University of Warsaw
Faculty of Philosophy and Sociology

Dariusz Kalociński

Learning the Semantics of Natural Language Quantifiers

Uczenie się semantyki kwantyfikatorów języka naturalnego

Ph.D. Thesis

supervisor:
prof. Marcin Mostowski

assistant supervisor:
Ph.D. Nina Gierasimczuk

Warsaw
May 5, 2016

Abstract

Traditional view of natural language semantics conceives meanings of expressions as being constant from speaker to speaker. This idealistic hypothesis has often been considered as an explanation of how language users are able communicate. The traditional paradigm to learning semantics (which we refer to as learning by recognizing) shows how such a uniformity may be achieved, provided that there is an externally imposed semantic standard which does not change during learning. These paradigms have resulted in plethora of interesting approaches and practical applications. It seems, however, that the uniform view of semantics does not always correspond to facts. To illustrate this, we mention some evidence of within-language synchronic semantic variation. This sort of variation consists in co-existence of non-equivalent semantics within the same language community at a given time. On these grounds we abandon both the uniform approach to semantics and learning by recognizing which is tightly connected to the static view of meaning. Learning by recognizing is also considered as inappropriate for the analysis of communication phenomena which defy description in terms of the traditional teacher-learner distinction.

To explain how communication is possible within a semantically non-uniform community of speakers connected by various social relations of authority, we consider a more general mechanism of learning which we call learning by coordination. We give our solution to the problem and define the coordination mechanism which is capable of handling inconsistent samples of language use and various social influences between communicating speakers. We apply this algorithm to learning the semantics of upward monotone proportional quantifiers. We consider simple models of symmetrically communicating dyads and show how to analyse coordination of such processes in terms of Markov chains. We observe various mathematical connections between the possibility of convergence, specific levels of agents authority and complexity of communication patterns.

Moreover, we study some natural language constructions which are deemed to express the existence of certain kinds of similarities between partial orderings. Specifically, we provide examples of natural language sentences and their plausible logical forms that express the existence of homomorphism, embedding and variations of those. Semantically, we interpret the constructions as polyadic generalized quantifiers. We examine some of the quantifiers in question with respect to their FO-definability over appropriate finite models. Since they are definable in the existential fragment of SO, we investigate their completeness in the class NP. We prove that among the quantifiers under investigation, only the homomorphism quantifier is tractable. We stress potential importance of our results for linguistics and discuss some connections between computational complexity, human comprehension and language evolution.

Streszczenie

Zgodnie z tradycyjnym podejściem do semantyki języka naturalnego, znaczenia wyrażeń są na tyle dobrze określone, że rozsądną idealizacją wydaje się być hipoteza, iż wszyscy użytkownicy języka przyporządkowują wyrażeniom te same znaczenia. Ten sposób ujmowania semantyki stosowano również do wyjaśnienia komunikacyjnej funkcji języka. Tradycyjne podejście do uczenia się, które nazywamy uczeniem przez rozpoznawanie, pokazuje, jak zuniformizowana semantyka może powstać w warunkach, gdy uczeń dostosowuje się do narzuconego z zewnątrz, niezmiennego standardu semantycznego. Wymienione podejścia do semantyki zaowocowały wieloma interesującymi odkryciami i praktycznymi zastosowaniami.

Wydaje się jednak, że wizja zuniformizowanej semantyki nie zawsze jest zgodna z faktami. Przedstawiamy wybrane świadectwa na rzecz wewnątrzjęzykowego synchronicznego zróżnicowania semantyki. Zróżnicowanie to polega na współistnieniu w danej społeczności językowej nierównoważnych znaczeń tych samych wyrażeń. Na podstawie podanych świadectw odchodzimy od wizji zuniformizowanej semantyki oraz uczenia przez rozpoznawanie, które jest ściśle powiązane ze statycznym ujmowaniem znaczenia. Ponadto, uczenie przez rozpoznawanie wydaje się nieadekwatne w zastosowaniu do sytuacji, których nie da się opisać za pomocą tradycyjnego rozróżnienia nauczyciel-uczeń.

By wyjaśnić, jak możliwa jest komunikacja w obrębie semantycznie niejednorodnej społeczności, której członkowie charakteryzują się dowolnie ustalonym autorytetem, podajemy mechanizm uczenia przez uzgadnianie, pozwalający na przetwarzenie napływających komunikatów wraz z uwzględnieniem autorytetu ich nadawców. Zaproponowany mechanizm badamy w odniesieniu do problemu uczenia się semantyki kwantyfikatorów proporcjonalnych. Rozważania ograniczamy do symetrycznej komunikacyji między dwoma osobnikami. Pokazujemy, jak reprezentować uzgadnianie za pomocą łańcuchów Markowa. Na tej podstawie tej możemy zaobserwować rozmaite matematyczne zależności między możliwością uzgodnienia semantyki, poziomami autorytetu rozmówców oraz złożoności sytuacji komunikacyjnych.

Ponadto, badamy pewne konstrukcje języka naturalnego, które wyrażają istnienie różnego rodzaju podobieństw między częściowymi porządkami. Podajemy przykłady zdań języka naturalnego oraz ich możliwe formy logiczne, które stwierdzają istnienie homomorfizmu, zanurzenia, itp. Konstrukcje interpretujemy w terminach uogólnionych kwantyfikatorów poliadycznych. Badamy definiowalność tak zdefiniowanych kwantyfikatorów w logice elementarnej nad odpowiednimi modelami skończonymi. Jako że kwantyfikatory te są definiowalne w egzystencjalnej części logiki drugiego rzędu, pytamy o ich zupełność w klasie NP. Dowodzimy, że tylko kwantyfikator wyrażający homomorfizm jest praktycznie obliczalny, natomiast pozostałe są NP-zupełne. W części bardziej filozoficznej, dyskutujemy, jakie wnioski płyną z tych badań dla lingwistyki oraz zarysowujemy pewne związki między złożonością obliczeniową, przetwarzaniem informacji przez ludzi i lingwistyką ewolucyjną.

Acknowledgements

First of all, I would like to thank my supervisor, Marcin Mostowski. Marcin taught me logic and computability theory. He has always been a great source of inspiration and insightful ideas. Some of these ideas are developed in this dissertation. Without him, I would probably be quite a different person, probably struggling at the moment with a piece of commercial code.

I am also indebted to my assistant supervisor, Nina Gierasimczuk. I am not able to count the number of suggestions she gave me to improve this thesis. I have profited a lot from her guidance, knowledge and ideas she brought into this project. Moreover, one of the chapters of this dissertation is based on collaboration with Nina (and Marcin).

I would like to thank Jakub Szymanik who has provided a lot of important comments concerning the semantic complexity of quantifiers. Jakub has also made many important suggestions which are still living in my mind.

I am very grateful to Michał Tomasz Godziszewski. We worked together on computational complexity of similarity quantifiers. Collaboration with Michał is an unforgettable experience. I have profited a lot from discussions with him. Michał has been also very supportive during the preparation of this thesis.

I would like to thank Franciszek Rakowski for his open-mindedness and inspiring discussions on the Steels model for coordination of color categories.

There are also others who helped in various ways or participated in relevant exchanges of ideas, among them: Tadeusz Ciecierski, Marek Czarnecki, Jędrzej Kołodziejski, Mateusz Łełyk, Damian Niwiński, Antonio Maciej Matamoros Ochman, Tomasz Puczyłowski, Vinicius Macuch Silva, Tomasz Steifer, Jakub Uszyński and Bartosz Wcisło.

To my closest family: my children, my parents, my brother Dawid, my parents-in-law and Piotr. And to my beloved wife, Ania. Thank you, this is our common achievement.

Dariusz Kalociński
Warsaw, May 2016

Contents

Said Yate Fulham: “And just how do you arrive at that remarkable conclusion, Mr. Mayor?”

“In a rather simple way. It merely required the use of that much-neglected commodity - common sense. You see, there is a branch of human knowledge known as symbolic logic, which can be used to prune away all sorts of clogging deadwood that clutters up human language.” “What about it?” said Fulham. “I applied it. Among other things, I applied it to this document here. I didn’t really need to for myself because I knew what it was all about, but I think I can explain it more easily to five physical scientists by symbols rather than by words.”

Hardin removed a few sheets of paper from the pad under his arm and spread them out. “I didn’t do this myself, by the way,” he said. “Muller Holk of the Division of Logic has his name signed to the analyses, as you can see.”

Pirenne leaned over the table to get a better view and Hardin continued: “The message from Anacreon was a simple problem, naturally, for the men who wrote it were men of action rather than men of words. It boils down easily and straightforwardly to the unqualified statement, when in symbols is what you see, and which in words, roughly translated, is, ‘You give us what we want in a week, or we take it by force.’”

There was silence as the five members of the Board ran down the line of symbols, and then Pirenne sat down and coughed uneasily.

Hardin said, “No loophole, is there, Dr. Pirenne?”

“Doesn’t seem to be.”

“All right.” Hardin replaced the sheets. “Before you now you see a copy of the treaty between the Empire and Anacreon - a treaty, incidentally, which is signed on the Emperor’s behalf by the same Lord Dorwin who was here last week - and with it a symbolic analysis.”

The treaty ran through five pages of fine print and the analysis was scrawled out in just under half a page. “As you see, gentlemen, something like ninety percent of the treaty boiled right out of the analysis as being meaningless, and what we end up with can be described in the following interesting manner:

“Obligations of Anacreon to the Empire: None!

“Powers of the Empire over Anacreon: None!”

— Isaac Asimov, *Foundation*, 1951

Preface

This dissertation is devoted to the problem of learning natural language semantics and is primarily focused on semantics of quantifiers. The general ideas underlying our approach may be familiar to the audience coming from research in evolutionary linguistics. However, these ideas are less likely to be familiar to logicians or philosophers.

The approach we present here is not congruent with a quite common acknowledgement of the existence of fixed, purely linguistic meanings of the expressions of natural language. According to this traditional claim, natural language semantics may be described in an orderly fashion by ascribing to linguistics constructions their well determined meanings.

It was quite a challenge for me, a person so deeply entrenched in the tradition, to abandon this conviction and accept a totally different view of language that was first presented to me three years ago by Marcin Mostowski who, not accidentally, is the supervisor of this thesis. Now, when this new view of language seems so familiar to me, it is difficult to comprehend how could I be so mistaken.

The crucial idea lying behind this thesis is simple as that: there is no such a thing as semantics of natural language in the traditional sense. Semantics should be viewed as a system of individuated meanings distributed among members of the language community. These meanings, conceived as algorithms for computing denotations of linguistic constructions, enable speakers to effectively use language in particular situations of usage and convey information to others. However, if each interlocutor has different understanding of language, how they are going to communicate? Well, this is exactly the story you can read about on subsequent pages.

Computational results from Chapter II have been obtained in collaboration with Michał Tomasz Godziszewski [Kalociński and Godziszewski, 2016]. Results from Chapter III have been published in [Kalociński et al., 2015].

I

The Problem of Learning Semantics

I.1 Introduction

Alfred Tarski laid foundations for the scientific approach to semantics. His semantic theory of truth for formalized languages [Tarski, 1933] shows how one may define and use semantic notions without the risk of falling into paradoxes. One of the crucial concepts employed in Tarki's work is that of interpretation of a language. A language, understood as a set of meaningless symbols, is mapped to appropriate objects from a given universe of discourse. A familiar picture is that of elementary logic where each individual constant is mapped to a distinguished element, function symbol to a function and predicate to a relation. This mapping, referred to as interpretation, gives each symbol a unique meaning. This formalism enables us to extend the notion of meaning to sentences. Given a definition of truth for sentences of the language, one may identify the meaning of a given sentence with its truth-conditions, namely the class of models under which the sentence is true.

This idea of language and meaning has been adopted by some logically oriented theories of natural language and proved to be useful in answering various question related to semantics.

A great deal of research on natural language semantics abstracts away from actual language use, as it was initially proposed in [de Saussure, 1916], and investigates language in isolation. However, if language is put in its natural environment consisting of communicating language users then the picture becomes more complicated. One of the main problems is how different language users are able to understand each other in natural language. The overall process of communication is very complicated. Our focus is on the semantic aspect of this problem.

It has been often hypothesised that speakers are equipped with common semantic rules and they are able to communicate precisely in virtue of this commonality (and other mechanisms responsible for production, vocalization, interpretation, etc.). We discuss this view in Section I.2 where we formulate the Uniform Semantics Thesis (Thesis 1) according to which competent speakers of a given language community share the same mapping between linguistic constructions and meanings. This gives an easy account for linguistic communication. We also provide some rationale that the uniform view of semantics was a rather commonly accepted idealisation and present some explicit formulations of the thesis from the literature.

The Uniform Semantics Thesis does not touch upon the problem of how this commonality in semantics may actually be achieved. Some rescue comes from the concept of learning by allowing language users to adopt the semantics from other speakers. Traditionally, language learning is conceived as a process of adopting an externally imposed standard. In the case of semantics, this process is usually modelled in terms of a hypothetical learning mechanism which enables a subject to eventually internalize the correct meanings by mere observation of language use (see, e.g., [Gierasimczuk, 2007, Piantadosi et al., 2012]). This type of learning, which we call learning by recognizing, gives an answer about how the commonality in semantics may be achieved. Learning by recognizing is tightly connected with the Uniform Semantics Thesis – one may hypothesise that the overall process of learning within a community consists in acquiring correct semantic conventions from competent speakers. We discuss learning by recognizing in Section I.3

Section I.4 is devoted to the criticism of both uniformity of semantics and learning by recognizing. We provide some evidence of the fact that language users may systematically employ non-equivalent interpretations of the same expressions in the same context (Section I.4). We call this phenomenon synchronic semantic variation. Based on this evidence, we reject Thesis 1 (Section I.5) and discuss a weaker claim according to which in some cases a uniform description of semantics may serve as an approximation of the semantics of a given language community. We also observe that the abandonment of the uniform view of semantics has negative consequences for learning by recognizing (Section I.5). If a language community manifests significant semantic variation then learners encounter inconsistent linguistic usage. Traditional learning mechanisms are not well suited for handling such data.

In Section I.6 we shortly present another approach which we refer to as learning by coordination. The crucial assumption of this approach is the abandonment of the uniform semantics – there is no distinguished, externally imposed semantic standard. To account for linguistic communication within community which is allowed to manifest semantic variation, we need a different leaning mechanism, capable of handling inconsistent responses from other speaker and allowing adaptation of the individuated meanings so that efficient communication becomes possible. At the end of this chapter we mention various approaches from evolutionary linguistics where a very similar ideas of language learning have first originated.

I.2 Uniform Semantics Thesis

Consider a language community, i.e., a group of speakers who effectively use natural language for communication purposes. The simplest analysable fragment of communication within such a community is a single interaction between two individuals: the speaker and the hearer. For the sake of example, assume the speaker wants to draw attention of the hearer to a particular object of the shared context. The speaker picks up an expression which, according to the meaning he associates with this expression, refers to the intended object from the context. He conveys this expression to the hearer by producing an utterance. The utterance is interpreted by the hearer according to the meaning he assigns to the received expression. Finally, the communication is considered successful if the expression, as interpreted by the hearer, points exactly to the same object as the speaker has in mind.

The problem which arises in this simple situation is as follows: how the speaker and the hearer are able to communicate successfully? One of the simplest answers to this question, which has been often formulated, is that the speaker's and the hearer's meanings are the same. Obviously, this explanation, if not counting other factors which may intervene in communication (errors, implicatures, noise, etc.), gives an easy explanation of communication in natural language. In general, it seems that this sort of account has been often generalized to the whole language community, which leads to the following thesis:

Thesis 1 (Uniform Semantics Thesis) *The semantics of a given language used by a community for communication purposes at a given time may be described in terms of an association between expressions and their meanings which is shared by all members of the community.*

At this point we should make a reservation that Thesis 1 cannot be taken literally. Similarly to hypotheses from other scientific fields, one may treat it as an idealisation. Its idealistic character consists in putting down the occurrence of anomalous data to factors which do not contradict the thesis itself. In other words, if some facts turned out to be incompatible with this explanation then we would search for additional causes, not connected to semantic variation. A familiar example of an idealisation is the claim the actual meaning the natural language quantifier *some* corresponds to the meaning of the logical meaning of the existential quantifier. However, quite often *some* is used in such a way that the right interpretation is *some but not all*, as it may be the case with the sentence *I ate some pie*. This may seem as an evidence against the aforementioned claim, as our linguistic practice shows that *some* has two meanings, not one. However, this inconsistency may be resolved by the appeal to Grice's conversational maxims [Grice, 1975] which provide separate explanation for this apparently anomaly.[1]

Note that Thesis 1 is compatible with various well know semantic properties of natural languages such as ambiguity, polysemy, homonymy, context dependence, etc. For example, the fact that a given expression is ambiguous may be represented by a mapping from this expression to a set of its distinct meanings. Context dependency roughly means that the right interpretation

[1]Another familiar example of an idealisation is the law of universal gravity which predicts that all objects fall at the same rate. Deviation form the expected behaviour may be explained by positing additional factors such as air resistance.

of an utterance may be identified only when provided enough contextual information. For example the sentence *I went to the bank* is context dependent as it cannot be assigned any interpretation unless we are given enough information such as the identity of the speaker and whether the speaker is referring to a financial institution or a riverside. However, the word *bank* has two conventional meanings which, when included in the shared uniform semantics, are disambiguated in the presence of a particular context.[2]

I.2.1 Acknowledgement of Uniform Semantics

It this subsection we provide some rationale for the claim that the uniform view of natural language semantics might have been quite commonly accepted in the twentieth century. Let us start with a short reminder about approaches which are primarily concerned with the synchronic descriptions of natural languages.

Synchronic descriptions abstract away from the actual language use and therefore enclose various properties of natural language in a uniform system underlying the actual performance of the speakers from a given time. Synchronic approach to the analysis of natural language has become influential through the works of early structuralists (see, e.g., [de Saussure, 1916]). This approach, referred to as synchronic linguistics, was a dominating paradigm of the twentieth century and overshadowed the earlier diachronic approaches from the nineteenth century. The so-called logical theory of language, inspired by the successes of logic [Russell and Whitehead, 1910] and logical semantics [Tarski, 1933], has strengthened the position of the synchronic view of language, especially among logicians and philosophers. One of the first examples of such studies, concerned with an analysis of conversational language in terms of logical forms for various natural language constructions, may be found in in the last chapter of [Reichenbach, 1947]. This and further work along these lines have inspired a great deal of researchers interested in the analysis of natural language. We are not to give a historical survey here, but let us mention at least a few contributions. The idea that the semantic theory of natural language should be given a form of a theory of truth has been advocated in [Davidson, 1967]. The mathematical turn in linguistics [Chomsky, 1957] laid foundations for a formal descrip-

[2]Sometimes such examples are treated as homonymy.

tion of the linguistic competence and clearly delineated syntax and semantics in the analysis of natural language. Logical semantics and generative linguistics have been heavily employed in the linguistic literature [Vasiliu, 1981]. The model-theoretic approach presented in [Peters and Westerståhl, 2006] analyses in detail many logical properties of natural language quantifiers whose meanings are represented in terms of generalized quantifiers theory [Lindström, 1966]. Approaches developed in this vein have been very fruitful, resulting in a wealth of practical applications, e.g., in dialogue systems, automated reasoning, information retrieval and search. A refined view on semantics conceives meanings as algorithms for computing denotations of expressions [Tichy, 1969, Suppes, 1980] (see also [Szymanik, 2016]).

It seems that the wide acknowledgement of the uniform view of semantics is visible in the vivid and widespread reaction to the Quine thesis on indeterminacy of translation [Quine, 1960]. Indeed, this «has been among the most widely discussed and controversial theses in modern analytical philosophy» [Wright, 1997]. Quine considers a problem of constructing a translation manual for an unknown language by observing linguistic usage of native speakers and rejects an absolute standard of right and wrong in doing so. Observe that if semantics of natural language is conceived as system which may be different from speaker to speaker then the Quine thesis is trivial. Obviously, no translation would work all the time, as there would be no well determined meanings of expressions among native speakers. However, if semantics of natural language is conceived as uniform from speaker to speaker then the Quine thesis becomes non-trivial and very controversial. The Quine thesis was in fact considered as non-trivial. It seems that this provides some rationale for the claim that at that time the uniform view of semantics was a rather commonly accepted conviction, for otherwise the Quine thesis could have been rejected on the more fundamental grounds.

The claim as to the wide acknowledgement of the existence of fixed, purely linguistic meanings of the expressions of natural language is expressed in [Stanosz, 1974a]. According to Stanosz, this claim is accepted by logicians, whose education is shaped by research in formalized languages. Moreover, it is accepted, at least tacitly, by practical linguists such as language teachers, the authors of dictionaries, etc. Stanosz writes: «When they coin definitions, or find the synonymity of expressions, they affirm the conviction that natural language can be described semantically in a general manner, and not only in particular instances of the usage of its expressions». This claim was true in the twentieth century and to all appearances it remains true in the early twenty first century. According to [Wacewicz et al., 2016] contemporary academic teaching hardly touches upon recent developments in evolutionary linguistics and is still performed within the traditional, out-dated paradigms and largely inadequate conceptual frameworks. Textbooks are mostly concerned with traditional synchronic linguistics approaches aimed at descriptions of grammar and sound structure of natural languages. This fact renders the wide acknowledgement of Thesis 1 as highly plausible.

Quantifiers and their denotations. There are various ways of defining uniform association between expressions and meanings. One of the problems in giving such an association is that we need a way to represent meanings. Let us mention approach presented in [Peters and Westerståhl, 2006] where the connection between natural language quantifier expressions and their meanings is presented in a clear and intuitive way. Meanings of quantifier expressions are modelled as generalized quantifiers. In the first approximation one may say that a generalized quantifier is a relation between sets of individuals (see Chapter II, Definition 2). The association between quantifier expressions and generalized quantifiers is referred to as the relation of denoting. Consider simple quantified sentences: *All boys entered the classroom* and *No girl watched the movie.* Here, quantifier expressions are *all* and *no.* Following [Peters and Westerståhl, 2006], we would say the English *no* and the Swedish *ingen* denote the relation which holds between two sets exactly when they are disjoint. Similarly, the English *all* and the Swedish *alla* denote the inclusion relation. Let us consider one example in more detail: *All boys entered the classroom.* Given that this sentence is used in a relevant context, we have two sets, namely the set of boys B and the set of people who entered the classroom C. Under the assumed interpretation of *all*, the sentence in question is true iff $B \subseteq C$. It seems that this reading overlaps with the conventional interpretation of *all*. We refer the reader to [Peters and Westerståhl, 2006] for more examples and details on this formal approach to natural language quantification.

Selected Formulations

In this section we provide excerpts of some authors who give a clear exposition of the fundamental assumptions underlying the uniform semantics approach. Such expositions are rather difficult to find in the literature. This section may be skipped without loosing the main points of this chapter.

Our examples are drawn from the philosophical and linguistic literature. It seems that the remarks of Polish authors form an important contribution in the exposition of assumptions involved in the uniform view of semantics, as they are by far the most clear and explicit. [Stanosz, 1974a] is very informative. The reader with an acquaintance of Polish may benefit from reading [Stanosz, 1999] which is a short response to a critique of the logical concept of language. [Suszko, 1957b] and [Suszko, 1957a], reprinted as a single chapter in [Suszko, 1998a], although concerned with an allegedly different topic, contains a few very informative remarks. In course of the main text, we recapitulate the most relevant ideas from the Polish articles. We also mention crucial ideas from [Vasiliu, 1972] where the logical theory of language and linguistics make an unprecedented fusion. Finally, we recapitulate some remarks from [Dummett, 1975] and [Katz, 1966].

An explicit attempt to defend Thesis 1 may be found in [Stanosz, 1974b]. Stanosz reflects upon the applicability of logical semantics to the analysis of natural language and tries to defend the view that context-dependence stays in accordance with the postulate that natural language expressions are mapped to their systematic meanings. Stanosz says that if we are to explain visual appearances of the physical solids we need to assign them definite, ideal shapes. For similar reason, if we are to explain interpretability of expressions in various contexts we need to assign them ideal meanings. Thesis 1 is formulated even more explicitly in [Stanosz, 1974b] where the logical theory of language is considered as a part of the formal theory language whose main goal is to explain human linguistic communication.[3]. The same idea is formulated in [Suszko, 1957b, Suszko, 1957a] (reprinted in [Suszko, 1998b]) where one can find a clear and simple exposition of the assumptions involved in modelling human linguistic communication by means of logical semantics. We shortly present this account as no English translation is available, and

[3]An independent value of [Stanosz, 1974b] is the presentation of Chomsky's linguistics and comparing it with the logical theory of language, especially with semantic theory proposed by [Katz, 1972].

as the exposition is straight to the point.

Suszko claims that the communicative functioning of language of a given group of people is explained by the fact that the expressions of the language have some intersubjective meanings for the members of the group. The existence of such meanings is tightly coupled with the activity of individuals, especially with the interaction between them and the surrounding reality (unfortunately, Suszko does not develop the idea about the role of individuals any further). It is ascertained that the fundamental and necessary condition of the communicative function of language, and thus of intersubjectivity of the meaning of its expressions, is a fixed semantic function of language. Suszko explains this in a more detail as follows. Let L be a formal language which serves as an abstract formalisation of the natural language of a given community A. Suszko assumes there is a model $\mathbb{M}$ of L such that each expression of L is assigned by $\mathbb{M}$ exactly the same semantic function as it has in the real communication between the members of A.[4] Suszko proposes to refer to such a model as to the proper model of L for the language community A. It is obvious that such a reconstruction of the human linguistic competence, assumed to be shared by all members of A, implies Thesis 1 as this is exactly how the interpretation works in model theory: each extra-logical symbol is unequivocally paired with its fixed denotation. Now, following [Stanosz, 1974b], one knows what information is conveyed by a sentence because one knows the interpretation of extra-logical expressions, the truth value of atomic sentences and how the truth value of compound sentences depends on the truth value of their components.

Another exposition of Thesis 1, accompanied with a direct linguistic motivation, may be found in [Vasiliu, 1972] where the author merges transformational grammar [Katz and Fodor, 1963] and symbolic logic [Carnap, 1958] in order to give a general form of a semantic theory of some fragments of natural language.[5] A system of this kind includes a formalized language which is understood as a translation of a given fragment of natural language. Denotational rules give the meaning of extra-logical symbols. Vasiliu, following [Carnap, 1947], says the rules in question yield a mapping from extra-logical symbols to natural language expressions. Given such rules each expression and linguistic construction has assigned a well deter-

[4]For an exposition of the rudiments of model theory, see, e.g., [Mendelson, 2009].

[5]The Polish reader may benefit if she consults [Vasiliu, 1981].

mined meaning. For instance, let H, L be 1- and 2-place predicates and x, y individual variables. Examples of such rules when applied to English, may be as follows: the logical form of *x is a human* is $H(x)$ whereas the logical form of *x loves y* is $L(x, y)$. These rules give the interpretations of the predicates H and L. Then the logical form of *Everybody loves somebody* is $\forall x(H(x) \Rightarrow \exists y(H(y) \wedge L(x, y)))$. Another type of semantic rules is concerned with the definition of truth, developed in the spirit of [Tarski, 1933] which enable speakers to derive the truth values of compound sentences (so, for example, the sentence *Everybody loves somebody* is true if and only if its logical form is a true sentence under the interpretation given by the logical forms of the predicates H and L). It seems Vasiliu assumes that everyday language expressions have well defined systematic meanings which he is able to ascribe to the extra-logical symbols of a formal language. This is quite in accordance with the intended meaning of Thesis 1.

Let us mention also some remark form [Katz, 1966]. Katz considers a situation of conveying a message by the speaker. A message is «encoded in the form of a phonetic representation of an utterance by means of the system of linguistic rules with which the speaker is equipped». The speaker vocalizes an utterance which is picked up by the hearer who recognizes the phonetic representation. This representation is decoded into the same message which was intended to convey by the speaker. «Hence, because the hearer employs the same system of rules to decode that the speaker employs to encode, an instance of successful linguistic communication occurs.»

We summarize this subsection with some remarks from [Dummett, 1975]. Dummett – recapitulating Frege's arguments in favour of the claim that the knowledge of the language by the speaker is better explained in terms of sense and not reference – gives a conditional statement of Thesis 1: «If we suppose that an account of the use of language in communication demands that each sentence possesses a common cognitive content for all speakers, then this argument does provide a ground for ascribing to each expression a sense constant from speaker to speaker». The advantage of this formulation is that it provides some ground for a description of a plausible cognitive equipment of language users, an equipment which enables them to communicate. We shall get back to this issue when introducing algorithmic theory of meaning which equates the sense of an expression with a procedure for computing its denotation [Tichy, 1969]. This concept is briefly introduced in the next section and extensively employed in on many occasions throughout this dissertation.

I.3 Learning by Recognizing

Thesis 1 provides a simple explanation about how communication in natural language is possible. However, it leaves aside a more foundational problem of how the commonality of semantics may actually be achieved. A natural approach is to appeal to the concept of learning and show how language users may adopt the semantics from other speakers and communicate successfully later on. Traditionally, language learning is conceived as a process of adopting an externally imposed standard. In the case of semantics, this process is usually modelled in terms of a hypothetical learning mechanism which enables a subject to eventually internalize the correct meanings by mere observation of language use. This type of learning, which we call learning by recognizing, gives an answer about how the commonality in semantics may be achieved.

Sometimes learning is modelled as an interaction between the learner and the teacher. [6] The teacher knows the semantics to be acquired by the learner and generates linguistic usage based on the predefined semantic standard. An important assumption is that the predefined semantics is fixed throughout learning. The learner, equipped with an appropriate algorithmic mechanism, observes more and more samples of language use and gradually adapts his guesses about the underlying semantics so that eventually he is able to stabilize on the right hypothesis which corresponds exactly to the concept being presented to him. We shall refer to this type of learning as to learning by recognizing. We justify our terminology by the fact that the learner's task is to acquire the ability to recognize the predefined semantic concept.

The research conducted in this vein is largely connected with the algorithmic theory of meaning – a conception according to which the meaning of an expression may be identified with a procedure for computing its denotation [Tichy, 1969, Suppes, 1980]. If we identify the extension of a sentence with the class of (finite) models in which this sentence is true, then the meaning of the sentence, conceived as an algorithm, computes the truth value of

[6] Some models do not mention the teacher. Instead, they employ certain modes of data presentation.

the sentence in a given model. The idea of meaning as algorithm has been widely used and applied in linguistics, philosophy and cognitive science [van Benthem, 1986, van Benthem, 1987, Mostowski, 1998, Mostowski and Wojtyniak, 2004, Moschovakis, 1990, Moschovakis, 2006, Lambalgen and Hamm, 2005, Szymanik, 2016].[7]

Let us give an illustrative example without going into too much details which are postponed to Chapters II and III. Consider simple cardinality quantifiers such as *more than k*, for any given natural number k. For the sake of this example, let us observe that for any given k, the denotation of the quantifier *more than k* may be identified with the class of finite structures – call it $\mathcal{Q}_{>k}$ – of the form (U, R), where R is a unary relation, such that for every such structure $\mathbb{M} = (U, R)$, $\mathbb{M} \in \mathcal{Q}_{>k}$ iff $|R| > k$. Let us provide an intuitive justification of the fact that for each $\mathcal{Q}_{>k}$ there is an intuitive procedure which on input $\mathbb{M} = (U, R)$ outputs the answer to the question whether $\mathbb{M} \in \mathcal{Q}_{>k}$. An informal justification of this fact is as follows. Fix k. Use a number variable n for counting. Initially $n := 0$. Look through all elements of U one by one. Each time you see an element belonging to R, increment n. After exhausting the whole (finite) universe, output *yes*, if $n > k$, and *no* otherwise.

The concept of meaning as algorithm suits well in the context of learning. Usually, the idea is that the learner is an algorithmic procedure which is fed with incoming samples of language use and outputs hypotheses which are often also algorithmic procedures (or some other computational formalisms such as grammars or automata). To translate this into the aforementioned cardinality quantifiers, let the hypotheses available to the learner be various algorithms for computing such quantifiers. Fix a cardinality quantifier $\mathcal{Q}$ –this is going to be the predefined semantic concept to be acquired by the learner. Incoming samples of language use are models $\mathbb{M} = (U, R)$ such that $\mathbb{M} \in \mathcal{Q}$. This is so-called positive presentation (we employ this mode of presentation for simplicity reasons). The learner observes more and more samples. To learn the concept being presented, the learner may use the following algorithm. Let $(U_1, R_1), \ldots, (U_t, R_t)$ be the sample observed so far. Keep as a current hypothesis the algorithm which computes $\mathcal{Q}_{>k}$, where $k = min\{|R_i| : i = 1, 2, \ldots, t\} - 1$. Of course, if each $\mathbb{M} \in \mathcal{Q}$ is presented at some point, the learner will stabilize on the correct hypothesis.

[7] This idea goes back to [Frege, 1892].

There have been many approaches to learning which fall within this category. They are mostly inspired by the identifiability in the limit framework, introduced in [Gold, 1967]. The research conducted in this and related frameworks has provided us with a plethora of interesting models of learning as well as theoretical and practical insights (see, e.g., [Tiede, 1999, Costa Florêntio, 2002, Gierasimczuk, 2007, Clark, 2010, Piantadosi et al., 2012]).

Observe that learning by recognizing is tightly connected to Thesis 1. One may hypothesise that the overall process of learning semantics within a community consists in acquiring by unskilled speakers the correct semantic conventions from competent language users. This concept of learning roughly corresponds to our intuitive idea of what does it mean to learn language. It simply consists in acquiring the skill of the right usage of linguistic expressions by observing competent speakers which represent the standard of right and wrong in semantics. Learning by recognizing may be an adequate model of learning natural language semantics if applied to scenarios which satisfy the underlying assumptions, namely a clear distinction between learners who are to adapt and teachers who hold the uniform semantic standard.

I.4 Synchronic Semantic Variation

In this section we present some evidence against Thesis 1. We show that the semantics of a given language community at a given time may not be uniform. Non-uniformity of semantics may be understood as the existence of different speakers that use significantly different meanings for the same expressions in the same context. We refer to this phenomenon as interindividual synchronic semantic variation. Another possible variation, referred to as intrandividual, occurs at the level of one language user – on different occasions the same speaker uses the same expression in exactly the same context in non-equivalent ways.

First we observe that the discussion on the meaning of the Hintikka sentence undermines the reliability of linguistic intuitions in this particular case. We mention a more promising methodology, namely experimentation with linguistic usage which overcomes (at least to some extent) the problems linguistic intuitions. Next, we give some theoretical support for the hypothesis that semantic change implies synchronic variation [Kay, 1975] Lastly, we discuss experiments in color naming which show that even such basic expressions reveal a surprising within-language semantic diversity.

I.4.1 Problems with Linguistic Intuitions

The discussion on the meaning of the Hintikka sentence begins with the so called Hintikka thesis [Hintikka, 1973] according to which an adequate interpretation of sentences such as *Some relative of each villager and some relative of each townsman hate each other* requires branching quantification (Hintikka considers also other examples).[8]. The Hintikka thesis has provoked a lively debate in philosophical and linguistic journals (see, e.g., [Guenthner and Hoepelman, 1974, Gabbay and Moravcsik, 1974, Fauconnier, 1975, Stenius, 1976, Hintikka, 1976, Barwise, 1979]). To get the gist of the discussion see [Barwise, 1979, Mostowski, 1994, Mostowski and Wojtyniak, 2004].

The debate is concerned with the correct logical form of the Hintikka and Hintikka-like sentences. The discussion reveals that linguistic intuitions of various authors diverge as they assign different, non-equivalent logical forms to the same linguistic construction. Moreover, many authors seem to adhere to the claim that there is one correct interpretation of the discussed sentence. Those reports, when taken together, are inconsistent. Certainly, they cannot be all true. Hence, the question is which of them are true. If we abandon the view that the Hintikka sentence have one systematic meaning across native speakers then the reported intuitions provide an evidence of within-language semantic variation (and evidence against Thesis 1). If we are to save the uniqueness of meaning across speakers then some reports about the logical form of the Hintikka sentence are false. It seems that there is little hope for advancement in this matter by further appeal to mere intuition. The reliability of this source of knowledge seems to be limited in this case.

I.4.2 Testing Linguistic Usage

Interestingly, the discussion on the meaning of the Hintikka sentence gives rise to a more direct empirical approach to the problem. [Barwise, 1979] reports on an experiment which consists in performing a model-checking

[8] Branching quantifiers [Henkin, 1961] strengthen the expressive power of elementary logic. In the early seventies, the so called first-order thesis – according to which natural language sentences have logical forms expressible in elementary logic – was considered reasonable. Extensive defence of this thesis may be found in [Quine, 1970] Now, this thesis is rather commonly rejected.

task by individual human subjects. An instance of the task is composed of two elements: the Hintikka sentence and a graphical representation of a finite model with a precise interpretation of non-logical symbols. Instances of the task are given to human subjects who are asked to answer whether a given sentence is true in a given model. The Barwise test is an important step forward as it bypasses the appeal to intuitions and relies on a different cognitive task: the truth-value judgement (verification).

According to Barwise, the majority of participants of his experiment accept the Hintikka sentence in models where only the linear interpretation of the Hintikka sentence is true, whereas the branching interpretation is false. Hence, the test shows that there are different language users which output different truth values for exactly the same task. The question is what are the sources of the lack of complete unanimity. One of the possible explanations is that differences result from processing errors during verification. Observe that this explanation renders it plausible that the meanings are actually shared. Another possible explanation is simple as that – the meaning of the Hintikka sentence is not fully shared. In other words, there are individuals who ascribe different meanings to the sentences under investigation. If this explanation is correct the strong reading of Thesis 1 must be rejected. Interestingly, based on his experiment, Barwise concludes that the Hintikka thesis is false. The reasoning behind this conclusion seems to be as follows. We are satisfied with high unanimity between informants (not necessarily a full one). In other words, we allow for negligible amount of deviations. Sufficient amount of informants outputs answers incompatible with the Hintikka thesis. Hence, the thesis is false.[9] Formulated in this way, the conclusion drawn by Barwise is still in accordance with Thesis 1. Nevertheless, the results of the Barwise test render Thesis 1 a bit doubtful and certainly open new ways in empirical investigations of semantics.

The first hypothesis that the semantics of the Hintikka sentence may be ambiguous (contrary to previous intuitions reported in the discussion) was formulated in [Mostowski, 1994]. Mostowski hypothesised that the inferential and referential meaning of the Hintikka sentence may be actually different. This hypothesis was based on the observation that there are good arguments

[9]Further experiments [Gierasimczuk and Szymanik, 2009] show that informants tend to interpret Hintikka-like sentences in a way consistent with a previously unforeseen linear interpretation (so-called two-way interpretation). This provides another counter-argument to the Hintikka thesis.

for the Hintikka thesis coming from inferential contexts and – at the same time – experimental evidence against it, coming from verification tasks. This is quite a peculiar intra-individual semantic variation. It depends essentially on the kind of the task a human subject is to perform with a sentence. If involved in inference, she uses one meaning. If performing verification, she employs another one.

Since [Mostowski and Wojtyniak, 2004] and other works on semantic complexity, it has slowly become evident that there are other important factors which may drastically shape the semantics of natural language constructions. Mostowski and Wojtyniak show that the problem of recognizing the truth value of the Hintikka sentence in finite models is NP-complete. Following their complexity results, the authors present some reasonable hypotheses, namely that $P \neq NP$ and that real time cognitive capacities of the human mind are limited to $PTIME$-computable functions (see also [Frixione, 2001, van Rooij, 2008]). Based on this, they conclude that the Barwise test simply could not be solved by human informants if they use semantics compatible with the Hintikka thesis. The results from [Mostowski and Wojtyniak, 2004] lead to an interesting hypothesis, namely that language users may unexpectedly switch between non-equivalent interpretations of the same construction depending on whether they perform inference task or a difficult verification task. A more extensive discussion of this observation, though concerned with a different example, is presented in Chapter II.

I.4.3 Kay's Principle

The claim according to which a within-language semantic diversity actually takes place may be inferred from the synchronic empirical data. This sort of reasoning is in line with the principle formulated by Kay in the context of color terms: «diachronic change implies synchronic variation» [Kay, 1975]. Intuitively, the argument goes as follows. Suppose we have two synchronic descriptions of the same language from two different time periods. Assume some fragment of the language underwent a semantic change between these two periods. Taking a reasonable assumption that a semantic change is not a one-step process but rather a gradual one, we infer that between the two observed periods, language manifested a significant semantic diversity, at least with respect to expressions which eventually changed their meaning.

This idea can be made more precise by investigating the proportion of language users which ascribe a particular meaning to a given expression. For the sake of simplicity, suppose we have two different meanings m_1 and m_2 and one expression that has undergone a semantic change. At time t all speakers of a given community use m_1 whereas later, at some time $t' > t$ all speakers use m_2 instead m_1. Since conventions spread through communication between members of the community, we may posit that the semantic change is a gradual process consisting of adopting the new convention m_2 by more and more language users. We may visualize this process by plotting the proportion of the adherents of the new meaning in moments between t and t'. If the population is large, we can make an idealistic assumption that the resulting plot is a continuous function. Hence, by the intermediate value theorem, for every $p \in [0, 1]$ there is a point in time $s \in [t, t']$ such that the proportion of adherents of m_2 is equal to p. What it actually means is that there is a period of time between t and t' such that the two meanings co-exist resulting in a within-language semantic variation.

I.4.4 Color Terms

A more direct evidence of synchronic semantic variation is presented in [Lindsey and Brown, 2009] where the authors investigate the semantics of color-naming terms in the languages examined during the World Color Survey (WCS) (`http://www.icsi.berkeley.edu/wcs`). WCS database gathers semantic information collected from informants of 110 mostly unwritten indigenous languages. The analysis reveals that there are a few non-equivalent color-naming systems (referred to as motifs) which occur in the lexicons of the WCS informants from all over the world. From our point of view, a more important result is that most of the WCS languages exhibit a significant internal semantic variation. In other words, multiple motifs coexist within majority of each of these languages. In more illustrative terms, given a language with multiple color-naming motifs, there are native speakers of that language who use the same color terms in a significantly non-equivalent manners.

A natural question to ask is what is the reason for such semantic variation. This question cannot be answered with respect to the WCS informants as the experiment was performed a few decades ago. This is why a new experiment has been performed with informants being native speakers of the Somali language. Results are reported in [Brown et al., 2016]. The experiments have revealed analogous within-language semantic variation: «the same names are often used in fundamentally different ways and thus, we infer, have different meanings for different individuals». No significant correlation between the used motif and individual variation in color vision has been observed. The authors have analysed associations between individual motifs and various demographic variables: gender, occupation, geographic location in Somalia, age and the number of years spent in the United States. Only the effects of age and gender turned out to be significant.

What is the reason for this within-language diversity? One explanation, considered in [Brown et al., 2016], is that the color lexicon of Somali undergoes a semantic change. The idea is that at the present stage this fragment of the Somali language is at the intermediate stage of an ongoing process of transformation from one motif to another. This, of course, might be the case, but other explanation may come to mind as well such as the influence of U.S. culture (where the informants have been living for some time) and the existence of different dialects in Somalia, associated with different regions. Brown and colleagues provide convincing arguments against these hypotheses. However, another hypothesis is that different motifs may be associated with other factors (for example social ones). However, yet another hypothesis is that the semantics of color terms need not be uniform to satisfy the communication demands of the society and thus the lexicon may divide into a few non-equivalent ones.

I.5 Abandoning Uniform Semantics

As we have shown in the previous section, there is theoretical and empirical evidence of significant variation in speakers' responses resulting from referential and inferential usage of various natural language expressions. More importantly, this variation occurs when the expression and corresponding context remain fixed across different instances of usage.

One may try to resolve this inadequacy by the appeal to the idealistic character of Thesis 1 and claim that it arises due to mistakes made by language users who engage in communication, or some other factors, for example pragmatic ones (as briefly discussed in Section I.2). However, this seems unlikely as the experimental settings seem to be constructed in such a way that they reduce the influence of other factors to the minimum (see, e.g. [Brown et al., 2016]). Moreover, the uniform semantics account lead to weird consequences if we consider situations of the ongoing semantic change. Take, for example, an argument presented in Section I.4.3. If the initial semantics m_1 is conceived as the correct semantic standard at time t then at time t' it is completely mistaken. We are not likely to say that at time t' all language users are mistaken about the right meaning of the underlying expression. A more reasonable view is that the standard has changed and now, at time t', it is m_2. However, in between, i.e. after t and before t', there is actually no correct uniform semantic description that could serve as an idealisation of the semantics of the community.

It seems that the evidence shows that divergent responses of informants come from their adherence to non-equivalent semantic interpretations. This renders the uniform semantics approach inadequate for the description of natural language semantics in situations of semantic variation. Moreover, observe that non-uniform semantics does not have to hinder the communicative functioning of language of the community. It simply may be the case that communication demands require sufficiently common semantics, not necessarily a completely shared one. Hence, we are likely to admit that universally shared semantics is does not have be considered as a distinctive feature of a language community which successfully employs their language in communication purposes. Communication may still work even if the community manifests some variation. Therefore, we opt to abandon Thesis 1.

Observe that the abandonment of Thesis 1 may be problematic for the approach of learning by recognizing (Section I.3). If a language community manifests significant semantic variation then learners frequently encounter inconsistent samples of language usage. Traditional learning mechanisms are not well suited for handling such data, as they are constructed to operate on samples derived from predefined semantics which does not change during learning. In Section I.3 we have shortly discussed the view according to which the semantic is learned from competent speakers that are supposed to represent the uniform semantic standard. However, this approach is also problematic, as experiments suggest that speakers of the same community which are likely to be considered as equally competent, are subjected to semantic variation in the same sense as less competent or unskilled language users. Moreover, this approach does not take into account communication phenomena which cannot be captured by the teacher-learner distinction, for example interactions between the speakers who consider themselves equally authoritative and thus may tend to learn from each other.

I.5.1 Uniform Semantics as an Approximation

Intuitively, we are tempted to say that uniform semantics actually provides an instantaneous picture of natural language semantics. It should be in principle possible to say something more about the accuracy of this picture.

The experiments reported in [Barwise, 1979,Gierasimczuk and Szymanik, 2009] and other experimental works exploring the semantic content of natural languages (see, e.g., [Schlotterbeck and Bott, 2013], probably the earliest such work is [Lenneberg and Roberts, 1955], a cross-linguistic study concerned with the meanings of color terms in Zuni and English) suggest that what we can realistically hope for is at most a high degree of unanimity or conventionality. Certainly not an absolute degree, as the naive reading of Thesis 1 suggests. The property of high conventionality points to another approach, different from the strong claim according to which the semantics is completely shared: uniform semantics is not to be treated as an idealistic assumption but rather as an approximation. And still, this a hypothesis which may be verified or disproved through experimentation.

Assuming this hypothesis, the approximative version of the Hintikka thesis may be formulated as follows: language users tend to ascribe to the Hintikka sentence the branching interpretation (in referential contexts). Data from the Barwise test [Barwise, 1979] and the evidence for the two-way interpretation [Gierasimczuk and Szymanik, 2009] seem to legitimate the rejection of the approximative version of the Hintikka thesis and support the hypothesis that the referential semantics of the Hintikka sentence may by approximated by a linear reading.

To put our view in more general terms, uniform semantics is sometimes a useful theoretical approximation of the actual state of the language of a given community. From a purely theoretical point of view, it seems very unlikely for a community to enter a pure steady state understood as semantics shared by every language user. At the other extreme, to little uniformity in semantics would hinder the communicative efficacy of language, contrary to our linguistic experience. It seems that a realistic picture is away from these two extremes and lies somewhere in between and may be characterized by the sufficient amount of semantic commonalities among the speakers. This is how we understand the claim that a synchronic account of language may serve as an approximation.

I.6 Learning by Coordination

The conclusion which arises from our considerations is that semantics (or more broadly – language) of a given community is better viewed not in terms of a universally shared semantics but rather as an system of individual semantic functions of particular language users who – driven by the need for effective communication – continuously adapt their current individuated semantics by taking into account the experience which they obtain through interaction with other members of the community. Within this picture, semantics may be viewed as an emergent property of the system, a property which by no means has to take the form prescribed by any predefined semantic standard. In fact, one may think of coordination as if there were many competing meanings some of which may dominate the population and result in a widespread convention.

The idea of language as convention is not new to the philosophical inquiry. For example, it has been extensively presented in [Lewis, 1969]. Lewis conceives language as a mapping from expressions to meanings. There are many possible languages satisfying the communicative needs of a given community. In game-theoretical terms, members of the community are faced with a coordination problem – there are many solutions (languages) which actually solve the underlying problem of communication.[10] The Lewis idea of language as convention seems accurate in positing that language conventions may be viewed as solutions to the communicative coordination problem. There have been some doubts whether the problem of arriving at common semantic conventions may be handled by the scientific method [Davidson, 1986,Chomsky, 1986]. However, recent developments in evolutionary linguistics (which we shortly present at the end of this chapter) and the approach presented in this dissertation (Chapter III) seem to vindicate the claim that we may approach to this problem in a systematic and rigorous way.

We refer to our approach as to learning by coordination. Its main assump-

[10] To illustrate this idea on a simple example, consider a traffic coordination problem: how to move to maintain flow and avoid collisions? Driving on the right is one solution adopted in Poland. However, there is another solution, namely driving on the left, which has been adopted in the United Kingdom.

tion is the abandonment of the claim that there is a distinguished, externally imposed semantic standard. To account for linguistic communication within community which is allowed to manifest semantic variation, we need a different leaning mechanism, capable of handling inconsistent responses from other speakers and various degrees of social influence that speakers may exert on each other. The mechanism is to be construed in such a way that it allows for adaptation of individuated meanings and thus makes efficient communication possible. Learning by coordination seems to shift the "logic" of uniform semantics to the realm of individual cognitive strategies which, fuelled with experience obtained via interaction with other speakers, drive the overall dynamics of interaction to a state which allows for efficient communication.

The approaches similar to ours have been developed since around the nineties of the twentieth century and are usually referred to as cultural language evolution or – more broadly – evolutionary linguistics. According to a recent survey [Steels, 2011] contemporary research in the subject may be classified in two categories. One encompasses theories developed within the paradigm of biolinguistics which investigate language in connection with our physical, biological, and ecological embodiment. The other one includes explanations which conceive language as a system shaped mainly through various cultural forces that come into play during online communication, learning and language processing. The first successful scientific endeavours in the field came from artificial intelligence. For example, Steels has formulated a hypothetical agent-based mechanism for the formation and coordination of spatial terms [Steels, 1995]. Simulations show that the activity of communicating software agents which adapt according to the specified mechanism results in the emergence of common lexicons. A major insight from this study is that language may be viewed as a self-adaptive system. In [Steels and Belpaeme, 2005] we find an approach to coordination of color terms which is based on an ingenious formalisation and application of the idea of language games and representation of color categories in terms of radial neural networks.

Other types of language evolution models is iterated learning (see [Kirby et al., 2014] for a recent survey). Iterated learning paradigm is similar to learning by recognizing in the respect that both are concerned with learning from the teacher. However, there is a substantial difference between these two approaches. Learning by recognizing is primarily concerned with the mechanisms of learning which enable a learner to fully accommodate the semantics of the teacher. In contrary, iterated learning purposely allows for partial learning – the learner is exposed to linguistic data which is far from complete. So the model is constrained to account for the poverty of stimulus. However, iterated learning goes even further and arranges teachers and learners in succession, where the learner from a given iteration becomes a teacher in the next iteration. Running such iterated simulations for several generations allows for observing various properties interesting from the evolutionary linguistics perspective. For example, one of the first results obtained within this paradigm was the demonstration that compositionality may emerge through such an iterated learning process [Kirby, 2000].

Recently, there has been a growing interest in a more experimental approach which was initiated in [Galantucci, 2005]. This empirical approach to investigating the emergence of novel forms of communication is referred to as experimental semiotics. The idea is to gather human subjects in laboratory where they are in a position to communicate with each other in the absence of any pre-established communicative conventions. This sort of experimental evidence is an ideal source of knowledge which may fill the gap between mathematical or computation modelling and natural phenomena. For more details, we refer the reader to a recent survey [Galantucci and Garrod, 2011].

II

Complexity of the Barwise Sentence and Similar Natural Language Constructions

In this chapter we elaborate on selected logical and computational complexity properties of some everyday language constructions. Our examples are based on the construction presented in [Barwise, 1979]:

The richer the country the more powerful some of its officials. (1)

Hereinafter, (1) is referred to as the Barwise sentence. Barwise observes that (1) expresses some way of embedding one ordering to another. Our constructions are analogous in this respect: they state the existence of various similarities between partial orderings, such as homomorphism, embedding and variations of those.

Semantically, we analyse the constructions as polyadic generalized quantifiers which are conceived as appropriate classes of models (see Definition 2 in Section II.1). The quantifiers which correspond to our constructions are referred to as similarity quantifiers. Our motivation is to provide for them some initial logical classification. We pursue this task by asking about the definability of those properties in elementary logic over finite models. Our work in this respect contributes to investigations of polyadic quantification in natural language [van Benthem, 1989, Keenan, 1992, Keenan, 1996, Peters and Westerståhl, 2006, Szymanik, 2010]. Our examination starts with some negative definability results which extend the theorem provided in [Barwise, 1979], according to which homomorphism is not expressible in elementary logic over arbitrary models.

Further questions are concerned with semantic complexity. Our aim to establish which constructions are tractable. We pursue this line of investigation in the spirit of algorithmic theory of meaning and semantic complexity of natural language quantifiers [Szymanik, 2016]. We show that the homomorphism quantifier is tractable, whereas other quantifiers under investigation are NP-complete. Our results on the complexity of similarity quantifiers supplement the work on complexity of polyadic quantification (see, e.g., [Szymanik, 2010]).[1]

Finally, we discuss how our results relate to empirical investigations about the influence of semantic complexity on human comprehension. We also make some remarks about the interplay between semantic complexity of natural language and evolutionary linguistics.

***

The chapter is organized as follows. In Section II.1 we define some polyadic generalized quantifiers that appear in everyday language and express the existence of various kinds of similarities between partial orders. The definitions are followed by examples of natural language sentences that are deemed to express the desired properties. In Section II.3 we study definability of similarity quantifiers over finite models. In Section II.4 we provide the complexity classification. Section II.5 revolves around a more philosophical discussion and the influence of complexity pressures on semantic preferences of human subjects. Finally, in Section II.6, we make some remarks on the interplay between computational complexity, human cognition and evolutionary linguistics.

[1]Generalized quantifiers were first investigated from the computational perspective in [Blass and Gurevich, 1986].

II.1 Similarity Quantifiers

Let $\sigma = (A, <_A, B, <_B)$ and $\sigma_R = (A, <_A, B, <_B, R)$ be relational vocabularies, where A, B are 1-place and $<_A$, $<_B$, R are 2-place predicates. We interpret the symbols so that $(A, <_A)$, $(B, <_B)$ are strict partial orders on A and B respectively, and R, called the coupling relation, satisfies $R \subseteq A \times B$ (observe that the required properties are easily expressible in FO). We refer to such σ-models and σ_R-models as to double partial orders and coupled partial orders, respectively. In the present work, we consider models with finite universes. This restriction seems reasonable in applications to natural language [Westerståhl, 1984].

One partial order may be similar to another one in various ways. The similarities we consider are: homomorphism, 1–1 homomorphism and embedding. The following second-order σ-sentences express the existence of homomorphism, 1–1 homomorphism and embedding from $(A, <_A)$ to $(B, <_B)$:

$$\exists f : A \to B \; \forall x, y \in A \; (x <_A y \Rightarrow f(x) <_B f(y)) \tag{2a}$$

$$\exists f : A \overset{1-1}{\to} B \; \forall x, y \in A \; (x <_A y \Rightarrow f(x) <_B f(y)) \tag{2b}$$

$$\exists f : A \overset{1-1}{\to} B \; \forall x, y \in A \; (x <_A y \Leftrightarrow f(x) <_B f(y)) \tag{2c}$$

Definition 2 *Let $t = (n_1, n_2, \ldots, n_k)$ be a k-tuple of positive integers, $k \geqslant 1$. A Lindström generalized quantifier of type t is an isomorphism-closed class $\mathcal{Q}$ of structures such that if $M \in \mathcal{Q}$ then $M = (U, R_1, R_2, \ldots, R_k)$, where U is the universe and R_i is an n_i-placed relation on U, for $i = 1, 2, \ldots, k$. A generalized quantifier of type $(n_1, n_2, \ldots, n_k)$ is called polyadic, if $n_i > 1$, for some $1 \leqslant i \leqslant k$.*

Observe that the classes of (finite) double partial orders in which the (2a), (2b) and (2c) conditions hold are isomorphism-closed. Hence, the respective classes are generalized quantifiers, hereinafter denoted by $\mathcal{H}$, $\mathcal{H}^{1-1}$, $\mathcal{E}$, accordingly. The type of these quantifiers is $(1, 2, 1, 2)$.

The concept of a coupled partial order is richer. The coupling relation $R \subseteq A \times B$ makes it possible to introduce new types of similarities between partial orders. In what follows, we define similarities that are restricted by the coupling relation. Given a coupled partial order, let R_a denote the set $\{b \in A : R(a,b)\}$. R induces an indexed family of sets $\{R_a\}_{a \in R}$. Let us formulate the following requirement: for each $a \in A$, the similarity function is not allowed to assume values outside R_a. In other words, we require that the similarity function, say f, is such that each $a \in A$ is in the relation R with the value $f(a)$. The following second-order σ_R-sentences express the existence of homomorphism, 1–1 homomorphism and embedding that are restricted by the coupling relation:

$$\exists f : A \to B \; \forall x, y \in A \; [R(x, f(x)) \wedge (x <_A y \Rightarrow f(x) <_B f(y))] \tag{3a}$$

$$\exists f : A \stackrel{1\text{--}1}{\to} B \; \forall x, y \in A \; [R(x, f(x)) \wedge (x <_A y \Rightarrow f(x) <_B f(y))] \tag{3b}$$

$$\exists f : A \stackrel{1\text{--}1}{\to} B \; \forall x, y \in A \; [R(x, f(x)) \wedge (x <_A y \Leftrightarrow f(x) <_B f(y))] \tag{3c}$$

The classes of (finite) coupled partial orders in which (3a), (3b), (3c) hold are isomorphism-closed and thus are generalized quantifiers. We denote them by $\mathcal{H}_F$, $\mathcal{H}_F^{1-1}$ and $\mathcal{E}_F$, respectively. The $_F$ subscript comes form the first letter of the word *family*. The family we have in mind is $\{R_a\}_{a \in A}$.

Additionally, we consider a natural condition that may be imposed on R. We refer to it as to the disjointness condition. It says that $R_a \cap R_b = \emptyset$, for every $a, b \in A$ such that $a \neq b$. This is expressed by

$$\forall x \forall y \forall x' \forall y' (x \neq x' \wedge R(x,y) \wedge R(x',y') \Rightarrow y \neq y') \tag{disjointness}$$

If each of the conditions (3a), (3b), (3c) is taken in conjunction with the disjointness requirement, we obtain the generalized quantifiers which we denote by $\mathcal{H}_{DF}$, $\mathcal{H}_{DF}^{1-1}$ and $\mathcal{E}_{DF}$, accordingly. The $_{DF}$ subscript comes form the first letters of the words *disjoint family*. The type of this quantifiers is $(1,2,1,2,2)$. Trivially, $\mathcal{H}_{DF} = \mathcal{H}_{DF}^{1-1}$. Observe that $\mathcal{E}_{DF}$ might as well be defined by the conjunction of the disjointness condition and

$$\exists f : A \to B \; \forall x, y \in A \; [R(x, f(x)) \wedge (x <_A y \Leftrightarrow f(x) <_B f(y)] \tag{4}$$

II.2 Examples

We provide everyday language sentences which are interpretable as polyadic generalized quantifiers defined in Section II.1. We build upon the Barwise sentence (see Example 1) [Barwise, 1979]. In the more purely linguistic literature the structures involved in such sentences are called comparative correlatives [Culicover and Jackendoff, 1999, Den Dikken, 2005], comparative conditionals [McCawley, 1988, Beck, 1997] and correlational comparatives [Keenan and Ralalaoherivony, 2014]. Obviously, some instances of this construction are first-order:

The richer the country the more powerful its ruler. (5)

Let A stand for countries and B for rulers. Countries are ordered by being richer $<_A$. Rulers are ordered by being more powerful $<_B$. Let f denote the function which assigns rulers to countries. The logical form of (5) is

$$\forall x, y \in A \ (x <_A y \Rightarrow f(x) <_B f(y)). \tag{6}$$

If each country has exactly one ruler, (6) is an obvious interpretation. Otherwise, (6) is acceptable if f, denoted by the noun phrase *its ruler*, is previously mentioned or easily identified. Now, consider *officials* instead of *rulers*. Each country has most likely more than one official. Additionally, change the number of the noun phrase in the second clause to plural:

The richer the country the more powerful its officials. (7)

Interpret A and $<_A$ as previously. Let B stand for officials. Officials are ordered by being more powerful $<_B$. Let $R(x, y)$ mean that y is an official of x. Assuming that countries have many officials, the relation $R \subseteq A \times B$ is one-to-many. Given some additional information, one may read (7) as (6), where the expression *its officials* corresponds to the already known function f which assigns officials to countries. If no such information is available, the meaning of (7) is rather ambiguous. Perhaps, the most natural interpretation of *its officials* is *all of its officials.* Instead of *all* we may use other quantifier determiners such as *some* or *most.* We restrict our analysis to *some.* Further examples use *can* to expresses the possibility of certain arrangement of objects.

Homomorphism quantifiers

A plausible reading of (1) says we may assign each country one of its officials in such a way that the richer the country the more powerful is the assigned official. Logically speaking, (1) reads similarly to (6) except that f is not a function symbol, but a second-order functional variable, quantified existentially. Additionally, we add the explicit use of R because for any given country a, f is allowed to assume values only from R_a. Assuming countries have disjoint sets of officials, the logical form of (1) is (3a) taken in conjunction with the disjointness condition. Hence, the semantics of (1) is interpretable as the quantifier $\mathcal{H}_{DF}$.

To drop the disjointness condition, consider the following sentence:

The wiser the professor the smarter some of his students. (8)

Let A stand for professors, B for students and $R \subseteq A \times B$ for the relation of being professors' student. Professors are partially ordered by being wiser $<_A$. Students are partially ordered by being smarter $<_B$. A plausible reading of (8) says we may assign each professor one of his students in such a way that the wiser the professor, the smarter is the assigned student. It is perfectly possible that two professors have the same students, so the chosen relation does not force the disjointness of $\{R_a\}_{a \in A}$. Ideally, $\{R_a\}_{a \in A}$ could be any family of sets of students whatsoever. Hence, the semantics of (8) is interpretable as the quantifier $\mathcal{H}_F$.

To get rid of R and allow the similarity function to assume its values from the whole B, consider the following sentence:

The smarter the student the wiser some of the professors. (9)

Let A stand for students in one high school (or university) group, B for professors of this group. Having such a domain of interpretation, assume that students are partially ordered by being smarter $<_A$. Professors are partially ordered by being wiser $<_B$. A plausible reading of (9) says one can assign each student a professor in such a way that the smarter the student, the wiser the assigned professor. Hence, the semantics of (9) is interpretable as the quantifier $\mathcal{H}$.

Injective homomorphism

Two people cannot buy the same product at the same time. The relation of buying is thus appropriate for obtaining injective homomorphism between two sets of objects: consumers and products. Consider the following sentence.

The richer the consumer, the better some of the products he can buy. (10)

Let A stand for consumers, B for products. Consumers are partially ordered by being richer $<_A$. Products are partially ordered by being better $<_B$. A plausible reading of (10) says each consumer can be assigned a product in such a way that the assignment is injective and the richer the consumer, the better is the assigned product. Hence, the semantics of (10) is interpretable as the quantifier $\mathcal{H}^{1-1}$. To obtain the quantifier $\mathcal{H}_F^{1-1}$, we simply add some relation R which may arbitrarily connect consumers and products:

The richer the consumer, the better some of his favourite products he can buy. (11)

Embedding quantifiers

Switching from homomorphism to embedding consists on enforcing injectivity and preserving the ordering in two directions. One approach is to use anaphoric expressions - imagine that the example below is used by some of parents of a member of a high school group (during some kind of a parental small talk), and refers to the group and their teachers:

The smarter the student, the wiser are some of the professors and whenever these professors are wiser, the students are smarter. (12)

If the anaphoric expression *these professors* is interpreted as the professors selected by the similarity function and *the students* as referring to the arguments of this function, then the semantics of (12) is $\mathcal{E}$. To obtain $\mathcal{E}_F$, apply this trick to sentences like (11). To get a natural example of $\mathcal{E}_{DF}$, consider a sentence like (11), where the coupling relation induces a disjoint family.

Anaphora allows a wide range of quite natural examples which may result in other quantifiers, not formally defined in this paper:

The wiser the professor, the smarter are some of his students, and whenever these students are smarter, the professors are wiser. (13)

This sentence does not require that the similarity function is injective. A reasonable reading says there is function from professors to students which respects the ordering in both directions and such that for any given professors the value of the function belongs to the set of his students.

There is another way to guarantee that the ordering is preserved in both directions. In some contexts, the adverb *exactly* is used for expressing the fact that the information we are giving is precise. For example, *exactly* may express the equivalence when used in the context of implication. When we say that the objects of type A are the objects of type B, we say that for every object x, if x is of type A then x is of type B. However, if we say that objects of type A are *exactly* the objects of type B, we mean that for every object x, x is of type A if and only if x is of type B. Since two-way order preservation is essentially expressed by the equivalence, we may simply use *exactly* to obtain $\mathcal{E}$:

Those consumers, who can buy better products, are exactly the richer ones. (14)

To get $\mathcal{E}_{DF}$, we use the disjointness condition to enforce injectivity:

Those countries, some of whose officials are more powerful, are exactly the richer ones. (15)

One may dispute whether the logical form of (15) is precisely (3c) in conjunction with the disjointness condition. The explicit requirement that the similarity function is injective (as it is expressed in (3c)) seems to strong. However, the injectivity is enforced by the coupling relation $R \subseteq A \times B$ which induces a disjoint family $\{R_a\}_{a \in R}$. The semantics we obtain is extensionally the same as $\mathcal{E}_{DF}$. To get $\mathcal{E}_F$, we cannot relay on the disjointness condition:

$$\textit{Those professors, some of whose students are smarter, are exactly the wiser ones.} \tag{16}$$

The semantics of (16) is not $\mathcal{E}_F$. (16) does not require that the similarity function is injective. For example, consider a model consisting of two professors incomparable with respect to their wisdom. Assume there is only one student who is taught by the two professors. (16) is true in such a model. However, this model does not belong to $\mathcal{E}_F$. To obtain $\mathcal{E}_F$, we use a statement which requires injectivity:

$$\textit{Those consumers, who can buy better products they like, are exactly the richer ones.} \tag{17}$$

Let A, B, $<_A$ and $<_B$ be interpreted as in the case of (10). Let $R \subseteq A \times B$ connect consumers with the products they like. A plausible reading of (17) says each consumer can be assigned a product he likes in such a way that the assignment is injective and the ordering is preserved in both directions. Hence, the semantics of (17) is interpretable as the quantifier $\mathcal{E}_F$.

Linear readings

It is perfectly possible that some of the examples from this section, if not all of them, allow other interpretations. For example, (1) may be read in at least one of the following ways:

$$\forall x, y \in A\, \exists z, w \in B\, [x <_A y \Rightarrow (z <_B w \wedge R(x,z) \wedge R(y,w))] \tag{18a}$$

$$\forall x, y \in A\, \exists z, w \in B\, [R(x,z) \wedge R(y,w) \wedge (x <_A y \Rightarrow z <_B w)] \tag{18b}$$

We may assign similar logical forms to other examples. The difference between second- and first-order readings becomes apparent in the model from Figure II.1. $A = \{a_1, a_2, a_3\}$ are countries, solid arrows represent the ordering by richeness $<_A$ (the richest country is a_3). Officials are represented by natural numbers enclosed in frames. The ordering by power $<_B$ is understood as the standard ordering of natural numbers. Dashed lines connect countries with their officials. In this model (3a) is false, whereas (18a), (18b) are true.

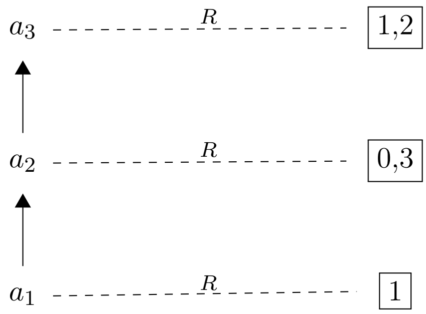


Figure II.1: Second-order vs first-order interpretation.

What about the difference between (18a) and (18b)? Of course, (18b) implies (18a) but not the other way round (in the model with two countries a_1, a_2 such that a_1 does not have any officials and $<_R$ is empty, (18b) is false whereas (18a) is true). Perhaps this simple example is appropriate to show that, to all appearances, our intuitions may be inconsistent. One the one hand, one may argue that *the richer the country* is the false antecedent of the implication and thus the entire sentence (1) is true (reading (18a)). On the other hand, one may believe that (1), when read in its entirety, presupposes that every country has an official. It seems that our intuition is not the best advisor in this respect.

II.3 Definability

In this section we study the definability of the generalized polyadic quantifiers introduced in Section II.1. We show they are not FO-definable over double and coupled partial orders. We use two techniques from finite model theory: reductions and locality [Libkin, 2004].

In [Barwise, 1979] one may find a model-theoretical proof that homomorphism between partial orders is not FO-expressible over arbitrary models. In what follows, we prove that

Theorem 3 *$\mathcal{H}$, $\mathcal{H}^{1-1}$, $\mathcal{E}$ are not expressible in FO over finite double partial orders.*

Proof: Let $\mathcal{Q} \in \{\mathcal{H}, \mathcal{H}^{1-1}, \mathcal{E}\}$. The proof is by contradiction. Assume $\mathcal{Q}$ is FO-expressible over double partial orders. We show that the parity is expressible in FO over linear orders. This yields a contradiction as the parity is not FO-expressible over linear orders [Libkin, 2004].

Let φ be a first-order σ-sentence expressing $\mathcal{Q}$. Observe that φ contains occurrences of A and B. Let $\tau = (<)$ be a vocabulary of linear orders. We construct an FO τ-sentence ψ which expresses the parity over linear orders. Let x be a variable not occurring in φ. In what follows, $X \rightsquigarrow Y$ means that expression X is to be substituted by expression Y. Construct $\psi_1(x)$ and $\psi_2(x)$ by applying to φ the substitutions rules (19a) and (19b), respectively:

$$\begin{array}{ll} A(y_i) \rightsquigarrow x < y_i & y_i <_A z_j \rightsquigarrow x < y_i < z_j \\ B(y_i) \rightsquigarrow y_i < x & y_i <_B z_j \rightsquigarrow y_i < z_j < x \end{array} \tag{19a}$$

$$\begin{array}{ll} A(y_i) \rightsquigarrow y_i < x & y_i <_A z_j \rightsquigarrow y_i < z_j < x \\ B(y_i) \rightsquigarrow x < y_i & y_i <_B z_j \rightsquigarrow x < y_i < z_j \end{array} \tag{19b}$$

Now, let ψ denote the τ-sentence $\exists x\ (\psi_1(x) \wedge \psi_2(x))$. We claim $\neg\psi$ expresses the parity over linear orders. It suffices to observe that ψ expresses its complement. Observe that ψ says there is an element x such that all elements above x may be homomorphically mapped to all elements below x and vice versa. Since we work in strict linear orders, it simply means that the set of all elements above x has the same cardinality as the set of all elements below x. Hence, together with x, the whole universe is odd. □

Theorem 4 *$\mathcal{H}_F$, $\mathcal{H}_F^{1-1}$, $\mathcal{E}_F$ are not expressible in FO over coupled partial orders.*

Proof: The argument is almost the same as for the Theorem 3. The only difference is that we extend (19a) and (19b) with additional rules $R(y_i, z_j) \rightsquigarrow x < y_i \wedge z_j < x$ and $R(y_i, z_j) \rightsquigarrow y_i < x \wedge x < z_j$, respectively. □

The technique for proving Theorem 3 and Theorem 4 does not apply easily for showing FO-inexpressibility of the quantifiers $\mathcal{H}_{DF}$, $\mathcal{H}_{DF}^{1-1}$, $\mathcal{E}_{DF}$. Instead, we use the Hanf-locality technique. We sketch the methodology and then prove the theorem.

Let ρ be a relational vocabulary. Let M be a ρ-model. The Gaifman graph of M is defined as $\mathcal{G}(M) = (|M|, E)$, where for every $a, b \in |M|$, $E(a, b)$ iff $a = b$ or there is a predicate R in ρ such that for some tuple $\bar{t} \in R^M$, a and b occur in $\bar{t}$. The distance $d_M(x, y)$ is understood as the length of the shortest path from x to y in $\mathcal{G}(M)$. The ball of radius r around $a \in |M|$ is defined as $B_M^r(a) = \{x : d_M(a, x) \leqslant r\}$. The r-neighborhood of a in M is the ρ-structure $N_r^M(a)$ such that:

- the universe is $B_M^r(a)$,
- each k-ary relation R is interpreted as R^M restricted to $B_M^r(a)$, i.e. $R^M \cap (B_M^r(a))^k$.

Now let $\mathbb{A}$ and $\mathbb{B}$ be two ρ-structures. If there exists a bijection $f : A \to B$ such that for every $c \in A$

$$N_d^{\mathbb{A}}(c) \cong N_d^{\mathbb{B}}(f(c)),$$

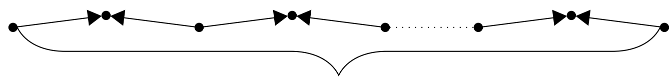
then we write $\mathbb{A} \leftrightarrows_d \mathbb{B}$ which means that the two structures *locally look the same*. We say that a class of ρ-structures Q is Hanf-local if there exists a number $d \geqslant 0$ such that for every ρ-structures $\mathbb{A}$ and $\mathbb{B}$:

$$\mathbb{A} \leftrightarrows_d \mathbb{B} \quad \Rightarrow \quad [\mathbb{A} \in Q \Leftrightarrow \mathbb{B} \in Q].$$

The smallest d for which the above holds is called the Hanf-locality rank of Q. Using Hanf-locality to proving that a class of structures Q is not definable in a logic $\mathcal{L}$ amounts to proving the every $\mathcal{L}$-definable class of structures is Hanf-local and that Q is not Hanf-local. It is known that every FO-definable class of structures is Hanf-local. We use it below. For details on locality, see [Libkin, 2004].

Theorem 5 *The quantifiers $\mathcal{H}_{DF}$, $\mathcal{E}_{DF}$ are not FO-expressible over coupled partial orders.*

Proof: The proof is by contradiction. Let $\mathcal{Q}$ be one of the above quantifiers. Assume that $\mathcal{Q}$ is FO-expressible over coupled partial orders. Hence, $\mathcal{Q}$ is Hanf-local [Libkin, 2004]. Let L_n, for $n > 0$, denote the type of strict partial order in Figure II.2. We call it a bridge. For each $n > 0$, we define the length of L_n, denoted by $|L_n|$, as $2n$. Let d be the Hanf-locality rank of $\mathcal{Q}$. Set



$\rightarrow\!\bullet\!\leftarrow$ repeated n times

Figure II.2: Bridge L_n.

$m = min\{|L_n| : n > 0 \wedge |L_n| > 2d\}$. We define two coupled partial orders $\mathbb{C} = (A \cup B, A, <_A, B, <_B, R)$ and $\mathbb{C}' = (A' \cup B', A', <_{A'}, B', <_{B'}, R')$. We set $|A \cup B| = |A' \cup B'| = 9m$ so that $|A| = |A'| = 3m$ and $|B| = |B'| = 6m$. $(A, <_A)$ is isomorphic with $(A', <_{A'})$ and is easily visualized as a triangle, wherein each face is a bridge of length m. $(B', <_{B'})$ consists of two separate orders $(B'_1, <_{B'}\restriction B'_1)$ and $(B'_2, <_{B'}\restriction B'_2)$, both isomorphic to $(A, <_A)$. Now, $(B, <_B)$ is easily visualized as a hexagon, wherein each face is a bridge of length m. To define the relation R, consider the following labelling of the vertices from A and B. Set $i := 1$ and choose an arbitrary vertex in A (B) with two outgoing edges in the ordering $<_A$ $(<_B)$. This is your current position. $(\star)$ If the current position is not labelled, label it with $a_i(b_i)$. Change your current position to the incident vertex in the clockwise direction, set $i := i + 1$ and go to $(\star)$.

Set $R(a_i, b_j)$ iff $i = j$ or $j = i + 3m$. Observe that $R_a \cap R_{a'} = \emptyset$, for each $a, a' \in A$ such that $a \neq a'$. To define R', consider two isomorphisms $f_1 : (A', <_{A'}) \to (B'_1, <_{B'}\restriction B'_1)$ and $f_2 : (A', <_{A'}) \to (B'_2, <_{B'}\restriction B'_2)$. Set $R'(a, b)$ iff $f_1(a) = b$ or $f_2(a) = b$. Observe that the family $\{R'_a\}_{a \in A'}$ is disjoint. This ends the definition of $\mathcal{C}$ and $\mathcal{C}'$.

Define a bijection $f : A \cup B \to A' \cup B'$ as follows. Label A' with $a'_1, a'_2, \ldots, a'_{3m}$, B'_1 with $b^1_1, b^1_2, \ldots, b^1_{3m}$ and B'_2 with $b^2_1, b^2_2, \ldots, b^2_{3m}$ in the same way as A with $a_1, a_2, \ldots, a_{3m}$ (see previous paragraph). Set $f(a_i) = a'_i$, $f(b_i) = b^1_i$ and $f(b_{3m+i}) = b^2_i$, for $1 \leqslant i \leqslant 3m$.

We prove that $\mathbb{C} \leftrightarrows_d \mathbb{C}'$. Let $c \in A \cup B$. We show that $N^{\mathbb{C}}_d(c) \cong N^{\mathbb{C}'}_d(f(c))$. Overall, there are at most four isomorphism types of d-neighbourhoods of points in $\mathbb{C}$ and $\mathbb{C}'$. The types correspond to the following conditions (by an edge we mean $<_A$-edge, $<_B$-edge or R-edge): *a)* c has 4 out-edges ($c \in A$, c has 2 out-$<_A$-edges and 2 out-R-edges), *b)* c has 2 in-edges and 2 out-edges ($c \in A$, c has 2 in-$<_A$-edges and 2 out-R-edges), *c)* c has 3 in-edges ($c \in B$, c has 2 in-$<_B$-edges and 1 in-R-edge), *d)* c has 1 in-edge and 2 out-edges ($c \in B$, c has 2 out-$<_B$-edges and 1 in-R-edge). Clearly, these are all possible immediate surroundings of an arbitrary point $c \in A \cup B$ in models $\mathbb{C}$ and $\mathbb{C}'$.

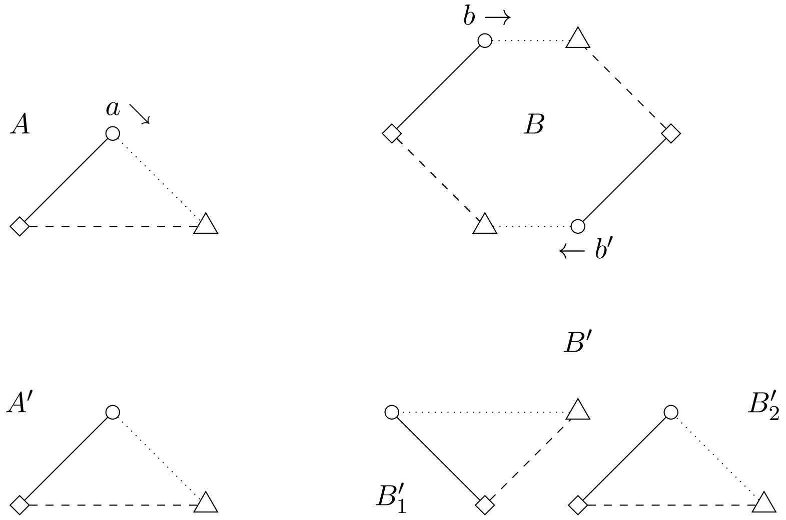


Figure II.3: Coupled partial orders $\mathbb{C}$ and $\mathbb{C}'$ consisting of $(A, <_A)$, $(B, <_B)$ and $(A', <_{A'})$, $(B', <_{B'})$, respectively.

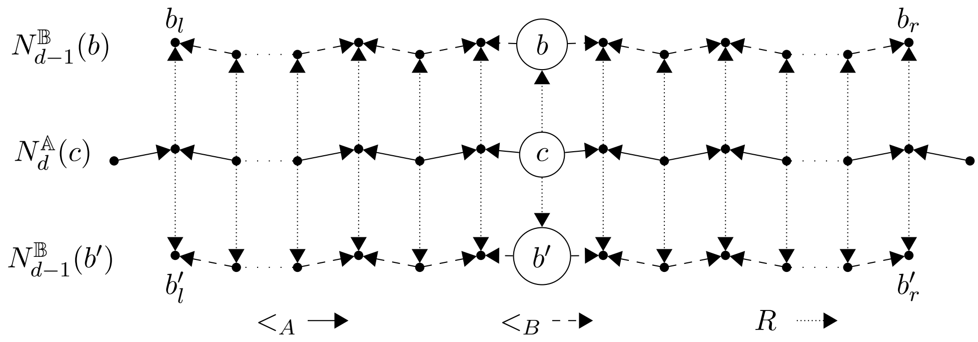


Figure II.4: d-neighbourhood of c with 4 outgoing edges, for even d. $\mathbb{A} = (A, A, <_A)$ and $\mathbb{B} = (B, B, <_B)$. The horizontal bridge in the middle is $N^{\mathcal{A}}_d(c)$. The top horizontal structure is $N^{\mathbb{B}}_{d-1}(b)$. The bottom horizontal structure is $N^{\mathbb{B}}_{d-1}(b')$

We prove $N^{\mathbb{C}}_d(c) \cong N^{\mathbb{C}'}_d(f(c))$ only for the case a). One demonstrates other cases analogously. Assume c satisfies a). f sends c to $f(c) \in A'$ with 2 out-$<_{A'}$-edges and 2 out-R'-edges. Hence, the immediate surroundings of c and $f(c)$ are isomorphic. To visualize the structure $N^{\mathbb{A}}_d(c)$, where $\mathbb{A} = (A, A, <_A)$, consider a movement which starts from c and goes along edges in $G(\mathbb{A})$ at most at distance d. One can mirror this movement in $G(\mathbb{A}')$, where $\mathbb{A}' = (A', A', <_{A'})$, starting from $f(c)$. Let s be the distance between the two points one can reach in this way. Clearly $s \leqslant 2d$. Another possible path between these points (either in $G(\mathbb{A})$ or $G(\mathbb{A}'))$ has length $3m - s \geqslant 3m - 2d > 3m - m > 4d$. Hence, d-neighbourhoods of c and $f(c)$, restricted to $\mathbb{A}$ and $\mathbb{A}'$ respectively, are simply isomorphic parts of bridges with c and $f(c)$ as their central points, where the endings of the bridges do not connect or overlap because the distance d is to small to make this happen.

Now, consider how we move from c at distance d using R-edges and $<_B$-edges as well. c has 2 out-R-edges which send it to different points $b, b' \in B$ separated by the distance $3m$ in $G(\mathbb{B})$, where $\mathbb{B} = (B, B, <_B)$. This is easily visualized in Figure II.3, where b and b' lie on the opposite sides of the hexagon. Now, the movement from c at distance k in a given direction within $\mathbb{A}$ is mirrored by the relation R by the two movements from b and b' in the same direction at distance k in $\mathbb{B}$ (see Figure II.3). The same remark applies to the movement from $f(c) \in A'$ except that R' sends $f(c)$ to the points in B' which lie on different triangles and thus the distance between them in $G(\mathbb{B}')$ is ∞. Hence, it is clear that $N_d^{\mathbb{C}'}(f(c))$ is isomorphic to the structure in Figure II.4. Getting back to $\mathbb{C}$, we see that if we start from c, we explore $\mathbb{B}$ at most at distance $d-1$ from b and b'. Let s be the distance between the two points that one can reach from c in $G(\mathbb{C})$ by going at most at distance d. Clearly, $s \leqslant 2d$. There are other paths connecting the two points, but they are not included in $N_d^{\mathbb{C}}(c)$, because the length of these paths is greater than $2d$. Hence, $N_d^{\mathbb{C}}(c)$ actually looks like in Figure II.4.

Going through other cases, we eventually show that $\mathbb{C} \leftrightarrows_d \mathbb{C}'$. However, $\mathbb{C} \notin \mathcal{Q}$ and $\mathbb{C}' \in \mathcal{Q}$, which contradicts the Hanf-locality of $\mathcal{Q}$. $\square$

II.4 Complexity

In this section we study the computational complexity of the generalized polyadic quantifiers introduced in Section II.1. In Section II.3 we show that the quantifiers in question are not FO-expressible. However, they are expressible in the existential fragment of SO. Hence, by Fagin's theorem [Fagin, 1974], they are in NP. This observation provokes natural questions about $PTIME$-computability and NP-completeness. The results are to some extent surprising. We show that $\mathcal{H}$ is in P, whereas other quantifiers are

NP-complete. For background information on computational complexity, see e.g. [Arora and Barak, 2009].

Definition 6 *Let $\mathcal{A} = (A, <_A)$ be a finite strict partial order. The height of $\mathcal{A}$, denoted by $h(\mathcal{A})$, is the number of vertices in the longest chain in $\mathcal{A}$.*

Lemma 7 *h is in P.*

Proof: The argument uses the Sedgewick strategy [Sedgewick and Wayne, 2011] for finding the longest paths in a directed graph with the help of the Ford-Bellman algorithm. The strategy rests on the observation that the shortest path (in terms of weights) in an acyclic directed graph with all weights multiplied by -1 is the same as the longest path in the original graph. The Ford-Bellman algorithm is designed for acyclic weighted directed graphs. Strict partial orders are acyclic directed graphs, so the shortest path in a strict partial order with negative weights always exists. Now, the application of the Sedgewick strategy to our problem is straightforward. Take a strict poset $\mathcal{X} = (X, <_X)$ as input. Construct a weighted strict poset $\mathcal{X}' = (X, <_X, w)$ by labelling all edges in $\mathcal{X}$ with -1 (which means we set $w(x, y) = -1$ for all $x, y \in X$ such that $x <_X y$). For every $v \in X'$, run the Ford-Bellman algorithm and find the shortest path from v to all other vertices in X'. Return the number of edges of the shortest path in $\mathcal{X}'$. The construction of $\mathcal{X}'$ from $\mathcal{X}$ is in P, the Ford-Bellman algorithms works in polynomial time, hence h is in P. $\square$

Lemma 8 *Let $\mathcal{A}$, $\mathcal{B}$ be strict posets. There is a homomorphism from $\mathcal{A}$ to $\mathcal{B}$ iff $h(\mathcal{A}) \leqslant h(\mathcal{B})$.*

Proof: Let $\mathcal{A}$, $\mathcal{B}$ be strict posets.

($\Rightarrow$) Let f be a homomorphism from $\mathcal{A}$ to $\mathcal{B}$. Observe that f maps every path in $\mathcal{A}$ into an isomorphic path in $\mathcal{B}$. Hence $h(\mathcal{A}) \leqslant h(\mathcal{B})$.

($\Leftarrow$) Assume $h(\mathcal{A}) \leqslant h(\mathcal{B})$. Let $\mathcal{A}'$ be the strict poset obtained from $\mathcal{A}$ by adding new vertex a, drawing edges from a to every source vertex in $\mathcal{A}$ (a source vertex is a vertex with no ingoing edges) and taking the transitive closure of the resulting relation. For $i = 1, 2, \ldots, h(\mathcal{A})$ define

$$A_i = \{v \in \mathcal{A} - \{a\} : \text{the longest path from } a \text{ to } v \text{ in } \mathcal{A}' \text{ has length } i\}. \quad (20)$$

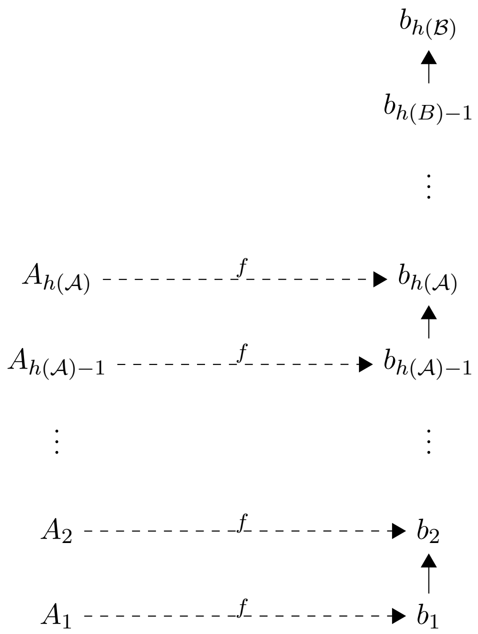


Figure II.5

We show that $\{A_i\}_{1\leqslant i\leqslant h(\mathcal{A})}$ is a partition of A. Of course, $\bigcup_{1\leqslant i\leqslant h(\mathcal{A})} A_i = A$, since all paths from a to any vertex in $A - \{a\}$ have lengths belonging to $\{1, 2, \ldots, h(\mathcal{A})\}$. Obviously, $A_i \cap A_j = \emptyset$, for every $i \neq j$, $1 \leqslant i, j \leqslant h(\mathcal{A})$. To prove that each A_i is non-empty, let $a_1a_2 \ldots a_{h(\mathcal{A})}$ be a path in $\mathcal{A}$. We claim $a_i \in A_i$, for $i = 1, 2, \ldots, h(\mathcal{A})$. For suppose the contrary. Choose j such that $1 \leqslant j \leqslant h(\mathcal{A})$ and $a_j \notin A_j$. Of course, $a_j \notin A_k$, for $k < j$, since $aa_1a_2 \ldots a_j$ has length $j > k$. So $a_j \in A_m$, for some $m > j$. Choose such a number m and let $aa'_1a'_2 \ldots a'_m$ be a path in $\mathcal{A}'$ with $a'_m = a_j$. But then $aa'_1a'_2 \ldots a'_m a_{j+1} \ldots a_{h(\mathcal{A})}$ is a path of length greater than $h(\mathcal{A})$ which is impossible. This proves that $a_i \in A_i \neq \emptyset$, for $i = 1, 2, \ldots, h(\mathcal{A})$ and concludes the proof that $\{A_i\}_{1\leqslant i\leqslant h(\mathcal{A})}$ is a partition of A. Now let $i \in \{1, 2, \ldots, h(\mathcal{A})\}$. We show that for all $u, v \in A_i$, it is not the case that $u <_A v$. Suppose the contrary and choose $u, v \in A_i$ such that $u <_A v$. Let $au_1u_2 \ldots u_i = u$ and $av_1v_2 \ldots v_i = v$ be paths in $\mathcal{A}'$. But then $au_1u_2 \ldots u_i v$ is a path in $\mathcal{A}'$ from a to v and it has length $i + 1$, so $v \in A_{i+1}$, contrary to our assumption.

We show that for every $u, v \in A$, if $u <_A v$, then there are i, j such that $u \in A_i$, $v \in A_j$ and $1 \leqslant i < j \leqslant h(\mathcal{A})$. Let $u, v \in A$ and $u <_A v$. Since $\{A_i\}_{1\leqslant i\leqslant h(\mathcal{A})}$ is a partition, $u \in A_i$ and $v \in A_j$, for some unique

$i, j \in \{1, 2, \ldots, h(\mathcal{A})\}$. We show $i < j$. To prove it, assume $i \geqslant j$. But then we have a path $au_1u_2 \ldots u_i = u$ of length i and $au_1u_2 \ldots u_i v$ of length $i+1$, so $v \in A_{i+1}$. Hence $v \in A_j \cap A_{i+1}$, where $j \neq i+1$, which contradicts the fact that $\{A_i\}_{1 \leqslant i \leqslant h(\mathcal{A})}$ is a partition.

We are ready to define a homomorphism from $\mathcal{A}$ to $\mathcal{B}$. Let $b_1, b_2, \ldots, b_{h(\mathcal{A})}$ be a path in $\mathcal{B}$. The desired homomorphism f is defined as follows: $f(A_i) = \{b_i\}$, for $i = 1, 2, \ldots, h(\mathcal{A})$ (see Figure II.5). Since $\{A_i\}_{1 \leqslant i \leqslant h(\mathcal{A})}$ is a partition of A, f is a function. We show f is indeed a homomorphism. Let $x, y \in A$. Assume $x <_A y$. Then there are unique i, j such that $x \in A_i$, $y \in A_j$ and $1 \leqslant i < j \leqslant h(\mathcal{A})$. By the definition of f, $f(x) = b_i$ and $f(y) = b_j$. We have $b_i <_B b_j$, so $f(x) <_B f(y)$. □

Theorem 9 *$\mathcal{H}$ is in P.*

Proof: Corollary from Lemma 7 and Lemma 8. □

$\mathcal{H}$ differs from other quantifiers defined in Section II.1 in two aspects. $\mathcal{H}^{1-1}$, $\mathcal{H}_F^{1-1}$, $\mathcal{E}$, $\mathcal{E}_F$, $\mathcal{E}_{DF}$ require that the similarity function is injective. $\mathcal{H}_F$, $\mathcal{H}_{DF}$, $\mathcal{H}_F^{1-1}$, $\mathcal{E}_F$, $\mathcal{E}_{DF}$ have an additional non-fixed relation $R \subseteq A \times B$ which restricts the behavior of the similarity function. The next theorem shows that these two aspects make a significant difference in the computational complexity (if $P \neq NP$).

Theorem 10 *$\mathcal{H}_F$, $\mathcal{H}_{DF}$, $\mathcal{H}^{1-1}$, $\mathcal{H}_F^{1-1}$, $\mathcal{E}$, $\mathcal{E}_F$, $\mathcal{E}_{DF}$ are NP-complete.*

Proof: Let $\mathcal{Q} \in \{\mathcal{H}^{1-1}, \mathcal{E}\}$ and $\mathcal{Q}_R \in \{\mathcal{H}_F, \mathcal{H}_{DF}, \mathcal{H}_F^{1-1}, \mathcal{E}_F, \mathcal{E}_{DF}\}$. $\mathcal{Q}$ and $\mathcal{Q}_R$ are definable by existential SO-sentences over the signature $\sigma = (A, B, <_A, <_B)$ and $\sigma_R = (A, B, <_A, <_B, R)$, respectively (see Section II.1). By the Fagin theorem [Fagin, 1974], $\mathcal{Q}$ and $\mathcal{Q}_R$ are in NP.

We prove that $\mathcal{Q}$ and $\mathcal{Q}_R$ are NP-hard. We do this by showing that 3SAT is polynomially reducible to $\mathcal{Q}$ and $\mathcal{Q}_R$. We demonstrate polynomial constructions which, given an arbitrary $3CNF$-formula φ, output a double partial order $\mathcal{D}_\varphi$ and a coupled partial order $\mathcal{C}_\varphi$ such that:

$$\varphi \in 3SAT \Leftrightarrow \mathcal{D}_\varphi \in \mathcal{Q} \tag{21a}$$

$$\varphi \in 3SAT \Leftrightarrow \mathcal{C}_\varphi \in \mathcal{Q}_R \tag{21b}$$

Let $\varphi := \bigwedge_{i=1}^{k}(a_i \vee b_i \vee c_i)$ be an arbitrary $3CNF$-formula, where k is the number of clauses, a_i, b_i, c_i are literals, for $i = 1, 2, \ldots, k$. We construct a double partial order $\mathcal{D}_\varphi = (A \cup B, A, <_A, B, <_B)$. Let $A = \{v_i : 1 \leqslant i \leqslant k\} \cup \{v_{ij} : 1 \leqslant i < j \leqslant k\}$ consist of $k + \frac{k^2-k}{2}$ elements. Let $<_A \subseteq A^2$ be such that

$$(\forall\, x, y \in A)\, \{x <_A y \Leftrightarrow \exists\, i, j\ [x = v_i \wedge (y = v_{ij} \vee y = v_{ji})]\}.$$

$(A, <_A)$ is easily visualised as a k-clique on $\{v_i : 1 \leqslant i \leqslant k\}$, where each edge $\{v_i, v_j\}$, $i < j$, is replaced by edges $v_i v_{ij}$, $v_j v_{ij}$. Obviously, $(A, <_A)$ is a strict partial order and verifying whether given two elements of A are in $<_A$ is polynomial. To construct $(B, <_B)$, let

$$B = \bigcup_{i=1}^{k} \{v_{a_i}, v_{b_i}, v_{c_i}\} \cup$$

$$\bigcup_{1 \leqslant i < j \leqslant k} \{v_{a_i a_j}, v_{a_i b_j}, v_{a_i c_j}, v_{b_i a_j}, v_{b_i b_j}, v_{b_i c_j}, v_{c_i a_j}, v_{c_i b_j}, v_{c_i c_j}\}$$

consist of $3k + 9\frac{k^2-k}{2}$ elements. The idea is that we construct vertices from all occurrences (tokens) of the literals in the formula φ and we add *artificial* vertices for all (unordered) pairs of literals such that the elements of the pair belong to different clauses. The size of B is polynomial with respect to k. Let $<_B \subseteq B^2$ be such that for all $u, w \in B$:

$$u <_B w \Leftrightarrow \exists\, x, y \in \{a, b, c\}\ \exists\, i, j \in [k]$$

$$[u = v_{x_i} \wedge (w = v_{x_i y_j} \vee w = v_{y_j x_i}) \wedge \neg(x_i \Leftrightarrow \overline{y_j})].$$

$(B, <_B)$ is easily visualized as follows. Let x_i and y_j be two tokens of literals from different clauses of φ, $i < j$. Formally, it means that $x_i \in \{a_i, b_i, c_i\}$, $y_j \in \{a_j, b_j, c_j\}$, $i < j$. If x_i and y_j are consistent, we add edges $v_{x_i} v_{x_i y_j}$, $v_{y_j} v_{x_i y_j}$. It is easy to see that $<_B$ is a strict partial order on the set B and computing the characteristic function of the relation $<_B$ is polynomial with respect to k. This ends the construction of $\mathcal{D}_\varphi$. $\mathcal{C}_\varphi$ is constructed in the same way as $\mathcal{D}_\varphi$, except that we add the coupling relation R.

$$R = \bigcup_{i=1}^{k} \Big(\{v_i\} \times \{v_{a_i}, v_{b_i}, v_{c_i}\}\Big) \cup \bigcup_{1 \leqslant i < j \leqslant k} \Big(\{v_{ij}\} \times \bigcup_{x,y \in \{a,b,c\}} \{v_{x_i y_j}\}\Big) \quad (22)$$

We prove (21a) and (21b).

($\Rightarrow$) Assume $\varphi \in 3SAT$. Hence, there is a valuation t of propositional variables in φ and a sequence of literals $l_1, \ldots, l_k$ such that $l_i \in \{a_i, b_i, c_i\}$, for $i = 1, 2, \ldots, k$, and $\forall\, i \leqslant k\ t(l_i) = 1$. Hence, the literals $l_1, l_2, \ldots, l_k$ are consistent. Let $f : A \to B$ be as follows:

$$f(x) = \begin{cases} v_{l_i} & x = v_i \\ v_{l_i l_j} & x = v_{ij}, i < j. \end{cases}$$

Clearly, f is an embedding of $(A, <_A)$ into $(B, <_B)$ and hence an injective homomorphism from $(A, <_A)$ to $(B, <_B)$. Therefore, $\mathcal{D}_\varphi \in \mathcal{Q}$. Observe that $\{R_a\}_{a \in A}$ is a disjoint family and $\forall\, a \in A\ f(a) \in R_a$. So $\mathcal{C}_\varphi \in Q_R$.

($\Leftarrow$) Observe that $\mathcal{D}_\varphi \in \mathcal{H}^{1-1}$ iff $\mathcal{D}_\varphi \in \mathcal{E}$. Assume $f : A \to B$ is an injective homomorphism from $(A, <_A)$ to $(B, <_B)$. Then $f(A)$ cannot have any additional edges apart from those that are mapped by f (we prove $\mathcal{C}_\varphi \in \mathcal{H}_{DF} \Leftrightarrow C_\varphi \in \mathcal{E}_{DF}$ in a similar way). Now, observe that the conditions $\mathcal{C}_\varphi \in \mathcal{H}_F$, $\mathcal{C}_\varphi \in \mathcal{H}_{DF}$, $\mathcal{C}_\varphi \in \mathcal{H}_F^{1-1}$, $\mathcal{C}_\varphi \in \mathcal{E}_F$, $\mathcal{C}_\varphi \in \mathcal{E}_{DF}$ are equivalent. The disjointness of $\{R_a\}_{a \in A}$ in $\mathcal{C}_\varphi$ settles the equivalences $\mathcal{C}_\varphi \in H_{DF} \Leftrightarrow \mathcal{C}_\varphi \in \mathcal{H}_F^{1-1}$, $\mathcal{C}_\varphi \in \mathcal{H}_F \Leftrightarrow \mathcal{C}_\varphi \in \mathcal{H}_{DF}$, $\mathcal{C}_\varphi \in \mathcal{E}_F \Leftrightarrow \mathcal{C}_\varphi \in \mathcal{E}_{DF}$. The equivalence $\mathcal{C}_\varphi \in \mathcal{H}_{DF} \Leftrightarrow C_\varphi \in \mathcal{E}_{DF}$ is already established. This ends the argument for all equivalences for $\mathcal{C}_\varphi$. Now, observe that $\mathcal{C}_\varphi \in \mathcal{Q}_R$ implies $\mathcal{D}_\varphi \in \mathcal{Q}$. For the converse, assume $\mathcal{D}_\varphi \in \mathcal{Q}$. Let f be an appropriate injective homomorphism. Let $v_i, v_j \in A$, $i < j$. Observe that f maps v_i, v_j to literals from different clauses (as literals form the same clause cannot point to the same vertex). Hence, f is easily rearranged to satisfy $\forall\, a \in A\ f(a) \in R_a$. By definition, $\{R_a\}_{a \in A}$ is a disjoint family. Therefore $\mathcal{C}_\varphi \in \mathcal{Q}_R$.

We are ready to prove the ($\Leftarrow$) part of (21a) and (21b). Assume $\mathcal{C}_\varphi \in \mathcal{Q}_R$. Observe $f(v_i) \in \{a_i, b_i, c_i\}$, for every $1 \leqslant i \leqslant k$ and $f(v_{ij}) \in \bigcup_{x,y \in \{a,b,c\}} \{v_{x_i y_j}\}$, for every $1 \leqslant i < j \leqslant k$. Consider literals $f(v_1), f(v_2), \ldots, f(v_k)$. Let $1 \leqslant i < j \leqslant k$. Since f is a homomorphism, we have $f(v_i) <_B v_{f(v_i) f(v_j)}$ and $v_{f(v_i) f(v_j)} <_B f(v_j)$. By definition of $<_B$, $f(v_i)$ and $f(v_j)$ are consistent. Hence, $f(v_1), f(v_2), \ldots, f(v_k)$ are all pairwise consistent. Hence, φ is satisfiable. $\square$

II.5 Discussion

Recent developments at the intersection of semantic complexity and cognitive science have vindicated the empirical relevance of various computational models of human comprehension [Szymanik, 2016]. For example, it has been demonstrated that various complexity distinctions are reflected in human processing during cognitive tasks such as verification. In this chapter we provide a logical and computational analysis of some everyday language quantifier constructions that are likely to serve as an empirical material for further research in this direction.

We are particularly interested in the connection between semantic complexity and semantic variation. Specifically, our focus is on the conjecture that computational complexity pressures may result in a specific kind of synchronic semantic variation. This hypothesis, anticipated in [Mostowski and Wojtyniak, 2004], is closely connected to another one which we refer to as the Shift from Complexity Hypothesis (SCH) according to which human subjects, when confronted with a hard problem, are likely to switch from the initially preferred complex interpretation to an easy one [Szymanik, 2010]. To our knowledge, only one work has attempted to test SCH experimentally [Schlotterbeck and Bott, 2013].

One way to test these conjectures consists in confronting results from inference and verification [Mostowski and Wojtyniak, 2004, Gierasimczuk and Szymanik, 2009]. Inference task amounts to recognizing whether one sentence follows from another one. Verification task consists in recognizing the truth value of a sentence in a model. If a human subject accepts (rejects) a sentence in a given model (situation) and accepts its consequence which is readily false (true) in this model (situation) then clearly the subject employs different meanings in verification and inference. In compliance to SCH, such a behaviour may be rooted in pressures coming from computational complexity, for example when a subject realizes that a verification task with an initially preferred meaning is too difficult, she may change the initial interpretation to an easier one. We present this argument in a greater detail in Section II.5.

We show that the results from Sections II.3 and II.4 may provide a theoretical basis for investigating the effects of semantic complexity on cognitive processing. It has been conjectured that human subjects may switch from hard to easy semantics when confronted with an intractable cognitive task [Szymanik, 2010]. This is what we have called the Shift from Complexity Hypothesis (SCH). In [Szymanik, 2010] it has also been pointed out that a similar behaviour is predicted by the so+called $PTIME$-Cognition Hypothesis [Frixione, 2001,Mostowski and Wojtyniak, 2004,van Rooij, 2008], according to which real time human cognitive capacities are limited to $PTIME$-computable functions. The SCH has already received some confirmation [Schlotterbeck and Bott, 2013]. If SCH is actually true, we expect that computational complexity pressures influence the interpretative preferences of human subjects in verification tasks involving hard examples from Section II.2. In the case of easy examples, namely instances of the homomorphism quantifier, there should be no significant effect of complexity.

Some predictions of the SCH allow for drawing conclusions which may be of interest for evolutionary linguistics. Specifically, we outline how computational complexity pressures may result in a specific kind of synchronic semantic variation. As an example, consider the original Barwise sentence (1) whose strong reading is:

$$\exists f : A \to B \; \forall x, y \in A \; [R(x, f(x)) \wedge (x <_A y \Rightarrow f(x) <_B f(y))] \tag{23}$$

We omit the disjointness condition as it does not bring nothing new to the present discussion. Recall that a weak first-order reading is one of the following:

$$\forall x, y \in A \; \exists z, w \in B \; [x <_A y \Rightarrow (z <_B w \wedge R(x, z) \wedge R(y, w))] \tag{24a}$$

$$\forall x, y \in A \; \exists z, w \in B \; [R(x, z) \wedge R(y, w) \wedge (x <_A y \Rightarrow z <_B w)] \tag{24b}$$

We are interested in finding out whether a semantic change occurs at the inter- or intra-individual level during verification. Hence, we need some empirical cues about the semantics employed by human subjects. This may be identified through picture completion experiments as in [Schlotterbeck and Bott, 2013]. Another approach consists in investigating the inferential meaning [Barwise, 1979,Mostowski, 1994]. Inferential semantics may be identified with the net of inferential connections between a given sentence and other sentences. To illustrate this, observe that (23) implies the following first-order sentence:

$$\begin{aligned}&\forall x, y, z \in A\, \exists u, v, w \in B\; [R(x, u) \wedge R(y, v) \wedge \\ &\wedge R(z, u) \wedge (x <_A y <_A z \Rightarrow u <_B v <_B w)]\end{aligned} \tag{25}$$

However, (25) follows neither from (24a) nor (24b) (see discussion on linear readings in Section II.2. Now, if the subject accepts that a sentence whose logical form is (25) actually follows from the Barwise example then we may hypothesise that the subject uses the strong reading during the inference.[2] Note, however, that according to Theorem (10), the referential semantics of the strong reading is intractable. Hence, according to SCH, a subject is likely to change her interpretation during verification and accept the Barwise sentence in a model where the strong reading is false. This is precisely the way how a semantic change may occur at the intraindividual level.

II.6 Summary and Outlook

The Barwise sentence (1) is deemed to expresses the existence of some kind of homomorphism between partial orders [Barwise, 1979]. We develop new examples and suggest logical forms which state the existence of other types of similarities, such as injective homomorphism, embedding and their restricted versions. We also briefly discuss possible linear interpretations. One of the key points is that we interpret the semantics of our constructions in terms of polyadic generalized quantifiers. We call them similarity quantifiers as they express various similarities between partial orders. We analyse some logical properties of the similarity quantifiers. Specifically, we show that they are not FO-expressible over finite models. If unexpectedly $P = NP$ then our result provides at least some logical classification. However, assuming that $P \neq NP$ (which is what our contemporary experience suggests), the result gives no clue whether the similarity quantifiers are practically computable.[3] Hence, taking into account SCH which predicts semantic shifts caused by complexity pressures, there is a need of a more fine-grained complexity classification. We show that the homomorphism quantifier is in P, whereas the

[2]Though it may be challenging to come up with a natural example of a sentence whose logical form is (25).

[3]If the similarity quantifiers were FO-definable, they would be in P [Immerman, 1999] and by the Edmonds thesis – practically computable.

other quantifiers are NP-complete. It amounts to a good classification of the quantifiers with respect to practical computability. Based on our complexity results we outline how synchronic semantic variation may occur as a consequence of computational complexity pressures.

Probably one of the most interesting observations that results from this and similar research is that there are everyday language constructions which cannot be given a unique meaning. The present work shows how this may happen in the context of the Barwise sentence and similar natural language constructions (see Section II.5). This peculiar semantic variation readily defies description in terms of Thesis 1 discussed in Chapter I. Taking into account other evidence of synchronic semantic variation, we have decided to abandon the naive view of semantics and adopt a more refined concept of language by positing the existence of cognitive mechanisms that allow language users to coordinate their individuated semantics through communication. In the next chapter, we shall see how one may approach this problem.

A natural continuation of the present work is to consider similarity quantifiers in a purely logical setting. There have been a great deal of research concerning generalized quantifiers and branching quantifiers, in particular. [Henkin, 1961] mentions the *Ehrenfeucht sentence* which states that there are infinitely many objects (which is a property inexpressible in elementary logic). Moreover, it has been demonstrated that first order logic with Henkin quantifiers is actually equivalent to the existential fragment of second-order logic [Enderton, 1970, Walkoe, 1970]. Further mathematical results are obtained in [Krynicki and Mostowski, 1992, Gottlob, 1997, Kołodziejczyk, 2002, Sevenster, 2006]. Therefore, the general question is what are the properties of logics with similarity quantifiers.

There is also another interesting point that connects semantic complexity with evolutionary linguistics. Obviously, with a little invention, one may formulate natural language sentences that have semantics of arbitrary complexity [Kontinen and Szymanik, 2008]. However, there is a part of the natural language, namely the everyday language, which we use in communication with ordinary people in everyday life. It has been argued on theoretical grounds that the class NP provides a strict semantic bound for everyday language [Mostowski and Szymanik, 2012] (see [Ristad, 1993] for a related hypothesis). By Fagin's theorem [Fagin, 1974], NP coincides with Σ_1^1-definable classes of finite models. The bound is strict as there are examples of everyday language sentences with NP-complete semantics (for example the Barwise sentence). So far we do not know any everyday language constructions that could serve as a counterexample to this Σ^1_1-thesis. Getting back to language evolution, it would be desirable to find out why the semantics of everyday language that emerges through communication does not exceed the Σ^1_1-notions.

III

Coordination of Proportional Quantifiers

III.1 Introduction

Therefore, on the most general level, the question we ask is how language users are able to create semantic conventions based on local linguistic interactions with other individuals.

In Chapter I we have provided some grounds for abandoning the uniform view of semantics and thus its explanatory role in the account of the communicative functioning of language in situations involving synchronic semantic variation. We have raised some doubts about the applicability of the learning by recognizing approach to semantically non-uniform populations. Moreover, we have observed that learning by recognizing seems to be restrained to learning phenomena which are restrained by the traditional teacher-learner distinction. If we accept that the semantics of a language community may be less uniform then we are faced with the problem of explaining how individuals are able to use their language as an effective tool for communication. Our main motivation in this chapter is to provide a plausible solution to this problem in the form of a mathematically precise and generic model.

Contemporarily, this type of research is conceived as a part of evolutionary linguistics and follows various methodological paradigms [Steels, 2011]. Linguistic approach searches for cognitive operations accounting for language change as evidenced by empirical data. Psychological approach consists on performing experiments with human subjects so as to find out what kind of language formation strategies are used in nature. Modelling approach formulates theoretical models for computer simulation, robotic experiments or mathematical analysis. The present work contributes to the modelling approach.

At the present stage of the field, we have several frameworks which differ in many respects (see Chapter I for a short overview). Each of these frameworks seems to bring new insights into language evolution and change. The relation between various modelling approaches is not obvious as they are often based on different assumptions or overlapping sets variables and parameters. It seems that this is one of the difficulties anticipated in [Chomsky, 1986]. For example, one of the obvious examples include the relation between collaborative and individualistic models. In the former language evolution and change is modelled as a dynamical system of intertwined interactions between a number of individuals in population. In the later language evolution is modelled as a chain of several generations of a teacher-learner scenario where a learner from a given generation becomes a teacher in the next generation [Kirby et al., 2014]. It is not immediately obvious what we miss or gain when considering one of these models instead of another one. Only just recently some authors have attempted to answer such questions [Fay et al., 2010, Spike et al., 2016].

Our approach is essentially collaborative. The model includes a population of individuals which we refer to as agents. Each agent is equipped with various hypotheses about the meaning of language expressions. Agents are communicating with each other in shared contexts. They want to adapt their semantic interpretations to minimize communication failures. Hence, the goal of the learning process is to converge on meaning that allows communication, i.e., to learn by coordination. For the coordination of semantics to take place, agents need some distributed mechanisms that allow them to introduce uniformity in semantics by performing local linguistic interactions. This mechanism does not only allow the agents to learn to communicate, but also forces the language change over time so that at some point a global semantic standard may emerge. In this work we provide a generic mechanics of semantic coordination and show how this mechanics work in the case where the semantics being coordinated is interpretable as a proportional quantifier. The emergence of quantifier semantics has already been tackled too some extent by other authors. The most recent work [Pauw and Hilferty, 2012] is concerned with the emergence absolute and scalable quantifiers in a population of grounded agents.

The plan of the present chapter is as follows. In Section III.5 we recall the concept of a language game and describe our approach to linguistic interaction. In Section III.6 we elaborate on the representation of our model and some simplifying assumptions that we apply. Section III.7 outlines the general principles of our solution to the proportional description game. Section III.8 is devoted to the analysis of simple instances of our model is in terms of Markov chains. In Section III.9 we draw some conclusions and mention possible extension of our model to account for distance factor.

III.2 Proportional Quantifiers

Let $\sigma = (R)$ be a vocabulary, where R is 1-place predicate. According to Definition 2 from Chapter II, a generalized quantifier of type (1) is an isomorphism-closed class of σ-structures. Recall that we restrict our attention only to finite models.

Definition 11 *A generalized quantifier* $\mathcal{Q}$ *of type* (1) *is upward monotone proportional if there is a rational number* $h \in [0,1]$ *such that for every* σ*-structure* $\mathbb{M} = (U, R)$*:*

$$\mathbb{M} \in \mathcal{Q} \Leftrightarrow |R|/|U| > h. \tag{26}$$

In what follows, we write proportional instead upward monotone proportional.

The prominent example of everyday language quantifier interpretable in this way is the English *most* and the Polish *większość*. One of the conventionalized meanings of these determiners corresponds to the proportional quantifier determined by $h = 1/2$. For more details on applications of a generalized quantifier theory in natural language see [Peters and Westerståhl, 2006].

From the logical point of view, proportional quantifiers are non-trivial as they are not definable in elementary logic (except $\exists$ and $\forall$ which may be interpreted as proportional). However, they are definable in the existential fragment of second order logic and have $PTIME$-computable model checking. These properties of proportional quantifiers, along with their naturalness, make them an interesting example of everyday language concepts.

Our choice of proportional quantifiers is partially motivated by mathematical and computational simplicity. As we shall see in Section III.8, our model is simple enough to provide a comprehensive analysis of two interacting agents (dyads) in terms of Markov chains. However, as it will become evident from Section III.7, the mechanics of coordination is generic and applicable to virtually any linguistic construction.

III.3 Cognitive Structure of Agents

For the quantifier to emerge, there must be some language in place. We assume agents use a vocabulary consisting of some unary predicate symbols. For simplicity, we assume the vocabulary is shared and thus cannot be a source of misunderstanding. Mathematically speaking, given a finite model, two agents and a predicate symbol, the interpretation of the symbol in the model is the same according to both agents. We assume that language contains a symbol Q and the following construction: given a predicate symbol R, $QxRx$ is a sentence. This syntactic construction is shared. However, the semantics of the construction Q is not shared and is precisely the object of coordination. Technically speaking, agents treat Q as an upward monotone proportional generalized quantifier.

The question is how to include a concept of a generalized quantifier in the cognitive equipment of agents. Certainly, manipulation with infinite classes of structures is unrealistic. We adopt a refined view on semantics according to which the meaning of an expression may be identified with an algorithm for computing it's truth-conditions or – more generally – denotation (see Section I.3 from Chapter I for appropriate references). To see how this view applies in our case, let $\mathcal{Q}$ be a proportional quantifier. We say an algorithm computes $\mathcal{Q}$ if for every finite σ-structure $\mathbb{M} = (U, R)$, the computation of the algorithm on input $\mathbb{M}$ outputs the correct answer to the query $\mathbb{M} \in \mathcal{Q}$. Algorithms are finite objects and thus may be easily included in the cognitive equipment of agents. We posit that agents share a meaning generation mechanism that allows them to use various algorithms for computing proportional quantifiers.

In what follows, we use H to denote the set of meanings that agents may generate. We often refer to meanings or algorithms from H as semantic hypotheses or simply hypotheses.

A typical cognitive structure for representing semantics is an associative map between expressions and meanings (see, e.g., [Steels and Belpaeme, 2005]). Each connected pair is assigned a weight designating the strength of the coupling. One may approach this problem in a more simplistic way. In our case, at any given stage t, each agent associates with Q only one preferred meaning which he effectively uses while communicating with others. At the population level, any such association may be described by a function $s_t : A \to H$. We refer to s_t as synchronic description at stage t. In simple words, $s_t(a)$ is the meaning that agent a associates with Q at stage t. We assume each agent is acquainted only with his own semantics, not the semantics of others. Observe that within this framework, it is perfectly possible that different agents ascribe different preferred meanings to the same expression. This is how the synchronic interindividual variation may be represented at the cognitive level.

We assume learners prefer simpler hypotheses, according to some computationally justifiable notion of simplicity. We model this by introducing an ordering of hypotheses. For $x, y \in H$, $x \preceq y$ means that x is simpler than or as simple as y. It seems that the notion of quasi-ordering suits well in this context as it ascertains that simplicity is transitive and reflexive. To give an example, consider natural language quantifiers such as *more than k*, for $k \in \mathcal{N}$. They are interpretable as type (1) generalized quantifiers and are recognizable by finite automata [van Benthem, 1986]. Suppose H consists of finite automata recognizing such quantifiers. Then we may define a quasi-order of simplicity as follows: for $A, B \in H$, $A \preceq B$ iff A has at most as many states as B.[1] Observe that, in this case, $\preceq$ is also well-founded. In the present work we take H to be finite so well-foundedness is not our worry.

III.4 Society of Agents

In what follows, a population is a finite set, on the most occasions denoted by A. To reflect the relative social standing of agents we introduce authority

[1] Curiously, such a simplicity classification is reflected in human comprehension [Szymanik and Zajenkowski, 2010].

functions. Specifically, we assume each agent $a \in A$ is equipped with a function $w_a : A \to \mathbb{R}_+$. Given two agents, $a, b \in A$, the intuitive sense of $w_a(b)$ is that a perceives b as having the authority equal to $w_a(b)$, whereas $w_a(a)$ reflects how much weight agent a assigns to his own opinion. This framework gives us a way to analyse various configurations of authority and its influence on coordination. In the present account we shall restrict ourselves to the authority functions which satisfy the following: for every $a, b, c \in A$, $w_a(c) = w_b(c)$. This condition roughly corresponds to a situation where members of the population react similarly to other agents. Thus, we shall always assume there is one authority function $w : A \to \mathbb{R}_+$ which ascribes to each agent his authority within a population. The role of social impact in the adoption of conventions has already been explored to some extent in the literature (see, e.g., [Kosterman and Gierasimczuk, 2015, Blythe et al., 2016]).

The multi-agent model of semantic coordination has to specify interaction topology, i.e. structure of relations between agents. There are several ways of structuring a society of agents. The paradigm of social networks allows thinking of a communication through pre-established channels between agents [DeGroot, 1974]. Another approach just assumes a probability of an encounter among any two agents. In the present chapter we consider dyads interacting in the round-robin fashion.

III.5 Proportional Description Game

In the philosophical literature, the notion of a language game has been introduced and discussed in [Wittgenstein, 1953]. Intuitively, a language game is a routinised symbolic interaction between agents embedded in a common context and having some communicative goal. Language games take different forms. We will be chiefly concerned with model-checking games – a group of agents verifies an expression against a given scenario, some of them judge the expression to be true, others to be false. They announce their individual judgements. Another relevant communication structure is the so-called signalling games, where upon an encounter one agent assumes the role of a speaker, the other of a receiver [Skyrms, 2010]. The goal of the speaker is to choose a signal (an expression), which will allow the receiver to correctly interpret what is that the sender has in mind. In a similar way, one can imagine studying other communication games such as games of reference, action games or description games [Steels, 2012].

Model-checking game involves a number of players that interact according to a realized communication pattern which may be represented in the form of a sequence of several interactions. A single interaction is a quadruple $(a, b, \mathbb{M}, v)$, where $a \in A$ is a speaker, $b \in A$ is a hearer, $\mathbb{M}$ is a context (which we refer to as the topic) and v is a truth value. An intended meaning of the realisation of $(a, b, \mathbb{M}, v)$ at stage t is that a communicates to b the output of $s_t(a)$ on input $\mathbb{M}$. Note that this is the only knowledge agents have about the semantics of others. We often denote games by uppercase greek letters Ω, Θ.

In this chapter, we focus on games with topics of the form $\mathbb{M} = (U, R)$ where U is a finite universe and R is a unary relation. Moreover, according to our previous assumptions, given an interaction $(a, b, \mathbb{M}, v)$ at stage t, the hypothesis $s_t(a)$ is an algorithm computing some proportional quantifier. We refer to this type of game as the proportional description game. Intuitively, a proportional description game is driven by the need to agree upon the meaning of a quantifier construction which essentially expresses that there are more objects satisfying a given property than some proportion. Players engaging in such games, if equipped with appropriate alignment strategies, may eventually develop semantics that satisfy the motivational need.

III.6 Simplifying Assumptions

In this section we simplify many aspects of the proportional description game. We assume agents use practical algorithms in the sense of the Edmonds thesis [Edmonds, 1987] according to which practical computability overlaps with $PTIME$, i.e. the class of functions computable by deterministic Turing machines in polynomial time [Arora and Barak, 2009]. We do not differentiate between practical algorithms for computing the same proportional quantifier. Hence, we represent algorithms for Q by completely reduced fractions $r \in [0, 1]$ which – in a number theoretic literature [Hardy and Wright, 1979] – are referred to as Farey fractions. Furthermore, we put some limitations on algorithms that agents may actually choose. We exclude the possibility of computations involving very large denominators. Hence, we restrict admissible fractions to denominators not greater than some given positive integer.

In this context, the notion of a Farey sequence is very handy [Hardy and Wright, 1979].

Definition 12 *Let $k > 0$ be an integer. A Farey sequence of order k is the set consisting of completely reduced fractions from $[0, 1]$ whose denominators do not exceed k. The Farey sequence of order k is denoted by F_k.*

Having this notation, a space of possible meanings that agents may actually choose is limited to F_k, for some $k \in \mathcal{N}$. Thus, at any give stage t, the synchronic description is given by a function $s_t : A \to F_k$.

Observe that for any given σ-structure $\mathbb{M} = (U, R)$, only the proportion $|R|/|U|$ is essential for computing $\mathbb{M} \models QxR(x)$ (see Equation 26 in Definition 11). Hence, we represent $\mathbb{M}$ by a Farey fraction equal to $|R|/|U|$. We can do that without loss of generality because any two models $\mathbb{M} = (U, R)$ and $\mathbb{M}' = (U', R')$ such that $|R|/|U| = |R'|/|U'|$ make exactly the same sentences $QxR(x)$ true. Now, the representation of a single interaction $(a, b, \mathbb{M}, v)$ reduces to (a, b, r, v), where $r = |R|/|U|$.

On many occasions, situational contexts we encounter in our everyday conversations are unpredictable. This is why we decide to model the environment by a random variable X. Given the fractional representation of finite σ-structures, we define X so that it assumes values in Farey fractions from $[0, 1]$. Discrete random variables are sufficient for our analysis (generally, one may consider continuous random variables supported on $[0, 1]$). Intuitively, X is a random variable on finite models of the form (U, R) and for any such model, the value of X is $|R|/|U|$.

Playing a language game consists in gathering at least two parties. In general, realization of particular interactions is largely shaped by the underlying social network and the spatial relations between agents. In the present work, we focus on the most fundamental and simple interactions performed within dyads, i.e. groups of two agents. Dyads are assumed to be static in the sense that agents are not replaced during learning. Furthermore, we assume that all interactions are symmetrical, i.e., if (a, b, r, v) is an interaction in the game then so is (b, a, r, v').

III.7 Solution

In this section we present a generic solution to the problem of adopting the semantic convention of a given expression through playing model-checking

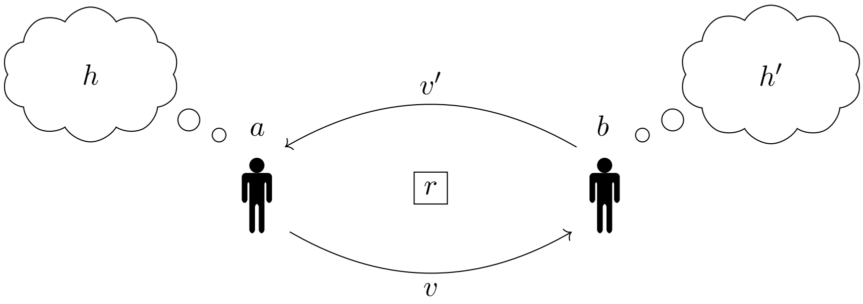


Figure III.1: Symmetrical dyadic communication on the topic r. Hypotheses of a and b are h and h', respectively. Agent a announces v (the truth value of $r > h$) whereas agent b announces v' (the truth value of $r > h'$).

games. The crucial part of the solution is the agent-based mechanism of coordination (see Algorithm 1).

The general picture of learning by coordination is as follows. The evolution of semantics is represented by a sequence of synchronic descriptions $s_0, s_1, s_2, s_3, \ldots$. At each stage t, agents play a model-checking game. During the game each agent gets acquainted with the announcements of the agents that speak to him. Moreover, the hearers of a given agent get acquainted with his judgements. Recall, that the announcements of truth values are based on the semantics given by the synchronic description s_t. Figure III.1 presents the simplest symmetrical dyadic communication game. At the end of stage t all agents simultaneously apply the coordination algorithm. Each agent, on the basis of the answers of his interlocutors and his own semantics, replaces his current hypothesis with a new one. Thus, the application of the coordination algorithm by all agents determines the synchronic description at stage $t + 1$.

Note that the coordination is entirely agent-based. There is no central control over the whole process whatsoever. Agents are also autonomous and (partially) independent. Autonomy means that they use their own cognitive resources. Independence ascertains that the coordination cannot be directly influenced by others. However, as we shall see, agents exert indirect influence due to their social authority. Perhaps the most important feature of this mechanism is that it is not transparent: neither the agent performing coordination nor the other agents have insight into the workings of the mechanism and its direct effects. The only observable effect of coordination manifests in the announcements of truth values.

We posit that the goal of the agents is to communicate successfully, using the quantifier construction Q. Intuitively, successful communication means that whenever agents are faced with a common context, they recognize the same sentences as true. To include this goal in the behaviour of agents, we assume successful communication is beneficial for interacting parties. We can measure the success rate of various hypotheses with respect to the amount of agreement they generate in a given situation and assume that compatibility earns most reward. This is explicitly represented by the reward function. Intuitively, given an agent and a communication game, the value of the reward function for a hypothesis h is to measure how much successful she would be if she adopted h in the present game.

Given an agent a, his current hypothesis h_0, his authority w_0 and a communication game, a runs the coordination algorithm on input consisting of exactly those interactions of the game where she is the hearer. We represent the input by providing separate lists of topics $r_1, r_2, \ldots, r_m$ and truth values $v_1, v_2, \ldots, v_m$. Instead of giving the list of the speakers we provide corresponding authorities $w_1, w_2, \ldots, w_m$. Now, given an arbitrary hypothesis h, let $\overline{z} = z_1, z_2, \ldots, z_m$ be a binary sequence defined as follows: $z_j = 1$ iff $(r_j > h \Leftrightarrow v_j = 1)$. The value of the reward function is then defined as

$$reward(h) = \begin{cases} \Sigma_{i=1}^{m} z_i \cdot w_i & \text{if } h \neq h_0 \\ w_0 + \Sigma_{i=1}^{m} z_i \cdot w_i & \text{otherwise} \end{cases} \tag{27}$$

Observe that if all agents have authority 1, the reward of h is simply the amount of successful interactions that a would participate as a hearer if she used h. Additionally, if h is equal to the current hypothesis of a, the reward is increased by the value w_0 to reflect the fact that a is to some extent independent from external social influences.

To mimic our natural preference for simple solutions we introduce a partial ordering on hypotheses.

Definition 13 *We say a Farey fraction h is simpler than or as simple as h' ($h \preccurlyeq h'$) if the denominator of h is lesser than or equal to the denominator of h'. Given a set $X \subseteq F$, the set of all $\preccurlyeq$-minimal elements of X is designated by $S(X)$.*

Algorithm 1 Coordination

Agent: current hypothesis h_0, authority w_0
Input: topics $r_1, r_2, \ldots, r_m$
authorities $w_1, w_2, \ldots, w_m$
answers $v_1, v_2, \ldots, v_m$
Output: hypothesis from H

1: **for all** $h \in H$ **do**
2: **for all** $i = 1, 2, \ldots, m$ **do**
3: $z_i :=$ the truth value of $(r_i > h \Leftrightarrow v_i = 1)$
4: **end for**
5: $reward(h) := \begin{cases} \Sigma_{i=1}^{m} z_i \cdot w_i & \text{if } h \neq h_0 \\ w_0 + \Sigma_{i=1}^{m} z_i \cdot w_i & \text{otherwise} \end{cases}$ ▷ see Equation 27
6: **end for**
7: $M := \{h \in H : \forall h' \; reward(h) \geqslant reward(h')\}$
8: **return** random element from $S(M)$

The agent computes the reward for every hypothesis h from the space of hypotheses H (lines 1–6). This amounts to computing, for each h, which interactions would be successful, if he used h in the present situation (lines 2–4) and assigning a reward to h, according to Equation 27 (line 5). Next, the agent considers only those hypotheses that have the highest value of the reward function (line 7). Intuitively, such hypotheses guarantee maximal compatibility with interlocutors in current situation. The agent rejects too complicated hypotheses (line 8) and, finally, changes his current semantics to a random hypothesis from what is left (line 8). We assume the probability of drawing an element from $S(M)$ such that $|S(M)| = m$ equals $1/m$.

III.8 Semantic Evolution as a Markov Chain

At this point, it is not obvious how the semantics evolves according to our model and how it is affected by various parameters such as authorities or the structure of communication games. We approach this problem by providing a Markov chain representation of the evolution of semantics for dyadic symmetric interactions. We define and analyse three models which we denote by: M_1, M_2 and M_3. In M_1 equal-authority dyads communicate symmetrically on a single topic per game (however, different games may have different topics). M_2 is the same as M_1, except that dyads have differentiated authority. M_2 extends M_2 by allowing agents to communicate in scenarios involving multiple topics.

Let A be a set of agents, w an authority function, H a space of hypotheses. The system evolves stage by stage. Suppose the semantics at stage t is given by the synchronic description $s_t : A \to H$ and agents play a communication game. Each agent takes the data from the present game as input and updates his semantics according to the Algorithm 1. This results in a new synchronic description $s_{t+1} : A \to H$. Observe that s_{t+1} depends only on s_t and the present game. In other words, the evolution of semantics implied by our solution is a memoryless process – the past is irrelevant for the current behaviour of the system. Discrete-time Markov chains are well suited for description of this kind of processes [Feller, 1968].

One way of defining a Markov chain is to specify a set S of all possible states of the system and a stochastic matrix $[p_{ss'}]_{s,s' \in S}$ of transition probabilities. (To obtain a complete description we need initial probabilities p_i, for every $i \in S$. However, p_i's are not important for our purposes.) A stochastic matrix is a square matrix of non-negative reals such that each row sums up to 1. The transition probability $p_{ss'}$ is to measure the likelihood that the state of the system changes from s to s' in one stage. We give an example of such system in Figure III.2 which illustrates a three-state Markov chain of a very simple model describing the probabilities of weather conditions given the weather at the preceding day.

To apply these considerations to the evolution of semantics, let A be a set of agents, H a space of hypotheses, w an authority function. We posit that the set of states S is the set of all synchronic descriptions, namely all functions from A to H or, equivalently, all $|A|$-tuples assuming values in H. Given $s, s' \in S$, the value $p_{ss'}$ designates the probability that the synchronic description changes from s to s' in one step of coordination according to the Algorithm 1.

To calculate the transitions we need a probabilistic description of communication games. For simplicity, each model we consider assumes a fixed number of interactions per game (this number is always even as the the communication is symmetrical). An important part is the probability that a given fraction becomes a topic of an interaction. This is given by the random variable X with an associated probability function P (see Section III.6). This information suffices to calculate the transition probabilities for each of the subsequent models.

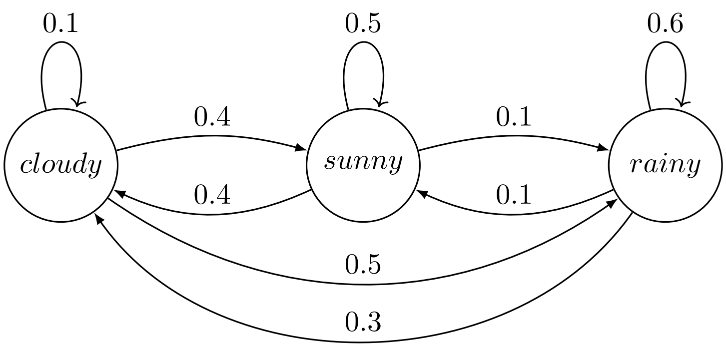


Figure III.2: Simple Markov chain. Labelled arrows designate transition probabilities, e.g, the transition from *rainy* to *cloudy* has probability 0.3.

III.8.1 Equality

In this section we investigate evolution of semantics for symmetrically communicating dyads consisting of agents of equal authority. At each stage a communication game involves one topic.

Let $A = \{1, 2\}$ and $H = F_k$, for some $k > 0$. The authorities of agents are equal, so w is a constant authority function. The topic encountered in the environment is understood as a random variate of a random variable X with an associated probability function P. At each stage we generate a random variate r and produce an instance of the proportional description game. The instance of the game consists of two interactions: $(1, 2, r, v)$ and $(2, 1, r, v')$ (see Figure III.1). We denote such a model by M_1.

Before we state the representation theorem, let us fix some notation. We write $s(1)s(2)$ to designate the state $s \in S$. For example, the state s such that $s(1) = 0$ and $s(2) = \frac{1}{2}$ is designated by $0\frac{1}{2}$. The transition probability from s to s', normally denoted by $p_{ss'}$, is written as $p_{s(1)s(2)\to s'(1)s'(2)}$. For example, the transition probability from the state s such that $s(1) = 0$ and $s(2) = \frac{1}{2}$ to the state s' such that $s'(1) = 1$ and $s'(1) = 1$ is designated by $p_{0\frac{1}{2}\to 11}$. Let xx', yy', zz' be states. If $p_{xx'\to yy'} = p_{xx'\to zz'} = p$, we write $p_{xx'\to yy'|zz'} = p$ to denote the conjunction of $p_{xx'\to yy'} = p$ and $p_{xx'\to zz'} = p$.

Theorem 14 *Fix an instance of M_1, where $A = \{1, 2\}$, $H = F_k$, for $k \geqslant 1$, X is a random variable with a probability function P and w is a constant authority function. Then the model is represented by the Markov chain on*

$S = H^2$ *induced by the following probabilities:*

$$p_{01\to01} = P(X=0) + \frac{P(X>0)}{4} \tag{28a}$$

$$p_{01\to00|11|10} = \frac{P(X>0)}{4} \tag{28b}$$

For all $u \in H$:

$$p_{uu\to uu} = 1 \tag{28c}$$

For all $u \in H$ *such that* $0 < u < 1$:

$$p_{0u\to0u} = P(X=0) + P(X>u), \tag{28d}$$

$$p_{0u\to00|10} = \frac{1}{2}P(0 < X \leqslant u) \tag{28e}$$

$$p_{u1\to u1} = P(X \leqslant u), \tag{28f}$$

$$p_{u1\to10|11} = \frac{1}{2}P(X > u) \tag{28g}$$

For all $u, v \in H$ *such that* $0 < u < v < 1$:

$$p_{uv\to uv} = P(X \leqslant u) + P(X > v) \tag{28h}$$

$$p_{uv\to10} = P(u < X \leqslant v) \tag{28i}$$

Proof: Let $u, v \in H$. Without loss of generality, assume $u \leqslant v$. We shall determine possible ways of changing the state from uv to other states, including uv. We consider the following cases: a) $u = v$, b) $u = 0$, $v = 1$, c) $u = 0$, $0 < v < 1$, d) $0 < u < 1$, $v = 1$, e) $0 < u < v < 1$. We show how to calculate transition probabilities for b). The core of the reasoning is similar in other cases.

Let $u = 0$, $v = 1$. We consider all possible arrangements of the topic r that may affect the truth values exchanged in the game. These arrangements are: i) $r = 0$, ii) $0 < r \leqslant 1$. Let us designate by $M^{xy}_{a,r\in R}$ the set of hypotheses for which the value of the reward function is maximal, where the reward is computed relative to agent $a \in \{1, 2\}$, the state $s \in S$ such that $s(1) = x$, $s(2) = y$ and the topic $r \in R \subseteq [0, 1]$. We need to carefully work through Algorithm 1.

Let us compute $M^{01}_{1,r\in(0,1]}$. The situation is visualised in Figure III.3. The

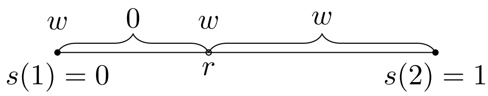


Figure III.3: Computing $M^{01}_{1,r\in(0,1]}$.

horizontal line is $[0,1]$. We have a topic $r \in (0,1]$. The semantics of agent 1 is 0 and that of agent 2 is 1. The numbers at the top indicate the values of the reward function relative to agent 1 for arguments $h \in H \subset [0,1]$. Below we show that in this situation the reward equals (i) w for $h = 0$, (ii) 0 for $0 < h < r$, (iii) w for $h \geqslant r$.

Case (i) Suppose agent 1 used $h = 0$. He would not agree with agent 2, since his answer is negative (as it is not the case that $r > 1$), whereas the answer of agent 1 would be positive (as $r > h = 0$). According to Algorithm 1, the reward relative to an agent for a given hypothesis is not increased by the interlocutor's authority when they disagree. However, $h = 0$ is the current hypothesis of agent 1, so the reward for h is promoted by the authority of agent 1, namely w, as indicated in Figure III.3.

Case (ii) Suppose agent 1 used h such that $0 < h < r$. Agent 1 would not agree with agent 2 for the same reason as in the case (i). Hence, the reward for h is not incremented by the authority of his interlocutor. The value of h is not promoted by the authority of a neither, since $h \neq s(1) = 0$. Therefore, the reward for h is 0.

Case (iii) Suppose a used $h \geqslant r$. Agent 1 would agree with agent 2, since the answer of agent 1 would be negative (as it is not the case that $r > h$). So the reward for h is increased by the authority of agent 2, namely w. Since $h \neq s(1)$, the reward for h is not further promoted by the authority of agent 1. Hence, the reward for h is w.

We have just proved that $M^{01}_{1,r\in(0,1]} = \{0\} \cup \{h \in H : h \geqslant r\}$. In a similar way we prove that $M^{01}_{1,r=0} = \{0\}$, $M^{01}_{2,r=0} = \{1\}$ and $M^{01}_{2,r\in(0,1]} = \{h \in H : h < r\} \cup \{1\}$.

Observe that under the condition that a random variate is $r = 0$, agent 1 chooses 0 from $M^{01}_{1,r=0}$ with probability 1 and agent 2 chooses 1 from $M^{01}_{2,r=0}$ with probability 1. Under the condition that a random variate is $0 < r \leqslant 1$, agent 1 chooses 0 from $M^{01}_{1,r=0} = \{0\}$ with probability 1/2 and agent 2 chooses 1 from $M^{01}_{2,r\in(0,1]}$ with probability 1/2. Observe that the

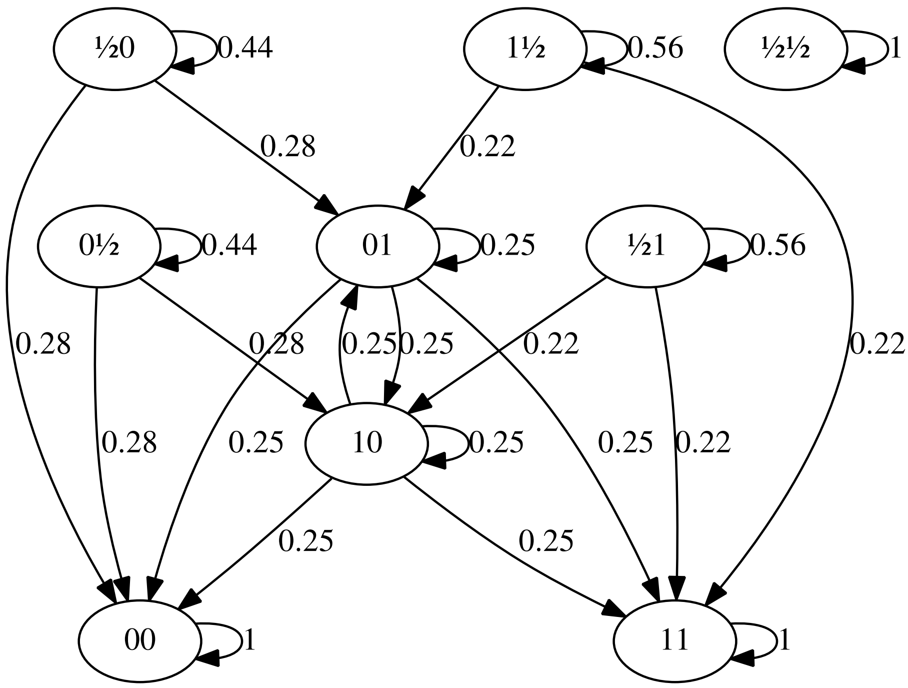


Figure III.4: Markov chain for M_1, where $A = \{1, 2\}$, $w_1 = w_2 > 0$, $H = \{0, \frac{1}{2}, 1\}$ and $X \sim B(50, 0.5)$.

events $X = 0$ and $0 < X \leqslant 1$ form a finite partition of a sample space. Hence, by the law of total probability, $p_{01\to01} = P(X = 0) + P(0 < X \leqslant 1)/4$. $\square$

By Theorem 14, agents never change to more complicated semantics (see Definition 13). They either retain their hypotheses or change to simpler ones. What is more, the actual change to a simpler hypothesis leads always to 0 or 1. 0 may be interpreted as the existential quantifier. 1 is a trivial (always false) quantifier—it can be read as *more than everything*. Hence, M_1 cannot explain how the semantics may evolve from simple to more complex forms. By Theorem 14, coordination processes of M_1 cannot stabilize on semantics other than 0 or 1, unless initial semantics is a constant function $s : A \to \{u\}$, for some hypothesis u such that $0 < u < 1$. In linguistic terms, only the existential quantifier or the trivial quantifier *more than everything* may emerge through such coordination processes. Hence, M_1 cannot explain how more complex semantics could emerge. In particular, *most* is not achievable in such a model, unless *most* was common to every agent from the start. The above observations are not favourable for M_1. We get back to these issues in

Section III.8.3.

In M_1 the evolution of semantics may proceed in subsequent rounds in the following way: $01, 10, 01, 10, \ldots$. This phenomenon is partially caused by equality and is easily observable in many everyday situations which require some kind of coordination. For example, two people approaching each other from the opposite directions must somehow coordinate their actions to avoid collision. They have two possibilities: to go right or left. If the choose different actions, namely either right-left or left-right, the will collide. However, very often people switch between two incorrect solutions: after choosing different actions they observe that this no good, so they both change their mind and again choose different actions, and so on.

III.8.2 Differentiated Authority

M_2 differs from M_1 in having non-constant authority function. Without loss of generality we assume that $w_1 > w_2$. Our intention is to investigate how the coordination processes of such models depend on authority.

Theorem 15 *Fix an instance of M_2, where $A = \{1, 2\}$, $H = F_k$, for $k \geqslant 1$, X is a random variable with a probability function P and w is an authority function such that $w_1 > w_2$. Then the model is represented by the Markov chain on $S = H^2$ induced by the following probabilities:*

$$p_{01\to01} = P(X = 0) \tag{29a}$$

$$p_{01\to00} = P(X > 0) \tag{29b}$$

$$p_{10\to11} = P(X > 0) \tag{29c}$$

For all $u \in H$:

$$p_{uu\to uu} = 1 \tag{29d}$$

For all $u, v \in H$ such that $0 < u < v < 1$:

$$p_{uv\to uv} = P(X \leqslant u) + P(X > v) \tag{29e}$$

$$p_{uv\to u0} = P(u < X \leqslant v) \tag{29f}$$

$$p_{vu\to v1} = P(u < X \leqslant v) \tag{29g}$$

For all $u \in H$ such that $0 < u < 1$:

$$p_{0u \to 0u} = P(X = 0) + P(X > u) \tag{29h}$$

$$p_{0u \to 00} = P(0 < X \leqslant u) \tag{29i}$$

$$p_{u0 \to u1} = P(0 < X \leqslant u) \tag{29j}$$

$$p_{u1 \to u1} = P(X \leqslant u) \tag{29k}$$

$$p_{u1 \to u0} = P(X > u) \tag{29l}$$

$$p_{1u \to 11} = P(X > u) \tag{29m}$$

Introducing differentiated authority significantly simplifies the whole process when compared to M_1. For example, in M_2 we cannot change from 01 to 10 (and hence from 10 to 01). Therefore, in M_2 the coordination cannot proceed in subsequent rounds in the following way: $01, 10, 01, 10, \ldots$. However, by Theorem 14, such phenomenon may occur in M_1.

Consider a population in the state 01. It is four times more probable in M_2 that the population will stabilize on the existential quantifier. Assuming the probability P of drawing values closer to 1/2 is greater than the probability of drawing values further from 1/2 (i.e. closer to either 0 or 1), the chances of changing from 01 to 00 are very high and increase while $P(X = 0)$ and $P(X = 1)$ are getting smaller. In the extreme case, $P(X = 0)$ is close to 0. Hence, by equations (29a) and (29b), a population starting from 01 changes to 00 with the probability close to 1.

In M_2, if an agent with the greatest authority starts with 0, then the semantics cannot stabilize on anything else than 00. Similarly, if an agent with the greatest authority starts with 1, then the semantics cannot stabilize on anything else than 11. In M_1, it does not matter whether initial hypothesis of an agent is 0 or 1 – he can always change to 0 or 1.

Observe that if an agent with the greatest authority starts with a hypothesis other than 0 and 1, then the semantics diverge forever. This effect is partially due to the simplicity criterion that tells agents to choose maximally simple hypotheses. However, another reason for this is a very low complexity of communication games. As we shall see in the next section, increasing the complexity of communication may drive less authoritative agents to adopt more complex semantics.

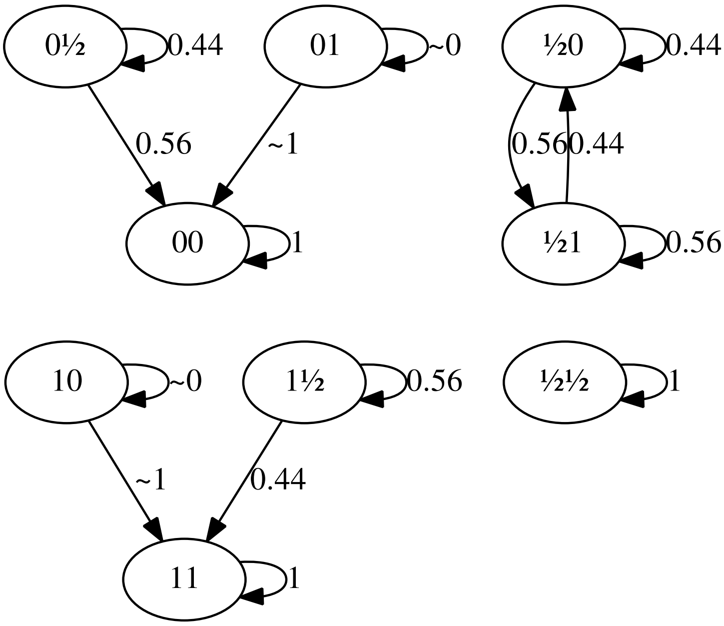


Figure III.5: Markov chain for M_1, where $A = \{1, 2\}$, $w_1 > w_2 > 0$, $H = \{0, \frac{1}{2}, 1\}$ and $X \sim B(50, 0.5)$. One topic per game.

III.8.3 Multiple Topics

A population that coordinates according to M_1 or M_2 cannot stabilize on semantics other than 0 or 1, unless the initial semantics is a constant function $s : A \to \{u\}$, for u such that $0 < u < 1$. In linguistic terms, such populations cannot arrive at semantic conventions different form the existential quantifier or the trivial quantifier *more than everything*. Proofs of Theorem 14 and Theorem 15 explain why this is the case. If agents differ in semantics and disagree on a given topic, then their maximum-reward hypotheses include either 0 or 1. The criterion of simplicity forces agents to choose the simplest hypotheses possible – hence they choose either 0 or 1.

One way of guaranteeing the emergence of more complex semantics is to relax the criterion of simplicity. We may posit that for all $h, h' \in H$, the probability of choosing h is positive and if h is simpler than h' then the probability of choosing h is greater than the probability of choosing h'. Such a relaxed criterion of simplicity may be adequate in certain contexts. We take a different direction and show that complex semantics may result from more complex communication games. We assume each communication game may contain many different topics. In some situations complex games force agents to change their current hypotheses to other than 0 or 1. Consequently, even if agents use the strong simplicity criterion, such coordination processes may stabilize on semantics other than 0 and 1. To keep things simple, we consider games consisting of two topics. At each stage of the coordination process we generate two random variates r, r' and produce an instance of the proportional description game. The instance of the game consists of four interactions: $(1, 2, r_1, v_1), (2, 1, r_1, v'_1), (1, 2, r_2, v_2), (2, 1, r_2, v'_2)$. Actually we have two symmetrical interactions (see Figure III.1), one with topic r_1 and another with topic r_2. We denote such a model by M_3.

We restrict our attention to $H = F_2$. The resulting Markov chain is simple enough to perform manual calculations and rich enough to observe some interesting features, such as complex semantics formation. In this case, complex semantics is 1/2, corresponding to the *most* quantifier.

We have seven types of authority functions that lead to different Markov chains representing instances of M_3: a) $w_1 > w_2 + w_2$, b) $w_1 = w_2 + w_2$, c) $w_1 < w_2 + w_2$, $w_1 > w_2$, d) $w_1 = w_2$, e) $w_2 > w_1 + w_1$, f) $w_2 = w_1 + w_1$, g) $w_2 < w_1 + w_1$, $w_2 > w_1$. We present the Markov chain for a).

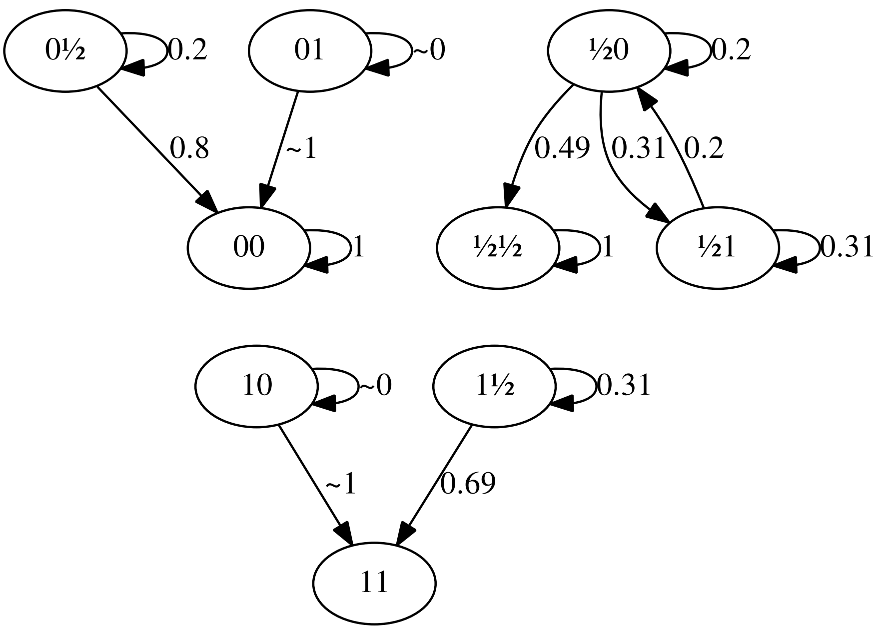


Figure III.6: Markov chain for M_3, where $A = \{1, 2\}$, $w_1 > w_2 + w_2$, $H = \{0, \frac{1}{2}, 1\}$ and $X\ B(50, 0.5)$. Two topics per game.

Theorem 16 *Fix an instance of the M_2, where $A = \{1, 2\}$, $H = F_2$, X is a* *random variable with a probability function P and w is an authority function satisfying a). Then the model is represented by the Markov chain on $S = H^2$ induced by the following probabilities:*

$$p_{uu \to uu} = 1, \text{ for all } u \in H \tag{30a}$$
$$p_{01 \to 01} = P(X = 0)^2 \tag{30b}$$

$$p_{01 \to 00} = 1 - P(X = 0)^2 \tag{30c}$$
$$p_{10 \to 10} = P(X = 0)^2 \tag{30d}$$
$$p_{10 \to 11} = 1 - P(X = 0)^2 \tag{30e}$$
$$p_{0\frac{1}{2} \to 0\frac{1}{2}} = [P(X = 0) + P(X > \frac{1}{2})]^2 \tag{30f}$$
$$p_{0\frac{1}{2} \to 00} = 1 - [P(X = 0) + P(X > \frac{1}{2})]^2 \tag{30g}$$
$$p_{\frac{1}{2}0 \to \frac{1}{2}0} = [P(X = 0) + P(X > \frac{1}{2})]^2 \tag{30h}$$
$$p_{\frac{1}{2}0 \to \frac{1}{2}1} = 2P(X = 0)P(0 < X \leqslant \frac{1}{2}) + P(0 < X \leqslant \frac{1}{2})^2 \tag{30i}$$
$$p_{\frac{1}{2}0 \to \frac{1}{2}\frac{1}{2}} = 2P(0 < X \leqslant \frac{1}{2})P(X > \frac{1}{2}) \tag{30j}$$
$$p_{\frac{1}{2}1 \to \frac{1}{2}1} = [P(X = 0) + P(0 < X \leqslant \frac{1}{2})]^2 \tag{30k}$$
$$p_{\frac{1}{2}1 \to \frac{1}{2}0} = 2P(X = 0)P(X > \frac{1}{2}) + P(X > \frac{1}{2})^2 \tag{30l}$$
$$p_{\frac{1}{2}1 \to \frac{1}{2}\frac{1}{2}} = 2P(0 < X \leqslant \frac{1}{2})P(X > \frac{1}{2}) \tag{30m}$$
$$p_{1\frac{1}{2} \to 1\frac{1}{2}} = [P(X = 0) + P(0 < X \leqslant \frac{1}{2})]^2 \tag{30n}$$
$$p_{1\frac{1}{2} \to 11} = 2P(X = 0)P(X > \frac{1}{2}) + P(X > \frac{1}{2})^2 + 2P(0 < X \leqslant \frac{1}{2})P(X > \frac{1}{2}) \tag{30o}$$

The main observation is that the coordination processes based on M_3 allow agents to change their semantics to more complex hypotheses. Since the coordination mechanism is essentially the same in all models we consider, the reason why agents may adopt more complex hypotheses lies in the complexity of communication patterns. However, this is not the only reason for this to happen. Another one is the type of authority function. It is worth noting that instances of M_3 having authority functions satisfying $w_1 = w_2$ (which we do not cover here) do not allow agents to adopt more complex hypotheses. Hence, coordination processes of M_3 combined with authority functions satisfying $w_1 = w_2$ cannot stabilize on the *most* quantifier, unless both agents have 1/2 as their hypotheses all along. However, when authority functions are differentiated, complex semantics may emerge. As one can see from equations (30j) and (30m), in some situations the quantifier *most* is achievable. The situations that allow this to happen are of peculiar nature: by Theorem 16, an agent may choose more complex semantics only if his interlocutor possesses complex semantics and has greater authority. We may conclude that agents choose more complex semantics as a result of coordination processes that combine differentiated authority functions and complex communication patterns.

If we assume that the probability P of drawing values closer to 1/2 is greater than drawing values further from 1/2 (for example, if the random variable X behaves similarly to a normally distributed variable with $\mu = 1/2$), the chances of changing from $\frac{1}{2}0$ or $\frac{1}{2}1$ to $\frac{1}{2}\frac{1}{2}$ are higher and increase while $P(X = 0)$ and $P(X = 1)$ are getting smaller.

III.8.4 Outlook: Spatial Separation

In [Allen, 1977] Thomas Allen reports on experiments aimed at finding out how distance between engineers' offices influences frequency of communication. The experiments show that there is a very strong negative correlation between these two factors. This effect is known as the Allen curve. At first sight, the discovery is not very surprising. Nevertheless, it came as a surprise, at least to enterprise engineers, that distance is prevailing when compared to other factors. Such effects scale up to larger populations and were already known to the regional scientists since the nineteenth century [Carey, 1858]. The so called gravity models of interaction behaviour become very popular since the work of Stuart [Stewart, 1941], who postulated that the intensity of interaction between populations is inversely proportional to the squared distance between their centres. Subsequently, gravity models have been adopted in such fields as economics, sociology and enterprise management (see, e.g., [Sen and Smith, 1995]).

If we want to describe the evolution of semantics for larger populations, we must account for the effect of the distance factor. It is not immediately obvious how to describe this dependency mathematically. Foundations of the subject indicate how this could be done [Sen and Smith, 1995]. However, this question is appropriate for a separate work and we do not develop our current proposal in this direction here. Instead, we propose a simple approximation of such a behaviour.

Assume each agent a is assigned a different point on the plane. For agents a, b, define the spacial separation c_{ab} as the euclidean distance between a and b. To describe the effects of distance, we posit that the chances that agents a and b communicate are proportional to $e^{-c_{ab}}$ (in many contexts, the inverse exponential function proved to approximate spatial interaction behaviour better than other functions [Sen and Smith, 1995], however applicability of this kind of function to situations of our interest is a matter for a separate study). Now, agents do not communicate with probability 1, but according to this new rule which results in communication patterns where distant agents occur less frequently. The analysis of the effects of distance on semantics evolution is left for future work.

III.9 Conclusions

We have presented a plausible mechanism which allows a population of communicating agents to arrive at a common semantic convention for a quantifier construction which corresponds to a proportional quantifier. We show that higher complexity of communication patterns may lead to the emergence of more complex semantics. Moreover, our solution includes authority of agents as an important factor of coordination. We observe a mathematical connection between the possibility of convergence and specific levels of agents authority. According to our model, appropriate differentiation of authority is important for the effectiveness of coordination, whereas equality makes coordination more difficult or, in some circumstances, even impossible. These properties agree with our experience. It would be desirable to investigate which types of authority functions make coordination more effective, especially for larger populations.

It is known that many natural properties obey the rule of normal distribution. Hence, the assumption that topics of conversations obey a similar rule seems reasonable. If this is so, the *most* quantifier should be used more frequently than other proportional quantifiers (apart from *some* and *all*) as it allows for more efficient information sharing. This frequency distribution is in fact evidenced by linguistic corpora [Thorne and Szymanik, 2015]. Hence, if learning by coordination plays a major role in acquisition of proportional quantifiers then a realistic coordination model for learning such quantifiers should tend to converge to 1/2. Out remarks about Theorem 16 indicate that our model may have this property. However, this matter requires further research.

Finally, we hypothesise that there is a close connection between learning by recognizing (Section I.3) and learning by coordination. Observe that the results on model M_3 show that the more authoritative agent seems to behave like a teacher in the learning by recognizing scenario (he does not change his semantics). This may indicate that recognition and coordination are not two different cognitive mechanisms. Instead, learning by recognizing may be seen as a specific manifestation of a more generic mechanism of coordination.

Bibliography

[Allen, 1977] Allen, T. J. (1977). *Managing the Flow of Technology : Technology Transfer and the Dissemination of Technological Information within the R&D Organization*. MIT Press, Cambridge, MA.

[Arora and Barak, 2009] Arora, S. and Barak, B. (2009). *Computational Complexity: A Modern Approach*. Cambridge University Press, New York, NY, USA, 1st edition.

[Barwise, 1979] Barwise, K. J. (1979). On branching quantifiers in English. *Journal of Philosophical Logic*, (1):47–80.

[Beck, 1997] Beck, S. (1997). On the semantics of comparative conditionals. *Linguistics and Philosophy*, 20(3):229–271.

[Blass and Gurevich, 1986] Blass, A. and Gurevich, Y. (1986). Henkin quantifiers and complete problems. *Annals of Pure and Applied Logic*, 32:1–16.

[Blythe et al., 2016] Blythe, R. A., Jones, A. H., and Renton, J. (2016). Spontaneous dialect formation in a population of locally aligning agents. In Roberts, S. G., Cuskley, C., McCrohon, L., Barceló-Coblijn, L., Fehér, O., and Verhoef, T., editors, *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11)*. Online at http://evolang.org/neworleans/papers/19.html.

[Brown et al., 2016] Brown, A. M., Isse, A., and Lindsey, D. T. (2016). The color lexicon of the Somali language. *Journal of Vision*, 16(5):14.

[Carey, 1858] Carey, H. C. (1858). *Principles of Social Science*. Lippincott, reprint edition.

[Carnap, 1947] Carnap, R. (1947). *Meaning and Necessity*. The University of Chicago Press.

[Carnap, 1958] Carnap, R. (1958). *Introduction to Symbolic Logic and its Applications*. New York, Dover Publications.

[Chomsky, 1957] Chomsky, N. (1957). *Syntactic Structures*. Mouton The Hague.

[Chomsky, 1986] Chomsky, N. (1986). *Knowledge of Language: its Nature, Origin, and Use*. Greenwood Publishing Group.

[Clark, 2010] Clark, R. (2010). On the learnability of quantifiers. In van Benthem, J. and ter Meulen, A., editors, *Handbook of Logic and Language*, pages 909–922. Elsevier. 2nd edition.

[Costa Florêntio, 2002] Costa Florêntio, C. (2002). Learning generalized quantifiers. In *Proc. 7th ESSLLI Student Session*.

[Culicover and Jackendoff, 1999] Culicover, P. W. and Jackendoff, R. (1999). The view from the periphery: The English comparative correlative. *Linguistic Inquiry*, 30(4):543–571.

[Davidson, 1967] Davidson, D. (1967). Truth and meaning. *Synthese*, 17(1):304–323.

[Davidson, 1986] Davidson, D. (1986). A nice derangement of epitaphs. In Lepore, E., editor, *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, pages 433–446. Blackwell.

[de Saussure, 1916] de Saussure, F. (1916). *Cours de Linguistique Générale*. Payot, Lausanne; Paris.

[DeGroot, 1974] DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.

[Den Dikken, 2005] Den Dikken, M. (2005). Comparative correlatives comparatively. *Linguistic Inquiry*, 36(4):497–532.

[Dummett, 1975] Dummett, M. A. E. (1975). What is a theory of meaning? In Guttenplan, S., editor, *Mind and Language*. Oxford University Press.

[Edmonds, 1987] Edmonds, J. (1987). Paths, trees, and flowers. In Gessel, I. and Rota, G.-C., editors, *Classic Papers in Combinatorics*, Modern Birkhäuser Classics, pages 361–379. Birkhäuser Boston.

[Enderton, 1970] Enderton, H. (1970). Finite partially ordered quantifiers. *Z. Math. Logik Grundlag. Math.*, 16:393–397.

[Fagin, 1974] Fagin, R. (1974). Generalized first-order spectra and polynomial-time recognizable sets. In Karp, R. M., editor, *Complexity of Computation, SIAM-AMS Proceedings*, volume 7, pages 43–73, Providence.

[Fauconnier, 1975] Fauconnier, G. (1975). Do quantifiers branch? *Linguistic Inquiry*, 6(4):555–567.

[Fay et al., 2010] Fay, N., Garrod, S., Roberts, L., and Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive Science*, 34(3):351–386.

[Feller, 1968] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Vol. 1.* Wiley, 3rd edition.

[Frege, 1892] Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*, 100:25–50.

[Frixione, 2001] Frixione, M. (2001). Tractable competence. *Minds and Machines*, 11(3):379–397.

[Gabbay and Moravcsik, 1974] Gabbay, D. M. and Moravcsik, J. M. E. (1974). Branching quantifiers, English, and Montague-grammar. *Theoretical Linguistics*, 1(1-3):141–157.

[Galantucci, 2005] Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive science*, 29(5):737–767.

[Galantucci and Garrod, 2011] Galantucci, B. and Garrod, S. (2011). Experimental semiotics: a review. *Frontiers in Human Neuroscience*, 5.

[Gierasimczuk, 2007] Gierasimczuk, N. (2007). The problem of learning the semantics of quantifiers. In *TBiLLC'05: 6th International Tbilisi Symposium on Logic, Language, and Computation. Revised Selected Papers*, volume 4363 of *Lecture Notes in Artificial Intelligence*, pages 117–126. Springer.

[Gierasimczuk and Szymanik, 2009] Gierasimczuk, N. and Szymanik, J. (2009). Branching quantification v. two-way quantification. *Journal of Semantics*, 26(4):367–392.

[Gold, 1967] Gold, E. M. (1967). Language identification in the limit. *Information and Control*, 10:447–474.

[Gottlob, 1997] Gottlob, G. (1997). Relativized Logspace and generalized quantifiers over finite ordered structures. *The Journal of Symbolic Logic*, 62(2):545–574.

[Grice, 1975] Grice, H. P. (1975). Logic and conversation. In Cole, P. and Morgan, J., editors, *Syntax and semantics*, volume 3. New York: Academic Press.

[Guenthner and Hoepelman, 1974] Guenthner, F. and Hoepelman, J. P. (1974). A note on the representation of branching quantifiers. *Theoretical Linguistics*, 3:285–291.

[Hardy and Wright, 1979] Hardy, G. and Wright, E. (1979). *An Introduction to the Theory of Numbers*. Oxford University Press.

[Henkin, 1961] Henkin, L. (1961). Some remarks on infinitely long formulas. In *Infinitistic Methods*, pages 167–183. Polish Scientific Publishers, Warsaw, Pergamon Press, New York.

[Hintikka, 1973] Hintikka, J. (1973). Quantifiers vs. quantification theory. *Dialectica*, 27(3-4):329–358.

[Hintikka, 1976] Hintikka, J. (1976). Partially ordered quantifiers vs. partially ordered ideas. *Dialectica*, 30(1):89–99.

[Immerman, 1999] Immerman, N. (1999). *Descriptive Complexity*. Springer Verlag.

[Kalociński et al., 2015] Kalociński, D., Gierasimczuk, N., and Mostowski, M. (2015). Quantifier learning: An agent-based coordination model. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 1853–1854, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

[Kalociński and Godziszewski, 2016] Kalociński, D. and Godziszewski, M. T. (2016). Computational complexity of the Barwise sentence and similar natural language constructions. submitted.

[Katz, 1966] Katz, J. J. (1966). *The Philosophy of Language*. New York, Harper & Row.

[Katz, 1972] Katz, J. J. (1972). *Semantic Theory*. New York,Harper & Row.

[Katz and Fodor, 1963] Katz, J. J. and Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 39(2):170–210.

[Kay, 1975] Kay, P. (1975). Synchronic variability and diachronic change in basic color terms. *Language in society*, 4(03):257–270.

[Keenan, 1992] Keenan, E. L. (1992). Beyond the Frege boundary. *Linguistics and Philosophy*, 15(2):199–221.

[Keenan, 1996] Keenan, E. L. (1996). Further beyond the Frege boundary. In van Eijck, J. and van der Does, J., editors, *Quantifiers, logic, and language*, CSLI Lecture Notes, pages 179–201. California: Stanford University.

[Keenan and Ralalaoherivony, 2014] Keenan, E. L. and Ralalaoherivony, B. (2014). Correlational comparatives (CCs) in Malagasy. The 21st Annual Meeting of Austronesian Formal Linguistics Association (AFLA 21).

[Kirby, 2000] Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In Knight, C., editor, *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, pages 303–323. Cambridge University Press.

[Kirby et al., 2014] Kirby, S., Griffiths, T., and Smith, K. (2014). Iterated learning and the evolution of language. *Current opinion in neurobiology*, 28:108–114.

[Kontinen and Szymanik, 2008] Kontinen, J. and Szymanik, J. (2008). A remark on collective quantification. *Journal of Logic, Language and Information*, 17(2):131–140.

[Kosterman and Gierasimczuk, 2015] Kosterman, S. and Gierasimczuk, N. (2015). Collective learning in games through social networks.

[Kołodziejczyk, 2002] Kołodziejczyk, L. A. (2002). *The expressive power of Henkin quantifiers with dualization.* Master's Thesis, Institute of Philosophy, University of Warsaw.

[Krynicki and Mostowski, 1992] Krynicki, M. and Mostowski, M. (1992). Decidability problems in languages with Henkin quantifiers. *Annals of Pure and Applied Logic*, 28(2):149–172.

[Lambalgen and Hamm, 2005] Lambalgen, M. V. and Hamm, F. (2005). *The Proper Treatment of Events.* Wiley-Blackwell.

[Lenneberg and Roberts, 1955] Lenneberg, E. H. and Roberts, J. M. (1955). The language of experience: a study in methodology. Technical report, [Cambridge, Mass.: Massachusetts Institute of Technology, Center for International Studies, 1955].

[Lewis, 1969] Lewis, D. (1969). *Convention.* Cambridge: Harvard University Press.

[Libkin, 2004] Libkin, L. (2004). *Elements of Finite Model Theory.* Springer.

[Lindsey and Brown, 2009] Lindsey, D. T. and Brown, A. M. (2009). World Color Survey color naming reveals universal motifs and their within-language diversity. *Proceedings of the National Academy of Sciences*, 106(47):19785–19790.

[Lindström, 1966] Lindström, P. (1966). First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195.

[McCawley, 1988] McCawley, J. D. (1988). The comparative conditional constructions in English, German and Chinese. In *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, pages 176–187.

[Mendelson, 2009] Mendelson, E. (2009). *Introduction to Mathematical Logic.* Chapman & Hall/CRC, 5th edition.

[Moschovakis, 1990] Moschovakis, Y. (1990). Sense and denotation as algorithm and value. In J. Oikkonen, J. V., editor, *Lecture Notes in Logic 2*, pages 210–249. Springer.

[Moschovakis, 2006] Moschovakis, Y. (2006). A logical calculus of meaning and synonymy. *Linguistics and Philosophy*, 29(1):27–89.

[Mostowski, 1994] Mostowski, M. (1994). Kwantyfikatory rozgałęzione a problem formy logicznej. In Omyła, M., editor, *Nauka i język*, pages 201–242. Biblioteka Myśli Semiotycznej.

[Mostowski, 1998] Mostowski, M. (1998). Computational semantics for monadic quantifiers. *Journal of Applied Non–Classical Logics*, 8(1-2):107–121.

[Mostowski and Szymanik, 2012] Mostowski, M. and Szymanik, J. (2012). Semantic bounds for everyday language. *Semiotica*, 188(1/4):363–372.

[Mostowski and Wojtyniak, 2004] Mostowski, M. and Wojtyniak, D. (2004). Computational complexity of the semantics of some natural language constructions. *Annals of Pure and Applied Logic*, 127(1-3):219–227.

[Pauw and Hilferty, 2012] Pauw, S. and Hilferty, J. (2012). The emergence of quantifiers. *Experiments in Cultural Language Evolution*, 3:277.

[Peters and Westerståhl, 2006] Peters, S. and Westerståhl, D. (2006). *Quantifiers in Language and Logic*. Oxford University Press, Oxford.

[Piantadosi et al., 2012] Piantadosi, S. T., Tenenbaum, J. B., and Goodman, N. D. (2012). Bootstrapping in a language of thought: a formal model of numerical concept learning. *Cognition*, 123(2):199 – 217.

[Quine, 1960] Quine, W. V. (1960). *Word and Object*. The MIT Press.

[Quine, 1970] Quine, W. V. (1970). *Philosophy of Logic*. Prentice-Hall.

[Reichenbach, 1947] Reichenbach, H. (1947). *Elements of Symbolic Logic*. Free press paperback. Macmillan Co.

[Ristad, 1993] Ristad, E. (1993). *The Language Complexity Game*. The MIT Press.

[Russell and Whitehead, 1910] Russell, B. and Whitehead, A. N. (1910). *Principia Mathematica*, volume 1. Cambridge University Press, Cambridge.

[Schlotterbeck and Bott, 2013] Schlotterbeck, F. and Bott, O. (2013). Easy solutions for a hard problem? The computational complexity of reciprocals with quantificational antecedents. *Journal of Logic, Language and Information*, 22(4):363–390.

[Sedgewick and Wayne, 2011] Sedgewick, R. and Wayne, K. (2011). *Algorithms*. Addison-Wesley.

[Sen and Smith, 1995] Sen, A. and Smith, T. (1995). *Gravity Models of Spatial Interaction Behavior*. Advances in Spatial and Network Economics. Springer Berlin Heidelberg.

[Sevenster, 2006] Sevenster, M. (2006). Henkin quantifiers: Logic, games, and computation. *Bulletin of the EATCS*, 89:136–155.

[Skyrms, 2010] Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. OUP Oxford.

[Spike et al., 2016] Spike, M., Stadler, K., Kirby, S., and Smith, K. (2016). Minimal requirements for the emergence of learned signaling. *Cognitive Science*.

[Stanosz, 1974a] Stanosz, B. (1974a). Meaning and interpretation. *Semiotica*, 11(4):347–358.

[Stanosz, 1974b] Stanosz, B. (1974b). Status poznawczy semantyki. *Studia Semiotyczne*, 5:101–115. Translation at `http://studiaes.pts.edu.pl/volumev/sses_V_3.pdf`.

[Stanosz, 1999] Stanosz, B. (1999). W sprawie krytyki teorii języka. In Pelc, J., editor, *Logika języka naturalnego*, volume 43 of *Biblioteka Myśli Semiotycznej*, pages 223–227. Znak-Język-Rzeczywistość, Polskie Towarzystwo Semiotyczne, Warszawa. No translation available.

[Steels, 1995] Steels, L. (1995). A self-organizing spatial vocabulary. *Artificial Life*, 2:319–332.

[Steels, 2011] Steels, L. (2011). Modeling the cultural evolution of language. *Physics of Life Reviews*, 8(4):339 – 356.

[Steels, 2012] Steels, L. (2012). Grounding language through evolutionary language games. In *Language Grounding in Robots*, pages 1–22. Springer US.

[Steels and Belpaeme, 2005] Steels, L. and Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: a case study for colour. *Behavioral and Brain Sciences*, 28(4):469–489.

[Stenius, 1976] Stenius, E. (1976). Comments on Jaakko Hintikka's paper "Quantifiers vs. quantification theory". *Dialectica*, 30(1):67–88.

[Stewart, 1941] Stewart, J. Q. (1941). An inverse distance variation for certain social influences. *Science*, 93(2404):89–90.

[Suppes, 1980] Suppes, P. (1980). Procedural semantics. In Haller, R. and Grassl, W., editors, *Language, Logic, and Philosophy: Proceedings of the 4th International Wittgenstein Symposium*, pages 27–35. Hölder-Pichler-Tempsy, Vienna.

[Suszko, 1957a] Suszko, R. (1957a). Logika formalna a niektóre zagadnienia teorii poznania. *Myśl Filozoficzna*, 2(28):27–56.

[Suszko, 1957b] Suszko, R. (1957b). Logika formalna a niektóre zagadnienia teorii poznania: Diachroniczna logika formalna. *Myśl Filozoficzna*, 3(29):34–67.

[Suszko, 1998a] Suszko, R. (1998a). Logika formalna a niektóre zagadnienia teorii poznania. In Omyła, M., editor, *Wybór Pism*, volume 42 of *Biblioteka Myśli Semiotycznej*, pages 57 – 108. Znak-Język-Rzeczywistość, Polskie Towarzystwo Semiotyczne, Warszawa. No translation available. This is a single chapter composed of two articles: [Suszko, 1957b] and [Suszko, 1957a].

[Suszko, 1998b] Suszko, R. (1998b). *Wybór pism*. Znak-Język-Rzeczywistość.

[Szymanik, 2010] Szymanik, J. (2010). Computational complexity of polyadic lifts of generalized quantifiers in natural language. *Linguistics and Philosophy*, 33(3):215–250.

[Szymanik, 2016] Szymanik, J. (2016). Quantifiers and cognition: Logical and computational perspectives. *Studies in Linguistics and Philosophy. Springer, forthcoming.*

[Szymanik and Zajenkowski, 2010] Szymanik, J. and Zajenkowski, M. (2010). Comprehension of simple quantifiers. Empirical evaluation of a computational model. *Cognitive Science: A Multidisciplinary Journal*, 34(3):521–532.

[Tarski, 1933] Tarski, A. (1933). *Pojęcie Prawdy w Językach Nauk Dedukcyjnych*, volume 34 of *Prace Towarzystwa Naukowego Warszawskiego*. Warszawa. Translation in [Tarski, 1936].

[Tarski, 1936] Tarski, A. (1936). The concept of truth in formalized languages. In Tarski, A., editor, *Logic, Semantics, Metamathematics*, pages 152–278. Oxford University Press.

[Thorne and Szymanik, 2015] Thorne, C. and Szymanik, J. (2015). Semantic complexity of quantifiers and their distribution in corpora. *IWCS 2015*, page 64.

[Tichy, 1969] Tichy, P. (1969). Intension in terms of Turing machines. *Studia Logica*, 24(1):7–21.

[Tiede, 1999] Tiede, H.-J. (1999). Identifiability in the limit of context-free generalized quantifiers. *Journal of Language and Computation*, 1:93–102.

[van Benthem, 1986] van Benthem, J. (1986). *Essays in logical semantics*. Reidel.

[van Benthem, 1987] van Benthem, J. (1987). Verisimilitude and conditionals. In Kuipers, T., editor, *What is closer-to-the-truth?*, pages 103–128. Amsterdam Rodopi.

[van Benthem, 1989] van Benthem, J. (1989). Polyadic quantifiers. *Linguistics and Philosophy*, 12(4):437–464.

[van Rooij, 2008] van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, 32:939–984.

[Vasiliu, 1972] Vasiliu, E. (1972). *Outline of a Semantic Theory of Kernel Sentences*. De Gruyter Mouton, Berlin, Boston.

[Vasiliu, 1981] Vasiliu, E. (1981). *Elementy Semantycznej Teorii Języków Naturalnych: Semantyka Logiczna Zdań Jądrowych*. Zakład Narodowy im. Ossolińskich Wydawnictwo.

[Wacewicz et al., 2016] Wacewicz, S., Żywiczyński, P., and Jasiński, A. (2016). Language evolution and language origins in teaching linguistics at the university level. In Roberts, S. G., Cuskley, C., McCrohon, L., Barceló-Coblijn, L., Fehér, O., and Verhoef, T., editors, *The Evolution of Language:*

Proceedings of the 11th International Conference (EVOLANGX11). Online at http://evolang.org/neworleans/papers/62.html.

[Walkoe, 1970] Walkoe, W. (1970). Finite partially-ordered quantification. *Journal of Symbolic Logic*, 35:535–555.

[Westerståhl, 1984] Westerståhl, D. (1984). Some results on quantifiers. *Notre Dame Journal of Formal Logic*, 25(2):152–170.

[Wittgenstein, 1953] Wittgenstein, L. (1953). *Philosophical Investigations.* Basil Blackwell, Oxford.

[Wright, 1997] Wright, C. (1997). The indeterminacy of translation. *A Companion to the Philosophy of Language*, pages 397–426.