

MONTAGUE REDUCTION, CONFIRMATION, AND THE SYNTAX-SEMANTICS RELATION

KRISTINA LIEFKE[†] AND STEPHAN HARTMANN[‡]

ABSTRACT. Intertheoretic relations are an important topic in the philosophy of science. However, since their classical discussion by Ernest Nagel, such relations have mostly been restricted to relations between pairs of theories in the natural sciences. In this paper, we present a model of a new type of intertheoretic relation, called *Montague Reduction*, which is inspired by Montague’s framework for the analysis of the linguistic syntax-semantics relation. To motivate the adoption of our new model, we show that this model extends the scope of application of the Nagelian (or related) models, and that it shares the epistemological advantages of the Nagelian model. The latter is achieved in a Bayesian framework.

Keywords Intertheoretic relations, Bayesian confirmation, Montague Grammar, Philosophy of linguistics, Reduction, Syntax-semantics relation.

1. INTRODUCTION

Epistemic relations between pairs of co-existing theories are an important topic in the philosophy of science. These relations involve a connection between the laws (or ‘propositions’) of two related theories – typically, the derivability of knowledge about the empirical domain of *one* of the two theories from knowledge about the domain of the *other* theory, cf. (Nagel, 1961). Historically, most examples of epistemic intertheoretic relations are taken from physics. They include the relation between chemistry and atomic physics, between rigid body mechanics and particle mechanics, and between thermodynamics and statistical mechanics. In the last thirty years, intertheoretic relations have also received increasing interest from other disciplines like biology (see e.g. (Schaffner, 1974; Weber, 2005)), ecology (see e.g. (Levins and Lewontin, 1980)), neuroscience (see e.g. (Bickle, 2006; Schouten and de Jong, 2012)), and economics (see e.g. (Hoover, 2010)).¹ However, this surge of interest has not been shared by linguistics.

The scarcity of work on linguistic intertheory relations cannot be attributed to linguists’ general disinterest in methodology. To the contrary: the availability of large computerized text corpora and the possibility of statistically probing and manipulating linguistic data sets have lately effected a boost of interest in linguistic methodology (see the recent textbooks (Litosseliti, 2010), (Rasinger, 2013), and (Podesva and Sharma, 2013)). The above-described scarcity can then only be

[†]CONTACT INFORMATION: KRISTINA LIEFKE, MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY, LUDWIG-MAXIMILIANS-UNIVERSITÄT MUNICH, GERMANY, K.LIEFKE@LMU.DE.

[‡]CONTACT INFORMATION: STEPHAN HARTMANN, MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY, LUDWIG-MAXIMILIANS-UNIVERSITÄT MUNICH, GERMANY, S.HARTMANN@LMU.DE.

¹The investigated theory-pairs include classical genetics and molecular genetics (or biochemistry), ecology and molecular biology, psychology and computational neuroscience, and macroeconomics and microeconomics.

explained by the fact that linguists' reasoning about intertheory relations presupposes the familiar models of these relations (assuming that the familiar models have a satisfactory fit with relations between theories in linguistics), or by the fact that linguists have little interest in developing or explicating² new models of these relations (assuming that the familiar models do *not* have a satisfactory fit with relations between theories in linguistics).

In this paper, we focus on a particular type of linguistic inter-theory relation which resists an analysis through the familiar models. This type is instantiated by the relation between Montague's theories of natural language syntax and semantics. Since the conjunction of these two theories is typically called *Montague Grammar* (Partee, 1973), cf. (Montague, 1970a; 1970b; 1973), we dub their relation *Montague Reduction*. The poor fit of the familiar models of intertheoretic relations with the syntax-semantics relation is due to the fact that Montague's theories of syntax and semantics describe the behavior of *different* target domains, and that the familiar models of intertheoretic relations are restricted to theories with the *same* (or largely overlapping) target domains. We expect that the development of our model of Montague Reduction will show the importance of investigating inter-theory relations in linguistics, that it will compensate for the absence of models for (one type of) these relations, and that it will yield new insight into the spectrum of intertheoretic relations.

Montague Reduction is related to the best-studied intertheoretic relation, Nagelian reduction³, cf. (Nagel, 1961), and to other undirected dependency relations by family resemblance. Like Nagelian reduction, Montague Reduction aims to derive a proposition of the reduced theory (here, Montague syntax, or *categorial grammar*, cf. (Ajdukiewicz, 1935)) from a proposition of the reducing theory (here, Montague's *model-theoretic semantics*, cf. (Tarski, 1933; Church, 1940)). As a result, our new type of intertheoretic relation shares the rationale of Nagelian reduction: The reduction of syntax to semantics promotes cognitive economy and simplicity, explains the success of the reduced theory in terms of the success of the reducing theory, establishes the theories' relative consistency, and effects a mutual flow of confirmation between the two theories.

This paper argues for the introduction of our model of Montague Reduction. To do this, we show that Montague Reduction shares the epistemological advantages of the Nagelian model of reduction.⁴ The paper is organized as follows: Section 2 presents Montague's formal framework for the analysis of natural language syntax and semantics, and contrasts its associated model with the model of reduction from (Nagel, 1961). To prepare the probabilistic analysis of our new model of reduction, Section 3 reviews the relevant concepts from Bayesian confirmation and network theory. Section 4 motivates the introduction of our model of Montague Reduction. To this aim, we give a Bayesian analysis of the syntax-semantics relation before and after the execution of a Montague Reduction, and

²We will see that there, in fact, *are* models of inter-theory relations in linguistics. Yet, since these models only play a peripheral role in linguistic practice, they are rarely explicitly discussed.

³For many years, Nagelian reduction has been considered a dead end. The present paper rejects this assumption. This stance is motivated by the observation (recorded in (Dizadji-Bahmani et al., 2010)) that Schaffner's (1974) revised model of Nagelian reduction overcomes the problems of Nagel's original model. For the present purposes, it will suffice to focus only on the Nagelian model. We outline a Schaffner-style extension of our model of Montague Reduction in Section 6.

⁴The epistemic advantages of Nagelian reduction are shown in (Dizadji-Bahmani et al., 2011).

show that, post-reduction, the two theories are confirmatory of each other. Section 5 identifies a problem with the generalization of our model of Montague Reduction to pairs of multi-proposition theories, and suggests a Montagovian solution. We close the paper by indicating how Montague Reduction can be incorporated into a sophisticated variant of Schaffner’s (1967) revised model of Nagelian reduction (cf. Sect. 6).

2. MONTAGUE REDUCTION

We first present the two linguistic theories that we aim to relate. Section 2.1 describes the elements of the two theories, and identifies the mechanism which connects these theories. Section 2.2 compares the Montagovian account of the syntax-semantics relation with the model of reduction from (Nagel, 1961). To allow a Bayesian analysis of our new type of intertheoretic relation, Section 2.3 identifies Montague’s rules for the formation of complex syntactic and semantic structures with the objects of probabilistic evaluations.

2.1. Montague’s ‘Two Theories’ Theory. Montague’s framework for the analysis and interpretation of natural language syntax⁵ constitutes a milestone in the understanding of linguistic syntax-semantics relations. At the end of the 1960s, there did not exist a sufficiently well-developed formal semantic theory which could be used for the systematic interpretation of natural language.⁶ Montague (1970b), cf. (Montague, 1970a; 1973), provides such a theory. The latter is a model of Church’s (1940) lambda calculus, which contains a designated domain of semantic objects for each syntactic category.

To enable the systematic interpretation of natural language, Montague assumes that the semantic objects in the model’s domains are associated with expressions from distinct syntactic categories, and that the model’s rules for the formation of semantic objects are associated with the familiar syntactic rules. He identifies categorial grammar (Ajdukiewicz, 1935), cf. (Moortgat, 1997), as the syntactic theory which best facilitates this correspondence. This theory describes syntax as an algebra over the set of linguistic expressions $\mathcal{E} = \{\mathcal{E}_N, \mathcal{E}_V, \dots\}$ which generates complex expressions (e.g. the sentence *Bill walks* $\in \mathcal{E}_S$) from simpler expressions (e.g. from the proper name *Bill* $\in \mathcal{E}_N$ and the intransitive verb *walk* $\in \mathcal{E}_V$) via syntactic operations like concatenation.

For the combination of proper names with intransitive verbs, the behavior of concatenation is governed by the rule \mathbb{G}_S , cf. (Montague, 1973, rule S4). In this rule, $[AB']$ is the result of concatenating the expressions A and B' (in that order), where A is a singular name and where B' is the result (e.g. *walks*) of replacing the verb B (here, *walk*) by its third person singular present form:

$$\mathbb{G}_S. \text{ If } B \in \mathcal{E}_V \text{ and } A \in \mathcal{E}_N, \text{ then } [AB'] \in \mathcal{E}_S.$$

To facilitate the presentation of Montague Grammar, we limit ourselves to a syntactically poor fragment of English, which only contains proper names, intran-

⁵For an introduction to Montague Grammar, the reader is referred to (Janssen, 2012), (Partee, 1997), and (Gamut, 1991).

⁶Thus, Montague (1970b) writes, “It is clear [...] that [with the exception of (Montague, 1970a)] no adequate and comprehensive semantical theory has yet been constructed” (p. 222). Chomsky (1971) supports this claim by stating, “In the domain of semantics there are [...] problems of fact and principle that have barely been approached, and there is no reasonably concrete or well-defined ‘theory of semantic representation’ to which one can refer.” (p. 183).

sitive verbs, and declarative sentences. As a result, the behavior of concatenation is only governed by the rule \mathbb{G}_s , such that the set of syntactic rules, \mathbb{G} , is identified with the singleton set $\{\mathbb{G}_s\}$. Our fragment is then identified with the closure of the set \mathcal{E} under the rule \mathbb{G}_s . By introducing other concatenation rules⁷, we can easily extend our fragment to syntactically more diverse subsets of English.

We next turn to Montague’s semantic theory: We have noted above that Montague’s models contain a designated semantic domain for each syntactic category. Thus, a model for our small fragment will include a domain of individuals \mathcal{D}_N , a domain of properties of individuals \mathcal{D}_V , and a domain of truth-values \mathcal{D}_s . The interpretation function I relates the domains in \mathcal{E} and $\mathcal{D} = \{\mathcal{D}_N, \mathcal{D}_V, \mathcal{D}_s\}$ by assigning, to each \mathcal{E}_k -expression, c , (where k designates a syntactic category) a model-theoretic object, C , in the semantic domain \mathcal{D}_k , such that $I(c) = C$. In this way, the function I will assign, to the name *Bill*, the individual *Bill* (i.e. λ), and will assign, to the verb *walk*, the property ‘walk’.

From the above interpretations, truth-values (here, the truth-value of the sentence *Bill walks*) are obtained via a semantic correlate, \mathbb{S}_s , of the rule \mathbb{G}_s . In the definition of \mathbb{S}_s , we abbreviate ‘ $I(c)$ ’ as ‘ $\llbracket c \rrbracket$ ’. ‘ $\llbracket B \rrbracket(\llbracket A \rrbracket)$ ’ is interpreted as the functional application of the interpretation of B to the interpretation of A :

\mathbb{S}_s . If $\llbracket B \rrbracket \in \mathcal{D}_V$ and $\llbracket A \rrbracket \in \mathcal{D}_N$, then $\llbracket B \rrbracket(\llbracket A \rrbracket) \in \mathcal{D}_s$.

As a result of the above, the semantics of our fragment constitutes an algebra $\langle \mathcal{D}, \mathbb{S} \rangle$ over the set of basic model-theoretic objects (where $\mathbb{S} = \{\mathbb{S}_s\}$). The described relation between the syntactic and the semantic algebra is captured in Figure 1:

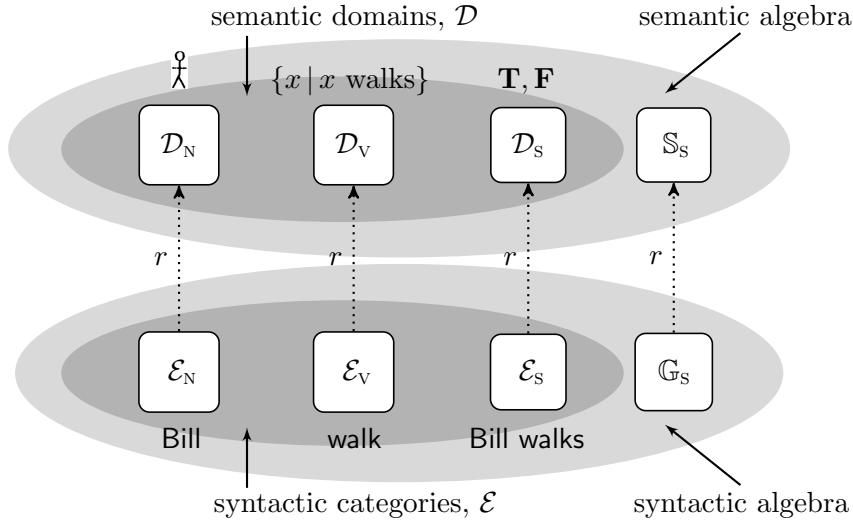


FIGURE 1. The syntax-semantics *map*.

Our previous presentation has suggested the existence of a *homomorphism* between Montague’s syntactic and the semantic algebra. However, in practice⁸, not

⁷These include other rules for the formation of sentences (e.g. the rule S9 from (Montague, 1973)), and rules for the formation of complex expressions from other categories.

⁸Montague’s interpretation of proper names in the domain of individuals, cf. (Montague, ’73, rule T1.(d)), and his interpretation of other noun phrases in the domain of generalized quantifiers over individuals, cf. (Montague, 1973, rule T2), reflect this observation.

every element of the syntactic algebra is associated with a *unique* element of the semantic algebra (s.t. the syntax-semantics relation is not a function). In particular, proper names are semantically ambiguous between objects in the domain, \mathcal{D}_N , of individuals and objects in the domain, \mathcal{D}'_N , of *generalized quantifiers* over individuals (i.e. of second-order properties of individuals). The interpretation of names as generalized quantifiers is required by the interpretation of universally quantified noun phrases (e.g. *every woman*) as generalized quantifiers over individuals, by the possibility of coordinating names with quantifier phrases (cf. the complex noun phrase *Bill and every woman*), and by the restriction of coordination (here, *and*) to expressions which receive an interpretation in the same semantic domain. While some occurrences of the name *Bill* will thus be interpreted as the individual *Bill* (here, $\llbracket \text{Bill} \rrbracket = \overset{\circ}{\lambda}$), others will be interpreted as the property of all of *Bill's* properties (here, $\llbracket \text{Bill} \rrbracket'$)⁹.

To preserve function-argument structure, intransitive verbs (e.g. *walk*) are made ambiguous between first-order properties of individuals ($\llbracket \text{walk} \rrbracket \in \mathcal{D}_V$) and properties of generalized quantifiers over individuals ($\llbracket \text{walk} \rrbracket' \in \mathcal{D}'_V$). To accommodate the obtaining of truth-values through the application of a property of generalized quantifiers to a generalized quantifier, we supplement the semantic rule \mathbb{S}_s by the rule \mathbb{S}'_s . The latter is specified below:

\mathbb{S}'_s . If $\llbracket B \rrbracket \in \mathcal{D}'_V$ and $\llbracket A \rrbracket \in \mathcal{D}'_N$, then $\llbracket B \rrbracket(\llbracket A \rrbracket) \in \mathcal{D}_s$.

Figure 2 illustrates the resulting relation between the elements of the syntactic and the semantic algebra.

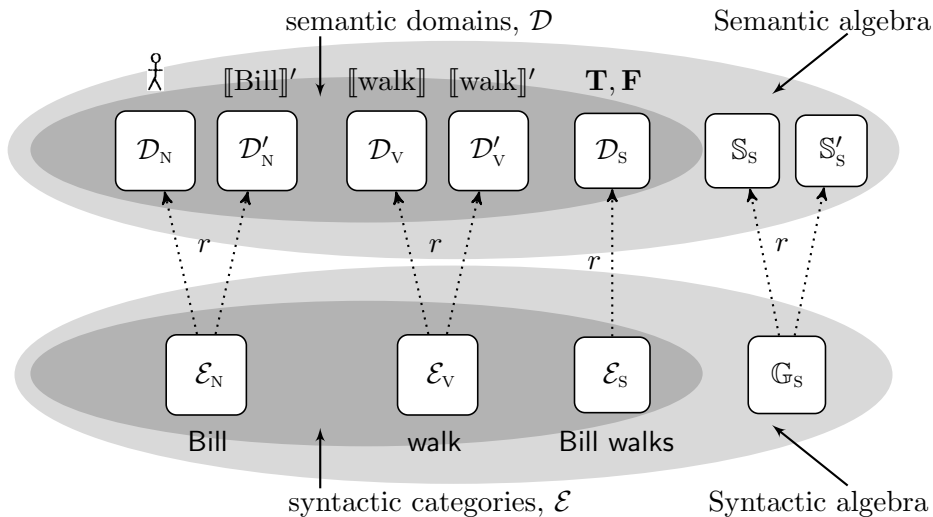


FIGURE 2. The syntax-semantics *relation*.

This completes our presentation of Montague’s account of the linguistic syntax-semantics relation. We next compare this account to the model of intertheoretic reduction from (Nagel, 1961). Since we are interested in *epistemic* intertheoretic relations, we will hereafter focus on the relation between the rules \mathbb{G}_s , \mathbb{S}_s , and \mathbb{S}'_s .

⁹Intuitively, $\llbracket \text{Bill} \rrbracket'$ abbreviates the interpretation of the term $\lambda P.P(\text{bill})$, where P is a variable over first-order properties of individuals, and where *bill* is the individual constant for *Bill*.

2.2. Montague’s Theory and Intertheoretic Reduction. In Section 1, we have described Montague’s model of the syntax-semantics relation as the instantiation of a specific type of intertheoretic relation. To emphasize the similarities and differences of the Montagovian model to the account of reduction from (Nagel, 1961, Ch. 11), we next describe the reduction between two theories on the Nagelian model.¹⁰ We then compare Montague’s account of intertheoretic relations to the Nagelian account.

In the following, we assume a *reduced* (or *phenomenological*) theory, \mathcal{T}_2 , and a *reducing* (or *fundamental*) theory, \mathcal{T}_1 .¹¹ As is well known, Nagelian reduction is a three-step process, which involves the establishment of connections (via bridge laws) between terms in the non-logical vocabulary of the theories \mathcal{T}_1 and \mathcal{T}_2 (step (i)), the substitution of the terms from \mathcal{T}_1 by their bridge-law correspondents from \mathcal{T}_2 (step (ii)), and the derivation (via auxiliary assumptions) of every proposition in \mathcal{T}_2 from the result of replacing the relevant terms in a proposition in \mathcal{T}_1 by their bridge-law correspondents (step (iii)), cf. (Nagel, 1961, pp. 353–354).

Accordingly, the use of the Nagelian model for the reduction of Montague’s syntax-semantics pair requires the formulation of bridge laws connecting the names for the elements in \mathcal{D} and \mathcal{E} and the designators of the syntactic resp. semantic operations in \mathbb{S} and \mathbb{G} (cf. step (i)), the substitution of the names for the elements of \mathcal{D} and the operations in \mathbb{S} by their bridge-law correspondents from \mathcal{E} and \mathbb{G} in the designators of the rules from \mathbb{S} (cf. step (ii)), and the derivation of every rule (or ‘proposition’) in \mathbb{G} from the corresponding proposition in \mathbb{S} (cf. step (iii)).

In particular, step (i) connects \mathcal{D}_N and \mathcal{D}'_N with \mathcal{E}_N , \mathcal{D}_V and \mathcal{D}'_V with \mathcal{E}_V , and \mathcal{D}_S with \mathcal{E}_S , and connects function application, $\lambda y \lambda x. y(x)$, with expression concatenation, $\{ \cdot \}$.¹² Step (ii) converts a copy of the rules \mathbb{S}_S and \mathbb{S}'_S (in (1a), (1b)) into the rule from (1c) by replacing every occurrence of ‘ \mathcal{D}_N ’ (resp. ‘ \mathcal{D}'_N ’) by ‘ \mathcal{E}_N ’, of ‘ \mathcal{D}_V ’ (resp. ‘ \mathcal{D}'_V ’) by ‘ \mathcal{E}_V ’, and of ‘ \mathcal{D}_S ’ by ‘ \mathcal{E}_S ’, and by replacing every occurrence of ‘ $\lambda y \lambda x. y(x)$ ’ by the operator ‘ $\{ \cdot \}$ ’. Step (iii) trivially derives a variant of the rule \mathbb{G}_S from the result of this conversion.

- (1) a. If $Y \in \mathcal{D}_V$ and $X \in \mathcal{D}_N$, then $Y(X) \in \mathcal{D}_S$.
- b. If $Y \in \mathcal{D}'_V$ and $X \in \mathcal{D}'_N$, then $Y(X) \in \mathcal{D}_S$.
- c. If $Y \in \mathcal{E}_V$ and $X \in \mathcal{E}_N$, then $\{YX\} \in \mathcal{E}_S$.

The above example shows that the Montagovian model of the syntax-semantics relation (hereafter, *Montague Reduction*, or MR) and the Nagelian model of intertheoretic reduction (*Nagel Reduction*, or NR) agree with respect to the connectability of the two theories (cf. step (i)). Yet, while Nagel Reduction satisfies this requirement through the formulation of syntactic bridge laws (which connect pairs of *terms* in the vocabulary of the two theories), Montague Reduction satisfies

¹⁰We will see below that – because of the interpretation of bridge laws as co-extensionality relations between \mathcal{T}_1 and \mathcal{T}_2 terms – the Nagelian model of reduction can, in fact, *not* be applied to the syntax-semantics pair. To allow a direct comparison of the Montagovian with the Nagelian model, we temporarily ignore this inapplicability.

¹¹The use of the adjectives *phenomenological* and *fundamental* (which suggests the directionality of the reduction relation) is in accordance with the treatment of theories in physics.

¹²To accommodate the word-order profiles of different languages (e.g. ‘*subject-verb-object*’ vs. ‘*verb-subject-object*’), we here use an order-invariant version, $\{ \cdot \}$, of the concatenation operation $[\cdot]$ from \mathbb{G}_S . The order-(in)variance of concatenation is discussed below and in Section 6.

this requirement through the assumption of a homomorphism $h := r^{-1}$ between the *objects* of the semantic and the syntactic algebra. This homomorphism generalizes the inverse image of the interpretation function I from Section 2.1, such that $h(\mathcal{D}_k) = \{I^{-1}(a) \mid a \in \mathcal{D}_k\} \subseteq \mathcal{E}_k$. Since the homomorphism h also establishes connections between *propositions* of the two theories (s.t. $h(\mathbb{S}_s) = h(\mathbb{S}'_s) = \mathbb{G}_s$), Montague Reduction obviates Nagel's substitution step (ii).

In Montague Reduction, the replacement of Nagelian bridge laws by (the inverse image of) the relation r is made necessary by the definition of bridge laws as co-extensionality relations between the terms in the theories \mathcal{T}_1 and \mathcal{T}_2 .¹³ As a result, it holds for all pairs of terms, t_1 and t_2 , from \mathcal{T}_1 , resp. \mathcal{T}_2 , that t_1 applies to all objects to which t_2 applies, and vice versa. Since model-theoretic semantics and categorial grammar have non-overlapping target domains¹⁴ (s.t. they do not satisfy the co-extensionality requirement on terms), their objects need to be connected in some other way. The assumption of a homomorphism between the algebraic formulations of \mathcal{T}_1 and \mathcal{T}_2 serves this purpose. Interestingly, in Montague's model, the replacement of bridge laws by the homomorphism h also obviates a relation between magnitude parameters, which is required in the Nagelian model, cf. (Schaffner, 1974): To accommodate the magnitudes of physical properties (e.g. the degrees Kelvin of a gas' temperature), Nagelian bridge laws specify the functional dependence relation f between the magnitudes, τ_1 and τ_2 , of the properties denoted by the terms t_1 and t_2 (s.t. $\tau_2 = f(\tau_1)$). Since the properties in the Montagovian theories are magnitude-free, Montague's model does not (need to) specify such a relation.

Notably, the semantic characterization of connectability is not the only salient property of Montague Reduction: Montague Reduction is also defined by the non-injectivity of the homomorphism h (and by the resulting *non-symmetry* of the reduction relation): the Nagelian and the Montagovian model both characterize reduction as a directed dependency relation. (This is reflected in the identification of one of the two theories with the phenomenological theory, and of the other with the fundamental theory, see the next paragraph). However, while the association of the semantic rules \mathbb{S}_s and \mathbb{S}'_s with the syntactic rule \mathbb{G}_s (cf. Sect. 2.1) makes the directedness of the syntax-semantics relation explicit, Nagel's formalization of bridge laws as biconditional statements conceals this property. As a consequence, the Nagelian model of reduction represents this relation as a *symmetric* relation. The Montagovian model represents this relation as an *asymmetric* relation. To emphasize the symmetric character of Nagelian reduction, we will sometimes describe this relation as an *undirected* relation.¹⁵

The identification of the semantics-syntax relation with a *non-injective* homomorphism motivates our identification of model-theoretic semantics with the fundamental theory, \mathcal{T}_1 , and of categorial grammar with the *phenomenological* theory, \mathcal{T}_2 . This identification is further justified by the fact that nearly all relevant

¹³This characterization of Nagelian bridge laws is only due to (Schaffner, 1974, pp. 614–615), cf. (Schaffner, 1993, pp. 411–477). However, since this characterization generalizes Nagel's categorization of bridge laws as meaning equivalences, conventional stipulations, or matters of fact, cf. (Nagel, 1961, pp. 354–355), we here treat it as a proper part of Nagel's model.

¹⁴While categorial grammar accounts for the well-formedness of syntactic structures, model-theoretic semantics accounts for the compositional properties of these structures' interpretations.

¹⁵For a discussion of this issue – and for a Nagelian solution –, the reader is referred to (Kuipers, 1982) and (Dizadji-Bahmani et al., 2010).

evidence¹⁶ directly supports categorial grammar (s.t. syntax fits the phenomena), and by Montague’s view of the primacy of semantics, cf. (Montague, 1970b, p. 223)¹⁷. Figure 3 compares Montague’s description of the syntax-semantics relation (right) with the Nagelian account of reduction (left):

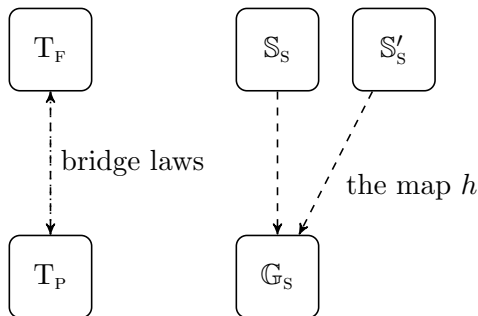


FIGURE 3. The Nagelian model (left) and the Montagovian model of reduction (right).

Note our use of dashed (rather than dotted) arrows in the above figure. This notation is required by the directedness of the syntax-semantics relation, such that the arrows from Figures 2 (resp. 1) and 3 have a different denotation: While the arrows from Figure 2 represent Montague’s relation r , the right arrow from Figure 3 represents its inverse r^{-1} (i.e. the homomorphism h).

For future reference, we define Nagel Reduction and Montague Reduction in terms of their salient properties:

Definition 1 (Nagel Reduction (NR)). A type of *undirected* (i.e. *symmetrically represented*) dependency relation, described in (Nagel, 1961), which is defined by the existence of intertheoretical connections between co-extensional *terms* in the non-logical vocabulary of the two related theories, and by the derivability of every proposition in the phenomenological theory from a corresponding proposition in the fundamental theory.

Definition 2 (Montague Reduction (MR)). A type of *directed* (or *non-symmetric*) dependency relation, implicit in (Montague, 1970b), which is defined by the existence of intertheoretical connections between *objects* of the two related theories, and by the resulting trivial derivability of every proposition in the phenomenological theory from a corresponding proposition in the fundamental theory.

The commonalities and differences between Nagel Reduction and Montague Reduction are captured in Figure 4 (next page).

It is clear from the above that the Montagovian model of the syntax-semantics relation instantiates only one particular type of intertheoretic relation. There are many others, ranging from ‘strict’ Nagelian Reduction (cf. Def. 1) via the ‘weaker’ Nagel-Schaffner reduction (Schaffner, 1967; 1974) (cf. Sect. 6), to undirected dependency relations, cf. (Darden and Maull, 1977; Hartmann, 1999; Mitchell, 2003).

¹⁶This evidence lies in strings of expressions (e.g. the sentence *Bill walks*) whose structure reflects the assumed formation process (here, the process described by the rule \mathbb{G}_S ; see Sect. 2.3).

¹⁷Thus, in (Montague, 1970b, p. 223, fn. 2), Montague writes, “[...] I fail to see any great interest in syntax except as a preliminary to semantics [...]”.

$$\text{Derivability} \left\{ \begin{array}{l} \text{NR} \\ \text{MR} \end{array} \right\} \left\{ \begin{array}{l} \textit{Syntactic} \text{ interth. connectability (via bridge laws),} \\ \textit{Undirected} \text{ dependency} \\ \textit{Semantic} \text{ interth. connectability (via the map } h), \\ \textit{Directed} \text{ dependency} \end{array} \right.$$

FIGURE 4. Montague Reduction vs. Nagel Reduction.

We expect that the relation between model-theoretic semantics and categorial grammar be found in the mid-range of this spectrum.

We close the present subsection with a number of caveats about the syntax-semantics relation: Our previous considerations have identified the Montague Reduction of syntax to semantics as a weak, i.e. directed, variant of Nagel Reduction. Significantly, however, Montague Reduction is even weaker than has been previously established. This is due to the greater structural richness of syntax, such that Montague’s syntactic rule \mathbb{G}_s contains *more* information than (the conjunction of) the semantic rules \mathbb{S}_s and \mathbb{S}'_s . Word order and agreement are a case in point: To ensure the ‘right’ formation of simple English sentences (in which the third person singular form of an intransitive verb combines with a proper name which occurs *on its left*), Montague’s rule \mathbb{G}_s uses the order-sensitive concatenation operation $[\cdot]$. Without this order-specification, nothing would prevent the concatenation of expressions which violate the language’s word-order profile (e.g. *subject-verb-object* for English). As a result, the rule from (1c) could yield either the expression *Walks Bill* or *Bill walks*. This observation motivates our description of Montague Reduction as a distinct type of intertheoretic *relation*, rather than strong Nagelian *reduction*.

Our characterization of the Montagovian syntax-semantics relation as a weak intertheoretic relation requires one further clarification: All popular accounts of reduction (incl. (Nagel, 1961)) assume that the reduced and the reducing theory have the same (or largely overlapping) target domains. On this account, the two theories both make more-or-less the same claims (e.g. about the behavior of a given physical system). We have argued above that this is not the case for our syntax-semantics pair:¹⁸ While categorial grammar accounts for the well- (or ill-)formedness of syntactic structures, model-theoretic semantics accounts for their interpretations’ compositional properties. Admittedly, the interpretation relation I establishes a firm connection between the objects of the two theories. However, this does not change the fact that the ‘reductive achievement’ of Montague Reduction will be comparatively weaker than the achievements of reductions between shared-domain theories.

The admonitions from the last two paragraphs all characterize our new type of intertheoretic relation. While some of them will be ignored in the rest of this paper, their neglect would distort our representation of the syntax-semantics relation. To enable a Bayesian analysis of the model of this relation, the next subsection discusses the use of probabilities in linguistic syntax and semantics. Section 3 gives a primer on Bayesian confirmation and network theory.

¹⁸Notably, relations between theories with different target domains may not be identified with *heterogeneous* reductions, whose constituent theories do not share the relevant predicates: Many heterogeneous reductions (e.g. the reduction of thermodynamics to statistical mechanics) still have a *common* target domain (here: observable phenomena in macroscopic systems like gases).

2.3. Montagovian Rules and Probabilities. Our presentation of Montague’s theory of the syntax-semantics relation has presupposed the existence of two sets of rules, \mathbb{G} and \mathbb{S} , for the formation of complex syntactic and semantic objects. Like hypotheses of any scientific theory, these rules are obtained by the scientific method (discussed, here, for the formulation of \mathbb{G}_s): Following the isolation of syntactically simple sentences in a given data-set (typically, an electronic text collection like the *British National Corpus*), linguists abstract information about the sentences’ structural properties and propose a hypothesis (here, \mathbb{G}_s) about their formation. Hypotheses are tested through the analysis of strings of expressions in other (new) corpora: A given string (e.g. the sentence *Bill walks*) is taken to *support* the hypothesis if its structure does, and to *question* the hypothesis if its structure does not reflect the assumed formation process (i.e. if it positively resp. negatively instantiates \mathbb{G}_s).

To enable a Bayesian analysis of our model of the syntax-semantics relation, we assign a probability to every syntactic and semantic rule. A rule’s probability is informed by the frequentist data which are available at the time. Thus, the probability of the truth of the hypothesized rule \mathbb{G}_s will be very high (or low) if a very large (resp. small) percentage of the expressions of the described form instantiates \mathbb{G}_s . We expect that the frequentist probability of a given rule will influence a linguist’s psychological confidence in the rule’s descriptive adequacy. In particular, if a very large (or small) percentage of the expressions of a given form instantiates \mathbb{G}_s , the linguist’s belief in the truth of \mathbb{G}_s will be similarly high (resp. low).

Our previous considerations have defined the probability of a given rule via the frequency of the rule’s positive instantiations in a given sample. Notably, the relation of direct instantiation by linguistic objects is restricted to *syntactic* rules. We will see below that this observation plays an important role in the two theories’ pre-reductive confirmation (cf. Thm. 1, Sect. 4.1). The semantic rule \mathbb{S}_s derives its support from the linguistic support of \mathbb{G}_s via the assumption of the homomorphism h . The probability of \mathbb{S}_s is thus obtained via the probability of its syntactic counterpart.

This concludes our discussion of the reductive and probabilistic aspects of Montague’s framework. We precede our introduction to Bayesianism with one final caveat: Importantly, our attribution of probabilities to Montagovian rules does *not* constitute a probabilistic extension of Montague Grammar. The central aim of this paper is methodological, not substantive. Consequently, we do not intend any revisions or additions to (our fragment of) Montague’s theory. The attribution of probabilities is only a means to an end, i.e. the possibility of providing a Bayesian analysis of Montague’s model of the syntax-semantics relation. To achieve this end, it will suffice to restrict ourselves to the use of probabilistic variables. While nothing prevents us from inserting actual values, the use of actual probabilistic values is not necessary for the success of our analysis.

3. A PRIMER ON BAYESIANISM

We analyze a rule’s evidential support via Bayesian confirmation theory: The central idea of this theory is the interpretation of confirmation as probability-raising, and the associated distinction between two notions of probability, relative to the receipt of a new piece of evidence: The initial, or *prior*, probability of a proposition H (for ‘hypothesis’) is the probability of H *before* the evidence E has been

considered. The final, or *posterior*, probability of H is the probability *after* E has been considered.

Bayesian conditionalization on E requires an update of the prior probability, $P(H)$, to the posterior probability, $P'(H)$, of H, where $P'(H)$ is typically expressed in terms of the original probability measure, i.e. $P'(H) = P(H|E)$, provided that $P(E) > 0$. Our use of Bayes' Theorem, a result from probability theory, yields the following expression for the posterior probability of H:

$$(1) \quad P(H|E) = \frac{P(E|H)P(H)}{P(E)} = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|\neg H)P(\neg H)}$$

$$= \frac{P(H)}{P(H) + P(\neg H)x}$$

In the above, the expression $x := P(E|\neg H)/P(E|H)$ is the likelihood ratio.

According to Bayesian confirmation theory, a piece of evidence E *confirms* the hypothesis H if the posterior probability of H (given E) is greater than the prior probability of H, i.e. if $P(H|E) > P(H)$. The piece of evidence, E, *disconfirms* H if $P(H|E) < P(H)$, and *is irrelevant for* H if $P(H|E) = P(H)$.¹⁹

While the case of two propositions is easy to compute, the confirmatory situation is often much more complicated. This is due to the fact that the respective hypothesis may have a fine structure, and that different pieces of evidence may stand in certain probabilistic relations to one another. As we will see is due course, the relation between linguistic syntax and semantics, upon which we focus in this paper, exhibits a similarly high degree of complexity.

Bayesian networks prove to be a highly efficient tool for the computation of the above-described scenarios.²⁰ A Bayesian network is a directed acyclical graph whose nodes represent propositional variables and whose arrows encode the conditional independence relations that hold between the variables. In the rest of this paper, we call *parent nodes* nodes with outgoing arrows, and call *child nodes* nodes with incoming arrows. *Root nodes* are unparented nodes; *descendant nodes* are child nodes, or the child of a child node, etc.

By the special choice of graph, paths of arrows may not lead back to themselves (s.t. the graph is acyclical). Variables at each node can take different numerical values, which are assigned by the probability function P . As a result, Bayesian networks do not only provide a direct visualization of the probabilistic dependency relations between variables, but come along with a set of efficient algorithms for the computation of whichever conditional or unconditional probability over a (sub-)set of the variables involved we are interested in.

We illustrate the use of Bayesian networks by framing the confirmatory relation between the hypothesis H and a piece of evidence E. To do so, we first introduce two binary propositional variables, H and E (printed in italic script).

¹⁹Bayesianism is presented and critically discussed in (Howson and Urbach, 2005) and (Earman, 1992). These texts also discuss Jeffrey conditionalization, which is an alternative updating rule. For an introduction to Bayesian epistemology, the reader is referred to (Hájek and Hartmann, 2010) and (Hartmann and Sprenger, 2010).

²⁰For an introduction to Bayesian networks, see (Neapolitan, 2003; Pearl, 1988). The monograph (Bovens and Hartmann, 2003) discusses applications from epistemology and the philosophy of science, and provides a short introduction to the theory of Bayesian networks.

Each of these variables has two values (printed in roman script): H or $\neg H$ (i.e. ‘the hypothesized rule is true’ resp. ‘false’), and E or $\neg E$ (‘the evidence obtains’ resp. ‘does not obtain’). The relation between E and H is represented in the graph in Figure 5:

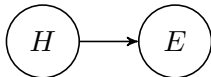


FIGURE 5. Bayesian network representation of the dependence between E and H .

The arrow from H to E denotes a direct influence of the variable in the parent node to the variable in the child node. The truth or falsity of the hypothesis affects the probability of the obtaining of E .

To turn the graph from Figure 5 into a Bayesian network, we further require the marginal probability distribution for each variable in a root node (i.e. the prior probability, $\mathbf{P}(H)$, of H), and the conditional probability distribution for every variable in a child node, given its parents. In the present case, the latter involves fixing the likelihoods $\mathbf{P}(E|H)$ and $\mathbf{P}(E|\neg H)$. From these distributions, we can then obtain all other probabilities via Bayesian networks. As will be relevant below, the graph’s probability distribution respects the Parental Markov Condition (PMC): A variable represented by a node in a Bayesian network is *independent* of all variables represented by its non-descendant nodes in the Bayesian network, and is *conditional* on all variables represented by its parent nodes.

4. MONTAGUE REDUCTION AND CONFIRMATION

Our previous efforts have restricted themselves to the presentation of Montague’s model of the syntax-semantics relation. To motivate the introduction of this model as a new model of intertheoretic relations (in addition to the established Nagelian model), we next provide a Bayesian analysis of this model.

To simplify the Bayesian analysis of our new model of reduction – and to facilitate its comparison with the Bayesian analysis of Nagel-Schaffner reduction from (Dizadji-Bahmani et al., 2011) –, we focus on the relation between the rules \mathbb{G}_s and \mathbb{S}_s .²¹ These rules are associated with the propositional variables G , respectively S . The reductive relation between syntax and semantics can then be represented via the graph in Figure 7 (next page). For simplicity, we assume that the rule G is supported by exactly one (set of) piece(s) of evidence. As has been explained in Section 2.3, we take evidence for a given syntactic rule to be an intuitively well-formed linguistic expression whose structure reflects the rule’s assumed formation process. The replacement of the arrows from Figures 1 to 3 by arrows of the form \longrightarrow (cf. Fig. 5) is motivated by our interest in probabilistic dependence relations between propositional variables (rather than in the relation r or the homomorphism h). Below, these arrows capture the dependence of the probability of the truth of the syntactic rule \mathbb{G}_s on the probability of the truth of the semantic rule \mathbb{S}_s . The conditional dependency of \mathbb{G}_s on \mathbb{S}_s enables us to obtain an aligned chain of arrows. As a result, we can represent a flow of evidence from the syntactic to the semantic theory.

²¹The resulting restriction to singleton sets of rules (s.t. $\mathbb{G} = \{\mathbb{G}_s\}$ and $\mathbb{S} = \{\mathbb{S}_s\}$) enables us to drop the subscript ‘s’ from G_s and S_s .

Figures 6 and 7 display the graphs associated with the dependence relations between S , G , and E before and after the establishment of the relation of Montague Reduction:



FIGURE 6. Pre-reductive dependence relations between S , G , and E .

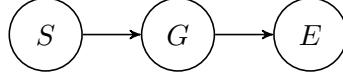


FIGURE 7. Post-reductive dependence relations between S , G , and E .

We determine the confirmation of S and G via their relevant probabilities, beginning with the pre-reductive situation (in Sect. 4.1, cf. Fig. 6). The comparison of this situation with the post-reductive situation (in Sect. 4.3; cf. Sect. 4.2, Fig. 7) will show that the Montague Reduction of syntax to semantics raises the joint (prior and posterior) probabilities of the two theories and improves the flow of confirmation between these two theories.

4.1. Pre-Reductive Confirmation. Let $P_1(S)$ and $P_1(G)$ be the marginal probabilities of the root nodes S and G of the Bayesian network in Figure 6, where P_1 is the relevant probability measure. Let $P_1(E|G)$ and $P_1(E|\neg G)$ be the conditional probabilities of the child node E . For convenience, we use the following abbreviation scheme:

$$(2) \quad \begin{aligned} P_1(S) &= \sigma \quad , \quad P_1(G) = \gamma, \\ P_1(E|G) &= \pi \quad , \quad P_1(E|\neg G) = \rho \end{aligned}$$

We assume a positive confirmatory relation between E and G , such that $\pi > \rho$.

From the network structure in Figure 6, we can read off the conditional and unconditional independences $E \perp\!\!\!\perp S|G$ and $S \perp\!\!\!\perp G$, such that $P_1(S|E) = P_1(S)$. Evidence E does not confirm (or disconfirm) S . Hence, there is no flow of confirmation from the syntactic to the semantic theory. In the absence of the homomorphism $h : \mathbb{S} \rightarrow \mathbb{G}$, the variables S and G are probabilistically independent before the reduction. This fact is captured by equation (3):

$$(3) \quad P_1(S, G) = P_1(S) P_1(G) = \gamma \sigma$$

By (3), the prior probability of the conjunction of S and G equals the product of the marginal probabilities of the positive instantiations of the root nodes. Using the methodology from (Bovens and Hartmann, 2003), we obtain the posterior probability of the conjunction of S and G as follows:

$$(4) \quad P_1^* := \frac{P_1(S, G, E)}{P_1(E)} = \frac{P_1(S, G, E)}{\sum_{S, G} P_1(S, G, E)} = \frac{\gamma \pi \sigma}{\gamma \pi + \bar{\gamma} \rho}$$

The denominator of the rightmost fraction in (4) is a convex combination of π and ρ weighed by γ , where $\bar{\gamma} := 1 - \gamma$.²²

We close the present subsection by assessing the degree of confirmation of the

²²We will hereafter abbreviate ‘ $1 - x$ ’ as ‘ \bar{x} ’.

conjunction of S and G. To do this, we use the difference measure d , cf. (Carnap, 1950), which is defined for our case as follows:²³

$$(5) \quad d_1 := \mathbf{P}_1(S, G|E) - \mathbf{P}_1(S, G)$$

Thus, E confirms G if the consideration of E raises the probability of the conjunction of S and G. By calculating d_1 , we show that this is indeed the case:

$$(6) \quad d_1 = \frac{\gamma \bar{\gamma} \sigma (\pi - \rho)}{\gamma \pi + \bar{\gamma} \rho}$$

If we assume that γ, π, ρ , and σ lie in the open interval $(0, 1)$, where $\pi > \rho$, the above fraction is always strictly positive. We summarize our observation in the following theorem:

Theorem 1. *E confirms S and G iff E confirms G.*

This completes our investigation of the joint probability of S and G *before* the execution of a Montague Reduction of G to S. We next investigate the joint probabilities of S and G *after* such a reduction has been performed.

4.2. Post-Reductive Confirmation. To determine the confirmation of S and G in the post-reductive situation (cf. Fig. 7), we must first restate the probability distributions from the previous subsection. In particular, since G is no longer a root node in Figure 7 (and is, thus, not assigned a prior probability), we replace the equation $\mathbf{P}_1(G) = \gamma$ from (2) by the equations from (7), where \mathbf{P}_2 is the new probability measure:

$$(7) \quad \mathbf{P}_2(G|S) = 1 \quad , \quad \mathbf{P}_2(G|\neg S) = 0$$

The equations from (7) are warranted by Montague's homomorphism h . All other assignments are as for \mathbf{P}_1 . Our introduction of the new measure \mathbf{P}_2 is motivated by the move to a different probabilistic situation, and the need to assign the received Montagovian propositions possibly distinct probabilistic values. Equality statements of the form $\mathbf{P}_2(S) = \mathbf{P}_1(S)$ ensure the possibility of comparing the confirmation of S and G in the different situations.

As is encoded by the arrow from S to G in Figure 7, Montague's homomorphism h effects a flow of evidence from syntax to semantics. The confirmation of S is defined simply as follows:

Theorem 2. *E confirms S iff $\pi > \rho$.*

According to the above theorem, the evidence E confirms the proposition S if (as has been assumed in Section 4.1) E supports G. The equations in (7) ensure a positive flow of confirmation from G to S.

On the basis of the above, the conjunction of S and G has the following prior and posterior probabilities: (All calculations are included in the Appendix.)

$$(8) \quad \mathbf{P}_2(S, G) = \sigma$$

$$(9) \quad \mathbf{P}_2^* := \mathbf{P}_2(S, G|E) = \frac{\pi \sigma}{\pi \sigma + \rho \bar{\sigma}}$$

²³As is discussed in (Fitelson, 1999), cf. (Eells and Fitelson, 2000), results may depend on our choice of confirmation measure. Whether (and to what extent) they do, will be a question for future research.

The degree of confirmation of the conjunction of S and G under the measure \mathbf{P}_2 is recorded below:

$$(10) \quad d_2 := \mathbf{P}_2(S, G|E) - \mathbf{P}_2(S, G) = \frac{\sigma \bar{\sigma} (\pi - \rho)}{\pi \sigma + \rho \bar{\sigma}}$$

This completes our investigation of the probabilities and confirmation of the conjunction of S and G in the post-reductive situation. To show the epistemic value of Montague Reduction, we next compare the conjunction's probabilities and confirmation in the two scenarios. We accept a reduction if it raises the conjunction's probabilities or evidential support, and reject (or ignore) it otherwise.

4.3. Comparing Situations. We begin by comparing the *prior* probabilities of the conjunction of S and G in the two situations from Sections 4.1 and 4.2. While the propositional variables S and G are independent *before* the reduction, they have become dependent *after* the reduction. This is due to the fact that G is no longer a root node in Figure 7. To compare the joint probabilities of S and G in the two scenarios, we assume the identity of $\mathbf{P}_2(G)$ and $\mathbf{P}_1(G)$, and of $\mathbf{P}_2(E|G)$ and $\mathbf{P}_1(E|G)$. By the first equality in (7), we further assume the equality in (11), such that $\gamma = \sigma$.

$$(11) \quad \mathbf{P}_2(G) = \mathbf{P}_2(G|S) \mathbf{P}_2(S) = \sigma$$

Using the above, we calculate the difference, Δ_0 , between the conjunction's pre- and post-reductive prior probabilities, and obtain

$$(12) \quad \Delta_0 := \mathbf{P}_2(S, G) - \mathbf{P}_1(S, G) = \sigma \bar{\sigma}.$$

Intuitively, the Montague Reduction of syntax to semantics is epistemically valuable if the prior probability of the conjunction of S and G is higher post- than pre-reduction, i.e. if $\Delta_0 > 0$. Since we assume that all non- h -based probabilities are non-extreme, we know that the former is indeed the case.

The difference, Δ_1 , between the conjunction's *posterior* probabilities under the measures \mathbf{P}_2 and \mathbf{P}_1 is also strictly positive:

$$(13) \quad \Delta_1 := \mathbf{P}_2(S, G|E) - \mathbf{P}_1(S, G|E) = \frac{\pi \sigma \bar{\sigma}}{\pi \sigma + \rho \bar{\sigma}}$$

To show the truth of this statement, we use the above assumptions together with the fact that $\pi > \rho$.

The post-reductive confirmation of our propositions witnesses a similar increase. To establish this, we calculate the difference between the conjunction's pre- and post-reductive degree of confirmation under the difference measure, and obtain

$$(14) \quad \Delta_2 := d_2 - d_1 = \frac{\sigma \bar{\sigma}^2 (\pi - \rho)}{\pi \sigma + \rho \bar{\sigma}}.$$

As can be read off from the expression in (14), the positivity of Δ_2 – and the attendant positive confirmatory impact of Montague Reduction – is conditional on the requirement that $\sigma \in (0, 1)$ and that $\pi > \rho$.

The above-observed increase in the joint probabilities and evidential support of the conjunction of S and G corresponds to the increase in a conjunction's probabilities and support after the execution of a *Nagelian* reduction, cf. (Dizadji-Bahmani et al., 2011). In particular, since Nagelian bridge laws and Montague's homomorphism h both set the posterior probability of the truth of the 'reduced' proposition (given the truth (resp. falsity) of the 'reducing' proposition) to 1

(resp. to 0) (cf. our (7), and equation (12) from (Dizadji-Bahmani et al., 2011)), the Montague Reduction of categorial grammar to model-theoretic semantics achieves an equally large²⁴ boost in confirmation as the Nagelian reduction of a ‘suitable’²⁵ proposition-pair. This observation is captured below:

Proposition 1. *For suitable pairs of propositions, Montague Reduction is epistemically equally advantageous as Nagel Reduction.*

We close our presentation by suggesting two possible extensions of our model of Montague Reduction. These include the adaptation of our model to fundamental theories with multiple propositions (in Sect. 5), and the adaptation of our model to a variant of Schaffner’s (1967) revised model of Nagelian reduction (in Sect. 6). We will see that the latter accounts for the structural richness of grammatical rules.

5. EXTENSION I: REDUCTIONS OF ‘LARGER’ THEORIES

The previous section has shown that Montague Reduction increases the joint probabilities and degree of confirmation of the conjunction of pairs of theories with a single propositional element (esp. of pairs of theories with a single-proposition reduced theory²⁶). Montague’s homomorphism from Section 2.1 suggests an easy generalization of this result to pairs with more comprehensive reduced theories (e.g. to syntactic theories with *more* categories of expressions) in which the behavior of the syntactic operations is governed by a larger set of rules.²⁷ The reductive relation between these ‘larger’ theories is represented via the graph from Figure 8 (next page). In the graph, we call the variables S , G , and E from Section 4 ‘ S_1 ’, ‘ G_1 ’, and ‘ E_1 ’, respectively. The variables S_i , G_i , and E_i (with $1 < i \leq n \in \mathbb{N}$) are associated with new (semantic resp. syntactic) rules and their supporting pieces of evidence.

As is captured in Figure 8, Montague’s homomorphism h effects a *pairwise reduction*, which reduces categorial grammar to model-theoretic semantics by reducing \mathbb{G}_1 to \mathbb{S}_1 , \mathbb{G}_2 to \mathbb{S}_2 , \mathbb{G}_3 to \mathbb{S}_3 , etc.²⁸ As a result, the probability of syntax reduced to semantics will correspond to the product of the probabilities of all individual proposition-pairs:

$$(15) \quad \mathbf{P}_2\left(\bigcap_k \langle S_k, G_k \rangle\right) = \mathbf{P}_2(S_1, G_1) \mathbf{P}_2(S_2, G_2) \mathbf{P}_2(S_3, G_3) \dots$$

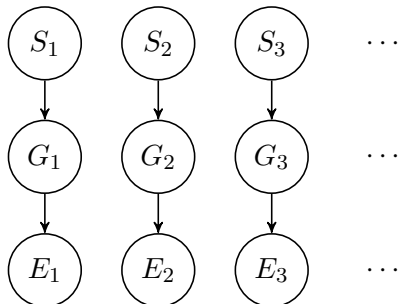
²⁴To ensure the comparability of the post-reductive situation from Section 4.2 with the post-reductive situation from (Dizadji-Bahmani et al., 2011) – which assumes Schaffner’s (1967) *revised* model of Nagel Reduction (cf. Sect. 6) –, we let the posterior probability of the truth, T_1^* , of the corrected version of the ‘reducing’ proposition (given the truth, T_1 (resp. falsity, $\neg T_1$), of the ‘uncorrected’ reducing proposition) and the truth, T_2 , of the uncorrected version of the ‘reduced’ proposition (given the truth, T_2^* (resp. falsity, $\neg T_2^*$), of the corrected ‘reduced’ proposition) both be 1 (resp. 0) (rather than p_1^* or p_2^* (resp. q_1^* or q_2^*), as in (Dizadji-Bahmani et al., 2011, (11) and (10))).

²⁵Here, *suitable* is defined as ‘allowing the application of the described reduction procedure’. As a result, suitable propositions for Nagelian reduction have a common target domain and contain bijectively related predicates.

²⁶Since the variables, S and S' , of the rules \mathbb{S}_s and \mathbb{S}'_s are only probabilistically independent before the reduction, Proposition 1 extends to the three-member set $\{\mathbb{G}_s, \mathbb{S}_s, \mathbb{S}'_s\}$.

²⁷Montague’s rule G9, cf. (Montague, 1973), is an example of such a new rule.

²⁸For simplicity, we again ignore ‘alternative’ semantic propositions \mathbb{S}'_k (with $k \in \mathbb{N}$) for which it holds that $h(\mathbb{S}'_k) = h(\mathbb{S}_k) = \mathbb{G}_k$.


 FIGURE 8. Post-reductive relations between pairs of propositions $\langle S_k, G_k \rangle$.

$$= P_2(S_1) P_2(S_2) P_2(S_3) \dots, \quad (\text{by (7)})$$

respectively

$$(16) \quad P_2\left(\bigcap_k \langle S_k, G_k | E_k \rangle\right) = P_2(S_1, G_1 | E_1) P_2(S_2, G_2 | E_2) P_2(S_3, G_3 | E_3) \dots \\ = P_2(S_1 | E_1) P_2(S_2 | E_2) P_2(S_3 | E_3) \dots$$

However, the stipulation of independent morphisms between all pairs $\langle S_k, G_k \rangle$ does not assign Montague's syntax-semantics relation an optimal epistemic value. This is due to the multiplication properties of real numbers in the open interval $(0, 1)$, such that the probability of the conjunction *decreases* in inverse proportion to the number of its conjuncts. But this contradicts our intuitions (reflected in much work in formal semantics, and in (Dizadji-Bahmani et al., 2011, p. 326)) that reductions between 'larger' (multi-proposition) theories share the epistemological advantages of reductions between 'smaller' theories.

Admittedly, the observed decrease in the joint probabilities of larger theories is also a problem for Nagelian reduction. However, Montague's framework provides a strategy for avoiding this problem. This strategy arises from the possibility of constructing certain semantic domains from other domains:²⁹ Our presentation of Montague's two theories from Section 2.1 has assumed that all semantic domains (especially, the domains \mathcal{D}_N , \mathcal{D}_V , and \mathcal{D}_S) are equally basic. Yet, in the Montagovian framework, this is in general not the case. In particular, to enable the *compositional* interpretation of natural language,³⁰ Montague's semantic models only contain basic domains for individuals (i.e. the set \mathcal{D}_N) and truth-values (i.e. the set \mathcal{D}_S), cf. (Montague, 1970a). From the elements of these domains, elements of all other domains (e.g. the members of the set \mathcal{D}_V) are obtained via a number of object-forming rules (here, via function-space formation, s.t. $\mathcal{D}_V \subseteq \{f \mid f : (\mathcal{D}_N \rightarrow \mathcal{D}_S)\}$).

We expect that these constructibility relations between domains will establish connections between objects of the reducing theory (and similarly, for the reduced theory), which will result in mutual probabilistic dependencies between *same-theory* propositions. A full development of this sophisticated variant of Montague Reduction, and an assessment of its epistemological merits, will be provided in a sequel to this paper.

²⁹A similar observation applies to syntactic categories, cf. (Ajdukiewicz, 1935).

³⁰s.t. the semantic value of an expression is a function of the semantic values of its constituents and their mode of combination, cf. (Partee, 1984).

6. EXTENSION II: SCHAFFNER-STYLE REDUCTION

Section 2.2 has identified Montague Reduction as a model of a *weak* type of inter-theoretic relation. We have attributed this weakness to the fact that the syntactic rule \mathbb{G}_s contains *more* structural information than the semantic rule \mathbb{S}_s (or \mathbb{S}'_s). We have used this fact to argue for the *lower* reductive strength of Montague Reduction vis-à-vis Nagelian reduction.

Yet, Nagelian reduction also does not enable the full reduction of phenomenological to fundamental propositions. For example, in the reduction of thermodynamics to statistical mechanics, Nagel's model is unable to derive the *exact* Second Law of thermodynamics, in which entropy does not fluctuate in equilibrium.³¹

The revised model of Nagelian reduction from (Schaffner, 1967; 1974), cf. (Nagel, 1977), improves the reductive accuracy of the Nagelian model by introducing a dedicated level of 'corrected' propositions of the two theories, and by demanding that the corrected version of each proposition of the fundamental theory be obtained from the original proposition via auxiliary assumptions, and that the corrected version of each proposition of the phenomenological theory be strongly analogous to the original proposition.³² The resulting model of reduction is captured in Figure 9. In the figure, T_1^* and T_2^* are the 'corrected' versions of the propositions T_1 and T_2 , respectively.

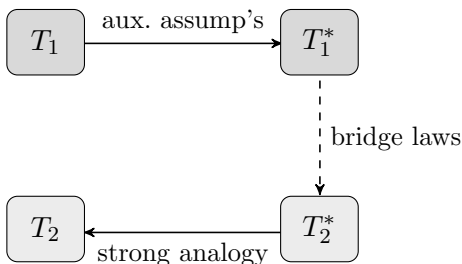


FIGURE 9. Generalized Nagel-Schaffner reduction.

We expect that the introduction of corrected propositions will also improve the success of our model of Montague Reduction. In Section 2.2, we have already noted that the syntactic correlate of function application will not contain any information about word order. We have attributed this observation to the fact that semantics contains much less structural information than syntax. However, since word order is a very stable property of a language, the specification of a language's word order-type (as an auxiliary assumption about the investigated language) will allow us to supplement this information: For example, once we have complemented \mathbb{S}_s with the information that English is a 'subject-verb-object' (SVO) language, we will be able to 'adjust' the rule from (1c) to the syntactic rule from (2), where $[\cdot]$ is the order-sensitive concatenation operation from \mathbb{G}_s . Since this operation respects the order of the constituent basic expressions, it combines intransitive verbs with a name which occurs *on their left*:

³¹Nagelian reduction can only derive a variant of this law, in which thermodynamic entropy *does* fluctuate in equilibrium.

³²For a detailed presentation and discussion of generalized Nagel-Schaffner reduction, the reader is referred to (Schaffner, 1967) and (Dizadji-Bahmani et al., 2010).

(2) If $Y \in \mathcal{E}_v$ and $X \in \mathcal{E}_N$, then $[XY] \in \mathcal{E}_s$.

Since the rule (1c) is strongly analogous to the rule (2) under this auxiliary assumption, a ‘Schaffner-style’ variant of our model of Montague Reduction will derive the ‘right’ concatenation rule for proper names and intransitive verbs.

The elaboration of this variant of our model of Montague Reduction, and a demonstration of its (expected) epistemic advantages, is left for another occasion. We close the paper with a summary of our main results.

7. CONCLUSION

In this paper, we have presented a model of a new type of intertheoretic relation, called *Montague Reduction*, which is inspired by Montague’s (1973) framework for the analysis and interpretation of natural language syntax. We have identified the commonalities of Montague Reduction with classical Nagelian reduction, cf. (Nagel, 1961), and have established their salient differences. In particular, we have observed that Montague Reduction can establish directed dependency relations between pairs of theories with non-overlapping target domains, which cannot be captured by Nagelian reduction. To show the epistemic rationale behind our new type of intertheoretic relation, we have demonstrated that – like its Nagelian counterpart – Montague Reduction raises the posterior probability of the conjunction of the two related theories, and increases the flow of confirmation between them. Finally, we have identified two strategies for the extension and improvement of our model of Montague Reduction.

APPENDIX: PROOFS AND CALCULATIONS FOR SECTION 4

We have calculated the pre-reductive probabilities of the conjunction of the positive instantiations of S and G in Section 4.1. The joint distribution, $P_2(S, G, E)$, of the (post-reductive) graph in Figure 7 is given by the expression

$$P_2(S) P_2(G) P_2(E|G).$$

Using the methodology from (Bovens and Hartmann, 2003), the prior probability of the conjunction of S and G is obtained as follows:

$$(17) \quad P_2(S, G) = \sum_E P_2(S, G, E) = \pi \sigma + \bar{\pi} \sigma = \sigma$$

We yield the posterior probability, $P_2^* := P_2(S, G|E)$, of the conjunction of S and G thus:

$$(18) \quad P_2^* = \frac{P_2(S, G, E)}{P_2(E)} = \frac{\pi \sigma}{\pi \sigma + \rho \bar{\sigma}}$$

To obtain the difference Δ_0 , we calculate

$$P_2(S, G) - P_1(S, G) = \sigma - \sigma^2 = \sigma \bar{\sigma}.$$

This proves the following proposition:

Proposition 2. $\Delta_0 = 0$ iff $\sigma = 0$ or 1 ; $\Delta_0 > 0$ iff $\sigma \in (0, 1)$.

The difference, Δ_1 , between the conjunction’s pre- and post-reductive posterior probabilities is obtained as follows:

$$(19) \quad \Delta_1 := P_2^* - P_1^* = \frac{\pi \sigma - \pi \sigma^2}{\pi \sigma + \rho \bar{\sigma}} = \frac{\pi \sigma \bar{\sigma}}{\pi \sigma + \rho \bar{\sigma}}$$

From the difference measure

$$(20) \quad d_2 := \mathbf{P}_2(S, G|E) - \mathbf{P}_2(S, G) = \frac{\sigma \bar{\sigma} (\pi - \rho)}{\pi \sigma + \rho \bar{\sigma}},$$

we calculate the difference, Δ_2 , between the conjunction's degree of confirmation before and after the execution of a Montague Reduction as follows:

$$(21) \quad \Delta_2 := d_2 - d_1 = \frac{\sigma \bar{\sigma} (\pi - \rho) - \sigma^2 \bar{\sigma} (\pi - \rho)}{\pi \sigma + \rho \bar{\sigma}} = \frac{\sigma \bar{\sigma}^2 (\pi - \rho)}{\pi \sigma + \rho \bar{\sigma}}$$

This completes our proofs and calculations for Section 4.

REFERENCES

- Ajdkiewicz, Kazimierz. 1935. *Die Syntaktische Konnexität*, *Studia Philosophica* **1**, 1–27.
- Bickle, John. 2006. *Reducing Mind to Molecular Pathways: Explicating the reductionism implicit in current cellular and molecular neuroscience*, *Synthese*, 411–434.
- Bovens, Luc and Stephan Hartmann. 2003. *Bayesian Epistemology*, Clarendon Press, Oxford.
- Carnap, Rudolf. 1950. *Empiricism, Semantics, and Ontology*, *Revue Internationale de Philosophie* **4**, 20–40; reprinted in Carnap, Rudolf. 1950. *Meaning and Necessity: A study in semantics and modal logic*, The University of Chicago Press, Chicago and London.
- Chomsky, Noam. 1971. *Deep Structure, Surface Structure, and Semantic Representation*, *Semantics: An interdisciplinary reader* (D. Steinberg and L. Jakobovits, eds.), Cambridge University Press, Cambridge, 1971.
- Church, Alonzo. 1940. *A Formulation of the Simple Theory of Types*, *Journal of Symbolic Logic* **5/2**, 56–68.
- Darden, Lindley and Nancy Maull. 1977. *Interfield Theories*, *Philosophy of Science* **44**, 43–64.
- Dizadji-Bahmani, Foad, Roman Frigg, and Stephan Hartmann. 2010. *Who's Afraid of Nagelian Reduction?*, *Erkenntnis* **73/3**, 393–412.
- . 2011. *Confirmation and Reduction: A Bayesian account*, *Synthese* **179/2**, 321–338.
- Earman, John. 1992. *Bayes or Bust?*, The MIT Press, Cambridge, Mass.
- Eells, Ellery and Branden Fitelson. 2000. *Measuring Confirmation and Evidence*, *The Journal of Philosophy* **97/12**, 663–672.
- Fitelson, Branden. 1999. *The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity*, *Philosophy of Science* **66**, 362–378.
- Gamut, L. T. F. 1991. *Intensional Logic and Logical Grammar*, *Logic, Language, and Meaning*, vol. 2, The University of Chicago Press, Chicago and London.
- Hájek, Alan and Stephan Hartmann. 2010. *Bayesian Epistemology*, *A Companion to Epistemology* (Jonathan Dancy, Ernest Sosa, and Matthias Steup, eds.), Blackwell Companions to Philosophy, Blackwell, Malden, Mass., 2010.
- Hartmann, Stephan. 1999. *Models and Stories in Hadron Physics*, *Models as Mediators* (Mary S. Morgan and Margaret Morrison, eds.), Cambridge University Press, Cambridge, 1999, pp. 326–346.
- Hartmann, Stephan and Jan Sprenger. 2010. *Bayesian Epistemology*, *Routledge Companion to Epistemology* (Sven Bernecker and Duncan Pritchard, eds.), Routledge, Malden, Mass.
- Hoover, Kevin D. 2010. *Idealizing Reduction: The microfoundations of macroeconomics*, *Erkenntnis* **73/3**, 329–347.
- Howson, Colin and Peter Urbach. 2005. *Scientific Reasoning: The Bayesian approach*, Open Court, La Salle.
- Janssen, Theo M.V. 2012. *Montague Semantics*, *The Stanford Encyclopedia of Philosophy*, Winter 2012 edition (Edward N. Zalta, ed.), 2012.
- Kuipers, Theo A. F. 1982. *The Reduction of Phenomenological to Kinetic Thermostatistics*, *Philosophy of Science* **49/1**, 107–119.
- Levens, Richard and Richard Lewontin. 1980. *Dialectics and Reductionism in Ecology*, *Synthese* **43**, 47–78.
- Litosseliti, Lia (ed.) 2010. *Research Methods in Linguistics*, Continuum, London and New York.
- Mitchell, Sandra D. 2003. *Biological Complexity and Integrative Pluralism*, Cambridge University Press, Cambridge and New York.

- Montague, Richard. 1970a. *English as a Formal Language*, Formal Philosophy: Selected papers of Richard Montague (Richmond H. Thomason, ed.), Yale University Press, New Haven and London, 1976.
- . 1970b. *Universal Grammar*, Formal Philosophy: Selected papers of Richard Montague (Richmond H. Thomason, ed.), Yale University Press, New Haven and London, 1976.
- . 1973. *The Proper Treatment of Quantification in Ordinary English*, Formal Philosophy: Selected papers of Richard Montague (Richmond H. Thomason, ed.), Yale University Press, New Haven and London, 1976.
- Moortgat, Michael. 1997. *Categorial Type Logics*, Handbook of Logic and Language (Johan van Benthem and Alice G.B. ter Meulen, eds.), Elsevier Science Publishers, Amsterdam, 1997.
- Nagel, Ernest. 1961. *The Structure of Science*, Routledge and Kegan Paul, London.
- . 1977. *Teleology Revisited*, The Journal of Philosophy **84**, 261–301.
- Neapolitan, Richard. 2003. *Learning Bayesian Networks*, Prentice Hall.
- Partee, Barbara H. 1973. *Some Transformational Extensions of Montague Grammar*, Journal of Philosophical Logic **2**, 509–534.
- Partee, Barbara. 1984. *Compositionality*, Varieties of Formal Semantics: Proceedings of the 4th Amsterdam Colloquium (Fred Landman and Frank Veltman, eds.), Groningen-Amsterdam Studies in Semantics, vol. 3, 1984.
- Partee, Barbara H. 1997. *Montague Grammar*, Handbook of Logic and Language (Johan van Benthem and Alice G.B. ter Meulen, eds.), Elsevier Science Publishers, Amsterdam, 1997.
- Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of plausible inference*, Morgan Kaufman, San Francisco.
- Podesva, Robert J. and Devyani Sharma (eds.) 2013. *Research Methods in Linguistics*, Cambridge University Press, Cambridge and New York.
- Rasinger, Sebastian M. 2013. *Quantitative Research in Linguistics: An Introduction*, Bloomsbury, London etc.
- Schaffner, Kenneth F. 1967. *Approaches to Reduction*, Philosophy of Science **34**, 137–147.
- . 1974. *Reductionism in Biology: Prospects and problems*, PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association **1974**, 613–632.
- . 1993. *Discovery and Explanation in Biology and Medicine*, Science and Its Conceptual Foundations, The University of Chicago Press, Chicago.
- Schouten, Maurice and Huib Looren de Jong (eds.) 2012. *The Matter of the Mind: Philosophical essays of psychology, neuroscience, and reduction*, Blackwell, Oxford.
- Tarski, Alfred. 1933. *The Concept of Truth in the Languages of the Deductive Sciences*, Logic, Semantics, Metamathematics: Papers from 1923 to 1938 (John Corcoran, ed.), Hackett Publishing Company, Indianapolis, 1983.
- Weber, Marcel. 2005. *Philosophy of Experimental Biology*, Cambridge Studies in Philosophy and Biology, Cambridge University Press, Cambridge and New York.