On Compositionality and Bidirectional Optimization*

Helen de Hoop, 1 Petra Hendriks, 2 and Reinhard Blutner 3

1Centre for Language Studies (CLS), Radboud University Nijmegen, H.deHoop@let.ru.nl
2Center for Language and Cognition Groningen (CLCG), University of Groningen, P.Hendriks@rug.nl
3Institute for Language, Logic and Computation (ILLC), University of Amsterdam, K.R.Blutner@uva.nl

In this paper we revisit the semantic principle of compositionality and argue that compositionality is bidirectional optimization. Underspecification approaches to natural language interpretation generally start with an underspecified or weak meaning, which is strengthened by contextual information. By contrast, the bidirectional optimization approach we advocate proceeds from the strongest possible meaning. This meaning can be changed or weakened by contextual information. Under this approach, the meaning of an utterance is composed in a functional rather than a concatenative way, while contextual sources of information play a major role. Yet, because the context of any utterance is in principle the same for the speaker and the hearer, composition and decomposition proceed hand in hand. Hence, bidirectional optimization ultimately guarantees (functional) compositionality.

Keywords: Bidirection, underspecification, weak compositionality, contextuality principle, concatenative compositionality, functional compositionality, optimality theory, recoverability.


* We thank three anonymous reviewers and Chungmin Lee for their stimulating comments and questions. Helen de Hoop and Petra Hendriks gratefully acknowledge the Netherlands Organisation for Scientific Research (NWO) for financial support of their research.
1. The Principle of Compositionality

One of the key principles in formal approaches to natural language interpretation is the principle of compositionality, which expresses the idea that the meaning of a complex expression can be derived from the meanings of its parts and the way these parts are syntactically linked. That is, a syntactic structure can be mapped onto a semantic structure in which the meanings of the elements of that structure are inserted. This principle of compositionality guarantees that syntax and semantics go hand in hand as each syntactic combinatorial rule is associated with a semantic one. For example, the meaning of a determiner is defined as a two-place relation between two sets of individuals. In the sentence *Most cats are black*, the set of cats (C) is thus related to the set of black individuals (B) by the determiner *most*. The sentence *Most cats are black* is true if and only if the cardinality of the intersection of these two sets exceeds the cardinality of their difference. That is, the sentence is true if there are more cats which are black than cats which are not black. In a formula: $\text{most CB}$ is true iff $|C \cap B| > |C-B|$. This semantic rule can be generalized to all sentences of the form *Most AB* where A and B represent the two sets of individuals related by the determiner *most*. Where do these sets of individuals come from? In accordance with the semantic principle of compositionality, they are assumed to come with the syntactic structure of these sentences. In sentences of the form *Most NP VP* (where NP stands for ‘noun phrase’ and VP for ‘verb phrase’, the predicate), the NP refers to set A (also called the domain of quantification) and the VP to set B. The sentence *Most black cats have some sort of white spot on them* is syntactically decomposed as *Most [black cats]NP [have some sort of white spot on them]VP* and thus set A is the set of black cats, while set B is the set of individuals that have some sort of white spot on them, and semantically the sentence is true if and only if there are more black cats which have some sort of white spot on them (set $A \cap B$) than black cats which do not have some sort of white spot on them (set $A-B$).