Dialectical text planning

Rodger Kibble
Dept of Computing, Goldsmiths College
University of London, UK
r.kibble@gold.ac.uk

Abstract. A key requirement for the automatic generation of argumentative or explanatory text is to present the constituent propositions in an order that readers will find coherent and natural, to increase the likelihood that they will understand and accept the author's claims. Natural language generation systems have standardly employed a repertoire of coherence relations such as those defined by Mann and Thompson's Rhetorical Structure Theory. This paper models the generation of persuasive monologue as the outcome of an "inner dialogue", where the author attempts to anticipate potential challenges or clarification requests. It is argued that certain RST relations such as Motivate, Evidence and Concession can be seen to emerge from various pre-empting strategies.

1 Introduction

A key requirement for the automatic generation of argumentative or explanatory text is to present the constituent propositions in an order that readers will find coherent and natural, to increase the likelihood that they will understand and accept the author's claims. Ideally, any objections or clarification requests that an audience might raise will already have been countered by elements of the author's argument. In fact this paper models the generation of persuasive monologue as the outcome of an "inner dialogue", where the author attempts to anticipate potential challenges or clarification requests. It will be argued that certain coherence relations can be seen to emerge from various strategies for pre-empting or "obviating" challenges or clarification requests.

This paper assumes a model of dialogue as updating participants' information states (IS), where an IS consists of a record of each interlocutor's propositional and practical commitments (cf [7, 2, 17]) rather than "mental states" such as belief and intention (cf [3]). This approach is motivated at greater length and contrasted with other commitment-based approaches such as [12] in [8, 9]; the key assumptions for the purposes of this paper are:

1. Each agent in a dialogue keeps a score of social commitments for all participants, including itself. Commitments can be classified into *practical* (commitments to act, corresponding to *intentions* in mentalistic accounts) and propositional or *doxastic* (commitments to justify an assertion, corresponding to *beliefs*).
2. Agents play one of three dynamically assigned roles at any given point in a dialogue: Speaker (**Sp**), Addressee (**Ad**), or Hearer (**He**) who is not directly addressed.
3. For an agent α to assert ϕ is to acknowledge commitment to ϕ; other agents may also attribute consequential commitments to α.
4. Additionally, a dialogue act constitutes an attempt to commit Addressee(s) to a proposition or a course of action, as detailed in the following section.
5. Addressee's options include accepting the proffered commitment, challenging it or requesting clarification.

This paper will focus on modelling persuasive monologue, or extended dialogue turns, as emerging from a process of internal argumentation, with the virtual agents Planner (**Pl**) in place of **Sp** and Critic **(Cr)** substituted for **Ad**. I will aim to show how a variety of Mann and Thompson's RST relations such as Motivate, Justify, Evidence, Concession and Elaboration can be seen to emerge from different text planning strategies [11, 16] . It might be argued that this is an essentially trivial exercise in shifting information from a predefined set of coherence relations to a pre-defined set of dialogue acts and moves. However, there are independent motivations for developing models for dialogue and argumentation, and the argument in this paper is that a (possibly partial) account of coherence relations in monologue emerges as a side-effect of these models. The paper will conclude by addressing some apparent differences between dialogue and monologue as discussed by [14] and [6].

2 Argumentation and discourse relations

The full framework will include specifications for the proto-speech acts listed below. Note that I use upper-case Greek letters such as Φ to represent speech acts themselves and lower-case letters such as ϕ for the propositional content of the speech acts.

assert(Sp, ϕ, Ad, He) undertake commitment to justify a propositional claim; attempt to bestow same commitment on **Ad**.

instruct(Sp, ϕ, Ad, He) attempt to bestow a practical commitment on Addressee.

endorse(Sp, ϕ, Ad, He) Speaker adopts a commitment specified by Addressee

challenge(Sp, Φ, Ψ, Ad, He): require agent to justify or retract a commitment offer Φ, with Ψ as an optional counter-commitment. Note that the challenge may be directed at the propositional content ϕ, or at the appropriateness of the speech act itself.

respond(Sp, challenge(Ad, Φ, Ψ, Sp, He), Ξ, Ad, He) respond to a challenge with a dialogue act Ξ which may be:

- asserting ξ as evidence for ϕ, or as justification for uttering Φ;
- retracting commitment to ϕ, the propositional content of Φ;
- withdrawing a claim to justification for the speech act Φ;
- challenging Ψ;

- requesting clarification of Ψ;
- ϵ - the null act. How this is interpreted will depend on the particular conventions currently in force: it may be understood at different times as implicit endorsement, implicit denial or non-committal.

retract(Sp, ϕ, Ad, He) withdraw a commitment to ϕ.
query(Sp, Φ, Ad, He) request clarification of Φ
respond(Sp, query(Ad, Φ, Sp, He), Ψ, Ad, He)
respond to request for clarification of Φ by uttering the speech act Ψ.

2.1 Examples of dialogue and monologue

The following examples consist of a short dialogue followed by two variants of a monologue expressing roughly the same content and exemplifying particular rhetorical structures.

Example (a)

A: You should take an umbrella.
B: Why?
A: It's going to rain.
B: It doesn't look like rain to me. It's sunny
A: Michael Fish predicted it.
B: Who's he?
A: He's a weather forecaster on the BBC.
B: OK.

In terms of the speech acts defined above, this exchange can be represented (somewhat simplified) as follows:

A: instruct(A, *take-umbrella*, B, _);
B: challenge(B, *take-umbrella*, _, A, _);
A: respond(A, challenge(B, *take-umbrella*, _, A, _), assert(A, *rain-later*, B, _), B, _)
B: challenge(B, *rain-later*, *sunny-now*, A, _);
A: respond(A, challenge(B, *rain-later*, *sunny-now*, A, _), assert(A,*fish*,B, _), B, _)
B: query(B, *fish, A,* _)
A: respond(A, query(B, *fish, A,* _), assert(A, *BBC-forecaster*, B, _), B, _)
B: endorse(B, {*BBC-forecaster*; *fish*; *rain-later*; *take-umbrella*}, A, _)

Example (b)

A: You should take an umbrella. It's going to rain. I heard it on the BBC.

A possible RST analysis of this example is:

Motivate
Nucleus *take-umbrella*
Satellite: Evidence

Nucleus *rain-later*

Satellite *BBC-forecast*

Example (b′)

A: You should take an umbrella. It's going to rain, even though it looks sunny right now. I heard it on Michael Fish's slot. He's a weather forecaster at the BBC.

Proposed RST analysis:

Motivate
Nucleus *take-umbrella*
Satellite

Evidence

Nucleus

Concession
Nucleus *rain-later*
Satellite *sunny-now*

Satellite

Background
Nucleus *fish*
Satellite *BBC-forecaster*

Example (c)

A: I listened to the weather forecast on the BBC. It's going to rain. You should take an umbrella.

Proposed RST analysis: same rhetorical structure as (b) but realised in a satellite-first sequence:

Motivate
Satellite: Evidence

Satellite *BBC-forecast*

Nucleus *rain-later*

Nucleus *take-umbrella*

2.2 Speaker strategies

In the above scenario, suppose A has the goal that B undertake a practical commitment to carry an umbrella. Examples (a - c) illustrate three different strategies:

(i) Issue a bare instruction; offer justification only if challenged.
(ii) Issue an instruction, followed by an assertion that **pre-empts** a potential challenge, and recursively pre-empt challenges to assertions.
(iii) **Obviate** the challenge by uttering the justification **before** the instruction, and recursively obviate potential challenges to assertions.

(The terms **pre-empt** and **obviate** are used with these particular meanings in this paper, which may not be inherent in their ordinary usage.) Note that examples (a) and (b′) exhibit the same sequence of propositions, which is consistent with the assumption that (b′) results from a process of internal argumentation with a virtual agent that raises **Ad**'s potential objections. The following section will sketch a formulation of strategies (i - iii) in terms of the Text Planning task of natural language generation.

3 Dialectical text planning

I will assume some familiarity with terms such as "text planning" and "sentence planning". These are among the distinct tasks identified in Reiter's "consensus architecture" for Natural Language Generation [15]; see also [1]:

Text Planning/Content Determination - deciding the content of a message, and organising the component propositions into a text structure (typically a tree). I will make a distinction between the **discourse plan** where propositions in the initial message are linked by coherence relations, and the **text plan** where constituents may be re-ordered or pruned from the plan.

Sentence Planning - aggregating propositions into clausal units and choosing lexical items corresponding to concepts in the knowledge base; this is the level at which the order of arguments and choice of referring expressions will be determined.

Linguistic realisation - surface details such as agreement, orthography etc.

3.1 Discourse planning

Text planning is modelled in what follows as the outcome of an inner dialogue between two virtual agents, the Planner (**Pl**) and the Critic (**Cr**). The Critic is a user model representing either a known interlocutor or a "typical" reader or hearer. A's options (i -iii) in Section 2.2 above can be seen to correspond to three different strategies which I will call *one-shot*, *incremental* and *global*. These strategies are presented in rather simplified pseudo-code below, in particular I only consider the assert action and selected responses to it.

One-shot planning

Speaker produces one utterance per dialogue turn which may be:

- a bare assertion ϕ;
- response to a challenge or clarification request from Addressee;
- challenge to Address's most recent or salient assertion, or request for clarification;
- ϵ

The message is passed directly to the text planner without being checked by the Critic. This strategy is appropriate when no user model is available.

Incremental Planning

Speaker generates the "nuclear" utterance and then calculates whether a challenge is likely, and recursively generates a response to the challenge if possible. This is the strategy of **pre-empting** challenges referred to in section 2.2. The response is immediately committed to the right frontier of Speaker's text plan.

procedure inc-tp(Φ)
where Φ is some speech act with propositional content ϕ;
send Φ to text planner;
assert(Pl, ϕ, Cr, _);
if challenge(Cr, ϕ, ψ, Pl, _)
then do inc-tp(respond(Pl, challenge(Cr, ϕ, ψ, Pl, _), Ξ, Cr, _);
else quit.

This strategy is appropriate when a suitable user model is available but resource limits or time-criticality make it desirable to interleave discourse planning, text planning and sentence generation.

Goal-directed Planning

The sequence is globally planned in order to rebut potential challenges by generating responses to them ahead of the nuclear proposition. This is the strategy I have dubbed **obviating** challenges in section 2.2.

procedure gd-tp(Φ)
where Φ is some speech act with propositional content ϕ;
initialise stack = [ ];
call gd-tp-stack(Φ);
do until stack = [ ]:
pop Ψ from stack;
add Ψ to text plan;
end gd-tp()

procedure gd-tp-stack(Φ)
stack = [Φ | stack];
assert(Pl, ϕ, Cr, _);
if challenge(Cr, ϕ, ψ, Pl, _)
then do gd-tp-stack(respond(Pl, challenge(Cr, ϕ, ψ, Pl, _), Ξ, Cr, _);
else quit gd-tp-stack
end gd-tp-stack()

This strategy is appropriate for applications where resources allow for the full discourse plan to be generated in advance of text planning so that constituents may subsequently be reordered or pruned to produce a possibly more "natural" and readable text.

3.2 Text planning and plan pruning

If we consider the examples in section 2.1: (b), (b′) are typical products of incremental planning and (c) of goal-directed planning. The former will result in **nucleus-first** structures, while the default ordering resulting from the latter will realise satellites **before** nuclei. Two refinements are discussed in this section: **plan pruning** and **reordering** of the text plan.

The differences between (b) and (b′) demonstrate that the text planner has a choice over whether to realise only the Planner's contributions or those of the Critic as well. The latter option, retaining the proposition *sunny-now*, results in instances of RST's Concession relation. This is a special case of **plan pruning** as described by [6], where a constituent may be removed if it is inessential to the speaker's purpose: for instance it may be inferrable from other material in the plan. Green and Carberry motivate this with the aid of the following example (their (13a-e)), illustrating how a question-answering system might decide how much unrequested information to include in an indirect answer to a yes-no question.

Example (d)

(i) Q: Can you tell me my account balance?
(ii) R: [No.]
(iii) [I cannot access your account records on our computer system.]
(iv) The line to our computer system is down.
(v) You can use the ATM machine in the corner to check your account.

Items (ii - iii), shown in square brackets, can be suppressed since (iii) is inferrable from (iv) and in turn implies (ii). This assumes that the user is aware, or can accommodate the fact that their account balance is kept on the computer system. This example is compared with an "imaginary dialogue" where each statement responds to a specific question from the user.

As stated above, the planning strategies outlined in section 3 produce texts that are uniformly either satellite-first or nucleus-first by default. There is a need to generalise the strategies so that the planner can dynamically switch from one to the other, in order to produce texts such as:

Example (e)

It's going to rain. I heard it on the BBC. You should take an umbrella.
RST analysis:

Motivate

Satellite: Evidence

Nucleus *rain-later*

Satellite *BBC-forecast*

Nucleus *take-umbrella*

By distinguishing between the **discourse plan** and **text plan** we allow for re-ordering of constituents at the level of the text plan, within the partial ordering defined by the discourse plan. For instance, a different ordering of propositions might improve the referential coherence of a text according to Centering Theory [10].

3.3 Summary

In contrast to approaches to text generation that carry out top-down planning using pre-defined coherence relations I have argued that certain RST relations can be seen to emerge from sequences of internalised dialogue moves that aim to pre-empt or obviate potential challenges or clarification requests, as follows:

instruct-challenge-respond underlies Motivation or Justify depending on the content of the challenge and response;

assert-challenge-respond underlies Evidence if the propositional content is challenged, or Justify if the appropriateness of the assert act itself is at issue.

<**any-speech-act**>**-challenge-respond** underlies Concession if the content of the challenge is realised in the text.

<**any-speech-act**>**-query-respond** underlies Background.

It remains to be seen if further RST relations can be modelled using the "dialectical" method.

4 Discussion and future work

4.1 Objections to "implicit dialogue"

Reed [14] argues against identifying a persuasive monologue with an implicit dialogue and emphasises the importance of distinguishing the *process* of creating a monologue from the *product*, the monologue itself. Now, it is not argued here that a monologue is nothing more than a trace of the dialogical process of constructing an argument. The "goal-directed" strategy allows for a phase of pruning and re-ordering the text plan (not described in detail here) although the default is for propositions to be realised in the sequence in which they are added to the discourse plan.

Reed puts forward an important argument: that a crucial difference is the fact that unlike a dialogue, a "pure" monologue must not contain a *retraction* in the sense of asserting a proposition and its negation. This has implications for the discussion of text planning strategies in section 3 above, since there is the possibility of a contradiction occurring in a sequence of responses to recursive challenges. On the one hand, goal-directed planning could be extended with a backtracking facility and consistency checking such that indefensible claims or even the nuclear proposition itself could be withdrawn before proceeding to sentence generation, if a challenge generated by the Critic shows up a contradiction in the existing plan. However, the essence of incremental planning is intended to be that each proposition is committed to the text plan, to be passed on to the sentence planner, *before* considering potential challenges. The algorithm as adumbrated above certainly allows the possibility that contradictory propositions will be added to the plan, as a consequence of limitations on speakers' memory and reasoning capabilities.

The proscription of overt retraction would certainly be a reasonable design feature for a computer system generating argumentative text. However, this paper is also concerned with modelling the ways in which human speakers might construct an argument, and so this comes down to an *empirical* question as to whether speakers delivering an extempore monologue will ever realise part-way through that there are insuperable objections to their initial claim (or a subordinate claim), and end up withdrawing it. For instance, the medium of communication might be an electronic "chat" forum such that all keystrokes are instantly and irrevocably transmitted to other logged-on users. It is not obvious that this possibility should be ruled out in principle, or even that it can be ruled out in a resource-limited system following "incremental planning" as defined here.

4.2 Future work

The following issues will be addressed in future research:

Coherence, user modelling and reasoning. It is assumed that for a text to be *coherent* as perceived by the intended audience means that there is an increased likelihood that they will endorse the proffered (practical or doxastic) commitments *and* that this will require less cognitive effort on the audience's part, by comparison with less coherent texts. The success of a dialectical, user-model oriented text planning regime will clearly depend crucially on the reliability of the user models and the validity of the reasoning processes by which the planner calculates potential challenges and suitable responses. Some important topics are:

- modelling *specific* users to whom a message is directed, versus *typical* readers of a text which is not directed at any particular individual;
- modelling information states of the virtual agents **Pl** and **Cr**, in view of arguments that speakers and hearers have asymmetric context models in dialogue [4].

Complexity. Goal-directed planning requires more computational resources on the part of the Speaker but arguably results in (satellite-initial or mixed) texts that are easier for Hearers to process. The question arises whether speakers optimise their utterances for the audience or follow a path of least effort. This is a topic of debate amongst

researchers in psycholinguistics, as evidenced by the claims put forward by [13] and the various responses collected together in the same journal issue.

Preempting clarification requests. This paper has modelled the Background relation as resulting from preemption of a clarification request (CR).) Studies including [5] have shown that CRs can be directed at various levels of linguistic representation or content. In the following example (constructed for this paper), the elliptical query *Maclean?* could have any of the responses shown:

Example (f)

(i)A: Maclean's defected to the USSR.
(ii) B: Maclean?
(iii) A: Yes, Maclean of all people.
(iv) A: Donald Maclean, head of the American desk at the FO.
(v) A: That's M - a - c - l - e - a - n.

This raises architectural issues since it has been assumed in this paper that preemptions are generated at the discourse planning stage, where details of linguistic realisation such as how to spell a proper name may not be available. Future work will address the question of whether and how clarifications at distinct levels of representation can be integrated into the dialectical planning model.

REFERENCES

[1] John Bateman and Michael Zock, 'Natural language generation', in *The Oxford Handbook of Computational Linguistics*, ed., Ruslan Mitkov, 284 – 304, Oxford University Press, Oxford, (2003).
[2] Robert Brandom, *Making it Explicit*, Harvard University Press, Cambridge, Massachusetts and London, 1994.
[3] Philip Cohen and Hector Levesque, 'Persistence, intention and commitment', in *Intentions in Communication*, eds., Philip Cohen, Jerry Morgan, and Martha Pollack, 33 – 69, MIT Press, Cambridge, Massachusetts and London, (1990).
[4] Jonathan Ginzburg, 'On some semantic consequences of turn taking', in *Proceedings of the 11th Amsterdam Colloquium*, (1997).
[5] Jonathan Ginzburg and Robin Cooper, 'Clarification ellipsis and the nature of contextual updates in dialogue', *Linguistics and Philosophy*, **27(3)**, 297 – 365, (2004).
[6] Nancy Green and Sandra Carberry, 'A computational model for taking initiative in the generation of indirect answers', *User Modeling and User-Adapted Interaction*, **9(1/2)**, 93–132, (1999). Reprinted in Computational Models of Mixed-Initiative Interaction, Susan Haller, Alfred Kobsa, and Susan McRoy, eds., Dordrecht, the Netherlands, 277-316.
[7] Charles Hamblin, *Fallacies*, Methuen, London, 1970.
[8] Rodger Kibble, 'Elements of a social semantics for argumentative dialogue', in *Proceedings of the Fourth Workshop on Computational Modelling of Natural Argumentation*, Valencia, Spain, (2004).
[9] Rodger Kibble, 'Reasoning about propositional commitments in dialogue', (2006). To appear in *Research on Language and Computation*.
[10] Rodger Kibble and Richard Power, 'Optimizing referential coherence in text generation', *Computational Linguistics*, **30 (4)**, 401–416, (2004).
[11] William C. Mann and Sandra A. Thompson, 'Rhetorical structure theory: A theory of text organization', Technical report, Marina del Rey, CA: Information Sciences Institute, (1987).
[12] Colin Matheson, Massimo Poesio, and David Traum, 'Modelling grounding and discourse obligations using update rules', in *Proceedings of NAACL 2000*, (2000).
[13] Martin Pickering and Simon Garrod, 'Toward a mechanistic psychology of dialogue', *Behavioral and Brain Sciences*, **27**, 169–225, (2004).
[14] Chris Reed, 'Is it a monologue, a dialogue or a turn in a dialogue?', in *Proceedings of the 4th International Conference on Argumentation (ISSA98)*, Amsterdam, (1998). Foris.
[15] Ehud Reiter, 'Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible?', in *Proceedings of 7th International Natural Language Generation Workshop*, pp. 163–170, (1994).
[16] Maite Tabaoda and William Mann, 'Rhetorical Structure Theory: Looking back and moving ahead', *Discourse Studies*, **8(3)**, (2006). To appear.
[17] Douglas Walton and Eric Krabbe, *Commitment in dialogue*, State University of New York Press, Albany, 1995.