

Defining the Right Frontier in Multi-Party Dialogue

Julie Hunter*, Nicholas Asher[†], Eric Kow*, Jérémy Perret*, Stergos Afantenos*

*IRIT, Université Paul Sabatier

[†] IRIT, CNRS

juliehunter@gmail.com, nicholas.asher@irit.fr, eric@erickow.com,
jeremy.perret@irit.fr, stergos.afantenos@irit.fr

Abstract

We develop a Right Frontier Constraint (RFC) for multi-party dialogue (“multi-logue”), after arguing that extant definitions of the RFC, and in particular that of SDRT, cannot be directly extended to multilogue. Our proposal is developed and tested on a corpus of chats from an online version of the game *The Settlers of Catan*.

Many theories of discourse structure posit a Right Frontier Constraint (RFC) on discourse attachment (Polanyi and Scha, 1984; Polanyi, 1985; Webber, 1988). The RFC restricts the attachment of newly processed units of a discourse to a small subset of the units in the structure already constructed for some portion of the discourse. The motivating hypothesis behind the RFC is that discourse structure plays a major role in controlling salience. A coherence relation R inferred between two bits of a discourse d will have a particular effect on the shape of the overall tree or graph used to represent d 's structure in a way determined by the semantics of R and the discourse theory in use. Relations thus determine what nodes are found along the tree or graph's Right Frontier (RF), a set that evolves dynamically as a discourse proceeds. The RF *constraint* captures the observation that new utterances are normally attached to these nodes, which are predicted to be the most salient.

The RFC constrains semantic phenomena like anaphora and topic, as antecedents for most anaphoric expressions and ellipses are hypothesized to be found along the RF (Polanyi, 1985; Webber, 1988; Asher, 1993). It is also potentially helpful for discourse parsing: restricting attachments to units on the RFC considerably reduces the search space for attachments for discourse units and thus has the potential to improve inter-sentential attachment scores, which are in general much lower than scores for intra-sentential

attachment (Joty et al., 2015). Note, however, that the RFC rarely on its own determines attachment, and it can be violated in certain discourse configurations (Asher, 1993), though violations are rare in our corpus study (§4.3). The RFC is a defeasibly necessary but not sufficient constraint.

More importantly, the RFC is practically the only structural constraint on discourse attachment that takes the overall *structure* into account. Most discourse parsing models optimize probabilities for attachments over pairs of elementary discourse units, based on features like textual distance or grammatical or lexical properties of the paired elements. While local features are useful, discourse parsing performance lags behind syntactic parsing, because it does not use global features, in the way syntactic methods have done since (Collins and Duffy, 2002). The RFC is just such a global feature: it says the overall structure of the discourse graph has to have a certain shape. Because of data sparseness and our current limitations to supervised learning, it is infeasible to learn probabilistic global constraints like the RFC from the data directly. So defining an appropriate RFC via symbolic methods is a necessary step to improve discourse parsing.

The RFC has in practice been developed for, and tested on, monologue, generally in the form of newspaper texts (Afantenos and Asher, 2010). It is expected to be helpful as a constraint on multilogue as well, though important differences between multilogue and monologue prevent a trivial extension of standard RFC definitions. In monologue, a speaker is uniquely responsible for the information presented in the discourse, and the RFC is a constraint on the way that information should be presented. In dialogue, we deal not only with how speakers present information but also how they pick up on information presented by others. One speaker might make multiple points, but her respondent might pick up on just one, or ignore

them all. Or one or more respondents might wish to discuss multiple points simultaneously, introducing multiple conversation threads.

This paper develops a modified RFC suitable for multilogue and makes precise the RFC as a general constraint on discourse parsing. §1 reviews one version of the RFC for monologue, §2 introduces the corpus that we will use to develop our modified RFC, and §3 explains our choice of theoretical framework. In §4, we first extend the RFC to handle certain phenomena found in our corpus that are independent of multilogue (§4.1), and then extend this modified RFC to one suitable for multilogue (§4.2). §4.3 describes some experimental results with this RFC on our corpus. §5 and §6 present open problems and related work.

1 Modelling the RFC for monologue

In general, when an utterance u is made, the content of the utterance immediately prior to u will be highly salient, but other contents might be salient as well. A speaker might linger on a topic—elaborating on it, providing background on it, or explaining it and so on. In such a case, the point that is being elaborated on or explained, etc. will remain salient, and potentially form a chain of salient and accessible contents underneath it.

On the other hand, when a speaker, say, lists a series of attributes or describes a sequence of events, the most recently described attribute/event will be more salient than the previously described ones, rendering the latter inaccessible to later utterances. Thus in (1), the content of π_1 is inaccessible to that of π_3 —we cannot infer the sequence $\pi_1 + \pi_3 + \pi_2$, even though that would yield a more coherent discourse (without further context).¹

- (1) Rose dumped the cookies on the floor. π_1
 (So) She was sent to her room. π_2 (And)
 She drew all over the kitchen wall. π_3

If we reverse the order of π_2 and π_3 , as in (2), we can group Rose’s two acts together, as desired.

¹Eliciting intuitions about examples like (1) is a delicate matter. While rhetorical theories hold that discourse structure and coherence are intimately related, this does not mean that other factors, such as intonation and word choice, do not affect coherence. In (1), it is important to read the example with a normal intonation. Were a speaker to preface π_3 with *and* and pronounce *and* with a certain intonation, it would be clear that she wanted to retroactively add π_3 to the list of reasons why Rose was sent to her room, i.e. π_3 could attach to π_1 . However, the special intonation would arguably be a signal that the speaker wanted to return to a less salient point.

What’s more, while π'_1 alone is inaccessible to π'_3 , the fact that π'_2 clearly describes an event in a series of related events makes the group $\pi'_1 + \pi'_2$ salient and accessible. That is, we understand Rose’s being sent to her room as the result of both acts, not just of the more recently described one.

- (2) Rose dumped the cookies on the floor. π'_1
 (And) She drew all over the kitchen
 wall. π'_2 (So) She was sent to her room. π'_3

To make this precise, let’s consider the RFC as defined in *Segmented Discourse Representation Theory* or SDRT (Asher and Lascarides, 2003). In SDRT, the structure for a discourse d is modelled as a rooted spanning DAG (SDAG), called an SDRS, $G = (V, E_1, E_2, Last)$. V is the set of *elementary discourse units* (EDUs; labelled π_0, \dots, π_n) and *complex discourse units* (CDUs) in d , where an EDU is a clausal or sub-clausal unit and a CDU is a collection of EDUs (and possibly other CDUs) that together serve as an argument to a discourse relation. $E_1 \subseteq V \times V$ is the set of edges or labelled discourse attachments between elements of V . $E_2 \subseteq V \times V$ is the parenthood relation that relates CDUs to their component DUs. We write $e(\pi_x, \pi_y)$ when e is an edge with initial point π_x and endpoint π_y . *Last* is the last EDU in V , following the linear ordering of EDUs determined by their order in d . An SDRS is “spanning” in that all elements of V other than the root have at least (and possibly more than) one incoming edge: $\forall \pi_x \in V. (\pi_x \neq \text{ROOT} \rightarrow \exists \pi_v \in V. ((\pi_v, \pi_x) \in E_1))$.

The set E_1 can contain two types of edges, *coordinating* and *subordinating*. Relations such as Explanation, Elaboration, and Background—in which the second argument extends the discussion about the first—are represented with subordinating (vertical) edges. Relations such as Continuation, Narration, and Result—in which the second argument shuts off the accessibility of the first—are represented with coordinating (horizontal) edges. Suppose we prefix (2) with π_0 , *We’ve been having a rough time*, so that $\pi'_1 - \pi'_3$ elaborates on π_0 . $\pi_0 + (2)$ would yield the graph $G_{\pi_0 + (2)}$:

- $V = \{\pi_0, \pi'_1, \pi'_2, \pi'_3\}$
- $E_1 = \{\langle \pi_0, C_1 \rangle, \langle \pi'_1, \pi'_2 \rangle, \langle C_0, \pi'_3 \rangle\}$
- $E_2 = \{\langle C_0, \pi'_1 \rangle, \langle C_0, \pi'_2 \rangle, \langle C_1, C_0 \rangle, \langle C_1, \pi'_3 \rangle\}$
- $Last = \pi'_3$.

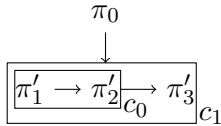


Figure 1: Graph of $\pi_0 + (2)$

For monologue, a node π_x is on the RF of a graph G , i.e. $\text{RF}_G(\pi_x)$, just in case π_x is *Last*, π_x is related to *Last* via a series of subordinating (*Sub*) edges, or π_x is a CDU that includes a node in RF_G :

Definition 1 Let $G = (V, E_1, E_2, \text{Last})$ be a discourse graph. $\forall \pi_x, \pi_y, \pi_z \in V$, $\text{RF}_G(\pi_x)$ iff

(i) $\pi_x = \text{Last}$, (ii) $\text{RF}_G(\pi_y)$ & $\exists e \in E_1$, $e(\pi_x, \pi_y)$ & $\text{Sub}(e)$, or (iii) $\text{RF}_G(\pi_y)$ & $\exists e \in E_2$, $e(\pi_x, \pi_y)$.

So the RF of $G_{\pi_0+(2)}$ is $\{\pi'_3, C_1, \pi_0\}$. Note that the RF is updated dynamically each time a new EDU is processed; the RF for (attachment of) an EDU π_n will be determined by the graph $G_{\pi_0-\pi_{n-1}}$. The RF for a CDU $\pi_m \dots \pi_n$, $m < n$, is the RF for π_m .

2 The Settlers Corpus

The Settlers of Catan is a win-lose game in which players trade resources (e.g. wood and sheep) to build roads and settlements. In the standard online version, players interact solely through the game interface, making trades and building roads, etc., without saying a word. In our online version, players were asked to discuss and negotiate their trades via a chat interface before finalizing them non-linguistically via the game interface. As a result, players frequently chatted not only to negotiate trades, but to discuss numerous topics, some unrelated to the task at hand.

The *Settlers* corpus is ideal for studying multilogue. The chats maintain the advantage of written text (no need for transcription) but they manifest phenomena particular to multilogue, such as multiple conversation threads. Also, the chats move quickly, which limits descriptively robust comments and forces players to exploit textual, discourse structuring clues.

The corpus consists of 59 games out of which 36 games (1027 dialogues, 9888 EDUs and 10181 relations) have so far been annotated for discourse structure in the style of SDRT, with a development subset of this corpus containing 9422 relations. This large annotation effort was carried out by 4 annotators who had no special knowledge of linguistics, but who received training over 22 negotiation dialogues with 560 turns. Because anno-

tating full discourse structures is a very complex task (using an exact match criterion of success, the inter annotator agreement score was a Kappa of 0.45 (Afantenos et al., 2012)), experts made several passes over the annotations from the naive annotators, improving the data and debugging it. The 4 naive annotators received no explicit instructions to obey SDRT’s RFC, and while expert annotators were aware of the constraint for monologue, they decided collectively not to make attempts to annotate in compliance with it; they picked attachment sites according to their best judgement.

3 Why SDRT?

We have chosen SDRT as the framework to develop an RFC for multilogue. The *Settlers* corpus is already annotated for discourse structure in the style of SDRT and in addition, SDRT’s RFC has been empirically validated on written monologue (newspaper articles and Wikipedia entries) using an annotation task in which annotators were not told about the RF, much less instructed to follow it (Afantenos and Asher, 2010). More importantly, however, SDRT deals easily with long distance attachments, which Ginzburg (2012) finds attested in multilogue, and has a semantics capable of dealing with fragments or non sentential utterances (Schlangen, 2003), which are frequent in our corpus. Also, it can model non-tree like structures, like that shown in Figure 2, which account for at least 9% of the links in our corpus. Such structures make theories that model discourse structures with rooted trees, like *Rhetorical Structure Theory* (RST) (Mann and Thompson, 1987) or simple dialogue models where attachments are always made to *Last*, cf. (Schegloff, 2007; Poesio and Traum, 1997), unsuitable. In Figure 2, QAP is the relation Question-Answer-Pair, ACK is Acknowledgement, and “kk” means “okay, cool”.²

From the perspective of discourse processing, the RFC could be key in solving the *attachment problem*—that of predicting where a discourse unit π_n will attach to the structure for $\pi_0-\pi_{n-1}$. If there are no constraints on a theory of attachment, the search space of solutions is very large making good attachment predictions impossible given the limited amount of data. So adding constraints is potentially crucial. Of course, if attachment is already very constrained, adding an RFC makes little

²To save space, we skip turns in examples when the turns are irrelevant to our main point.

234 gw anyone got wheat for a sheep?
 235 inca sorry, not me
 236 ccg nope. you seem to have lots of sheep!
 238 dmm i think i'd rather hang on to my wheat
 239 gw kk I'll take my chances then...

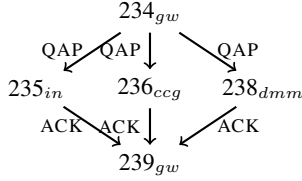


Figure 2: Example of a non-tree-like structure

to no difference. In RST, attachment is restricted to adjunction over trees from contiguous spans, so the attachment problem is comparatively easy to solve; attachment is even more trivial in a theory of dialogue where attachments must be made to *Last*. Such theories would gain little to nothing from an RFC.

SDRT is more liberal in its attachment principles than RST: though it incorporates constraints like connectedness, acyclicity and constraints on CDUs (Venant et al., 2013), non-adjacent and long distance attachments are common. Thus, adding an RFC to SDRT in principle greatly reduces the search space for attachment. When we combine this with the fact that SDRT’s graphs can deal with examples like Figure 2 and the examples of multiple threads discussed below, using SDRT to develop an RFC for multilogue is a natural choice.

4 Modifying the RFC

4.1 First modifications

SDRT’s RFC relies on an incremental construction procedure that ensures that each EDU π_n is attached at some point along the RF of a connected graph G for EDUs π_1, \dots, π_{n-1} before π_{n+1} is even considered. Before developing an RFC for multilogue, we first need to modify this procedure to handle two phenomena: CDUs and backwards links. This subsection treats these topics in turn.

The incremental construction procedure assumes that it is possible to tell where a CDU will attach to an incoming discourse structure even before the full content of the CDU is known. Given that a CDU is a group of DUs that function together to form a single argument to a discourse relation, the incremental procedure potentially introduces a fair amount of guesswork into the process of reasoning about attachment. Consider (3) and the two possible continuations, (a) and (b).

(3) Bill: I’m running late π_0 because my car broke down π_1 .
 Janet: If you call Mike π_2 , ...

- a. he might be able to pick you up and get you to the party on time π_3 .
- b. he might be able to come over and fix your car π'_3 .

In (3a), $\pi_2 + \pi_3$ intuitively attaches to π_0 , while (3b) suggests an attachment of $\pi_2 + \pi'_3$ to π_1 . Until Janet utters the consequent, we can’t tell where she is going with the antecedent.

There are two solutions to the problem posed by CDUs without resorting to a probabilistic version (which does not seem automatically learnable): (i) allow graphs to be corrected/repaid in light of new information (Asher, 1993) or (ii) wait to attach CDUs to an incoming discourse until the content of the CDU is complete. As an illustration, consider the graph G , shown in Figure 3. We can, as shown in (i), construct G by first drawing an edge e_1 from π_x to π_y and then adding an edge e_2 from π_y to π_z and correcting e_1 so that its endpoint is the CDU ($\pi_y + \pi_z$). Alternatively, as shown in (ii), we can wait to draw an edge with π_x as initial point until the CDU ($\pi_y + \pi_z$) has been constructed. (Relevant steps are separated by commas.)

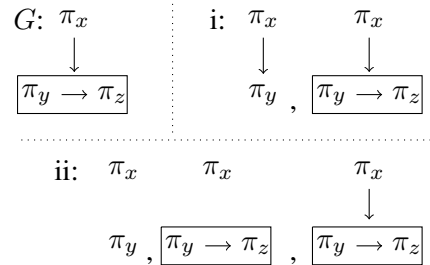
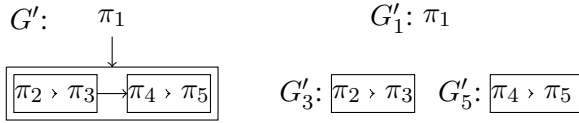


Figure 3: Corrected vs. delayed CDU construction

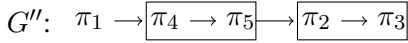
We adopt option (ii) and recast the RFC as a constraint on attaching subgraphs. This makes the construction of an SDRS more compositional and allows us to wed the RFC with standard, non-incremental discourse parsing models. Even the standard case of EDU attachment can be thought of in this way. Let π_5 be an EDU that needs to be attached to a connected discourse graph $G_1 = \langle \{\pi_1, \pi_2, \pi_3, \pi_4\}, E_1, E_2, \pi_4 \rangle$ and treat π_5 as the sole node in a graph $G_2 = \langle \{\pi_5\}, \emptyset, \emptyset, \pi_5 \rangle$. The problem of attachment for π_5 can be recast as the problem of attaching G_2 to G_1 .

To verify that a graph G contains no RF violations, we must be able to check for any

subgraph of G , whether that subgraph violates the constraint. And we must allow that a subgraph of G might contain further, unconnected subgraphs, G_1, G_2, \dots, G_n , each with its own *Last*. Let G be an SDRS over EDUS $\{\pi_1, \dots, \pi_j, \pi_{j+1}, \dots, \pi_k, \pi_{k+1}, \dots, \pi_n\}$ and suppose we have constructed three subgraphs $G_j = G \upharpoonright \{\pi_1, \dots, \pi_j\}$, i.e. G restricted to π_1, \dots, π_j in their textual order, $G_k = G \upharpoonright \{\pi_{j+1}, \dots, \pi_k\}$, and $G_n = G \upharpoonright \{\pi_{k+1}, \dots, \pi_n\}$. G_j, G_k , and G_n each has its own RF, open to attachment, which makes possible highly undesirable graphs. Consider G' below and its subgraphs G'_1, G'_3 , and G'_5 :



If we allow any subgraph to attach to the RF of any other subgraph, we could in theory, combine the subgraphs of G' to build a graph G'' as follows:



In fact, every EDU in any graph G could be considered a single-node subgraph, in which case allowing attachment on the RF of any graph would render an RFC pointless. An utterance could provide the output for a link to an arbitrarily later utterance, and speakers would be able to respond to points that haven't been salient for some time.

G'' is problematic because the CDU $\pi_2 + \pi_3$ is attached to $\pi_4 + \pi_5$, but neither π_4, π_5 , nor $\pi_4 + \pi_5$ is on the RF for π_2 . Moreover, the RF for a new EDU, π_6 , would be defined by π_5 (*Last* in G''), despite the the coordinating link from $\pi_4 + \pi_5$ to $\pi_2 + \pi_3$, which should block attachment to π_5 .

We need to constrain graph development. Let's return to our subgraphs G_j, G_k , and G_n of G , and let G_{jn} be the extension of G_j with G_n . We must eventually construct a graph that attaches G_k to G_{jn} ; call it $G_{jn} + G_k$. Such configurations can occur when G_k contains a parenthetical remark about G_{jn} or when it provides the topic. This means that G_k will be subordinate to G_{jn} or that $\text{RF}_{G_k} \cap \text{RF}_{G_{jn}+G_k} \neq \emptyset$. Let $\text{RFC}(G_{jn})$ mean that each edge in G_{jn} complies with the RFC in that each node π_n in G_{jn} attaches to a node on the RF for π_n as defined in Definition 1. The predicate OK, defined below, constrains the construction of

graphs like G_{jn} . Note that Axiom 1 requires G_k be non-empty.

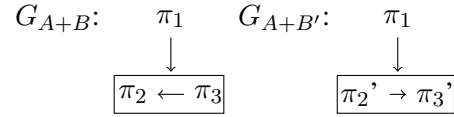
Axiom 1 Let $G = G_{jn} + G_k$, with G_j, G_n, G_k and G_{jn} as described above. Then $\text{OK}(G)$ iff:

- (a) $\text{RFC}(G_{jn}) \wedge \exists e(e(G_{jn}, G_k) \wedge \text{Sub}(e))$ or
- (b) $\exists \pi_x(\text{RF}_{G_k}(\pi_x) \wedge \text{RF}_{G_{jn}+G_k}(\pi_x))$

We apply this axiom below.

Another complication, given that edges in E_1 are directed, is that the direction of some edges reverses the textual order of their arguments.

- (4) A [Would anyone give me some clay?] $_{\pi_1}$
- B [I would,] $_{\pi_2}$ [if you give me a sheep] $_{\pi_3}$
- B' [if you give me a sheep] $_{\pi'_2}$ [I would,] $_{\pi'_3}$



$A+B$ yields a coherent SDRS, yet the backwards link $\pi_2 \leftarrow \pi_3$ violates the RF defined by Definition 1. The EDU π_1 is *Last* from the point of view of π_2 , and so defines the RF for π_2 ; π_3 will not figure in this RF, thus the edge from π_3 to π_2 is a violation.

Furthermore, while (4B) is truth conditionally equivalent to (4B'), they are not discourse equivalent because $(\pi_2 + \pi_3)$ and $(\pi'_2 + \pi'_3)$ do not have the same felicitous continuations; i.e., $(\pi_x \rightarrow \pi_y)$ and $(\pi_y \leftarrow \pi_x)$ make importantly different contributions to discourse structure.

- (5) [I would,] $_{\pi_2}$ [if you give me a sheep,] $_{\pi_3}$
 - a. [and an ore] $_{\pi_4}$
 - b. ??[with pleasure,] $_{\pi'_4}$
- (6) [if you give me a sheep] $_{\pi'_2}$ [I would,] $_{\pi'_3}$
 - a. ??[and an ore] $_{\pi_4}$
 - b. [with pleasure] $_{\pi_4}$

The examples above are noticeably more felicitous if the continuation targets the textually last EDU (π_3 or π'_3) despite the fact that these EDUs are the inputs for their respective conditional links.

To handle backwards links, we permit two graphs G_n and G_m to be attached with an edge in either direction. $\text{RFC}(G, e(\pi_x, \pi_y))$ means that the edge e complies with the RFC in G . We define an undirected RFC constraint over graphs G_n and G_m of an eventual graph G by extending Definitions 1 Axiom 1 with Axiom 2:

Axiom 2 $\forall \pi_x \in V^{G_n}, \forall \pi_y \in V^{G_m}$ such that $\neg \exists e \in E_1^{G_n} \cup E_1^{G_m}. (e(\pi_x, \pi_y) \vee e(\pi_y, \pi_x))$:
 $\text{RFC}(G_n + G_m, e(\pi_x, \pi_y))$ iff
 (a) $\text{RF}_{G_n}(\pi_x) \wedge \text{RF}_{G_m}(\pi_y)$ or
 (b) $\text{RF}_{G_n}(\pi_y) \wedge \text{RF}_{G_m}(\pi_x)$

The full definition of an undirected RFC, RFC_u , over the fusion of any two subgraphs now is:

Definition 2 $\text{RFC}_u(G_n + G_m)$ iff $\forall e \in (E_1^{G_n + G_m} \setminus (E_1^{G_n} \cup E_1^{G_m}))$: $\text{OK}(G_n + G_m) \wedge \text{RFC}(G_n + G_m, e)$

We can now handle examples (5)-(6). Consider (6). In constructing the graph for (6a), π'_2 and π'_3 potentially determine separate subgraphs. Suppose we attach π_4 to π'_2 to build the structure $[\pi'_2 \rightarrow \pi_4] \rightarrow \pi_{3'}$ (a felicitous combination of the EDUs in (6a)). $\pi_{3'}$ is the only node on the RF in the subgraph consisting only of $\pi_{3'}$, so by Axiom 1, it should remain on the RF once we attach it to $\pi'_2 + \pi_4$, but this will not be the case, as the RF will be defined by π_4 , the *Last* node. Hence we predict that (6a) is unacceptable while (6b) is acceptable. Reversing the links makes no difference; while the highest link is reversed in (5), *Last* is determined by textual order, so *Last* is π_3 not π_2 . Thus we cannot attach π'_4 to π_2 in (5b) for the same reason that we cannot attach π_4 to π'_2 in (6a).

4.2 Extending the modified RFC to multi-party dialogue

Our undirected RFC cannot yet handle structures like that in Figure 2 (as neither 235 nor 236 are on the RF for 239) or examples of “interleaved threads”, in which speakers juggle multiple conversations simultaneously. Both types of example are common in our corpus; the example in Figure 4 involves (at least) three interleaved threads.

165	lj	anyone want sheep for clay?
166	gw	got none, sorry :(
167	gw	so how do people know about the league?
168	wm	no
170	lj	i did the trials
174	tk	i know about it from my gf
175	gw	[yeah me too,] _a [are you an Informatics student then, lj?] _b
176	tk	did not do the trials
177	wm	has anyone got wood for me?
178	gw	[I did them] _a [because a friend did] _b
179	gw	lol william, you cad
180	gw	afraid not :(
181	lj	no, I'm about to start math
182	tk	sry no
183	gw	my single wood is precious
184	wm	what's a cad?

Figure 4: Example of interleaved threads

To handle such examples, we assign each speaker s in a multi-party dialogue a textual *Last*, i.e. the textually last EDU that s introduced into the chat. We call the RFC defined with individual speaker *Last*s RFC+MLAST. RFC+MLAST allows the discourse parser to attach turns 235, 236 and 238 in Figure 1 to turn 239 without violations, because for every edge with 239 as its endpoint, its initial point is *Last* for some speaker. For Figure 4, MLAST lets 168 (*no*) attach to 165 as an answer, even though GW has introduced a separate question on a completely different topic that attaches via a coordinating Continuation relation to 165. Similarly, MLAST allows us to attach 175b to LJ's turn in 170 and GW's in 178 to 176 in spite of WM's attempt to start a new bargaining session. Likewise for the attachment of 182 to 177. RFC+MLAST fails, however, to allow the intuitive attachment of 181 to 175b, because GW's *Last* is 180 not 175b (see §5 for discussion). Still, it yields considerable improvement over the modified RFC from §4.1. Table 1 shows the effect of MLAST on RFC violations on the development portion of the *Settlers* corpus. The manually annotated structures obey RFC+MLAST on 95% of the links, while only 83.5% of the links obey the RFC from §4.1.

4.3 Experiments and Results for MLAST

A dynamic calculation of restrictions to the search space for attachments using basic RFC and RFC+MLAST shows that RFC+MLAST has a positive effect on the search space for dialogue parsing in the *Settlers* corpus. As shown in Figure 5, the number of possible attachment points decreases dramatically with RFC+MLAST as the size of the dialogues in the corpus increases.

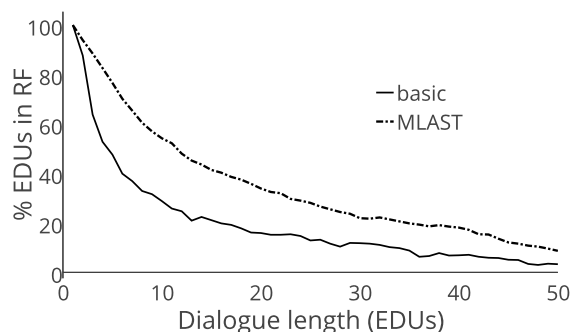


Figure 5: BASIC and MLAST versions of RFC

Using RFC+MLAST can have an important and beneficial effect on parsing. Yet just as the value of adding an RFC can vary depending on the dis-

Data	total links	RFC	MLAST	F-attachment
gold	9293	1536	447	100%
MST	8179	267	191	60.4%
ILP	17430	4342	2693	49.3%
LAST	8179	0	0	56%

Table 1: RFC violations

course theory in question, it can also vary depending on the discourse parser in question. We have developed and trained learner and decoder dialogue parsers for attachment on a simplified version of the *Settlers* chat corpus (without CDUs). The learner is a regularized maximum entropy (MaxEnt) model (Berger et al., 1996). Using standard, superficial features for discourse parsing of the sort found in e.g., Muller et al. (2012) and Li et al. (2014), we learn a probability distribution over pairs of EDUs as an input to several decoders. One decoder uses the MST algorithm (jin Chu and hong Liu, 1965; Edmonds, 1967). Another constructs first a maximal spanning DAG or MSDAG (McDonald and Pereira, 2006; Schluter, 2014) and then prunes it with constraints defined using ILP. The attachment F-scores for MST and ILP without the RFC are provided in Table 1.

As shown in Table 1, MST closely complies with the standard RFC; 96,7% of its predicted attachments obey the RFC while 97,7% comply with RFC+MLAST. Therefore, using RFC+MLAST as a filtering constraint on MST would have little effect. ILP on the other hand could benefit considerably from having RFC+MLAST as a constraint, gaining up to 10% in its attachment score.

The data on MST, however, raise questions about its value as a parsing algorithm for our corpus. Note how closely it complies with the RFC. This is surprising, because CDUs are important in calculating the RF in both monologue and multilogue, so we would expect a considerable amount of RFC violations with a decoder that ignores CDUs. This is especially so given that removing CDUs from the gold annotations on the *Settlers* corpus results in about a 10% increase in violations of the basic RFC; only 73% of the attachments in the manually annotated corpus obey RFC once we drop CDUs.

A baseline where we simply attach each EDU to the preceding one verifies the plain RFC at 100%. We call this baseline LAST. The RFC violations over our corpus suggest that MST is much closer to LAST than it is to the gold annotations. The figures suggest that tree construction algorithms such

as MST miss around 12% of the attachments in the gold corpus that are RFC violations but not violations on RFC + MLAST. Thus while MST might be a locally good strategy (with attachment F-scores at 0.81 within a sequence of consecutive turns by the same speaker), it is a globally mediocre strategy. This worsening echoes the difference reported by others between intra-sentential attachment scores and inter-sentential attachment scores in monologue (Joty et al., 2015). ILP, on the other hand, patterns more closely with the gold data and has many more long distance links.

5 Beyond MLAST

Double-tasking Recall that RFC+MLAST blocks the attachment of 181 to 175b in Figure 4, because GW’s *Last* is 180, and not 175b. This violation is interesting, because it illustrates a systematic pattern in which the same speaker carries on several interleaved threads, while others are talking. Such cases intuitively call for multiple *Lasts* for a single speaker; that is, a *Last* for speaker s for each thread in which s is engaged. This notion, in turn, calls for a criterion for distinguishing threads.

One possible, and simple, solution would be to individuate threads by their members. Then we could extend the RFC+MLAST to include a *Last* for each speaker for each subset of speakers that is engaged in a thread. This would solve the problem of attachment in Figure 4; however, it would not solve the problem in general, as we also have examples of multiple threads involving the very same subset of speakers. In Figure 6, LJ and GW

```

119  lj    gw did you take logic1 this year?
123  gw    anyone got more clay? I fancy another
124  gw    can offer a range of items
125  lj    i have clay
126  gw    no i didn't lj, I'm not a student :)
128  lj    would like wood
129  gw    1 for 1?
130  lj    ahhh ok, never mind
131  lj    sure

```

Figure 6: More interleaved threads in duologue

are engaged in both a trade negotiation, which takes place over turns 123-125, 127-129 and 131, and a thread about whether gw took logic, which takes place over turns 119,126 and 130. Even if we add a *Last* for each subgroup of speakers, 126, 128, and 130 will still give rise to RF violations.

It is difficult to define a thread precisely. And in fact, it’s not clear to us that 126, 128, and 130

shouldn't count as RFC violations, in the same way that "discourse subordinations" (Asher, 1993) in monologue text count as RFC violations. Violations involving multiple threads with the same two speakers can be coherent but they require more effort to understand. For instance, annotators and interpreters could argue about the attachment of 130 to 126; and if we imagine that GW had made a different offer in 129 (say, 2 for 2 or 2 for 1), then we could easily imagine 130 as a response to 129. Moreover, GW actually refers to LJ by name in 126. This is a funny thing to do given that LJ is his only interlocutor at this point; if we treat 126 as an example of discourse subordination, however, then we can imagine that the name is being used as a signal for a discourse subordination.

Turn internal violations While we have not found a significant number of such examples in our corpus, the RFC might ultimately need loosening to handle examples like the following.

- (7) B: Who has ore? **I have sheep to give.** *I could also give some clay.*
 A': How many sheep?
 B': ?? Three sheep even.
- (8) A: Anyone want ore for sheep?
 B: **I'm not giving up my sheep for now,** *but lj might want to give some of hers.*
 A': What if I offer you two ore?
 B': ?? Not for all the ore in the world.

Attachment possibilities for speakers are asymmetric. In (7)-(8), the boldface argument is related to the italicized argument by a coordinating relation (Alternation in (7), Contrast in (8)), which should block the accessibility of the boldface argument. Indeed, B cannot continue with a comment targeting this argument (B+B'), though B' *would have been* felicitous in the absence of the italicized argument. By contrast, if another speaker, A, responds to B's turn, both arguments of the coordinating relation are accessible, as shown by the felicity of the A' continuations (B+A').

The theoretical explanation of this has to do with the underlying semantics of contributions in multilogue. The meaning of a dialogue is a set of commitment slates, one for each speaker. Speakers commit to their own contributions in dialogue but not necessarily to the contributions of their interlocutors, *unless* the attachments they make of *their own contributions* requires also that they

take on board the commitments of the interlocutor (Hamblin, 1987; Lascarides and Asher, 2009). From this point of view, an asymmetry in the RFC is to be expected in multilogue.

6 Related Work

The RFC is related to projectivity in parsing (Nivre, 2003). Like projectivity, RFC compliance is a property of a graph with respect to textual order, and like projectivity, the RFC rules out crossing dependencies (relative to textual order) except in special cases. Unlike projectivity, however, the RFC depends on a semantic distinction between subordinating and coordinating relations, and a distinction between CDUs and EDUS. Projectivity and the RFC are thus not equivalent even on trees.

The RFC has been a topic of interest in theoretical work on discourse structure for a long time. But to our knowledge, we are the first to study how it fares for multilogue on a large discourse annotated corpus. With regard to empirical work on discourse parsing, Afantenos and Asher (2010) demonstrate the potential of this constraint, but we are not aware of any actual parsing results with the RFC for monologue or dialogue. Afantenos and Asher (2010) also conducted an empirical study on RFC for monologue. However, we have shown that the RFC for monologue is not suitable for multilogue and must be modified.

7 Conclusions

This paper has presented an account of the RFC in multilogue with complex segments, backwards links, and simultaneously running multiple threads of conversation. We have shown our corpus verifies our modified RFC+MLAST. Our experiments have shown that some discourse parsing methods can benefit substantially from the RFC as a processing constraint and that in general the RFC provides an important reduction in the search space of possible attachments. In future work, we will implement our modified RFC for parsing on multilogue data and investigate further the empirical effects of modified LAST to account for the difficulties mentioned in section 5.

Acknowledgements

We thank Semdial reviewers for their comments. This work was supported by ERC grant 269427.

References

- Stergos Afantenos and Nicholas Asher. 2010. Testing SDRT's right frontier. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, pages 1–9.
- Stergos Afantenos, N. Asher, F. Benamara, A. Cadilhac, C. Dégremont, P. Denis, M. Guhe, S. Keizer, A. Lascarides, O. Lemon, P. Muller, S. Paul, V. Popescu, V. Rieser, and L. Vieu. 2012. Modelling strategic conversation: model, annotation design and corpus. In *Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue (Seinedial)*, Paris.
- Nicholas Asher and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.
- Nicholas Asher. 1993. *Reference to Abstract Objects in Discourse*. Kluwer Academic Publishers.
- Adam Berger, Stephen Della Pietra, and Vincent Della Pietra. 1996. A maximum entropy approach to natural language processing. *Computational Linguistics*, 22(1):39–71.
- Michael Collins and Nigel Duffy. 2002. New ranking algorithms for parsing and tagging: kernels over discrete structures, and the voted perceptron. In Isabelle Pierre, editor, *ACL '02 Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 263–270.
- Jack Edmonds. 1967. Optimum branchings. *Journal of Research of the National Bureau of Standards*, 71B(233–240).
- Jonathan Ginzburg. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford University Press.
- Charles Hamblin. 1987. *Imperatives*. Blackwells.
- Yoeng jin Chu and Tseng hong Liu. 1965. On the shortest arborescence of a directed graph. *Science Sinica*, 14:1396–1400.
- Shafiq Joty, Giuseppe Carenini, and Raymond Ng. 2015. Codra: A novel discriminative framework for rhetorical analysis. *Computational Linguistics*. in press.
- Alex Lascarides and Nicholas Asher. 2009. Agreement, disputes and commitment in dialogue. *Journal of Semantics*, 26(2):109–158.
- Sujian Li, Liang Wang, Ziqiang Cao, and Wenjie Li. 2014. Text-level discourse dependency parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 25–35, Baltimore, Maryland, June. Association for Computational Linguistics.
- William C. Mann and Sandra A. Thompson. 1987. Rhetorical structure theory: A framework for the analysis of texts. *International Pragmatics Association Papers in Pragmatics*, 1:79–105.
- Ryan McDonald and Fernando Pereira. 2006. Online learning of approximate dependency parsing algorithms. In *Proceedings of EACL*, pages 81–88.
- Philippe Muller, Stergos Afantenos, Pascal Denis, and Nicholas Asher. 2012. Constrained decoding for text-level discourse parsing. In *Proceedings of COLING 2012*, pages 1883–1900, Mumbai, India, December. The COLING 2012 Organizing Committee.
- Joakim Nivre. 2003. An efficient algorithm for projective dependency parsing. In *Proceedings of the 8th International Workshop on Parsing Technologies (IWPT)*. Citeseer.
- Massimo Poesio and David Traum. 1997. Conversational actions and discourse situations. *Computational Intelligence*, 13(3):309–347.
- Livia Polanyi and Remko Scha. 1984. A syntactic approach to discourse semantics. In *Proceedings of the 10th International Conference on Computational Linguistics (COLING84)*, pages 413–419, Stanford.
- Livia Polanyi. 1985. A theory of discourse structure and discourse coherence. In P. D. Kroeber W. H. Eilfort and K. L. Peterson, editors, *Papers from the General Session at the 21st Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society.
- Emanuel Schegloff. 2007. *Sequence Organization in Interaction*. Cambridge University Press, Cambridge.
- David Schlangen. 2003. *A Coherence-based Approach to the Interpretation of Non-Sentential Utterances in Dialogue*. Ph.D. thesis, University of Edinburgh.
- Natalie Schluter. 2014. On maximum spanning dag algorithms for semantic dag parsing. In *Proceedings of the ACL 2014 Workshop on Semantic Parsing*, pages 61–65, Baltimore, MD, June. Association for Computational Linguistics.
- Antoine Venant, Nicholas Asher, Philippe Muller, and Pascal Denis Stergos D. Afantenos. 2013. Expressivity and comparison of models of discourse structure. In *Proceedings of Sigdial 2013*, Metz, France.
- Bonnie Webber. 1988. Discourse deixis: Reference to discourse segments. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, pages 113–122. Association for Computational Linguistics, Morristown, NJ.