

More or less Elaborate

Peter Kühnlein

Center for Language and Cognition
University of Groningen
PO Box 716, NL-9700AS Groningen

Abstract

The notion of structural scalarity of subordinating relations is introduced and semantic and pragmatic consequences pointed out. A measure for relevance of elaborations, as a special case of subordinating relation, is suggested.

Discourse structure and scalarity

Theorists (Grosz & Sidner 1986; Mann & Thompson 1988; Asher & Lascarides 2003) agree that coherent discourse is representable by a connected graph: for every vertex in a graph representing a discourse unit, there is a path connecting it to every other vertex in the representation. Moreover, these graphs are usually taken to be hierarchical in that two types of relations hold: coordinating and subordinating. Besides these general agreements, there are stark differences in the assumptions on graph representations. (E.g., only in *Rhetorical Structure Theory* (RST, (Mann & Thompson 1988)), the edges are assumed to be asymmetric in that they represent relations connecting a *satellite* to a *nucleus*.) Recently, it has been suggested to consider the assignment of relations between discourse units not to be an all-or-nothing decision. The difference in acceptability of examples like

(1) John loves sport. He hates football.

(2) John loves sport. But he hates football.

(taken from (Asher & Lascarides 2003, 169)) are said to be indicative for a stronger Contrast relation in example (2) as compared to (1). This is a case of what I call qualitative scalarity. Until now, scalarity has always been discussed in the context of co-ordinating discourse relations (like Contrast or Narration); in the present paper, I will shift the view and include sub-ordinating relations, my prime example being Elaboration.

Copyright © 2008, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Structural scalarity

Consider the following example, translated from a dutch newspaper¹

[¹According to the police, there was also some work for first aid teams and ambulance personnel.][²Six visitors had to be brought to the hospital because they incurred small injuries.][³Hurt knees, burns, that sort of things.][⁴Five others could be treated immediately.][⁵Blister and hyperventilation were at issue.]

I suggest to analyse the example as follows: unit 1 is elaborated by unit 2, which in turn is elaborated by unit 3; unit 4 is coordinated to unit 2, and further elaborated by unit 5. For a start, I follow (Asher & Vieu 2005, 595) and adopt the principle *Continuing Discourse Patterns* (CDP): If $R_1(\alpha, \beta)$ and $R_2(\beta, \gamma)$, and Subord(R_1) and Coord(R_2), then $R_1(\alpha, \gamma)$ and Continuation(β, γ) (in addition to $R_2(\beta, \gamma)$). From this principle it follows that unit 4, just like unit 2, elaborates unit 1.

I now suggest to drop the Continuation relations and generalize CDP to Structural Scalarity: If $R_1(\alpha, \beta, n)$ and $R_2(\beta, \gamma, 0)$, and Subord(R_1) and

- either Coord(R_2), then $R_1(\alpha, \gamma, n)$
- or Subord(R_2), then $R_1(\alpha, \gamma, n + 1)$

(in addition to $R_2(\beta, \gamma, 0)$). In case of a co-ordinating relation between β and γ , the relation to a superordinated unit α is just the same, as stipulated by CDP. The extension I propose captures the intuition from (Grosz & Sidner 1986) that subordinated (purposes of) discourse units serve to achieve the purposes of dominating units: the “inherited” relation $R_1(\alpha, \gamma, n + 1)$ if Subord(R_2) in the above definition is, as I want to put it, a *structurally weaker* version of R_1 as holding between α and β .

Structural scalarity is thus a syntactic measure for the distance between a node in a discourse representation and the

¹Source: Algemeen Dagblad 20 Aug 2007, www.ad.nl/rotterdam/stad/article1596957.ece

nodes “below” that one, connected with the claim that the subordinate nodes behave in a homogenous way wrt to the unit they are subordinate to. In order to preserve treeness of the representation, the subordinating relations implied by Structural Sclarity will not be represented as additional edges. (And they don’t have to, since they are implied anyway.) Let’s see what structural sclarity can be used for, and if it can be tied back to qualitative sclarity.

Forests out of trees — semantics

Suppose τ is a subtree from a representation of a discourse analysis conforming to Structural Sclarity: the top node in τ is attached to subordinate nodes via some scalar subordinating relation. (The example from the newspaper given above is just such an example, with Elaboration(1,2,0), Elaboration(1,3,1), Elaboration(1,4,0) and Elaboration(1,5,1); the fact that additionally Elaboration(2,3,0) and Elaboration(4,5,0) hold is just a property of the example chosen.). Depending on the rules which the theory of discourse relations involved employs, a forest of subtrees can be constructed from τ by (recursively) dropping terminal branches consisting of at least one leaf and an edge. In some cases, e.g., conditionals, the rules will not license dropping just one vertex, but will require dropping two: dropping an antecedent of a conditional and keeping the consequent will not be licensed. Likewise, dropping a correcting unit and keeping the corrected one will be ruled out.

Let \sqsubset denote subtreehood; i.e., $\tau_1 \sqsubset \tau_2$ just in case τ_2 contains some $R(1, \beta, n)$ which τ_1 doesn’t contain, but not vice versa. (And observe that the construction algorithm has to be in accordance to the theory of relations involved.) Except for the case Correction($\alpha, \beta, 0$), it will then be the case that τ_2 will be satisfied by less models M than τ_1 . So, if $\mathcal{M}_1 = \{M | M \models \tau_1\}$ and $\mathcal{M}_2 = \{M | M \models \tau_2\}$, it will generally (except for corrections) be the case that $\tau_1 \sqsubset \tau_2$ iff $\mathcal{M}_2 \subset \mathcal{M}_1$. Corrections are characterized by the fact that they cancel information, so they are an exception to the growth of informativity with complexity of subtrees.

For some applications it might be desirable to have just a measure of informativity of subtrees like that; it will not be so for other applications. E.g., for summarization tasks, it will not be the right strategy to choose the most informative subtree, because this will (corrections notwithstanding) always be the most complex subtree, i.e. τ . Rather, what one would be interested in here would be the most relevant subtree. I will now suggest a strategy for selecting the most relevant subtree(s) from the forest, thus returning to a quality measure.

Selecting trees from forests — relevance

That mere increase in informativity does not automatically mean an increase in relevance was already acknowledged in (Sperber & Wilson 1986). There, the authors claim that on the one hand to be informative for an action means to be rel-

evant, but on the other hand contend that processing effort counteracts that tendency: an increase in processing effort will reduce subjective relevance. Unfortunately, processing effort depends on many subjective factors and doesn’t lend itself easily to formalization.

(Rooy 2005), in reconstructing *Bi-Directional Optimality Theory* as formulated in (Blutner 2000), gives an elegant game theoretical definition of relevance. According to this definition, $R(A) > R(B)$ iff (i) $AV(A) > AV(B)$, or (ii) $AV(A) = AV(B)$ and $inf(A) < inf(B)$, where R returns the relevance, AV expresses the argumentative value, and inf is a measure for informativity. Argumentative value $AV(B)$ in turn is defined as $P(h/B) - P(h)$, i.e., the increase in acceptability of a proposition h that is due to another (set of) proposition(s) B . For the present paper, I will substitute “(set of) proposition(s)” by “subtree”. The definition for argumentative value then reads $AV(\tau_i) = P(1/\tau_i) - P(1)$, where “1” is short for the top-most node. Relevance accordingly is defined as $R(\tau_1) > R(\tau_2)$ iff (i) $AV(\tau_1) > AV(\tau_2)$, or (ii) $AV(\tau_1) = AV(\tau_2)$ and $inf(\tau_1) < inf(\tau_2)$. We end up with a semi lattice of subtrees ordered by their relevance, as desired.

Outlook & Acknowledgements

The account given will be tested for specific theories of discourse relations, and empirically evaluated on a corpus of dutch texts. The possibility to compare the relevance of conflicting analyses of stretches of discourse has to be explored. I want to thank Gisela Redeker and Markus Egg from CLCG, Univ. Groningen, for continuous discussions on that topic, my colleague Ildikó Berzlanovich for the examples, and the members of the Nijmegen Semantics Colloquium for the opportunity to discuss a(n) too) early version of the present paper. This research is funded by NWO.

References

- Asher, N., and Lascarides, A. 2003. *Logics of Conversation*. Cambridge University Press.
- Asher, N., and Vieu, L. 2005. Subordinating and coordinating discourse relations. *Lingua* 115:591–610.
- Blutner, R. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17:189–216.
- Grosz, B. J., and Sidner, C. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3):175–204.
- Mann, W. C., and Thompson, S. A. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3):243–81.
- Rooy, R. v. 2005. Relevance and Bi-Directional OT. In Zeevat, H., and Blutner, R., eds., *Optimality Theory and Pragmatics*, Studies in Pragmatics, Language and Cognition. Palgrave.
- Sperber, D., and Wilson, D. 1986. *Relevance*. Blackwell.