Pragmatic identification of the witness sets

Livio Robaldo* **and Jakub Szymanik**$^+$

*Department of Computer Science, University of Turin.
$^+$Institute of Artificial Intelligence, University of Groningen.
robaldo@di.unito.it, jakub.szymanik@gmail.com

Abstract

Among the readings available for NL sentences, those where two or more sets of entities are independent of one another are particularly challenging from both a theoretical and an empirical point of view. Those readings are termed here as 'Independent Set (IS) readings'. Standard examples of such readings are the well-known Collective and Cumulative Readings. (Robaldo, 2011) proposes a logical framework that can properly represent the meaning of IS readings in terms of a set-Skolemization of the witness sets. One of the main assumptions of Robaldo's logical framework, drawn from (Schwarzschild, 1996), is that pragmatics plays a crucial role in the identification of such witness sets. Those are firstly identified on pragmatic grounds, then logical clauses are asserted on them in order to trigger the appropriate inferences. In this paper, we present the results of an experimental analysis that appears to confirm Robaldo's hypotheses concerning the pragmatic identification of the witness sets.

Keywords: quantifiers, pragmatics, witness sets

1. Introduction

This paper is about the truth values of the Independent Set (IS) readings of NL sentences in the simple form 'Subject-Verb-Object'. IS readings are interpretations where two or more sets of entities are independent of one another. Four kinds have been identified in the literature, since (Scha, 1981):

(1)

a. **Branching Quantifier Readings**, e.g. *Exactly two students of mine have seen exactly three drug-dealers in front of the school.*

b. **Collective Readings**, e.g. *Exactly three boys made exactly one chair*.

c. **Cumulative Readings**, e.g. *Exactly three boys invited exactly four girls*.

d. **Cover Readings**, e.g. *Exactly three children ate exactly five pizzas*.

The preferred reading of (1.a) is the one where there are exactly two students and exacly three drug-dealers and each of the students saw each of the drug-dealers. (1.b) may be true in case three boys cooperated in the construction of a single chair. In the preferred reading of (1.c), there are three boys and four girls such that each of the boys invited at least one girl, and each of the girls was invited by at least one boy. Finally, (1.d) allows for any sharing of five pizzas between three children. In Cumulative Readings, the single actions are carried out by *atomic*[1] individuals only, while in (1.d) it is likely that the pizzas are shared among subgroups of children. For instance, the sentence is satisfied by the following extension of ate' ('$\oplus$' is the standard sum operator, from (Link, 1983)):

(2)

$$\|ate'\|^M \equiv \{\langle c_1 \oplus c_2, p_1 \oplus p_2\rangle, \langle c_2 \oplus c_3, p_3 \oplus p_4\rangle, \langle c_3, p_5\rangle\}$$

In (2), children c_1 and c_2 (cut into slices and) shared pizzas p_1 and p_2, c_2 and c_3 (cut into slices and) shared p_3 and p_4, and c_3 also ate pizza p_5 on his own.

Branching Quantifier Readings have been the more controversial (cf. (Beghelli et al., 1997) and (Gierasimczuk and Szymanik, 2009)), as many authors claim that those readings are always sub-cases of Cumulative Readings. Collective and Cumulative Readings have been largely studied; see (Link, 1983), (Beck and Sauerland, 2000), (Ben-Avi and Winter, 2003), and (Kontinen and Szymanik, 2008) to begin with. However, the focus here is on Cover readings. This paper assumes, following (van der Does and Verkuyl, 1996), (Schwarzschild, 1996), (Kratzer, 2007), that they are *the* IS readings, of which the three kinds exemplified in (1.a-c) are merely special cases. The name "Cover readings" comes from the fact that they are traditionally represented in terms of Covers, a particular mathematical structure. With respect to two sets S_1 and S_2, a Cover is formally defined as:

(3) A Cover Cov is a subset of $Cov_1 \times Cov_2$, where $Cov_1 \subseteq \wp(S_1)$ and $Cov_2 \subseteq \wp(S_2)$, s.t.

a. $\forall s_1 \in S_1, \exists cov_1 \in Cov_1$ s.t. $s_1 \in cov_1$, and $\forall s_2 \in S_2, \exists cov_2 \in Cov_2$ s.t. $s_2 \in cov_2$.

b. $\forall cov_1 \in Cov_1, \exists cov_2 \in Cov_2$ s.t. $\langle cov_1, cov_2\rangle \in Cov$.

c. $\forall cov_2 \in Cov_2, \exists cov_1 \in Cov_1$ s.t. $\langle cov_1, cov_2\rangle \in Cov$.

Covers may be denoted by 2-order variables called "Cover variables". We may then define a meta-predicate $Cover$ that, taken a Cover variable C and two unary predicates P_1 and P_2, asserts that the extension of the former is a Cover of the extensions of the latter:

[1] In line with (Landman, 2000), pp.129, and (Beck and Sauerland, 2000), def.(3), that explicitly define Cumulative Readings as statements among atomic individuals only.

(4)

$$Cover(C, P_1, P_2) \Leftrightarrow$$
$$\forall_{X_1X_2}[C(X_1, X_2)\rightarrow$$
$$\forall_{x_1x_2}[((x_1 \subset X_1) \wedge (x_2 \subset X_2))\rightarrow$$
$$(P_1(x_1) \wedge P_2(x_2))]] \wedge$$
$$\forall_{x_1}[\ P_1(x_1) \rightarrow \exists_{X_1X_2}[\ (x_1 \subset X_1) \wedge C(X_1, X_2)\ ]\ ] \wedge$$
$$\forall_{x_2}[\ P_2(x_2) \rightarrow \exists_{X_1X_2}[\ (x_2 \subset X_2) \wedge C(X_1, X_2)\ ]\ ]$$

Thus, it is possible to decouple the quantifications from the predications. This is done by introducing two relational variables whose extensions include the *atomic* individuals involved. Another relational variable that covers them describes how the actions are actually done. For instance, in (2), in order to evaluate (1.d) as true, we may introduce three variables P_1, P_2, and C such that:

$$\|P_1\|^M = \{c_1, c_2, c_3\}$$
$$\|P_2\|^M = \{p_1, p_2, p_3, p_4, p_5\}$$
$$\|C\|^M = \{\ \langle c_1 \oplus c_2, p_1 \oplus p_2\rangle, \langle c_2 \oplus c_3, p_3 \oplus p_4\rangle, \langle c_3, p_5\rangle\ \}$$

Among the Cover approaches mentioned above, an interesting one is (Schwarzschild, 1996). Schwarzschild discusses numerous NL sentences where the identification of Covers appears to be pragmatically determined, rather than existentially quantified. In other words, in the formulae the value of the Cover variables ought to be provided by an assignment g. One of the examples mostly discussed in (Schwarzschild, 1996) is:

(5) a. The cows and the pigs were separated.

b. The cows and the pigs were separated...
...*according to color*.

The preferred reading of (5.a) is the one where the cows were separated from the pigs. However, that is actually an implicature that may be rewritten as in (5.b), where the separation is not done by race. Schwarzschild claims that the NP in (5.a) must be denoted by a unary predicate whose extension is the set of *individual* cows and pigs, while the precise separation is described by a contextually-dependent Cover variable. Similarly, in (1.c) the Cumulative interpretation is preferred as in real contexts invitations are usually thought as actions among pairs of individual persons. But it may be the case that two or more boys *collectively* invited two or more girls. Analogously, in (1.a) the fact that each student saw each drug-dealer seems to be favoured by the low value of the numerals. If the sentence were *Almost all of my students have seen several drug-dealers*, the preferred reading appears to be Cumulative.

The next section illustrates a final component needed to build whole formulae for representing Cover readings. This is the requirement of Maximal participance of the witness sets, e.g. the Maximal participance of P_1 and P_2's extension in the formula denoting (1.d). Two possible approaches for maximizing the involved witness sets have been proposed in the literature: *Local* and *Global* Maximalization. The present paper argues in favour of Local Maximalization, provided that also Witness sets, besides Cover variables, are pragmatically interpreted.

2. The Maximality requirement

The previous section showed that, for representing IS readings, it is necessary to reify the witness sets into relational variables as P_1 and P_2. Separately, the elements of these sets are combined as described by the Cover variables, in order to assert the predicates on the correct pairs of (possibly plural) individuals. As argued by (Sher, 1990), (Sher, 1997), (Steedman, 2007) and (Robaldo, 2010) the relational variables must, however, be *Maximized* in order to achieve the proper truth values with any quantifier, regardless to its monotonicity[2] (cf. also (Dalrymple et al., 1998) and (Winter, 2001)). To see why, let us consider (6.a-c), taken from (Robaldo, 2010), that involve a single quantifier.

(6) a. At least two men walk.

b. At most two men walk.

c. Exactly two men walk.

In terms of reified relational variables, it seems that the meaning of (6.a-c) may represented via (7.a-c), where $^{\geq}2$, $^{\leq}2$, and $^{=}2$ are, respectively, an M↑, an M↓, and a non-M Generalized Quantifier.

(7) a. $\exists P[\ ^{\geq}2_x(\text{man'}(x), P(x)) \wedge \forall_x[P(x)\rightarrow \text{walk'}(x)]\ ]$

b. $\exists P[\ ^{\leq}2_x(\text{man'}(x), P(x)) \wedge \forall_x[P(x)\rightarrow \text{walk'}(x)]\ ]$

c. $\exists P[\ ^{=}2_x(\text{man'}(x), P(x)) \wedge \forall_x[P(x)\rightarrow \text{walk'}(x)]\ ]$

Only (7.a) correctly yields the truth values of the corresponding sentence. To see why, consider a model in which three men walk. In such a model, (7.a) is true, while (7.b-c) are false. Conversely, all formulae in (7) evaluate to true, as all of them allow to choose P such that $\|P\|^M$ is a set of two walking men. Therefore, we cannot allow a free choice of P. Instead, P must denote the Maximal set of individuals satisfying the predicates, i.e. the Maximal set of walking men, in (7). This is achieved by changing (7.b-c) to (8.a-b) respectively.

(8) a. $\exists P[\ ^{\leq}2_x(\text{man'}(x), P(x)) \wedge \forall_x[P(x)\rightarrow \text{walk'}(x)] \wedge$
$\forall'_P[(\forall_x[P(x)\rightarrow P'(x)] \wedge \forall_x[P'(x)\rightarrow \text{walk'}(x)])\rightarrow$
$\forall_x[P'(x)\rightarrow P(x)]\ ]\ ]$

b. $\exists P[\ ^{=}2_x(\text{man'}(x), P(x)) \wedge \forall_x[P(x)\rightarrow \text{walk'}(x)] \wedge$
$\forall'_P[(\forall_x[P(x)\rightarrow P'(x)] \wedge \forall_x[P'(x)\rightarrow \text{walk'}(x)])\rightarrow$
$\forall_x[P'(x)\rightarrow P(x)]\ ]\ ]$

The clauses $\forall'_P[\ \ldots\ ]$ in the second rows are Maximality Conditions asserting the non-existence of a superset P' of P that also satisfies the predication. There is a single choice for P in (8.a-b): it must denote the set of *all* walking men. Note that, for the sake of uniformity, the Maximality condition may be added in (7.a) as well: in case of M↑ quantifiers, it does not affect the truth values.

[2]See (Barwise and Cooper, 1981) for a survey on possible monotonicities of Generalized Quantifiers.

2.1. Local Maximalization

Let us term the kind of Maximalization done in (8) as *Local* Maximalization. The Maximality conditions in (8) require the non-existence of a set $\|P'\|^M$ of walkers *that includes* $\|P\|^M$. (Robaldo, 2010) proposed a logical framework for representing Branching Quantifier based on Local Maximalization. For instance, in (Robaldo, 2010), the *two* witness sets of students and drug-dealers in (1.a) are respectively reified into two variables P_1 and P_2, and the Maximality condition requires the non-existence of a *Cartesian Product* $\|P_1'\|^M \times \|P_2'\|^M$, that also satisfies the main predication and *that includes* $\|P_1\|^M \times \|P_2\|^M$:

(9)

$$\exists P_1 P_2 [\, {}^{=}2_x(\text{stud'}(x), P_1(x)) \wedge {}^{=}3_x(\text{drugD'}(y), P_2(y)) \wedge \\ \forall_{xy}[(P_1(x) \wedge P_2(y)) \rightarrow \text{saw'}(x, y)] \wedge \\ \forall_{P_1' P_2'}[\, (\forall_{xy}[(P_1(x) \wedge P_2(y)) \rightarrow (P_1'(x) \wedge P_2'(y))] \wedge \\ \forall_{xy}[(P_1'(x) \wedge P_2'(y)) \rightarrow \text{saw'}(x, y)]\,) \rightarrow \\ \forall_{xy}[(P_1'(x) \wedge P_2'(y)) \rightarrow (P_1(x) \wedge P_2(y))]\,]\,]$$

As extensively argued in (Robaldo, 2011), in order to extend (Robaldo, 2010) to Cover readings we cannot simply require the inclusion of $\|P_1\|^M \times \|P_2\|^M$ into the main predicate's extension. Rather, we require the inclusion therein of a pragmatically-determined Cover $\|C\|^{M,g}$ of $\|P_1\|^M$ and $\|P_2\|^M$. Furthermore, the (local) Maximality condition must require the non-existence of a superset of either $\|P_1\|^M$ or $\|P_2\|^M$ whose corresponding Cover is a superset of $\|C\|^{M,g}$ that is also included in the main predicate's extension. Thus, (1.d) is represented as[3]:

(10)

$${}^{=}3_x(\text{child'}(x), P_1(x)) \wedge {}^{=}5_y(\text{pizza'}(y), P_2(y)) \wedge \\ Cover(C, P_1, P_2) \wedge \forall_{xy}[C(x, y) \rightarrow \text{ate'}(x, y)] \wedge \\ \forall_{P_1'}[\, (\forall_x[P_1(x) \rightarrow P_1'(x)] \wedge \\ \exists_{C'}[Cover(C', P_1', P_2) \wedge \forall_{xy}[C(x,y) \rightarrow C'(x,y)] \wedge \\ \forall_{xy}[C'(x,y) \rightarrow \text{ate'}(x, y)]]) \rightarrow \forall_x[P_1'(x) \rightarrow P_1(x)]\,] \wedge \\ \forall_{P_2'}[\, (\forall_y[P_2(y) \rightarrow P_2'(y)] \wedge \\ \exists_{C'}[Cover(C', P_1, P_2') \wedge \forall_{xy}[C(x,y) \rightarrow C'(x,y)] \wedge \\ \forall_{xy}[C'(x,y) \rightarrow \text{ate'}(x, y)]]) \rightarrow \forall_y[P_2'(y) \rightarrow P_2(y)]\,]\,]$$

Note that there are *two* Maximality conditions, i.e. $\forall_{P_1'}[\dots]$ and $\forall_{P_2'}[\dots]$, rather than a single one. Contrary to what is done with Cartesian Products, in Cover readings P_1 and P_2 must be Maximized independently, as it is no longer required that *every* member of the former is related with *every* member of the latter. Note also that P_1 and P_2 are pragmatically determined, as it is done with Cover variables in Schwarzschild's, rather than being existentially quantified as in formula (9). In other words, their value is provided by an assignment function g. This is the main point addressed (below) in this paper.

[3]Without going down into further details, we simply stipulate that quantifiers are Conservative (Barwise and Cooper, 1981): for every quantifier Q_x, we require $\|P_x^B\|^{M,g} \subseteq \|P_x^R\|^{M,g}$.

2.2. Global Maximalization

The other kind of Maximalization of the witness sets, termed here as 'Global Maximalization' has been advocated by (Schein, 1993), and formalized in most formal theories of Cumulativity, e.g. (Landman, 2000), (Hackl, 2000), and (Ben-Avi and Winter, 2003). With respect to IS readings involving two witness sets $\|P_1\|^M$ and $\|P_2\|^M$, Global Maximalization requires the non-existence of other two witness sets that also satisfy the predication but *that do not necessarily include* $\|P_1\|^M$ *and* $\|P_2\|^M$. For instance, the event-based logic defined by (Landman, 2000) represents the Cumulative Reading of (1.c) as:

(11)

$$\exists e \in {}^*\text{INVITE}: \exists x \in {}^*\text{BOY}: |x|=3 \wedge {}^*\text{Ag}(e)=x \wedge \\ \exists y \in {}^*\text{GIRL}: |y|=4 \wedge {}^*\text{Th}(e)=y \wedge \\ |{}^*\textbf{Ag}(\bigcup\{e \in \textbf{INVITE: Ag}(e) \in \textbf{BOY} \wedge \textbf{Th}(e) \in \textbf{GIRL}\})| = \textbf{3} \\ \wedge \\ |{}^*\textbf{Th}(\bigcup\{e \in \textbf{INVITE: Ag}(e) \in \textbf{BOY} \wedge \textbf{Th}(e) \in \textbf{GIRL}\})| = \textbf{4}$$

Formula in (11) asserts the existence of a plural event e whose Agent is a plural individual made up of three boys and whose Theme is a plural individual made up of four girls. The two final conjuncts, in boldface, are Maximality conditions. Taken e_x as the plural sum of all inviting events having a boy as agent and a girl as theme, i.e.

$$e_x = \bigcup\{e \in \text{INVITE: Ag}(e) \in \text{BOY} \wedge \text{Th}(e) \in \text{GIRL}\}$$

the cardinality of its agent ${}^*\text{Ag}(e_x)$ is exactly three while the one of its theme ${}^*\text{Th}(e_x)$ is exactly four.

Therefore, Landman's Maximality conditions in (11) do not refer to the same events and actors quantified in the first row. Rather, they require that the number of the boys who invited a girl *in the whole model* is exactly three and the number of girls who were invited by a boy *in the whole model* is exactly four.

3. An experiment on IS readings

To summarize, in (Robaldo, 2011) witness sets are firstly identified on pragmatic grounds, then they are locally maximized. It is important to understand that, in Robaldo's, Maximality conditions are not thought as "constraints that must be satisfied in order to judge the formula as true in the context". Rather, they must be thought as "asserted knowledge needed to draw the appropriate inferences from the sentences' meaning". Conversely, the *evaluation* of the formula, i.e. the task of deciding whether a sentence is true or false in a certain model, is totally devolved upon the interpretation function g.

What could be "the pragmatic grounds" that may affect the identification of the witness sets? As mentioned above, the use of certain determiners seems to affect the interpretation of the main predicate (Cumulative rather than Collective or each-all), i.e. the value of the Cover variables. Analogously, (Geurts and van der Silk, 2005) provide evidence that M↑ quantifiers are simpler to reason with, and this seems to explain why the identification of their witness sets is usually oriented towards the whole set of individuals in the model, rather than to specific sub-groups.

On the other hand, several cognitive experimental results showed that many other factors besides monotonicity, e.g. expressivity/computability, fuzzyness, the fact that quantifiers are cardinal rather than proportional, etc., may affect the interpretation of IS readings (cf. (Sanford and Paterson, 1994), (Szymanik, 2009), (Bott and Rad, 2009), (Musolino, 2009), (Szymanik and Zajenkowski, 2010), and (Szymanik, 2010)). As it is clear to understand, however, extra-linguistic factors seem the ones that mainly affect the interpretation of the variables. For instance, knowing that certain individuals are friends or are member of a team could induce the identification of sub-groups of individuals. In order to attest these hypotheses on empirical data, we carried out an online questionnaire. The experiment and its results are presented below.

3.1. Instructions

In the questionnaire, we show a set of sentences, each together with a figure. The subjects are asked to tell whether the sentence is true or false in the context depicted by the figure. There are eight target sentences, i.e. sentences for which we collect the results, plus twelve fillers, i.e. sentences whose answers are rather obvious and so they are not registered in the database. Fillers were used to prevent subjects from using some simplified strategy that could only work with specific experimental target items. The eight target sentences are:

(12)

a. Exactly three boys ate exactly three pizzas.

b. Exactly one boy ate exactly one pizza.

c. Fewer than three boys ate exactly one pizza.

d. More than three boys ate most pizzas.

e. Fewer than half of the boys ate exactly three pizzas.

f. Exactly two boys ate exactly three pizzas.

g. More than five boys ate more than four pizzas.

h. Fewer than three boys ate exactly one pizza.

The figures describe boys eating pizzas. Boys and pizzas are represented with stylized drawings, while the eating actions with lines connecting boys to pizzas. When a boy is connected by a line to a pizza, we mean that he ate the pizza. When two or more boys are connected to the same pizza, we mean that they ate it together, by cutting it into slices and sharing the slices. Boys are grouped into teams. Boys belonging to different teams are shown in the figures by means of different colors.

Each target sentence is associated with four figures. One of the figures is randomly chosen and shown to the subject together with the sentence. The four figures associated with a sentence include the same boys, the same pizzas and the same connections. They differ to each other for the presence/absence of two "pragmatic factors". Some distance may be added between sub-groups of boys, and/or the boys may belong to to different teams rather than to a single one. Examples of the figures/scenarios used are shown below.

4. The questionnaire

We exploited the social network Facebook for inviting people to the questionnaire. We registered more than 23,000 participants.

Let us start by analyzing single experiment trials. The role of pragmatics in quantifiers' interpretation is strongly visible in the analysis of sentence (12.f). The sentence was tested with respect to the four scenarios shown in fig.1. As pointed out above, the scenarios differ for the occurrence of two "pragmatic factors": the subgroups of boys could have different colors and more distance may be added between the two pairs of witness sets. Each of the four scenarios corresponds to one of the available combinations: (A) does not include any pragmatic factor, (D) includes both, while (B) and (C) include only one of them. Obviously, our predictions were that the presence of pragmatic factors would induce the identification of the sub-structures, i.e., they would favor the local interpretation rather than the global one.

As said above, our predictions are met with respect to sentence (12.f) in the scenarios of fig.1 (see Table 1). Interestingly, also in scenario (A) a slight majority of subjects chose the local interpretation.

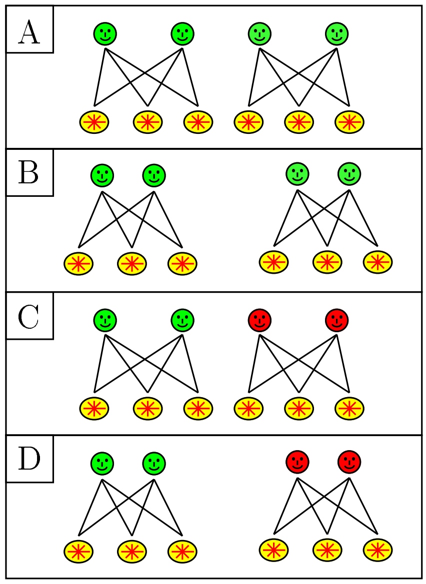

Figure 1: Four scenarios for sentence (12.f).

Table 1: Evaluation of (12.f) in scenarios fig.1.A-D

Scenario	Yes	No	Don't know	Yes%
A	2830	2143	292	56.91%
B	3316	1653	264	66.73%
C	3352	1569	260	68.12%
D	3525	1406	291	71.49%

In fig.2, we show the four scenarios associated with sentence (12.d). Those scenarios have been used to evaluate the sentence 'More than three boys ate most pizzas', that includes two M$\uparrow$ quantifiers. The results for the scenarios in fig.2 are shown in Table 2.

Also Table 2 appears to confirm our predictions. The sentence is logically true in all scenarios shown in fig.2, in line

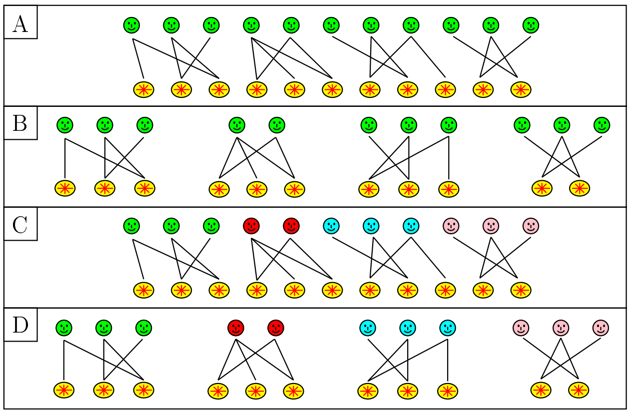

Figure 2: Four scenarios for sentence (12.d).

Table 2: Evaluation of (12.d) in scenarios fig.2.A-D

Scenario	Yes	No	Don't know	Yes%
A	1746	2247	1278	43.72%
B	1800	2101	1326	46.14%
C	1590	2284	1313	41.04%
D	1825	2115	1276	46.32%

with Schein's theory. Nevertheless, most subjects answered it is false. Note also the high number of 'Don't know' answers; in our view, many subjects simply found this example'confusing', due to the high number of boys and pizzas occurring in the figures and the two pragmatic factors we inserted therein.

Note that sentence (12.g) is very similar to (12.d) as it also includes two $M\uparrow$ monotone quantifiers. The results of the (12.g)'s evaluation are very similar to the ones of (12.d).

In fig.3 we show the four scenarios where the sentence (12.b) is evaluated. Note that we inserted a different pragmatic factor in place of the greater distance between sub-structures. The sub-structure including one boy and one pizza only is crossed with respect to the other (bigger and more complex) one. The goal of the crossing is to *avoid* the identification of the witness sets making true the sentence. Nevertheless, we observe that in most cases subjects do manage to identify these witness sets. The results[4] of fig.3 are shown in Table 3. In our view, these results may be explained by observing that the quantifier "Exactly one" has a very strong pragmatic preference towards the identification of sub-structures. Whenever a subject reads "Exactly one", s/he most likely look for a single individual isolated from the others. In other two tests including the quantifier "Exactly one", i.e. (12.c) and (12.h), the result are very similar.

In fig.4, we show the scenarios where sentence (12.e) has been evaluated. The crucial feature of this sentence is that it involves both a non-M quantifier ('Exactly three') and a $M\downarrow$ one ('Fewer than half of the boys'). In other words, it represents a mixed case. The results seem to confirm our hypotheses on the sentence's preferred meaning. Most subjects do appear to identify the sub-structure of boys and pizzas making true the two quantifiers.

[4]Surprisingly, the percentage of 'yes' in (A) is superior to the one in (D). The visual effect given by the vertical line connecting the last boy on the right with the pizza below him appears to be a pragmatic factor even stronger than colors.

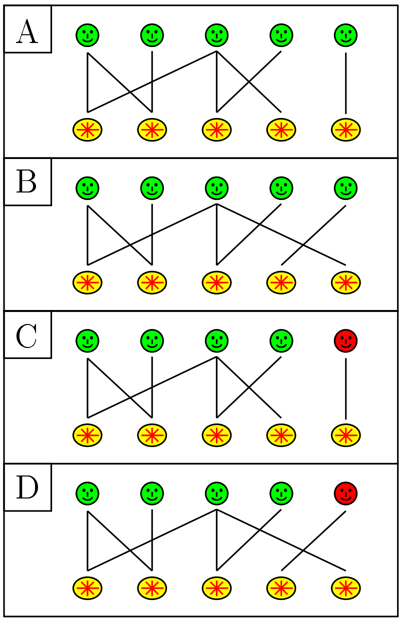

Figure 3: Four scenarios for sentence (12.b).

Table 3: Evaluation of (12.b) in scenarios fig.3.A-D

Scenario	Yes	No	Don't know	Yes%
A	3303	1395	130	70.30%
B	2765	1838	174	60.06%
C	3220	1367	132	70.19%
D	2950	1655	171	64.06%

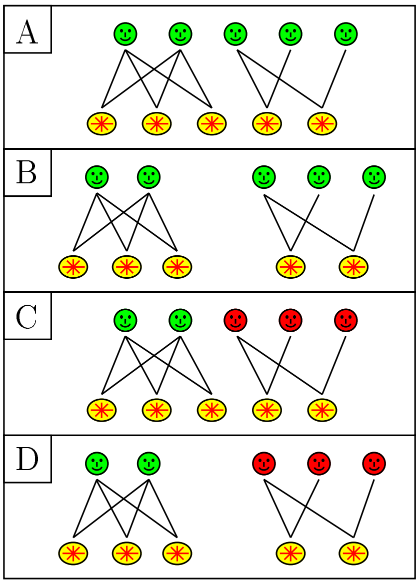

Figure 4: Four scenarios for sentence (12.e).

Finally, we show in fig.5 the single tuple of scenarios for which our hypotheses were *not* met. They are the four scenarios associated with sentence (12.a). The results are shown in Table 5. The percentages of 'yes' are very low in

Table 4: Evaluation of (12.e) in scenarios fig.4.A-D

Scenario	Yes	No	Don't know	Yes%
A	3021	1252	461	70.70%
B	3119	1247	395	71.44%
C	3118	1241	467	71.53%
D	3221	1166	392	73.42%

each of the four scenarios.

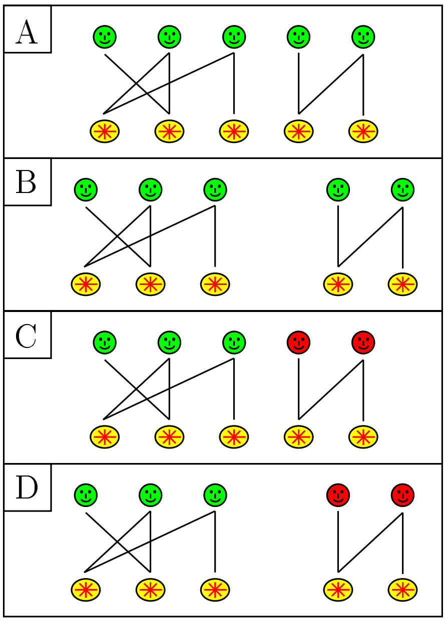

Figure 5: Four scenarios for sentence (12.a).

Table 5: Evaluation of (12.e) in scenarios fig.5.A-D

Scenario	Yes	No	Don't know	Yes%
A	1292	3189	257	28.83%
B	1282	3213	262	28.52%
C	1567	3015	239	34.19%
D	2044	2526	214	44.72%

In our view, these results are due to the fact that sentence (12.a) was the first sentence shown to the subjects. Note that (12.a) (and its associated scenarios) is very similar to sentence (12.f) (and its associated scenarios). But the results are quite different. Therefore, perhaps subjects had an initial inclination towards Global interpretation, but after evaluating some sentences for which the Local one is preferred, they tended to interpret also (12.f) locally.
The order along which sentences was evaluated can be obviously considered as a further pragmatic factors affecting the interpretation.

4.1. Statistical analysis of the results

Below we only present a preliminary statistical analysis indicating that overall the interpretation depends on the pragmatic factors. We focus on two independent variables: Color' with values 'Non-colored' and 'Colored', and 'Distance' with values 'No-distance' and 'Distance' (cf. figures). Let us describe the influence of those manipulations on the selection of local or global reading by our subjects.

Non-colored/Colored Under No-colored condition 37% of all responses were global, 54% local, and 9% undecided. Under Colored condition: 34% global, 57% local, 9% undecided. The value of the color condition and the reading preferred by the subject are dependent (χ^2=231; df=2; p<0.001). Therefore, in line with our predictions, sentences associated with pictures marking possible subgroups with different colors were more often interpreted locally.

Non-crossed/Crossed Under No-crossed condition 29% of all responses were global, 67% local, and 4% undecided. Under Crossed condition: 35% global, 60% local, 5% undecided. The value of the crossed condition and the reading preferred by the subject are dependent (χ^2=227; df=2; p<0.001). Therefore, in line with our predictions, sentences associated with pictures suggesting the whole group as a witness set were more often interpreted globally.

No-distance/Distance Under No-distance condition 34% of all responses were global, 56% local, and 9% undecided. Under Distance condition: 36% global, 53% local, 10% undecided. The value of the distance condition and the reading preferred by the subject are moderately dependent (χ^2=7; df=2; p<0.05). Therefore it seems that there is a statistical tendency towards the interpretation that added distance could trigger a preference for the local interpretation.

The main conclusion one can draw from our results is that the considered sentence do not have the absolute truth values. Their interpretation appears to be dependent on the possible pragmatic factors.

5. Conclusions

In this paper we presented an empirical study on Independent Set readings. The aim of the study was the one of comparing the two kinds of Maximalization proposed in the literature for handling the proper truth values of IS readings, termed here as 'Local' and 'Global' Maximalization respectively. The former requires the non-existence of any tuple of supersets of the witness sets that also satisfies the predication. The latter requires the witness sets to be the only tuple of sets that satisfies the predication.
The results of our experiment show that none of them suffices to properly handle the truth values of IS readings. The reason is that the identification of the witness sets appears to be highly subjective. Sometimes, subjects are able to focus on sub-structure of witness sets. Sometimes they are not, i.e. they consider all occurring individuals as a whole. Moreover, certain pragmatic factors, e.g. the knowledge that boys are divided into teams, a greater distance between sub-structures, the use of certain determiners, the oddity of certain sentences, etc., can affect the identification of the sub-structures.
Therefore, a logical framework designed to represent the proper truth conditions of these sentences should put at disposal suitable formal items where the pragmatic preferences may be taken into account and implemented. This is exactly what is done in (Robaldo, 2011), where pragmatics is formally kept separated from semantics. In Robaldo's,

an assignment function g identifies the witness sets the sentence refers to, then (local) Maximality Conditions are asserted on them.

6. Acknowledgements

Livio Robaldo was supported by the ATLAS project, that is co-funded by Regione Piemonte within the "Converging Technologies - CIPE 2007" framework (Research Sector: Cognitive Science and ICT). Jakub Szymanik was supported by a Vici grant NWO-277-80-001.

7. References

J. Barwise and R. Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2):159–219.

S. Beck and U. Sauerland. 2000. Cumulation is needed: A reply to winter (2000). *Natural Language Semantics*, 8(4):349–371.

F. Beghelli, D. Ben-Shalom, and A. Szabolski. 1997. Variation, distributivity, and the illusion of branching. In *Ways of Scope Taking*, pages 29–69. Kluwer Academic Publishers, Dordrecht.

G. Ben-Avi and Y. Winter. 2003. Monotonicity and collective quantification. *Journal of Logic, Language and Information*, 12:127–151.

O. Bott and J. Rad. 2009. How to provide exactly one interpretation for every sentence, or what eye movements reveal about quantifier scope. In *The fruits of empirical linguistics*. Walther de Gruyter, Berlin.

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

B. Geurts and F. van der Silk. 2005. Monotonicity and processing load. *The Journal of Semantics*, 22(17).

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

M. Hackl. 2000. *Comparative quantifiers.* Ph.D. thesis, Massachusetts Institute of Technology.

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

A. Kratzer. 2007. On the plurality of verbs. In *Event Structures in Linguistic Form and Interpretation*. Mouton de Gruyter, Berlin.

F. Landman. 2000. *Events and Plurality: The Jerusalem Lectures*. Kluwer Academic Publishers.

G. Link. 1983. The logical analysis of plurals and mass terms. In *Meaning, Use, and Interpretation in Language*, pages 302–323. de Gruyter, Berlin.

J. Musolino. 2009. The logical syntax of number words: Theory, acquisition and processing. *Cognition*, 111.

L. Robaldo. 2010. Independent set readings and generalized quantifiers. *The Journal of Philosophical Logic*, 39(1):23–58.

L. Robaldo. 2011. Distributivity, collectivity, and cumulativity in terms of (in)dependence and maximality. *The Journal of Logic, Language, and Information.*, 20(2):233–271.

Moxey L. M. Sanford, A. J. and K. Paterson. 1994. Psychological studies of quantifiers. *The Journal of Semantics*, 11(3):153170.

R. Scha. 1981. Distributive, collective and cumulative quantification. In *Formal Methods in the Study of Language, Part 2*. Mathematisch Centrum, Amsterdam.

B. Schein. 1993. *Plurals and Events*. MIT Press, Cambridge, MA, USA.

R. Schwarzschild. 1996. *Pluralities*. Kluwer, Dordrecht.

G. Sher. 1990. Ways of branching quantifiers. *Linguistics and Philosophy*, 13:393–422.

G. Sher. 1997. Partially-ordered (branching) generalized quantifiers: a general definition. *The Journal of Philosophical Logic*, 26:1–43.

M. Steedman. 2007. *The Grammar of Scope*. forthcoming. See 'Surface-Compositional Scope-Alternation Without Existential Quantifiers'. Draft 5.2, Sept 2007. Retrieved September 25, 2007 from ftp://ftp.cogsci.ed.ac.uk/pub/steedman/quantifiers/journal6.pdf.

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

Jakub Szymanik. 2009. *Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language*. Ph.D. thesis, University of Amsterdam, Amsterdam.

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

J. van der Does and H. Verkuyl. 1996. The semantics of plural noun phrases. In *Quantifiers, Logic and Language*. CSLI.

Y. Winter. 2001. *Flexibility principles in Boolean Semantics: coordination, plurality, and scope in natural language.* MIT Press, Cambridge, Massachussets.