Extended Discourse Plan Operators

Owen Rambow
CoGenTex, Inc. (Ithaca, NY) and University of Delaware

1 Introduction

For planning extended discourse, (at least) two types of approaches have been proposed in the literature. In the first approach, general knowledge about discourse and its structure is used to reason about the speaker's (S's) communicative goals in order to assemble a tree representing the content and structure of the discourse, which, when delivered to the hearer (H), will meet S's goals. We will call this the "rhetorical approach". In a second approach, the text structure is directly specified using a simple device, no intentions are modeled, and no reasoning takes place. We will call this the "schema approach".

While at first one may assume that the schema approach is just a “compilation” (in some sense to be made specific) of the rhetorical approach, we argued in Kittredge et al. (1991) that certain kinds of linguistic communications - namely, reports - cannot be adequately planned using general knowledge about discourse structure, because their structure has been, through historical precedent, fixed in some now arbitrary manner (even though historically, it may have been motivated). Thus, while the rhetorical approach is theoretically interesting and has explanatory power, it is empirically inadequate.

In this paper, we propose a way of integrating the two approaches into a single, theoretically interesting and empirically adequate framework. This framework is based on the notion of extended discourse plan operators. We propose that in planning discourse, the planning agent assembles large pieces of structure on several levels of representation in parallel. The proposed approach also solves another problem: if, as has been argued by Moore and Pollack (1992), discourse can be represented by parallel levels of structure which are not necessarily isomorphic, how can we plan such parallel structure?

The analysis of existing approaches is similar to that provided Moore and Paris (1993) and of Moore (1994). Like those authors, this paper proposes a way of integrating different types of discourse-related information in a planning system. However, the approach chosen is different.

2 Review, Summary, and Interpretation: Rhetorical Structure Theory

In the following, we will assume a theory of discourse structure which is based on the Rhetorical Structure Theory of Mann and Thompson (1987), with additions and comments which have been made to the theory since then (we retain the name “RST”). In RST, discourse relations are defined in terms of an effect on H’s cognitive states (belief, desire, intention). Specifically, the effect is obtained if the two segments being related meet certain conditions (which is also formulated as effects on H’s cognitive state), and if the two segments are juxtaposed. The case in which there is no relation between text spans is accounted for by the Joint schema, which is not a relation (i.e., the juxtaposition has no effect on H’s cognitive state).

We claim that the major theoretical contribution of RST (and what makes it a theory) is the claim that only certain relations exist, and others, logically possible, do not. Consider a hypothetical relation Counter-Evidence. Any purported examples, we claim, turn out to be cases of Evidence and irony. The claim that only certain discourse relations exist and others do not is a powerful one, since it means that in achieving communicative goals, S must follow certain rules in assembling a larger discourse.

Mann and Thompson (1987) make a by now well-known distinction among the relations between presentational relations (intentional relations) and subject-matter relations (informational relations). A presentational relations does not correspond (in and of itself) to real-world fact, but rather exists only in the moment of communication. There do

---

1This paper is substantially based on joint work with Benoit Lavoie and Mike White. In addition, it has profited from input and feedback from Ted Caldwell, Richard Kittredge, Tanya Korelsky, Kathy McCoy, Daryl McCullough, and two LORID reviewers.
not appear to be cue words for all intentional relations. A subject-matter relation, in contrast, does correspond to a real-world relation that exists independently of the discourse. Crucially, Moore and Pollack (1992) have pointed out that subject-matter relations appear to coexist with intentional relations. Their example is reproduced here:

(1) (S1) Come home by 5pm.
    (S2) Then we can go to the hardware store before it closes.
    (S3) That way we can finish the bookshelves tonight.

The two analyses, subject-matter and presentational, are shown in Figure 1. They are both valid at the same time, and in fact the structures are not isomorphic.²

3 Discourse Planning

What we will call the "rhetorical approach" rests on a very simple idea. Since the RST relations are defined in terms of H's cognitive states (both for the effect and for the constraints fields), we can "invert" the definition, interpret the effect field as a communicative goal (to effect H's cognitive state in some way) and the constraints as preconditions.

The basic algorithm is as follows. We start with single discourse goal to affect H's cognitive state in some manner. We then refine goals iteratively using plan operators, or we specify speech acts to satisfy (atomic) goals.

This rhetorical approach (essentially a straw man) was outlined in (Hovy, 1988). However, as has been pointed out by Kittredge et al. (1991) and Moore and Paris (1993), the simple approach did not work very well (if at all) when subject-matter relations were used in planning (as was attempted in (Hovy, 1988)) and required additional mechanisms, such as including "growth points" or a "focus tree".

Kittredge et al. (1991) argue that a domain-independent approach to planning reports (weather, statistical reports) is impossible. Instead, we need Domain Communication Knowledge (DCK), which perhaps would better be called genre communication knowledge, which is knowledge specifically about how to communicate in a given domain and in a given genre. While it may have grown historically in explainable patterns, it is now fixed. It has important computational ramifications: it eases text understanding and especially text production. Reports can be easily generated using schemas, which can be seen as a parametrized context-free text grammar.

4 Extended Discourse Plan Operators

4.1 EDPO: Parallel Chunks of Structure

In this section, we motivate the two key structural aspects of our proposed extended discourse plan operators (EDPOs) from problems that existing approaches face.

In *John committed suicide*, *committed* and *suicide* form a collocation. This is difficult to represent concisely in a context-free grammar. Similarly, in a weather report, certain topics always appear together in a particular textual

²We have changed the subject-matter analysis given in (Moore and Pollack, 1992) to a multi-nuclear one, which better captures the data. The difference can be seen most easily in a causality relation, where there is a difference between a causal chain ( multinuclear) and a causal relation itself being the cause of another fact (say, a mental attitude in an agent). This distinction has also been discussed by Claire Gardent.
pattern, forming a "textual collocation". As in syntax, a large number of nonterminal symbols or goal representations (each corresponding to a schema or a plan operator) is needed, even though actually we have a single extended base structure for the report. Following proposals made for syntax (TAG), we therefore choose a tree-based representation that exploits larger elementary structures (following (Claire and Webber, 1998) and related work).

The second problem is that of parallel representations. We consider again an analogy from syntax first. If we wish to use a foreign phrase, such as *Sic transit gloria mundi* ("thus passes the glory of the world") in the generated text, we will not require a Latin realizer. Instead, the sentence planner should be able to choose directly the surface string, bypassing the realizer.

For discourse planning, an obvious example is conventionalized greetings. Here, the communicative goal could be defined as greet(H). (Note that this goal is a communicative goal even if accomplished by non-verbal means such as a raised eyebrow.) H is merely expected to understand S's intention to greet H, not any propositional content. Thus, the sentence planner should not be involved, and instead the intentional representation should directly produce a syntactic representation for the realizer. The realizer then handles tense, aspect, number agreement, and other factors according to which standardized greetings may vary according to the standard grammar of the target language.

Thus, each EDPO represents extended structure not only at one level of representation, but at several levels of representation in parallel. These representations include an intentional representation (IntentR) with nodes related by presentational RST relations, a conceptual-informational representation (InformR), with nodes related by RST subject-matter relations, a syntactic representation (SyntR), a surface representation (SurfaceR), and perhaps a formatting representation (FormR). Some samples – relating to the examples discussed in Section 4.1 – can be found in Figures 2 and 3.
4.2 The Rhetorical Approach and the Schema Approach as Special Cases

One of our motivations for developing EDPOs has been to overcome the dichotomy between a schema-based approach needed for generating reports and the rhetorical approach, which, while not appropriate for most discourse planning tasks, appears to capture some important constraints on discourse structure. In fact, we claim that for our approach, the schema approach and the rhetorical approach are special, if extreme, cases. A DCK-schema is simply an EDPO that has a ConcR of depth one with an associated IntentR structure consisting of a Joint, while an RST-based plan operator is an EDPO that has an IntentR of depth one with no associated InformR structure.

4.3 Discourse Generation with EDPOs

To perform discourse generation with EDPOs, we extend traditional pipeline model to allow for "shortcuts". If an EDPO specifies structure at a higher level of representation, then the modules that perform the intermediate transformations (content determiner, sentence planner, realizer, and/or formatter) need not be invoked. The process is driven by communicative goals (i.e., intentions represented in the IntentR).

- The text planner finds a plan operator that matches a current communicative intention; higher levels of representation (InformR, SyntR, SurfaceR, FormR) are assembled as much as possible. Note that nodes at different levels of representation are linked; when one structure is expanded, these links show where to assemble related chunks of structure.

- A separate module of the text planner, the content planner, chooses content to communicate in order to meet an intention if the EDPO has not specified any. This content is represented in the InformR. In the most general case (i.e., when the EDPO is simply an RST-style plan operator), this step is very complex and requires extensive knowledge and reasoning about the world and about the hearer. As shown in Figure 4 below, the EDPO can constrain the search for adequate content by specifying specific subject-matter relations, and/or by constraining the contents of the nodes in the InformR.

- The sentence planner finds SyntRs (lexico-syntactic representations) for InformRs, incorporating pre-assembled chunks provided by EDPOs as required.

- The realizer determines the SurfaceR for a SyntR, incorporating pre-assembled chunks provided by EDPOs as required.

- The formatter determines the FormR (formatting instructions) for a SurfaceR if needed, incorporating pre-assembled chunks provided by EDPOs as required.

Let us consider an example. Suppose a speaker (S) wants to make the hearer (H) strongly believe that John overslept. Let's call that proposition \( p \). We assume that H is not S's communicative slave, meaning H will not necessarily believe what S says. In fact, let us assume that if S simply states \( p \), H will believe \( p \) too weakly; instead, S needs to convince H. S searches in his data base of EDPOs for one that meets the goal (of increasing H's belief in a proposition \( p \)) and finds the one shown in Figure 4.

H searches for a fact causally related to \( p \) in his knowledge base, and finds proposition \( q = \) John is unshaven. Let us suppose further that H determines that H won't believe \( q \) sufficiently, either. H tries using the same EDPO again; he searches for another fact causally related to \( q \), but fails.

H then tries the next EDPO that matches the goal (of increasing H's belief in \( q \)), shown in Figure 5. The information at the informational level is simply related by a Joint, so there is no need to search the domain data base for facts related to \( q \). Now H can attach the structures in conformance with the dotted lines; the result is shown in Figure 6.

At this point, \( p \) and \( q \) are instantiated propositions, but \( r \) is only a variable. S determines (presumably through reasoning about communication) that the proposition that Bridget saw John will achieve the associated communicative intention (namely, to establish S as expert in this matter in H’s ears). Let's call that proposition \( r \). S then determines that \( p \), \( q \), and \( r \) leaves can be generated as locutionary acts. Thus, no further discourse planning is needed and S invokes his sentence planner, and then the realizer. Finally, S utters a discourse such as John has overslept again! He hasn't shaved - Bridget saw him.

Let us consider, as another example, the discourse form (Moore and Pollack, 1992), shown above in (1). To plan this discourse, we use the EDPO in Figure 7. This EDPO captures the intuition that in order to motivate H to intend to perform an act \( A_1 \), S can point out that it is a precondition for another act \( A_2 \) which S already knows H to firmly
Figure 4: EDPO for convincing H of a proposition p using causation

Figure 5: EDPO for convincing H of a proposition p using expert status

Figure 6: Derived structure after one planning step
intend to perform. Thus, S can search her knowledge base for preconditions for A, which is a constrained task. If S finds a precondition A₂ but is not sure whether H intends with a sufficient degree of firmness to perform A₂, then S can apply the same EDPO recursively. At the IntentR, this will result in expanding the satellite node of the first Motivation relation. At the InformR level, the effect of a Condition (namely, the proposition that a condition relation holds) does not match the precondition of the second segment. However, the precondition of the first segment of the second application of this EDPO is identical to the precondition of the second segment of the first application (namely, a description of A₂). Thus, at InformR we unify the two nodes, and obtain the structure shown above in Figure 1 for discourse 1.

As we have seen in these two examples, the action associated with plan operator is a complex action of assembling connected chunks of trees into larger chunks. (This should be contrasted with the case of the simple rhetorical approach, where the action associated with an RST-derived plan operator is always simply one of juxtaposing). It is this fact that allows us to derive two non-isomorphic structures. We will return to the issue of exactly what operations are needed for combining these chunks in the next subsection.

4.4 EDPOs and Discourse Competence

A collection of extended discourse plan operators represents the knowledge that lets S plan discourse, or, put differently, a speaker's competence in planning discourse. We will call this knowledge rhetoric. Note that rhetoric complements other forms of communicative competence, including grammatical, semantic, and pragmatic competence.

The size of an EDPO (roughly speaking), both in terms of the size of its representations at each level, and in terms of the number of levels at which it specifies structure, is a measure of the degree to which the EDPO represents conventionalized rhetoric. The more structure is specified, the easier discourse planning is, but the less variation in discourse exists. Conventionalized rhetoric permits agents to perform adequately in a vast variety of contexts. Report generation is a case in which highly conventionalized rhetoric is in fact required. At the other extreme, purely intentional EDPOs provide constraints on discourse structure but require extensive processing in order to determine content that meets the constraints (see Section 4.3). For example, advertising copy writers may spend weeks finding an optimal two sentence discourse. However, in many cases, neither approach will be sufficient, and intermediate EDPOs will be used. For example, to convince H that something is true, a good strategy for S is to provide a cause (see Figure 4). This knowledge is directly available and need not be derived each time S wants to convince H. Thus, rhetoric serves to constrain search in discourse planning.

4.5 Practical Issues

Simplified versions of the approach proposed here have been implemented at CoGenTex, Inc. (Lavoie and Rambow, 1998; White and Caldwell, 1998). Neither implementation currently includes all proposed levels of representation (the closer to market an application, the less useful the deeper levels of representation), and (Lavoie and Rambow, 1998) does not allow for larger chunks of structure.

The experience with these implementations in applications points to another interesting aspect of the EDPOs approach: because it allows us to link different amounts of information at different levels of representation, EDPOs can bridge the gap between templates and full generation. Templates (=pure surface representations) can be used when possible, 

[^3]: Unlike Moore and Paris (1993), we do not use this term in opposition to “intentions”.
and deeper representations can be accessed when needed. In this respect, the proposed work is similar to the IDAS system (Reiter et al., 1995), in which action representations can be hybrid and contain bits of structure from different levels (canned text templates with pointers to the knowledge base, and case frames with canned text).

5 Conclusion

We have presented a new kind of representation for discourse planning operators, extended discourse plan operators. EDPOs allow for a uniform representation in one formalism of both pure RST-style plan operators, and of DCK-schemas. Like the approach of Moore and Paris (1993), we can (and in some sense must) retain notion of communicative intention, but unlike Moore and Paris (1993), we keep it on a separate level from subject-matter relations. EDPOs represent a coherent notion of rhetoric as knowledge about how to achieve communicative goals in varying contexts of conventionalization of discourse.

Bibliography


