotated dialogue (authored by Twain) and its annotated monologue translation. Each terminal node of the RST structure corresponds to a part of a monologue snippet. All nodes with the same *id* correspond to a complete snippet and are linked to the dialogue act(s) with the same *id*s. The relation between monologue snippets and dialogue act segments is one-to-many. In other words, one snippet (e.g. snippets with id=0, id=2) can be expressed by multiple dialogue act segments.

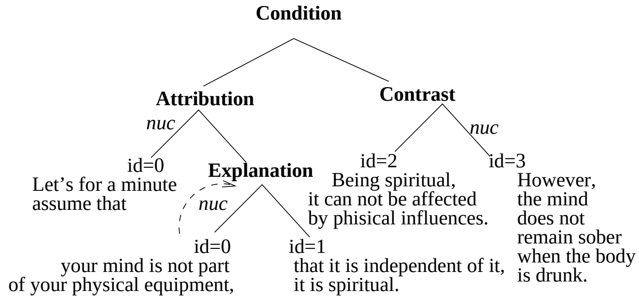

Figure 2: RST structure for the translation of dialogue in Table 3

span	rule
0-0	Attribution(0, 0)
0-1	Attribution(Explanation(0, 1))
2-3	Contrast(2, 3)
0-3	Condition (Attribution(Explain(0, 1)), Contrast(2, 3))

Table 4: RST sub-structures: LHS of monologue-to-dialogue mapping rules

Rules are extracted as follows: For each (automatically extracted) sub-structure of the RST structures on the monologue side, a rule is created (see Table 4). Two constraints restrict extraction of sub-structures: 1) spans of the structure's terminal nodes must be consecutive and 2) none of the *ids* of the terminal nodes are shared with a node outside the sub-structure.

For example, *Explanation(0, 1)* is not extracted because the node with *id=0* appears also under the *Attribution* relation which is not a part of this sub-structure.

Additionally, rules are generated by removing a relation and its satellite node and moving a nucleus node one level up. *Attribution(0, 0)* was extracted from a tree that had the Explanation relation and its satellite child *1* pruned. This operation relies on the validity of the following principle for RST (Marcu, 1997): 'If a relation holds between two textual spans of the tree structure of a text, that relation also holds between the most important units of the constituent subspans.'

The RST sub-structure is the LHS of a rule and dialogue act sequences are the RHS of a rule.

5 Results: Analysis of the Rules

In this section we describe the rules collected from the corpus. We compare the rules collected from the dialogues of different authors. We also compare the rules constructed manually in previous work with the rules collected from the corpus, specifically for the *attribution* relation.

5.1 Rule Statistics

relation	Twain	Gurev	Berk	all
simple	31 (33)	29 (38)	25 (26)	**81** (97)
complex	19	26	16	**61** (61)
null	15 (22)	9 (18)	9 (27)	**25** (67)
total	65	64	50	**167**
# turns	85	78	96	**259**

Table 5: Numbers of extracted distinct structural rules (total occurrences are parenthesized)

relation	Twain	Gurevich	Berkley
attribution	15%	2%	12%
contrast	18%	9%	17%
expl/elab	34%	47%	26%
eval	9%	6%	21%
other	24%	36%	24%
total	100%	100%	100%

Table 6: Proportions of relations expressed as rules

relation	Twain	Gurevich	Berkley
overall	2.4	1.9	2.9
contrast	2.3	2	2.6
elab/expl	2.7	1.7	3.3
eval	2	2	2.5

Table 7: Average number of turns in simple rules

Simple rules are the rules with one RST relation in the LHS. *Complex* rules are the rules with multiple RST relations in the LHS. In Table 4, rules for the LHS 0-0 and 2-3 are simple while the rules for 0-1 and 0-3 are complex. *Null* rules are the rules with no RST relation in the LHS.

From our sample of 259 translated and annotated dialogue turns from the corpus, we extracted 81 simple, 61 complex, and 25 null rules (null rules involve no RST structure and are discussed below). Table 5 shows the number of distinct rules per author.[4] In parentheses we show the number of actual (not necessarily distinct) rule occurrences in corpus. The majority of *simple* rules in the corpus (65 out of 81) occur only once.[5] This shows that the dialogue authors use a variety of dialogue act sequences when presenting their arguments in dialogue.

To compare dialogue *styles* we compare the rules across the dialogues of different authors. Table 6 shows the proportions of relation types in each author's dialogues that are mapped to a dialogue structure and produce a mapping rule.[6] Not all relations in monologue are mapped to a dialogue structure. For example, *Explain* moves may contain multiple clauses that are presented by a single character in the same turn. We find differences in distributions of relation types mapped to dialogue between the three authors (Fisher's exact test p<.01). Berkeley's dialogues produce more mapping rules with *Evaluation* and less with *Explanation/Elaboration* relations than the other two authors. Gurevich's dialogues produce less mapping rules with *Attribution* and *Contrast* relations than the other two authors. This difference between distributions of relation types mapped to dialogue has an important implication for dialogue generation. Dialogue generation programs may vary the style of a dialogue by choosing which discourse relations of the monologue are mapped to dialogue turns.

Another relevant property of a rule is the number of turns in the RHS of the rule. Number of turns in a rule shows how many times the dialogue characters switch to present information of the monologue corresponding to the LHS of the rule. The average numbers of turns in the RHS of all rules of the Twain, Gurevich, and Berkeley dialogues are 2.4, 1.9, 2.9 respectively (see Table 7). They are all pairwise significantly different (t-test $p < .05$) ranking the au-

[4]Two rules are distinct if either their LHS (relation in monologue) or RHSs (sequence of dialogue acts) are different.

[5]65=81-(97-81)

[6]This includes simple and complex rules

thors in the order $Gurevich < Twain < Berkeley$ according to the number of turns in the RHS of the rule. Similar ranking also appears as a trend for individual relations suggesting that this is the effect of the author's style rather than the relations (the distribution of relation types is different across the authors). This suggests that dialogue generation may affect the style of automatically generated dialogue by selectively choosing rules with longer (or shorter) RHS.

5.2 Null Rule

A *null* rule is a rule where a sequence of dialogue turns between two characters corresponds with a text segment with no rhetorical relation. A text segment without a rhetorical relation corresponds to a leaf node in the RST structure. A *null rule* typically creates a dialogue fragment consisting of a yes/no question (*Init-YN-Info-Req*) followed by yes/no answer, or a complex information request (e.g. What is your opinion on X?) followed by an *Explain* dialogue act, or a presentation of an argument (*Explain* dialogue act) followed by a response that signals agreement (*Resp-Agree*). *Null* rules create more interactivity in the dialogue.

The monologue segment corresponding to the LHS of a *null* rule may be in a rhetorical relation with another segment, such that the LHS of the *null* rule is embedded into another rule. Table 8 shows an example of a *null rule* embedded in a *contrast rule*. Turns 1 - 3 correspond to the RHS of the *Null* rule and 1 - 4 correspond to the RHS of the *Contrast* rule.

Null rules can be used to turn information into dialogue, even when there is no RST relation. For example, we may want to convey the piece of information A,B,C,D,E in that order, with rel1(A,B) and rel2(D,E). Whereas a simple rule may apply to relations and turn them into dialogue, C is left untouched. However, a *null* rule can be applied to C, to also turn its presentation into a dialogue exchange.

5.3 Case Study: the Attribution Rule

In this section we present a comparison of manually created rules for the RST attribution relation and rules extracted from the CODA corpus.

Hernault et al. manually construct two surface-level rules for the **Attribution (S,N)**[7] relation (see Table 9). In the *Dialogue Act* column we show the dialogue act representation of the corresponding surface-level rules. The first rule converts attribution relation into a *Complex-Info-Request* by the Layman followed with the *Explain* by the Expert. The second rule converts the attribution relation into *Explain* by the Expert, *Factoid-Info-Request* by the Layman and *Factoid-Response* by Expert. In both rules, the Expert is the one providing information (N) to the Layman and information is presented in *Explain* dialogue act

Table 10 shows six attribution rules we collected from phrases with attribution relation in the corpus (Twain1-4,Berkeley1,Gurevich)[8]. We notice several differences with the manually constructed rules:

- The variety of dialogue act sequences: each RHS of the rule (or dialogue act sequence) is different.
- Main information (N) can be presented by either the expert (Twain1, Twain2, Twain3, Berkeley1) or by the layman (Twain4, Gurevich1).
- Main information (N) can be presented in different dialogue acts: *Explain* dialogue act (Twain1, Twain4, Berkeley), *YN-Info-Request* (Twain2, Twain3), or *Complex-Info-Request* (Gurevich).
- Contextual information is part of the rule and may be used when choosing which rule to apply.

6 Conclusions and Further Work

In this paper, we have introduced a new approach to creating resources for automatically generating expository dialogue. The approach is based on extracting high-level rules from RST relations to Dialogue Act sequences using a parallel Monologue-Dialogue corpus. The approach results in rules that are reusable across applications and based on known expert dialogue authors.

After examining differences between the dialogues in the corpus in order to obtain prima facie evidence for differences in style, we conducted a detailed evaluation of the rules that were extracted

[7] N is a nucleus phrase that carries main information and S is a satellite phrase that contains the entity to whom N is attributed

[8] These are all the rules for attribution RST relation from 50 annotated turns for each author

Turn	Speaker	Dialogue act	Dialogue
Contrast rule. Segment with contrast relation: *[He never does anything for any one else's comfort , spiritual or physical.] [EXCEPT ON THOSE DISTINCT TERMS – that it shall FIRST secure HIS OWN spiritual comfort].*			
Null rule. Segment without rhetorical relation: *He never does anything for any one else's comfort , spiritual or physical*			
1	Layman	decorative	Come!
2	Expert	Init-YN-Request	He never does anything for any one else ' s comfort , spiritual or physical ?
3	Expert	Resp-Answer-No	No
4	Expert	Explain	EXCEPT ON THOSE DISTINCT TERMS – that it shall FIRST secure HIS OWN spiritual comfort .

Table 8: **Contrast rule** example containing **null rule** from Twain dialogue.

Speaker	Surface-level Rule	Dialogue act	Example Dialogue
Rule 1			
Layman	What did + *GetSubject(S+N)* + *Getmain-VerbLemma(S+N)*	Complex-Info-Request	What did S say?
Expert	*AddifNotPresentIn(N, That)* + N	Explain	N
Rule 2			
Expert	*RemoveIfPresentIn(N, That)* + N	Explain	N
Layman	Who *GetMainVerb(N)* that?	Factoid-Info-Req	Who said that?
Expert	*GetSubjectFromSentence(S+N)*	Factoid-Response	S did

Table 9: Manually created rules for **Attribution(S,N)** relation (Hernault et al., 2008)

from the corpus. We extracted 167 distinct rules and discussed the three types of rules: null, simple and complex (depending on the number of RST relation in the LHS: 0, 1 or more).

We found differences between authors in several respects, specifically:

- number of turns per simple rule
- number of dialogue acts per simple rule
- combination of speaker roles and dialogue acts

A detailed comparison between our automatically extracted attribution rule and the hand-crafted rules used by Hernault et al. showed up a number of differences. Apart from the fact that the corpus yielded many more rules than the two manually created ones, there were differences in which interlocutor presented particular information and which dialogue acts were being used.

The current work has focussed on high-level mapping rules which can be used both for generation from databases and knowledge representations and also for generation from text. In future work, we will focus on mapping text (in monologue form) to dialogue. For this we need to combine the high-level rules with rules for paraphrasing the text in the monologue with text for the dialogue acts that express the same information in dialogue form. For automatically extracting these surface level mappings we will draw on the approach to learning paraphrases from a corpus that is described in Barzilay and McKeown (2001). An important component of our future effort will be to evaluate whether automatically generating dialogues from naturally-occurring monologues, following the approach described here, results in dialogues that are fluent and coherent and preserve the information from the input monologue.

Acknowledgements

We would like to thank the anonymous reviewers of INLG2010 for their helpful comments and our colleagues in the Open University's NLG group for stimulating discussions on the content of this paper. The research reported in this paper was carried out as part of the CODA project (COherent Dialogue Automatically generated from text; see http://computing.open.ac.uk/coda/) which is funded by the UK Engineering and Physical Sciences Research Council under grant EP/G/020981/1.

References

E. André, T. Rist, S. van Mulken, M. Klesen, and S. Baldes. 2000. The automated design of believable dialogues for animated presentation teams. In *Em-*

Speaker	Dialogue act	Dialogue
Twain1 *I will put that law into words, keep it in your mind:* FROM HIS CRADLE TO HIS GRAVE A MAN NEVER DOES... **Satellite of Summary**		
Layman	Init-YN-InfoReq	Will you put that law into words?
Expert	Resp-Answer-Yes	Yes.
Expert	Resp-Explain	This is the law, keep it in your mind. FROM HIS CRADLE TO HIS GRAVE A MAN NEVER DOES...
Twain2 *I can not imagine* that there is some other way of looking at it. **Satellite of Explanation**		
Expert	Init-Complex-InfoReq /clarify	What makes you think that?
Layman	*decorative*	Pray what else could I think?
Expert	Init-YN-InfoReq	Do you imagine that there is some other way of looking at it?
Twain3 *One cannot doubt* that he felt well.**Satellite of Evaluation-Conclusion**		
Expert	Init-YN-InfoReq	He felt well?
Layman	Resp-Answer-Yes	One cannot doubt it.
Twain4 *As I said a minute ago* Hamilton fought that duel to get PUBLIC approval. **Nucleus of Explanation**		
Layman	Init-Explain/contradict	A minute ago you said Hamilton fought that duel to get PUBLIC approval.
Resp-Agree	Resp-Agree	I did.
Berkeley1 *You can not conceive* a vehement sensation to be without pain or pleasure.		
Expert	Init-Explain	Again, try in your thoughts, Hylas, if you can conceive a vehement sensation to be without pain or pleasure.
Layman	Resp-Contradict	You can not.
Gurevich *I will explain* what static algebras are exactly. **Nucleus of Statement-response**		
Layman	Init-Complex-InfoReq	Please explain to me what static algebras are exactly.
Expert	Resp-Agree	Gladly.

Table 10: Attribution Examples. Satellite is *italicised*.

bodied Conversational Agents, pages 220–255. MIT Press, Cambridge, Mass.

R. Barzilay and K. McKeown. 2001. Extracting Paraphrases from a Parallel Corpus. In *Proceedings of the ACL*, Toulouse, France.

A. Belz and E. Kow. 2009. System Building Cost vs. Output Quality in Data-to-Text Generation. In *Proceedings of the 12th European Workshop on Natural Language Generation (ENLG'09)*, Athens, Greece.

J. Carletta, A. Isard, and J. C. Kowtko. 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23:13–31.

L. Carlson and D. Marcu. 2001. Discourse tagging reference manual. Technical Report ISI-TR-545, ISI, September.

M. Core and J. Allen. 1997. Coding dialogs with the damsl annotation scheme. In *Working Notes: AAAI Fall Symposium on Communicative Action in Humans and Machine*.

D. DeVault, D. Traum, and R. Artstein. 2008. Making Grammar-Based Generation Easier to Deploy in Dialogue Systems. In *Procs SIGdial 2008*, Ohio, June.

E.Reiter and S. Williams. 2008. Three approaches to generating texts in different styles. In *Proceedings of the Symposium on Style in text: creative generation and identification of authorship*.

J. Henderson, O. Lemon, and K. Georgila. 2008. Hybrid Reinforcement / Supervised Learning of Dialogue Policies from Fixed Datasets. *Computational Linguistics*, 34(4):487–511.

H. Hernault, P. Piwek, H. Prendinger, and M. Ishizuka. 2008. Generating dialogues for virtual agents using nested textual coherence relations. In *IVA08: 8th International Conference on Intelligent Virtual Agents*.

William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

D. Marcu. 1997. From Discourse Structures to Text Summaries. In *The Proceedings of the ACL'97/EACL'97 Workshop on Intelligent Scalable Text Summarization*, pages 82–88, Madrid, Spain.

D. McDonald and J. Pustejovsky. 1985. A computational theory of prose style for natural language generation. In *Proceedings of the second conference on European chapter of the Association for Computational Linguistics*, pages 187–193, Geneva, Switzerland.

A. Oh and A. Rudnicky. 2002. Stochastic natural language generation for spoken dialog. *Computer Speech and Language*, 16(3/4):387–407.

P. Piwek, B. Krenn, M. Schroeder, M. Grice, S. Baumann, and H. Pirker. 2002. RRL: A Rich Representation Language for the Description of Agent Behaviour in NECA. In *Proceedings of the AAMAS workshop "Embodied conversational agents - let's specify and evaluate them!"*, Bologna, Italy, July.

P. Piwek, H. Hernault, H. Prendinger, and M. Ishizuka. 2007. T2D: Generating Dialogues between Virtual Agents Automatically from Text. In *Intelligent Virtual Agents*, LNAI 4722, pages 161–174. Springer Verlag.

P. Piwek. 2008. Presenting Arguments as Fictive Dialogue. In *Proceedings of 8th Workshop on Computa-*

tional Models of Natural Argument (CMNA08), Patras, Greece, July. ISBN 978-960-6843-12-9.

E. Reiter and R. Dale. 2000. *Building Natural Language Generation Systems*. Cambridge University Press, Cambridge.

K. van Deemter, B. Krenn, P. Piwek, M. Klesen, M. Schroeder, and S. Baumann. 2008. Fully generated scripted dialogue for embodied agents. *Artificial Intelligence Journal*, 172(10):1219–1244.

S. Williams, P. Piwek, and R. Power. 2007. Generating Monologue and Dialogue to Present Personalised Medical Information to Patients. In *Procs ENLG 2007*, pages 167–170, Schloss Dagstuhl, Germany.