STRONG MEANING HYPOTHESIS FROM A COMPUTATIONAL PERSPECTIVE

JAKUB SZYMANIK

Institute for Logic, Language and Information
University of Amsterdam
szymanik@science.uva.nl

In this paper we study the computational complexity of reciprocal sentences with quantified antecedents. We observe a computational dichotomy between different interpretations of reciprocity and its connection with Strong Meaning Hypothesis.

1. Introduction

The English reciprocal expressions *each other* and *one another* are common elements of everyday English. In this paper we study the computational complexity of reciprocal sentences with quantified antecedents. We bring attention to possible cognitive consequences of complexity issues in semantics. Particularly, by observing a computational dichotomy between different interpretations of reciprocity we shed some light on the epistemological status of the so-called Strong Meaning Hypothesis (proposed in Dalrymple et al. 1998).

Our results also give an additional argument for the robustness of semantic distinction established by Dalrymple et al. 1998. Moreover, we present NP-complete natural language quantifiers which occur frequently in everyday English. As far as we are aware, all other known NP-complete semantic constructions are based on ambiguous and artificial branching operations.

1.1. Basic examples

We start by recalling examples of reciprocal sentences versions of which can be found in the corpus of English (see footnote 1 in Dalrymple et al. 1998). Let us consider the sentences (1)–(3).

(1) An even number of parliament members refer to each other indirectly.

(2) Most Boston pitchers sat alongside each other.

(3) Some Pirates of the Caribbean were staring at each other in surprise.

The possible interpretations of reciprocity exhibit a wide range of variations. In this paper we will restrict ourselves to these three possibilities. Sentence (1) implies that there is a subset of parliament members of even cardinality such that each parliament member in that subset refers to some statement of each of the other parliament members in that subset. However, the reciprocals in the sentences (2) and (3) have different meanings. Sentence (2) entails that each of most of the pitchers is directly or indirectly in the relation of sitting alongside with each of the other pitchers from a set containing most pitchers. Sentence (3) says that there was a group of pirates such that every pirate belonging to the group stared at some other pirate from the group. Following Dalrymple et al. 1998 we will call the illustrated reciprocal meanings *strong*, *intermediate*, and *weak*, respectively.

2. Reciprocals as polyadic quantifiers

Monadic generalized quantifiers provide the most straightforward way to define the semantics of noun phrases in natural language. Sentences with reciprocal expressions transform such monadic quantifiers into polyadic ones. We will analyze reciprocal expressions in that spirit by defining appropriate lifts on monadic quantifiers. These lifts allow us to express the meanings of sentences with reciprocals in the compositional way with respect to monadic quantifiers occurring in sentences. For the sake of simplicity we will restrict ourselves to reciprocal sentences with monotone increasing quantifiers in the antecedent. However, our definitions can be extended to cover also sentences with decreasing and non-monotone quantifiers, for example following the strategy of bounded composition as explained by Dalrymple et al. 1998.

In order to define the meaning of the strong reciprocity we make use of well-know operation on quantifiers called *Ramseyfication*. Let Q be a monadic monotone increasing quantifier, we define:

$$\mathsf{Ram}_{\mathsf{S}}(\mathsf{Q})AR \iff \exists X \subseteq A[\mathsf{Q}(X) \wedge \forall x, y \in X(x \neq y \Rightarrow R(x,y))].$$

The result of such a lift is called *Ramsey quantifier*.

In an analogous way we define two other lifts to express intermediate and weak reciprocity. For intermediate reciprocity we have the following:

$$\mathsf{Ram_I}(\mathsf{Q})AR \iff \exists X \subseteq A[\mathsf{Q}(X) \wedge \forall x, y \in X$$
$$(x \neq y \Rightarrow \exists \text{ sequence } z_1, \ldots, z_\ell \in X \text{ such that}$$
$$(z_1 = x \wedge R(z_1, z_2) \wedge \ldots \wedge R(z_{\ell-1}, z_\ell) \wedge z_\ell = y)].$$

For weak reciprocity we take the following lift:

$$\mathsf{Ram_W}(\mathsf{Q})AR \iff \exists X \subseteq A[\mathsf{Q}(X) \wedge \forall x \in X \exists y \in X(x \neq y \wedge R(x, y))].$$

All these lifts produce polyadic quantifiers of type (1, 2). We will call the values of these lifts strong, intermediate and weak reciprocity, respectively. The linguistic application of them is straightforward. For example, formulae (4)–(6) give readings to the sentences (1)–(3).

(4) $\mathsf{Ram_S}(\mathsf{EVEN})$MP Refer.

(5) $\mathsf{Ram_I}(\mathsf{MOST})$Pitcher Sit.

(6) $\mathsf{Ram_W}(\mathsf{SOME})$Pirate Staring.

3. Complexity of the reciprocal lifts

3.1. Strong reciprocity

We will restrict ourselves to finite models. We identify models of the form $M = (U, A^M, R^M)$, where $A^M \subseteq U$ and $R^M \subseteq U^2$, with undirected graphs. In graph theoretical terms we can say that $M \models \mathsf{Ram_S}(\mathsf{Q})AR$ if and only if there is a complete subgraph in M of size bounded by the quantifier Q. For example, to decide whether some model M belongs to $\mathsf{Ram_S}(\exists^{\geq k})$ we must solve the CLIQUE problem for M and k. A brute force algorithm to find a clique in a graph is to examine each subgraph with at least k vertices and check to see if it forms a clique. This means that for every fixed k the computational complexity of $\mathsf{Ram_S}(\exists^{\geq k})$ is in PTIME. However, in general — for changing k — this is a well-known NP-complete problem.

Let us define a unary counting quantifier $\exists_y^{\geq k}$ — expressing the statement $\mathsf{At\ least\ k}$ for a natural number k — as follows:

$$M \models \exists^{\geq k} y\varphi(y)[v] \iff card(\varphi^{(M,y,v)}) \geq v(k).$$

Then it is obvious that:

Proposition 1 *The Quantifier* $\mathsf{Ram}_\mathsf{S}(\exists^{\geq k})$ *is NP-complete.*

Therefore, strong reciprocal sentences with counting quantifiers in antecedents are NP-complete.

We can give one more general example of strong reciprocal sentences which are NP-complete. Let us consider the following sentences:

(7) Most members of the parliament refer to each other indirectly.

(8) At least one third of the members of the parliament refer to each other.

(9) At least $q \times 100\%$ of the members of the parliament refer to each other.

We will call these sentences the *strong reciprocal sentences with proportional quantifiers*. Their general form is given by the sentence schema (9), where q can be interpreted as any rational number between 0 and 1. These sentences say that there is a clique A in M such that $\frac{card(A)}{card(U)} \geq q$.

For any rational number $q \in]0, 1[$ we say that a set $A \subseteq U$ is q-big if and only if $\frac{card(A)}{card(U)} \geq q$. In this sense q determines a proportional Ramsey quantifier $\mathsf{R_q}$ of type (2) such that $M \models \mathsf{R_q} xy\ \varphi(x, y)$ iff there is a q-big $A \subseteq |M|$ such that for all $a, b \in A$, $M \models \varphi(a, b)$. Obviously such quantifiers might be used to express meanings of sentences like (7)–(9). It was observed by Mostowski and Szymanik 2007 that:

Proposition 2 *Let* $q \in]0, 1[\cap \mathbb{Q}$*, then the quantifier* $\mathsf{R_q}$ *is NP-complete.*

In fact one can show much more general results, but we leave this rather technical enterprise for the full paper.

4. The intermediate and weak lifts

Analogously to the case of strong reciprocity we can also express the meanings of intermediate and weak reciprocal lifts in graph-theoretical terms. We say that $M \models \mathsf{Ram_I}(\mathsf{Q})AR$ if and only if there is a connected subgraph in M of size bounded by the quantifier Q. And $M \models \mathsf{Ram_W}(\mathsf{Q})AR$ if and only if there is a subgraph in M of size bounded by the quantifier Q without isolated vertices.

We prove that the class of PTIME quantifiers is closed under intermediate lift and weak lift.

Proposition 3 *If* Q *is in PTIME, then* $\mathsf{Ram_I}(\mathsf{Q})$ *is in PTIME.*

Proposition 4 *If* Q *is in PTIME, then* $\mathsf{Ram_W}(\mathsf{Q})$ *is in PTIME.*

These results show that intermediate and weak reciprocal lifts do not increase the complexity of quantifier sentences in such drastic ways as in the case of strong reciprocal lifts. In other words, in many natural language situations intermediate and weak interpretations are relatively easy as opposed to the strong reciprocal reading.

5. The complexity perspective on SMH

Dalrymple et al. 1998 proposed a pragmatic principle, called Strong Meaning Hypothesis, to predict the proper reading of sentences containing reciprocal expressions. According to SMH the reciprocal expression is interpreted as having logically strongest truth conditions that are consistent with a given context. Therefore, if it is only consistent with specified facts, then the statement containing *each other* will be interpreted as strong reciprocal sentence. Otherwise, the interpretation will shift towards logically weaker readings, intermediate or weak, depending on the context.

SMH is a quite effective pragmatic principle. We will discuss shifts it predicts from a computational point of view using the results provided in the previous section.

Let us first think about the meaning of a sentence in the intensional way — identifying the meaning of an expression with an algorithm recognizing its denotation in a finite model. Such algorithms can be described by investigating how language users evaluate the truth-value of sentences in various situations. On the cognitive level it means that subjects have to be equipped with mental devices to deal with meanings of expressions. Moreover, it is cognitively plausible to assume that we have one mental device to deal with most instances of the same logical notion. For example, we believe that there is one mental algorithm to deal with counting quantifiers in most of the possible contexts. In the case of logical expressions, as quantifiers, this analogy seems uncontroversial.

However, notice that some sentences are too hard for identifying their truth-value directly. Programming experience suggests that we can claim a sentence to be difficult when it can not be computed in polynomial time (see Mostowski and Szymanik 2005 for a more detailed discussion). Despite the fact that some sentences are too hard for direct comprehension we can identify their inferential relations with relatively easier sentences. For instance, knowing that φ implies ψ and that ψ is not true we can easily decide that φ is false, no matter how complex is φ.

According to SMH any reciprocal sentence, if it is only possible, should be interpreted as strong reciprocal sentence. We showed that strong interpretation is sometimes NP-complete. Therefore, it is reasonable to suspect that in some linguistic situations strong reciprocal interpretation is cognitively much more difficult than intermediate or weak interpretation. If it happens to be too hard, then the subject will try to establish the truth-value of a sentence indirectly, by shifting to the accessible inferential meanings. They are — depending on context — the intermediate or the weak interpretation. Summing up, our descriptive complexity perspective on reciprocity is consistent with SMH. Moreover, it gives a cognitively reasonable argument explaining some of SMH predictions.

Bibliography

Dalrymple, M., Kanazawa, M., Kim, Y., Mchombo, S., and Peters, S.: 1998, Reciprocal expressions and the concept of reciprocity, *Linguistics and Philosophy* 21, 159–210

Mostowski, M. and Szymanik, J.: 2005, Semantical bounds for everyday language, *Semiotica* to appear(see also ILLC Preprint Series, PP-2006-40)

Mostowski, M. and Szymanik, J.: 2007, Computational complexity of some Ramsey quantifiers in finite models, *The Bulletin of Symbolic Logic* 13, 281–282