Generating Presupposition-Sensitive Discourse Connectors in Dialogue

When Discourse Structure is not Enough for Generating Connectors in Dialogue

There are known situations in discourse when the rhetorical relations between two text spans (or utterances, in a conversational setting) are not sufficient for predicting the appropriate surface connector between them (e.g., a punctuation mark or a discourse connector). Consider, for instance: 'If you went to school today, *then* you learned the multiplication table.', versus 'You went to school today, *so* you learned the multiplication table.' In these examples, *then* and *so* are not interchangeable, although they both mark a *Consequence* rhetorical relation – e.g. in SDRT (Segmented Discourse Representation Theory (Asher, 1993)) between the two text spans (Jayez, 2004). Hence, the sole rhetorical information does not suffice for predicting the adequate connector. Starting from these observations, Jayez (2004) showed that some connectors are sensitive to *presuppositions*, i.e., in a language production perspective, they indicate that their utterer wants to convey an *informative* presupposition to its addressee. However, this idea does not suffice either for explaining the two examples above, because the two connectors are both a hint that their utterer presupposes a consequence relation between the two text spans. This is why Jayez (2004) puts forth the notion of *pedigrees* for explaining such a behavior: essentially, a pedigree is a *trace* of the illocutionary contribution of the utterance placed before the connector. In the first example above, the pedigree of the first utterance is of being **hypothetical** (hyp), whereas in the second example, the pedigree of the first utterance is of being **taken for granted** (gr). Then, Jayez (2004) shows that *so* requires that the pedigrees of the two utterances (that are in a presupposed *Consequence* relation) be the same, whereas this is not the case for *then*. By using this concept, Jayez (2004) manages to predict adequately the selective usage of different connectors, such as *then*, *so in this case* and *indeed*. We will give a more formally precise account on Jayez (2004)'s findings, extending them to other connectors as well, in the context of language production in dialogue.

Basic Model

To keep our model general, we assume that we have 2 dialogue participants, L_1 and L_2, and that L_1 addresses to L_2 a speech act composed of two utterances, χ and ψ. We denote this by $L_1 \rightarrow_{L_2} \chi * \psi$, where "$*$" stands for the appropriate connector between χ and ψ. If, accessing the logical forms $K(\chi)$ and $K(\psi)$ for χ and ψ respectively, the rhetorical relation $\rho(\chi, \psi)$ is inferable, then we assume that $\rho \Rightarrow \Gamma$, where $\Gamma = \{c_1, ..., c_n\}$ is a set of potential connectors between χ and ψ. In a dialogue setting, we replace Jayez (2004)'s notion of worlds, with that of *commitment store* of each dialogue participant (Lascarides and Asher, 2009), which contains the facts that a dialogue participant has committed to in conversation: for agent L_i, we denote by CS_{L_i} its commitment store. This structure is updated as in (Lascarides and Asher, 2009), with the logical forms of the (SDRT) discourse structures that correspond to the dialogue acts: for example, in a dialogue scenario, if L_1 utters 'If you went to school today$^{\pi_1}$, then you learned what happened to the principal$^{\pi_2}$', then L_2 may infer that $@_{CS_{L_1}}(Consequence(\pi_1, \pi_2))$, where π_1 and π_2 are utterance labels. In the context of Jayez (2004)'s presupposition-based model, if $L_1 \rightarrow_{L_2} \chi * \psi$, and L_2 infers that there is a rhetorical relation $\rho(\chi, \psi)$, then if $CS_{L_1} \models (\chi \wedge \psi)$, i.e., the commitment store of L_1 entails that $\chi \wedge \psi$, and if L_1 wants, by means of $\chi * \psi$, to make L_2 aware that she is committed to a presupposition ϕ (we denote this by using the "at" operator in hybrid logic (Blackburn, 2001): $@_{CS_{L_1}}(\phi)$), then L_1 will choose one connector $c_i \in \Gamma$ such that $(L_1 \rightarrow_{L_2} \chi c_i \psi) \models @_{CS_{L_1}}(\phi)$, so that L_2 can update its commitment store with this information: $@_{CS_{L_2}}(@_{CS_{L_1}}(\phi))$.

Refinements for Dialogue Settings

In order to further refine this model in the context of utterance production in dialogue, we first need to formally define what a presupposition is. In line with Jayez (2004), we assume that, according to L_i, ψ presupposes ϕ iff: if CS_{L_i} admits ψ, then CS_{L_i} accepts ϕ; in other words, if updating CS_{L_i} with ψ does not result in an empty set of facts in CS_{L_i}, then no fact needs to be removed from CS_{L_i} when ϕ is added in CS_{L_i}: denoting by $\oplus$ the commitment stores updating operation, $(CS_{L_i} \oplus \psi \neq \emptyset) \Rightarrow CS_{L_i} \oplus \phi = CS_{L_i}$. This update operation generalizes Jayez (2004)'s definition: for a clause ϕ, L_i maintains her/his commitment store consistent, hence she/he withdraws certain commitments γ in CS_{L_i} such that $\gamma \wedge \phi = \perp$ (logical contradiction). In general, in set-theoretic terms, $\forall \phi \; CS_{L_i} \oplus \phi \subseteq CS_{L_i}$. If successful (i.e., CS_{L_i} at least admits ϕ), such an update yields that $CS_{L_i} \models \phi$ (ϕ is entailed by CS_{L_i}). However, this does not mean, in our view, that ϕ *is* in CS_{L_i}, i.e., in semantic terms: $@_{CS_{L_i}}(\phi)$. This is why Jayez's update operation should not be confused with the commitment store update operation, as defined by Lascarides and Asher (2009): first $CS_{L_i} \oplus \phi$ is performed, then all γ such that $\gamma \wedge \pi = \perp$ are withdrawn from CS_{L_i} and, finally, if $CS_{L_i} \neq \emptyset$, then ϕ is added in CS_{L_i}; this is marked by a set union operation, $CS_{L_i} \cup \{\phi\}$. Another important nuance, not caught by Jayez's account, concerns the distinction between presupposition – L_i already believed ϕ by the moment she/he utters $\chi * \psi$, and *conventional implicature* (Potts, 2003) – L_i believes ϕ at the moment she/he utters $\chi * \psi$. This is important, because this model needs L_i to presuppose, not to conventionally implicate, ϕ, in order for an appropriate discourse marker c_i to be chosen between χ and ψ. For this distinction, we perform a temporal reasoning on the assertion and truth times for utterances: we denote by $[@^t_{CS_{L_i}}(\phi)]@\tau$ the fact that at the utterance time τ it is true that $@_{CS_{L_i}}(\phi)$ was true at time t. Hence, ψ, uttered at time τ, presupposes ϕ if $\exists t < \tau : \forall \theta \in [t; \tau] \; [@^\theta_{CS_{L_i}}(\phi)]@\tau$. Conventional implicatures are obtained for $\theta = \tau$, being thus a special case of presupposition, and not vice versa, as proposed by Jayez (2004).

Prospects: (S)DRT-based connector generation procedure

In the full paper, we will provide a formal account on the production of several connectors, including concessives (*but*, *yet*, *although*, etc.), along the lines of the present treatment, by using, when needed, pedigrees attached to the DRSs (Asher, 1993), in a richer rhetorical and argumentative context (Moeschler, 1989). For example, *yet* triggers a presupposition of the form $\overline{K} = [\Xi : \mathit{Correction}(\Xi, {}^{\mathbf{gr}}K_\psi)]$, where $\overline{K}$ is a handle of a presupposed DRS (Asher, 1993), ${}^{\mathbf{gr}}K_\psi$ is the handle for the DRS corresponding to utterance ψ with pedigree **gr**, when $L_i \rightarrow_{L_j} \chi * \psi$, and Ξ is a non-empty list of DRSs such that at least one is of pedigree **gr**. However, *although* does not impose this constraint on the presupposed content, although the same *Correction* rhetorical relation is involved; instead, at least one DRS in Ξ refers to an action anterior to the one reported in ψ, regardless of the pedigrees involved. The case of *but* is even more interesting, given its two usages (Moeschler, 1989): *direct* – 'The wheather is fine, *but* I have work to do [hence, I am not coming]', and *indirect* – 'The wheather is fine, *but* [I have work to do, hence] I am not coming.' (the utterance fragments in square brackets are not overtly produced).

Main References Asher, N. (1993) *Reference to Abstract Objects in Discourse*. ◆ Blackburn, P. (2001) Representation, reasoning and relational structures: A hybrid logic manifesto. *Logic Journ. IGPL* 8: 339-365. ◆ Jayez, J. (2004) Presuppositions and pedigrees for discourse markers. *Papers at CCSP 2003*: 89-110. ◆ Lascarides, A. and Asher, N. (2009) Agreement, Disputes and Commitments in Dialogue. *Journ. Semantics* 26(2): 88-128. ◆ Moeschler, J. (1989) *Modélisation du dialogue*. ◆ Potts, C. (2003) *The Logic of Conventional Implicatures*. Ph. D., U. C. Santa Cruz.