Discourse Processing

Keywords: language interpretation#language generation#discourse segmentation#discourse relations
Contents list: Theories of discourse structure; Interpretation of discourse; Generation of discourse; Empirical approaches to discourse

Discourse Processing concerns the computational processes underlying the interpretation and production of text encompassing more than one sentence, i.e., discourse. Discourse is generally taken to be written, and often, but not always, monologic.

1 Introduction

Discourse processing refers to the computational processes underlying the interpretation and production of text encompassing more than one sentence, i.e., discourse. Discourse is generally taken to be written, and often, but not always, monologic. Two phenomena are considered intrinsically pertaining to discourse processing: a) the interpretation and production of phrases and utterances whose meaning depends on the discourse context; b) the fact that a sequence of two or more utterances almost always conveys a meaning that is more than the sum of meanings of the individual utterances. Consider the following example:

(1) As soon as they got to the beach, Karin jumped into the water. She was so hot from the long drive.

Example (1) illustrates the issues most closely associated with a) and b): respectively, reference resolution and production, and text coherence.

Reference resolution concerns the interpretation of those noun phrases speakers use to refer to what are called *discourse entities*, i.e., entities in the model of the discourse—for example, *Karin, she, the long drive.* Because reference resolution is closely related to the notion of processing of anaphora discussed in Chapter 245, we will not elaborate on it here. However, we will discuss the converse problem of *referential expression generation*, namely, how to choose a specific referential expression among all those that can potentially be used to refer to a discourse entity.

It is difficult to define text coherence exactly. We could define it as the quality of a text that is "tied" together just right. It is text that can be readily comprehended by the hearer, apparently without effort; at the same time, it is text where relations between different sentences are not so explicit as to make it uninteresting. Example (1) is coherent; however, consider (2):

(2) As soon as they got to the beach, Karin jumped into the water. She hates ice-cream.

Example (2) sounds incoherent: it is likely that the hearer will wonder about the connection between hating ice-cream and jumping into the water. The hearer may in fact proceed to make up scenarios in which (2) makes sense: e.g., Karin had ice-cream on the way to the beach, it gave her a stomach ache, and her way to deal with stomach aches is to swim. Scenario building of this sort is an exercise precisely in accounting for text coherence, i.e., in explaining text in terms of sentence connections to one another. However, this does not mean that every possible link between the sentences should be explicit. Expanding Example (1) as in (3) results in a tedious text, not a clearer one. Namely, text coherence appears to obey Grice's maxims of quantity (see Chapter 233 on Implicature).

(3) As soon as they got to the beach, Karin jumped into the water. She was so hot from the long drive, so she wanted to cool down. Because the temperature of the sea is generally much lower than that of the air, going for a swim accomplished her goal.

Text coherence encompasses more than appropriate connections between individual sentences. The discourse processing community agrees that discourse has a hierarchical structure: sentences are part of *segments* which in turn are part of superordinate segments. Informally, a segment can be seen as a group of locally coherent utterances (see below for a more formal definition). Consider the following discourse:

(4) (a) Georgia called Jeffrey on the phone.
    (b) She wanted to wish him happy birthday.
    (c) She also asked him if she could borrow his tent.
        (d) She had bought a tent herself a few months back.
        (e) However, it got torn on her summer hikes.
  (f) After the phone call, she went out for a jog.

The discourse in (4) is about Georgia's activities. Intuitively, we recognize that sentences (a) through (e) form a subsegment S_{a-e} of the whole discourse, as they pertain to Georgia's phone call to Jeffrey. In turn, (c), (d) and (e) form the subsegment S_{c-e} of S_{a-e} that concerns the request for the tent; and (d) and (e) form subsegment S_{d-e} of S_{c-e}, because together they provide a justification for the request in (c).

We will now discuss how different researchers account for text coherence in both its manifestations, connections between sentences and discourse segmentation, and how coherent discourses can be interpreted and generated.

2 Theories of discourse structure

Two main theories of discourse structure came to the fore in the mid eighties, and are still the most prominent fifteen years later: Grosz and Sidner's [1986] and Mann and Thompson's [1988].

Grosz and Sidner's theory (henceforth, G&S) sees discourse structure as the surface manifestation of the relationships among elements of the intentional structure underlying the discourse. In turn, the intentional structure is comprised of the intentions that a speaker brings to the discourse. There will be a primary intention, the *discourse purpose (DP)*, i.e., the intention that underlies engaging in that particular discourse. Further, a *discourse segment purpose (DSP)* is associated to each discourse segment, which is fully individuated by the corresponding DSP. Each DSP specifies how the specific segment contributes to achieving the overall DP. A plausible DP underlying the whole discourse in (4) is *Tell hearer about Georgia's activities*. The DSP associated to the subsegment S_{d-e} could be something like *Explain to hearer why Georgia needs to borrow Jeffrey's tent*. G&S does not specify which intentions can count as DPs or DSPs, other than noting that they are meant to be recognized (cf. Grice's notion of utterance-level intentions in Chapter 233 on Implicature). DSPs can be related to one another only via two relationships: dominance and satisfaction-precedence. DSP_1 dominates DSP_2 if DSP_2 is intended to provide part of the satisfaction of DSP_1. DSP_1 satisfaction-precedes DSP_2 if DSP_1 must be satisfied before DSP_2. Grosz and Sidner argue that the intentional structure of the discourse is intertwined to the attentional state as well, i.e., to the set of entities that are salient at any point in the discourse. Attentional state is modeled by a set of focus spaces, which are associated to discourse segments, and contain the entities salient within the corresponding discourse segment. The processing of focus spaces is modeled via a stack. Shifts to attentional state that are local to a discourse segment are outside the scope of G&S, but are accounted for by Centering Theory [Grosz, Joshi, and Weinstein1995; Walker, Joshi, and Prince1998].

Mann and Thompson [1988] propose Rhetorical Structure Theory (RST for short) as a descriptive framework that identifies hierarchical structure in text. RST is based on relations that relate two non-overlapping *text spans*, the *nucleus* and the *satellite*. The nucleus is the central member of the pair, the satellite is more peripheral. Relations include an effect and constraints on nucleus and satellites. Relations defined in the original paper include *Elaboration, Enablement, Evidence, Contrast* (comparable inventories of discourse relations have been proposed by a number of other researchers besides Mann and Thompson [Hobbs1979; Lascarides and Asher1993]). For example, the Evidence relation has as effect that the belief of the hearer in the nucleus is increased, and among its constraints, that the hearer will find the satellite believable. In Mann and Thompson's view, an analyst will first identify the minimal units of the analysis, which they assume to be clauses. Then, the analyst will start applying relation schemas to adjacent text spans, which are minimal units or, recursively, constituents of relations. In the end there will be one relation schema encompassing text spans that cover the whole text.

From these brief descriptions, we can see that G&S mainly accounts for the segmentation aspect of discourse coherence, but does not address how individual sentences are linked one to the other by domain or rhetorical relations. This is by choice. As Grosz and Sidner believe that the intentions underlying discourse are too diverse, they argue that it would impossible to enumerate the intentions that can serve as DSPs; hence, they conclude that enumerating a fixed number of relations as in RST is wrong. On the other hand, RST accounts for both individual relations between individual sentences, and for hierarchical segmentation of the discourse. The latter is a side-effect of how the RST analysis is conducted. Note that an analysis of a discourse according to G&S generally results in fewer and shallower segments than an RST analysis.

Grosz and Sidner present their theory as a computational account of discourse processing, but they do not provide much insight into the underlying computational processes other than proposing that the attentional state is modeled as a stack. Mann and Thompson do not make any computational claims, however their theory has been widely used in Computational Linguistics. The question thus arise, which processing paradigm is most appropriate for each theory. G&S lends itself to a top-down model of discourse processing: the hearer recognizes the DP, and then recursively the subordinate DSPs. RST, instead, lends itself more directly to a bottom up interpretation of discourse.

To conclude this section, we will note that a synthesis of G&S and RST has been proposed in [Moser and Moore1996]. The synthesis is based on the observation that the dominance relation between intentions in G&S closely corresponds to the nucleus versus satellite distinction between text spans in RST.

3 Interpretation of discourse

Discourse interpretation consists of the computational inferences that compute the extended meaning of discourse. We can divide the approaches into two main groups: logical approaches that compute domain and rhetorical relations between sentences in written texts [Lascarides and Asher1993], and plan inference approaches that compute the speech acts performed by participants in a dialogue [Perrault and Allen1980; Litman and Allen1990; Carberry and Lambert1999]. Plan inference approaches have been applied mainly to dialogue, nevertheless, this topic is considered part of discourse processing. Because of its inherent difficulty, not many researchers have tried to compute discourse segmentation as proposed in G&S, however see [Lochbaum1998] for such an attempt.

Traditionally, approaches to inferring relations between sentences make use of some type of logical inferencing, such as a variant on non-monotonic logic or abduction. We will briefly discuss approaches based on abduction [Hobbs et al.1993]. Abduction is an unsound inference rule that reasons from an effect to a potential cause: for example,

(5) the alarm went off $\Rightarrow$ there is a burglar in the house

Clearly, there may be other reasons why the alarm went off, e.g., the landlady forgot to switch it off. Abduction is a useful reasoning mechanism because it tries to find the *best* explanation for a fact. As far as discourse coherence is concerned, an abductive approach tries to find the most plausible coherence relation linking two utterances, on the basis of rhetorical, domain and world knowledge. Namely, an abductive approach will build a full explanation that supports that specific coherence relation. For example, to establish a *cause* relation between the two sentences in Example (2), an abductive approach would build an explanation, expressed in first order predicate logic, akin to the one in (3). The problem abduction has to face is how to choose the most plausible explanation among many possible ones. One can adopt heuristics such as choosing the explanation that uses the fewest assumptions, or compute the probabilities of each explanation and choose the most likely one. Both approaches have serious flaws: the former, that even plausible heuristics can fail fairly often; the latter, that it is unclear over which space of events to compute those probabilities.

The computational approaches just discussed are not explicitly based on cognitive findings on text comprehension. Nevertheless, questions addressed by cognitive scientists and psycholinguists have affected omputational models. Relevant issues include: inference control, i.e., which of the many possible inferences are made at comprehension, and which later, during recall; how the connectedness of sentences affects reading times and the accuracy of recall. For example it has been found that sentences that have a close causal connection are read faster and engender better content recall [Myers, Shinjo, and Duffy1987].

3.1 Plan Inference

The plan inference approach to discourse has been mainly applied to dialogues, although applications to monologic discourse that describes one or more agents' actions exist as well. It originated at the end of the '70s [Perrault and Allen1980], with the goal of providing an interpretation for indirect requests such as *I need to be in Boston on the 20th in the afternoon* (directed to a travel agent), or *The next train to Brighton* (directed to a clerk at the ticket booth). It rests on three components: the notion of speech acts from pragmatics; a theory of belief, desire and intentions from computational linguistics, which in turn owes much to philosophy of action; and planning models from artificial intelligence.

Every utterance counts as an action performed by the speaker, i.e. a *speech act*, such as *asking* or *promising* [Austin1962]. (We are oversimplifying here. In reality there are three acts associated with each utterance, *locutionary*, *illocutionary*, *perlocutionary*. The term *speech act* has come to refer mostly to the *illocutionary* act. Please see Chapter 230 on Formal Pragmatics for further details.) Utterances can perform speech acts directly, as Example (6a), or indirectly, as (6b).

(6a) Please find me a flight that arrives in Boston on the 20th in the afternoon

(6b) I need to be in Boston on the 20th in the afternoon

To explain how a statement such as (6b) can count as a request, proponents of the *inferential approach* [Searle1975] contend that indirect speech acts concern felicity conditions on the corresponding direct act. For example, a request such as (6a) is felicitous under the assumption that the speaker wants to fly to Boston on that specific date and time. (6b) then works because it explicitly states the speaker's mental attitude, once the hearer has recognized that the literal meaning of (6b) is inappropriate and must be "repaired" by some inference.

Planning is a computational technique from Artificial Intelligence that, given a goal s_g to achieve, builds a plan, i.e., a partially ordered sequence of actions whose execution will bring the agent from the initial state s_0 to s_g. Often the plan is built as a tree, whose leaves are the actions to be executed; the internal nodes represent actions at a higher level, that further decompose into lower level actions. For example, if an agent has the goal *Attend conference in Washington* and s/he lives in Chicago, the agent may build a plan that includes taking a flight from Chicago to Washington; in turn, to achieve taking the flight, the agent will need to buy a ticket, drive to the airport, and board the plane (see further details on planning in Chapter 46). Planners build plans on the basis of action operators, which as a minimum include: preconditions, the conditions that need to hold for the action to be executable; effects, what becomes true after performing the action; body, a decomposition into a partially ordered set of subactions whose execution will result into the execution of the action.

In the plan inference approach, speech acts are modeled as action operators from planning. However, the logical language in which the operators are expressed is augmented with mental attitudes such as knowledge, beliefs and desire. For instance, a formalization of *Request(S, H, α)* will include as a precondition that the speaker S wants the hearer H to perform action α (one of the felicity conditions on requests), and as an effect, that H wants to perform α. Such a formalization can be used to build the interpretation of an indirect speech act via plan inference rules that work backwards from the utterance to its interpretation. One such rule is: if γ is a precondition of action α and H believes S to want γ, then it is plausible that H believes S to want α. Note that the representation can also be used by a regular planner to produce a speech act, starting from a communicative goal to be achieved.

The plan inference approach has been extensively used in dialogue modeling. The original model was extended in various ways, such as introducing different levels of inferred plans, e.g., the discourse plan and the domain plan, that the speaker is pursuing [Litman and Allen1990; Carberry and Lambert1999].

As a final observation, approaches to discourse based on abduction, nonmonotonic logic or plan inferencing, albeit elegant, suffer from brittleness. One missing domain axiom may cause the model to fail as it is not able to find any complete explanation. Thus, many implemented systems nowadays, instead of a logical approach, use information that can easily be derived from the surface form of the utterance, such as intonation, connectives, idiomatic expressions, lexical associations between words [Reithinger and Maier1995; Qu et al.1997; Samuel, Carberry, and Vijay-Shanker1998]. These cues to the phenomenon of interest are derived from linguistic and corpus analysis (see last section of this Chapter).

4 Generation of discourse

From a computational point of view, discourse generation concerns the production of coherent, extended text. Whereas discourse processing is seen as the last stage in language interpretation, after parsing and semantic analysis, it is the first stage for language production. Computationally, language generation starts from a non-linguistic representation of information that we can consider parceled into messages to be conveyed. The first task to be performed is *discourse planning*, i.e., impose ordering and structure over the set of messages to be conveyed. Then comes *sentence planning and linearization*, which includes at least 1) sentence aggregation, i.e., grouping the elements of the discourse plan together into sentences; 2) the choice of referential terms to individuate the entities of interest. The final step is linguistic realization proper, namely, applying the rules of grammar in order to produce a text which is syntactically and morphologically correct.

All these topics are covered more in detail in Chapter 86 on Language Generation. Here, we concentrate on discourse planning, and on referring expression generation.

4.1 Planning and linearization

There are two main approaches used to generate a coherent discourse, *planning* and *schemata*.

The discourse planner is given a communicative goal to achieve such as *Intend S (Intend H α)*. Communicative goals represent the speaker's intentions to affect the beliefs or goals of the hearer. The planner will build a plan consisting of rhetorical actions to achieve the given communicative goal. For example, to achieve *Intend S (Intend H α)*, S may look for a β such that S expects H to want β, and then utter β as *Motivation* for α (Motivation is an RST relation):

(7) *Come to the party on Saturday. I will make your favorite deviled eggs*

The connection between discourse planning and the theories of discourse structure discussed earlier has mainly been achieved through RST. RST relations are recast in terms of planning operators [Moore and Paris1993]. The planner posts a high level communicative goal such as *Intend S (Intend H* α*)* in terms of the effect $\mathcal{E}$ the text should have on the reader. The planner will then search for a RST operator whose effect unifies with $\mathcal{E}$, and post the subgoals that correspond to constraints on nucleus and satellite of that rhetorical relation. These subgoals are then recursively expanded until the planner reaches the leaves of the rhetorical structure tree, those expressible as simple clauses.

Schemata are an alternative approach to using a planner. Schemata represent common patterns that texts in a certain domain or in a specific genre follow [McKeown1985]. A schema specifies how a particular discourse plan should be built using other schematas or messages, and the discourse relations that hold between different components of the discourse plan. Although schemata are not generally developed following a planning model, they can be considered as compilations of discourse plans produced by a planning system. As a mechanism for generation, schemata are less flexible, but easier to develop than a full fledged discourse planner. For example, because schemata lack information on the intentions of the speaker, they cannot be used if the system needs to replan, e.g. if the explanation of a certain p is not understood by the hearer and the discourse planner needs to build a different explanation for p [Moore and Paris1993].

Note that a discourse plan, whether built by a planner or as a schema instantiation, does not encode decisions regarding how the leaves should be parceled into individual sentences, and how these sentences should be connected. For instance, the two sentences in Example (7) could be linked in a variety of different ways, both paratactically and hypotactically, such as:

(8) *Come to the party on Saturday if you don't want to miss your favorite deviled eggs*

There are also more subtle decisions that need to be made. In Example (7) the adjective *favorite* in the second clause may well be derived from a full proposition in the discourse plan, such as *You like the deviled eggs I make a lot*. This is why many researchers consider *lexicalization* as part of sentence planning as well. Lexicalization pertains to choosing words to express concepts and relations.

The solutions proposed in the literature for sentence planning and linearization are diverse, although some general paradigms are beginning to emerge, and we do not have the space to discuss them here. For further details, see Chapter 86 on Language Generation.

4.2 Establishment of referential terms

The task of *generating referring expressions* concerns selecting words or phrases to identify discourse entities. The choice of referring expressions greatly affects the readability of a text. Compare text (9) to text (10), which always uses the nominal expression *Bill Gates*.

(9) When a Stanford University professor asked for volunteers to have their heads scanned, *Bill Gates* was the first to volunteer.

The billionaire CEO of Microsoft Corporation sat patiently while a laser scanner orbited *his* head several times. A short time later, a 3-D image of *Gates'* head floated on a screen.

(10) When a Stanford University professor asked for volunteers to have their heads scanned, *Bill Gates* was the first to volunteer.

Bill Gates sat patiently while a laser scanner orbited *Bill Gates'* head several times. A short time later, a 3-D image of *Bill Gates'* head floated on a screen.

In (10), the repeated use of the proper name *Bill Gates* makes the text sound clumsy. The much more fluent text in (9) makes use of different forms of proper names (*Bill Gates* or simply *Gates*), pronouns (*his*), and complex definite referring expressions such as *the billionaire CEO of Microsoft Corporation.*

The problem of generating referring expressions can be subdivided into:

1. initial introduction, i.e., how to perform the initial reference to a discourse entity
2. subsequent references. This includes choosing between a pronoun and a definite description; if the latter is chosen, then the issue is which features of the entity in question to include in the description.

The initial introduction and the choice of pronoun or definite description are generally performed by algorithms based on the *given/new* distinction [Prince1981] or on *centering* [Grosz, Joshi, and Weinstein1995; Walker, Joshi, and Prince1998] or a combination of the two (see Chapter 254 on processing of anaphora).

Regarding the choice of appropriate definite descriptions, initial approaches [Appelt1985] took a full planning approach to generating referring expressions. This means that in principle they could generate any description that satisfies a given communicative goal. As this approach was computationally inefficient, later approaches, most notably Dale's [1992], focused on the restricted problem of building a *distinguishing description*. A distinguishing description is true only of the entity being described and of no others among the currently salient discourse entities. These algorithms generally aim at finding the *minimal* distinguishing description. However, even computing a minimal distinguishing description is an inherently hard computational task. [Dale and Reiter1995] showed it is NP-hard by reducing it to a set cover problem (on the notion of NP-hardness, please consult Chapter 8 on Computational Complexity). Moreover, humans do not produce minimal distinguishing descriptions, either because also humans face computational limitations, or because they intend to achieve other goals beside identification [Jordan2000]. In Example (9), the complex noun phrase *the billionaire CEO of Microsoft Corporation* may be used to introduce information that the hearer is not expected to know, or, more likely in this context, to remind the hearer of Gates' position. Algorithms used today then try to strike a balance between conciseness of the definite description and trying to reproduce human behavior, as observed in corpus analysis (see next section).

5 Empirical approaches to discourse

To conclude this chapter, we will note that in the nineties there has been a shift in focus towards a rigorous empirical foundation for discourse processing work. The general methodology that has emerged comprises [Walker and Moore1997]:

- Development and evaluation of coding schemes. Coding schemes are used to annotate language corpora for features deemed likely to affect the phenomena under study, e.g., correlates of discourse segments, minimality of referential expressions with respect to providing distinguishing descriptions, etc. A necessary condition for a coding scheme to be useful is that it is reliable, namely, that two or more independent coders can use that coding scheme to annotate the same text in a "similar enough" way. Much interest has thus been devoted to measures of intercoder agreement [Carletta1996; Di Eugenio2000].
- Extraction of information from the annotated corpus. Researchers use either statistical measures or data mining tools on the annotated features [Di Eugenio, Moore, and Paolucci1997; Poesio and Vieira1998; Samuel, Carberry, and Vijay-Shanker1998; Jordan2000]. The purpose is to verify hypotheses (e.g., in naturally occurring texts, do speakers use minimal distinguishing descriptions?), and to find linguistic correlates of higher-level phenomena, such as intonation patterns and adverbs for discourse segmentation.
- Development of computational frameworks based on the information extracted from the corpus. For example, the result of an annotation for referring expressions is used to inform algorithms to generate referring expressions [Poesio and Vieira1998; Jordan2000].
- Evaluation. The computational models developed either theoretically or on the basis of corpus analysis need to be evaluated. This has motivated much interest in evaluation methodologies for computational

models and implemented systems, e.g. [Walker et al.1997]. However, it is still too early to report specific results that pinpoint which techniques, models or systems are the most promising. Systematic evaluations have only recently started to be the norm, and there is no standard testbed of problems and phenomena that can be used to make comparisons across systems and techniques.

Acknowledgements

This paper was partially supported by grant N00014-00-1-0640 from the Office of Naval Research, Cognitive, Neural and Biomolecular S&T Division.

References

[Appelt1985] Appelt, Douglas. 1985. Planning English referring expressions. *Artificial Intelligence*, 26(1):1–33. Also in *Readings in Natural Language Processing*, Grosz, Sparck Jones and Webber eds., Morgan-Kaufmann, 1986.

[Austin1962] Austin, John L. 1962. *How to Do Things With Words.* Oxford University Press, Oxford.

[Carberry and Lambert1999] Carberry, Sandra and Lynn Lambert. 1999. A process model for recognizing communicative acts and modeling negotiation subdialogues. *Computational Linguistics*, 25(1):1–53.

[Carletta1996] Carletta, Jean. 1996. Assessing agreement on classification tasks: the Kappa statistic. *Computational Lingustics*, 22(2):249–254.

[Dale1992] Dale, Robert. 1992. *Generating Referring Expressions.* ACL-MIT Series in Natural Language Processing. The MIT Press.

[Dale and Reiter1995] Dale, Robert and Ehud Reiter. 1995. Computational interpretations of the gricean maxims in the generation of referring expressions. *Cognitive Science*, 18:233–263.

[Di Eugenio2000] Di Eugenio, Barbara. 2000. On the usage of Kappa to evaluate agreement on coding tasks. In *LREC2000, Proceedings of the Second International Conference on Language Resources and Evaluation*, pages 441–444, Athens, Greece.

[Di Eugenio, Moore, and Paolucci1997] Di Eugenio, Barbara, Johanna D. Moore, and Massimo Paolucci. 1997. Learning features that predict cue usage. In *ACL-EACL97, Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, pages 80–87, Madrid, Spain.

[Grosz, Joshi, and Weinstein1995] Grosz, Barbara, Aravind Joshi, and Scott Weinstein. 1995. Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, 21(2):203–225.

[Grosz and Sidner1986] Grosz, Barbara and Candace Sidner. 1986. Attention, Intentions, and the Structure of Discourse. *Computational Linguistics*, 12:175–204.

[Hobbs et al.1993] Hobbs, Jerry, Mark Stickel, Douglas Appelt, and Paul Martin. 1993. Interpretation as Abduction. *Artificial Intelligence, Special Volume on Natural Language Processing*, 63(1–2):69–142.

[Hobbs1979] Hobbs, Jerry R. 1979. Coherence and co-reference. *Cognitive Science*, 3(1):67–82.

[Jordan2000] Jordan, Pamela W. 2000. *Intentional Influences on Object Redescriptions in Dialogue: Evidence from an Empirical Study.* Ph.D. thesis, Intelligent Systems Program, University of Pittsburgh.

[Lascarides and Asher1993] Lascarides, Alex and Nicholas Asher. 1993. Temporal interpretation, discourse relations, and commonsense entailment. *Linguistics and Philosophy*, 16(5).

[Litman and Allen1990] Litman, Diane and James Allen. 1990. Discourse Processing and Commonsense Plans. In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication*. MIT Press, pages 365–388.

[Lochbaum1998] Lochbaum, Karen E. 1998. A collaborative planning model of intentional structure. *Computational Linguistics*, 24(4):525–572.

[Mann and Thompson1988] Mann, William C. and Sandra Thompson. 1988. Rhetorical Structure Theory: toward a Functional Theory of Text Organization. *Text*, 8(3):243–281.

[McKeown1985] McKeown, Kathleen R. 1985. *Text generation. Using discourse strategies and focus constraints to generate natural language text.* Cambridge University Press.

[Moore and Paris1993] Moore, Johanna D. and Cécile L. Paris. 1993. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, 19(4):651–695.

[Moser and Moore1996] Moser, Megan and Johanna D. Moore. 1996. Towards a synthesis of two accounts of discourse structure. *Computational Linguistics*, 22(3):409–419.

[Myers, Shinjo, and Duffy1987] Myers, J. L., M. Shinjo, and S. A. Duffy. 1987. Degree of causal relatedness and memory. *Journal of Verbal Learning and Verbal Behavior*, 26:453–465.

[Perrault and Allen1980] Perrault, Raymond and James Allen. 1980. A Plan-Based Analysis of Indirect Speech-Acts. *American Journal of Computational Linguistics*, 6:167–182.

[Poesio and Vieira1998] Poesio, Massimo and Renata Vieira. 1998. A corpus-based investigation of definite description use. *Computational Lingustics*, 24(2):183–216.

[Prince1981] Prince, Ellen. 1981. Toward a Taxonomy of Given-New Information. In P. Cole, editor, *Radical Pragmatics*. Academic Press, pages 223–255.

[Qu et al.1997] Qu, Yan, Barbara Di Eugenio, Alon Lavie, Lori Levin, and Carolyn Penstein Rosé. 1997. Minimizing cumulative error in discourse context. In Elisabeth Maier, Marion Mast, and Susann LuperFoy, editors, *Dialogue Processing in Spoken Language Systems*, Lecture Notes in Artificial Intelligence. Springer Verlag.

[Reithinger and Maier1995] Reithinger, Norbert and Elisabeth Maier. 1995. Utilizing statistical dialogue act processing in Verbmobil. In *ACL95, Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics.*

[Samuel, Carberry, and Vijay-Shanker1998] Samuel, Ken, Sandra Carberry, and K. Vijay-Shanker. 1998. Dialogue act tagging with transformation-based learning. In *ACL/COLING 98, Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics* (joint with the 17th International Conference on Computational Linguistics), pages 1150–1156.

[Searle1975] Searle, John R. 1975. Indirect Speech Acts. In P. Cole and J.L. Morgan, editors, *Syntax and Semantics 3. Speech Acts.* Academic Press. Reprinted in *Pragmatics. A Reader*, Steven Davis editor, Oxford University Press, 1991.

[Walker, Joshi, and Prince1998] Walker, Marilyn, Aravind Joshi, and Ellen Prince, editors. 1998. *Centering Theory in Discourse.* Oxford University Press.

[Walker et al.1997] Walker, Marilyn A., Diane J. Litman, Candace A. Kamm, and Alicia Abella. 1997. PARADISE: A Framework for Evaluating Spoken Dialogue Agents. In *ACL-EACL97, Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, pages 271–280.

[Walker and Moore1997] Walker, Marilyn A. and Johanna D. Moore. 1997. Empirical studies in discourse. *Computational Linguistics*, 23(1):1–12. Special Issue on Empirical Studies in Discourse.

Glossary

Discourse Relation: Domain or rhetorical relation that links utterances or segments one to the other
Discourse Segment: A group of locally coherent utterances
Planning algorithms: Algorithms that build a plan, i.e., a partially ordered sequence of actions, to achieve a goal s_g
Referential Expression: Any expression that can be used to refer to a discourse entity
Speech Act: Action, such as request, promise, etc, performed by a speaker when producing an utterance