Compositional semantics of *same*, *different*, *total*

Oleg Kiselyov
Tohoku University, Japan

Abstract

We propose a strictly compositional and uniform treatment for the internal readings of the symmetric predicates such as *same* and *different* and for the summative phrases such as *a total of*. The analyses handle multiple occurrences of symmetric and summative predicates in the same sentence, alongside of multiple quantificational elements. We treat symmetric predicates and summative phrases as wide-scope generalized existential quantifiers for choice functions. Like Barker's (2007) we treat symmetric and summative expressions as generalized quantifiers and relate them to universal quantification. However, our quantifiers scope wide rather than parasitically and avoid Barker's non-standard interpretation of universal quantification.

We introduce the analyses in terms of the familiar Quantifier Raising framework, which we make more precise and treat as a somewhat informal version of a type-logical grammar.

Keywords: Quantification; QR; TLG; Dependent quantification; Quantifier ambiguity; Distribution

1 Introduction

The internal readings of symmetric predicates in (1-4) challenge the compositional semantics (Barker, 2007; Kubota and Levine, 2013, 2014):

(1) John and Bill checked the same book.
(2a) John and Bill checked the same book at the same store.
(2b) John and Bill checked the same book at different stores.
(2c) John and Bill checked the same book at the same store
on the same day for the same topic.
(3) John read and Bill reviewed {the same book/different books}.
(4) I gave the same book to John on Wednesday and to Bill on Friday.

All these sentences have (deictic) readings, referring to some book, store or day identified from the context. For example, when (1) appears in a passage after the sentence "Mary bought the latest Harry Potter novel" it may mean that John and Bill also checked that Harry Potter novel. However, (1) also has the meaning by itself: "the same book" refers to some unspecified book that John and Bill separately checked out. The sentence itself then provides the context of interpreting "the same". This is the internal reading of the sentence, which is the main subject of this paper.

Even in simple cases (1-2) the conjoined NPs seem to need the denotations that expose the conjuncts (so they can be accessed by distributivity operators). The meaning of a phrase then is determined not just from the meaning but also from a part of the structure of the components (see (Kubota and Levine, 2014, §2.1) for detailed discussion.) Problems accumulate when symmetric predicates are used with the non-canonical coordination (3-4).

The compositional semantic analysis of such internal readings has been a challenge, reviewed in (Barker, 2007). There has been even a proof that there can be no compositional analysis with generalized quantifiers (Keenan, 1992). Barker (2007) was first to give the strictly compositional account of the *same*, using the novel idea of "parasitic scope" and a non-standard quantifier. Alas, this analysis do not appear to be adequate (Kubota and Levine, 2014) in cases of multiple symmetric predicates.

Summative predicates such as 'total' in (5a) are just as challenging:

(5a) John gambled and Bill lost the total of 10,000.
(5b) The two men lost the total of 10,000.
(5c) John gave, and Bill lent, a total of 10,000 to {the same students/different students}

In (5a), 10,000 is the sum of John's gambling amount and Bill's losses. To truly complicate the matters, summative and symmetric predicates can also appear together, as in (5c).

We start in §3 with a simpler intuitive analysis, treating *same* as a property of belonging to an equivalence class. Hence *same* behaves like a wide-scope existential quantifier for the reference entity of the class. The internal and deictic *same* are analyzed uniformly. The simple analysis does not work for *different* and it does not account for the empirical fact that the internal reading is only available when there is some element in a sentence that entails multiple events or situations.

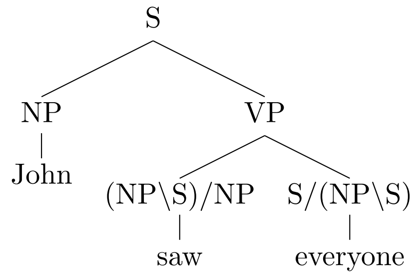
The generalized analysis in §5 overcomes both problems. It uniformly treats *same*, *different* and *total* – as wide-scope quantifiers, not over entities but over Skolem functions whose result is a property. The analysis is based on 'dependent quantification' explained in §4, which may be seen as an alternative to parasitic scope.

We introduce the analyses in terms of quantifier raising (QR) at LF in the style of Heim and Kratzer (1997) – solely because of the wide familiarity of that style and for the sake of comparison with (Barker, 2007). The analysis does not require positing LF. In fact, we treat QR as a somewhat informal version of a type-logical grammar.

2 QR, a less ad hoc version

For reference we briefly describe the QR formalism used in this paper. It is essentially the same as in Heim and Kratzer (1997), but with one less ad hoc construction and with the explicit marking of trace variables (hypotheses) in node types, bringing QR closer to type-logical grammars.

We introduce the formalism on the familiar example:

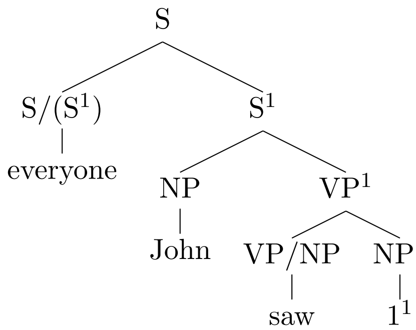


We take S, NP and N as basic categories and write VP as the abbreviation for $NP\backslash S$, *Adj* for N/N, *PP* for $N\backslash N$ or $VP\backslash VP$, *Det* for NP/N and *Prep* for PP/NP.

If we give everyone the lifted type $S/(NP\backslash S)$, type clash occurs: everyone can neither apply nor be applied to the verb saw. QR resolves this clash: the quantifier is raised and adjoined to its scope target, leaving at its former place a variable (to be called trace) of the simple NP type. We use numerals for such variables. Here is the result:

$$
\begin{array}{ccc}
X & \mapsto & \lceil X \rceil \\
\hline
NP & \mapsto & e \\
N & \mapsto & (et) \\
S & \mapsto & t \\
A\backslash B & \mapsto & (\lceil A \rceil \lceil B \rceil) \\
B/A & \mapsto & (\lceil A \rceil \lceil B \rceil) \\
X^{n:A,\Gamma} & \mapsto & (\lceil A \rceil \lceil X^{\Gamma} \rceil)
\end{array}
$$

Figure 1: Mapping Node types to semantic types. The type of hypothetical variables, shown as A, is most always NP. We will omit this type for brevity.



Two things are different from the Heim and Kratzer's (1997) presentation. First, we index all nodes with the variables contained therein. After raising, the object of *saw* is the variable 1 of the category NP^1. The VP node likewise gets the 1 superscript, because its meaning depends on that *assumed* NP. In other words, the traces of raised quantifiers are treated as hypotheses; each node of the tree is indexed with its hypotheses.

The raised quantifier changes type: it is no longer $S/(NP\backslash S)$ but S/S^1. In the new type, NP is dropped from the middle and its neighbor S gets the superscript 1, which is the name of the NP variable left at the quantifier's original place. The raised quantifier hence is the binder of that variable, the eliminator of the corresponding hypothesis. Our presentation of QR is less ad hoc, keeping the explicit track of variables and requiring no extra adjoining step with the non-standard interpretation. The former brings the QR analysis closer to the logically-motivated type-logical grammars (TLG).

The denotation of LF also becomes more precise on our analysis: the meaning of a branch X^Γ (where Γ is the list of variables within that branch, along with their types) is, informally, a function from the meaning of the hypotheses to the meaning of X. More precisely, the denotation is determined according to Fig. 1. For example, the denotation of the VP^1 node is $(e(et))$. The raised quantifier of the category S/S^1 gets the semantic type $(et)t$ with the obvious denotation $\lambda C\!:\!(et).\, \forall x\!:\!e.\, C\ x$.

The quantifier-raising procedure is thus a sequence of the following steps (compare with (Heim and Kratzer, 1997) and its presentation in (Barker, 2007, §5.4))

1. Replace the scope-taking expression (generalized quantifier) of the type $S/(X\backslash S)$ with the variable n of the type X;

2. Add the new variable as a superscript to all parent nodes;

3. Change the type of the scope-taking expression to S/S^n and adjoin to its scope target.

The type X is usually NP; in §3 we consider quantifiers over other types, for example, Adj. We will argue that *same* is such quantifier.

We stress that in our presentation of QR, in contrast to that in (Heim and Kratzer, 1997; Barker, 2007), there is no longer an ad hoc step of the adjoining the second occurrence of the trace variable. That second occurrence was crucial for Barker's analysis: that was the scope target for the parasitic scope. Since in our presentation there is no such step, our analysis hence is distinct from parasitic scope.

3 Simple compositional analysis of *same*

As the preface to the full analysis of symmetric and summative predicates we describe in this section a simplified version, which applies only to *same* and the similar *identical*, *nearly identical*, *similar*, etc.

Our analysis takes literally the remark in (Barker, 2007, §3) about the meaning of

(7) Everyone read the same book.

"In addition to the deictic reading (on which we have a specific book in mind, say, Emma), (7) has an internal reading on which it is true if there is any book such that everyone read that book." It appears therefore that the truth conditions of (7) are the same as

Everyone read some book.

on the reading where the existential takes the wide scope. The observation is in agreement with Barker's (2007) conclusion that "it seems inescapable that on the internal reading, some element in the sentence must in effect introduce an existential quantifier over names." We propose that *same* is such an element.

We further observe that the relation of 'being the same' (as well as 'similar', 'identical', 'nearly identical', etc) is an equivalence relation. The adjective 'same' thus signifies the property of an entity of being a member of an equivalence class. A class can be represented by one of its members, to be called reference entity. The first attempt at the analysis of *same* is to index it by the reference element. Thus we have for same$_i$:

Category	Adj
Semantic type	$(et)(et)$
Denotation	$\lambda P\!:\!(et).\, \lambda x\!:\!e.\, P(x) \wedge x = i$

Our example (7) will then read

(7') Everyone read the same$_i$ book.

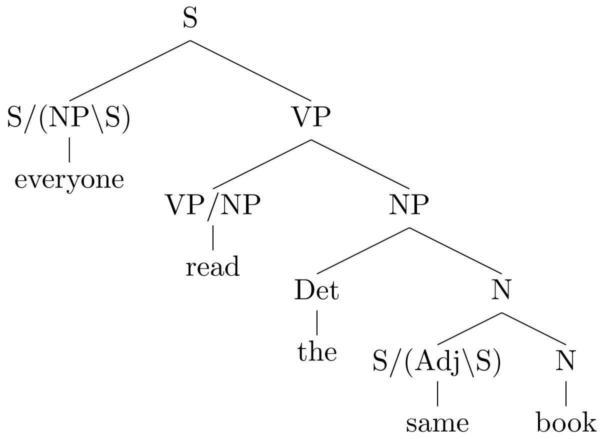
It has the internal reading if i is existentially quantified at the sentence boundary, and the deictic reading if i is left unbound, to be bound somewhere in the context. The symbol $=$ is a shorthand for the "sufficiently the same" relation.

We may as well make *same* to introduce the quantifier for the reference entity. This gives us the final analysis of *same* in this section, cast in terms of the QR framework of §2:

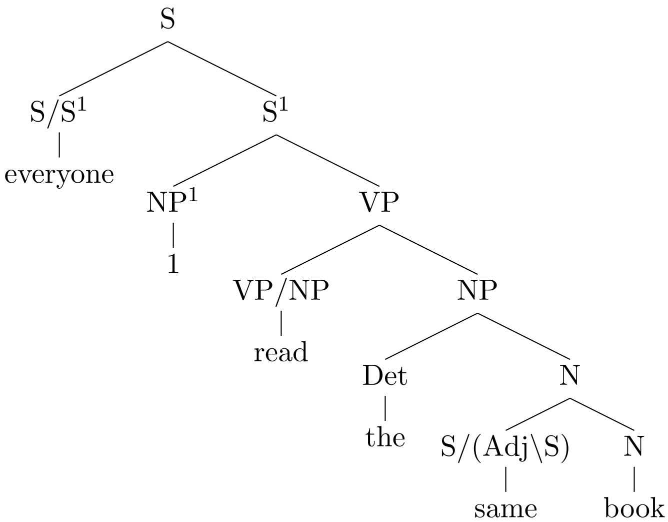
Category	$S/(Adj\backslash S)$
Trace named n	$Adj^{n:Adj}$
Category of raised *same*	$S/S^{n:Adj}$
Its semantic type	$(((et)(et))t)t$
Denotation	$\lambda C\!:\!((et)(et))t.\, \exists i\!:\!e.\, C\ (\lambda P\!:\!(et).\, \lambda x\!:\!e.\, P(x) \wedge x = i)$

We postulate that *same* takes the widest scope.

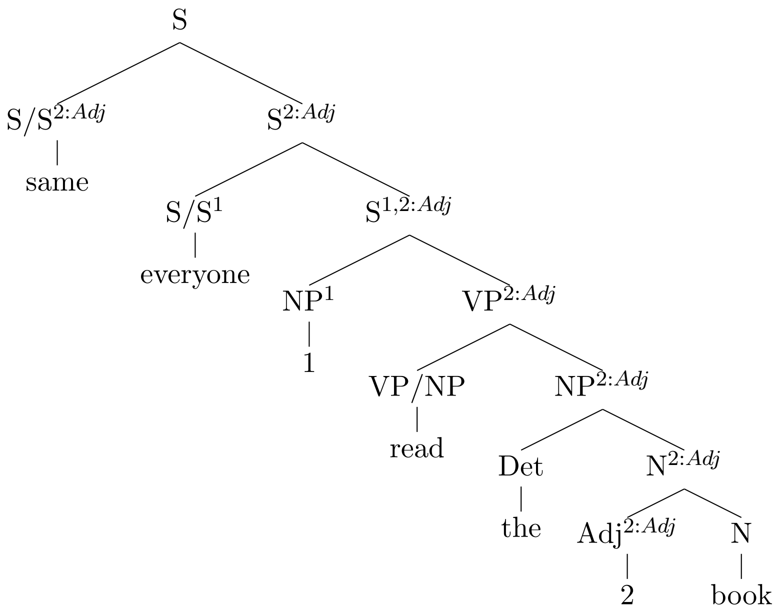
As an illustration, we give the complete analysis of (7). Before raising, (7) is represented by:



We raise *everyone*:



with *same* to follow:



The truth conditions are read off in the standard way:

$\exists i : e. \forall x : e. \iota y.\ read(y, x) \wedge book(y) \wedge y = i$

We have postulated that *same* scopes over any other quantifier that may be in the sentence. The difference between the internal and deictic readings now boils down to quantifying at the sentence vs. paragraph boundary. For example, Barker's (2007) example of the deictic *same* "My friend Heddy's name is highly unusual; nevertheless, two women in this room have the same name." will have the meaning

$$\exists i.\, (i = Heddy) \wedge name(Heddy) \wedge unusual(Heddy) \wedge$$
$$two(\lambda w.\, woman(w) \wedge inThisRoom(w) \wedge hasName(i))$$

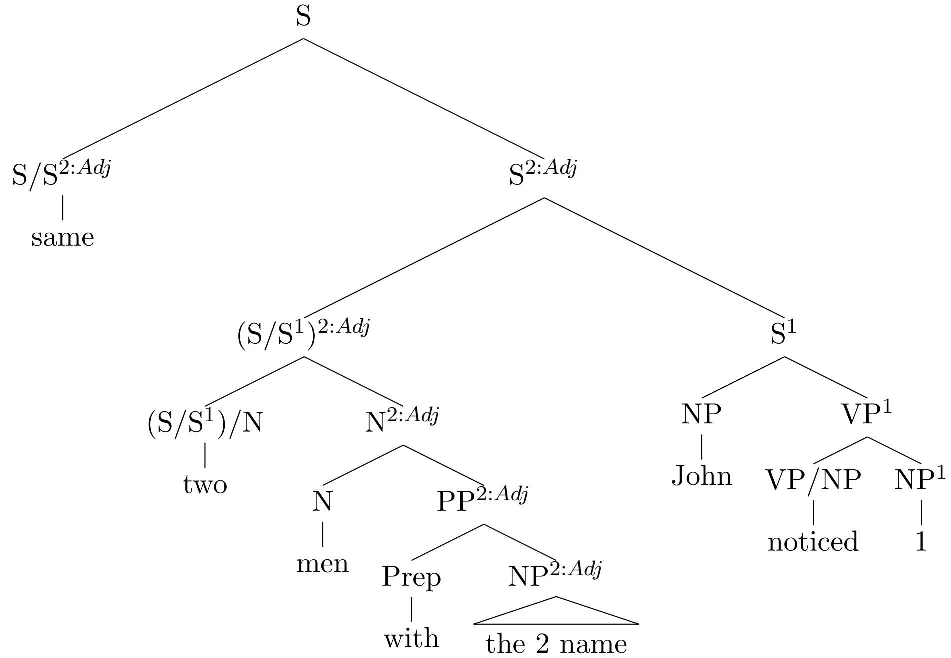
Barker (2007) likewise posits *same* to be a quantifier, which has the uncommon category $N/(Adj\backslash N)$ and takes scope at nominal boundary. It is this nominal boundary that gives rise to the parasitic scope. Our *same* is the common generalized quantifier, taking scope at sentence or clause boundary (or even beyond).

Our analysis easily accommodates several *same* in the same sentence, as in

(2a) John and Bill checked the same book at the same store.

Both *same* are raised, one after the other. As yet another example, here is the final analysis of

John noticed two men with the same name.



The analysis, albeit simple and intuitive, comes short in two respects. First, it does not extend to *different*, which does not denote an equivalence relation. Second, nothing prevents raising *same* in cases with no internal readings, such as (8a):

(8a) John read the same book.

(8b) John read a book.

On our simple analysis, (8a) has the same meaning as (8b), which is odd. We will overcome these drawbacks in the more general analysis in §5.

4 Dependent quantification

Quantifying one of term's variables can be done independently of the others. On the other hand, the other variables may depend on the quantified one. The dependency can be represented by Skolem (choice) functions. This process of eliminating one hypothesis that affects others underlies our general analysis of symmetric predicates and summative phrases. This section describes the process in detail.

The inspiration of our approach is the Kubota and Levine's (2014) observation that *same* and *different* are essentially disambiguators. Indeed, consider (9a) with an indefinite determiner:

(9a) John and Bill checked a book.
(9b) John and Bill checked the same book.
(9c) John and Bill checked a different book.

(9a) is ambiguous: either there is one book that John and Bill checked, or John and Bill each checked a book that may or may not be the same. Then *same* and *different* disambiguate.

We illustrate the mechanism of the disambiguation using first-order logic. Let P be a binary predicate.

(10a) $\exists x. \forall y. P(x, y)$
(10b) $\forall y. \exists x. P(x, y)$
(10c) $\exists f. \forall y. P(f(y), y)$

Formula (10a) asserts that there exists one particular x' such that $P(x', y)$ holds for any choice of y. In contrast, in (10b) the x that makes $P(x, y)$ hold *generally* varies with the choice of y. The choices of x may be all different, all the same, or some different and some the same.

One way of disambiguating (10b) immediately springs to mind: Skolemization, as in (10c). Choosing the Skolem function f to always return the same x' will disambiguate (10b) to its special case (10a). One can also make a Skolem function that ensures all choices are different from each other.

Let us discuss the introduction of Skolem functions more precisely. Let $P(x, y)$ be a formula with two free variables (in other words, derived with two hypotheses named x and y of the types A and B respectively). The formula can be written as

$$p_{xy} = \lambda x\!:\!A.\, \lambda y\!:\!B.\, P(x, y)$$

making the dependence on the hypotheses explicit. We wish to universally quantify the variable y, by applying the quantifier $qU : ((Bt)t)$

$$qU = \lambda F\!:\!(Bt).\, \forall z\!:\!B.\, F\ z$$

Quantifying without regard for x gives $\forall y. P(x, y)$. In more detail:

$$\lambda x\!:\!A.\, qU\ (p_{xy}\ x)$$

In general however the hypothesis x may be correlated with y. Therefore, the general quantification is

$$\lambda f\!:\!(BA).\, qU\ (\lambda y\!:\!B.\, p_{xy}\ (f\ y)\ y)$$

The Skolem (choice) function f expresses the possible correlation between the hypothesis. The simple quantification is obtained as a special case when f is a constant function. We call this general procedure dependent quantification and use it extensively in the analyses of the next section.

It should be clear that dependent quantification only makes sense for universal quantifiers, which asserted a formula for potentially multiple choices of the quantified variable.

5 General analysis of symmetric and summative predicates

This section presents the general and uniform analysis of *same*, *different* and *total*. The analysis is the generalization of the simple version of §3. Here is how we propose to analyze *same* now:

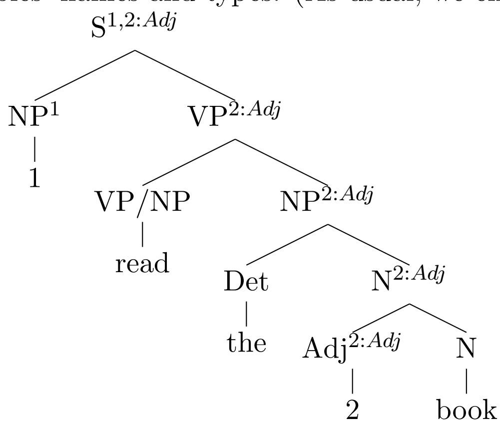
Category	$S/(Adj\backslash S)$
Trace named n	$Adj^{n:Adj}$
Category of raised *same*	$S/S^{n:Adj^{\Gamma}}$
Its semantic type	$((\Gamma((et)(et)))t)t$
Denotation	$\lambda C:(\Gamma((et)(et)))t.\, \exists i:e.\, C\ (\lambda y:\Gamma.\, \lambda P:(et).\, \lambda x:e.\, P(x) \wedge x = i)$

As before, the trace left by the raised *same* is of the type of an adjective $Adj^{n:Adj}$. The raised *same*, however, binds a variable of a higher-type: $n : Adj^{\Gamma}$ rather than $n : Adj$, where Γ is some type (usually NP). The denotation has hardly changed though: the new argument $y : \Gamma$ is ignored. Thus the meaning of *same* is still the property that an entity belongs to an equivalence class, regardless of the choice for other variables that may be in scope.

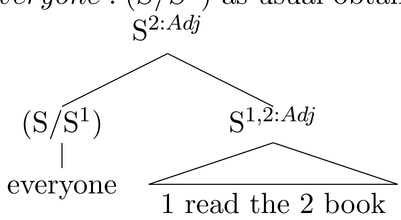
As an illustration, we re-do the analysis of (7) from §3

(7) Everyone read the same book.

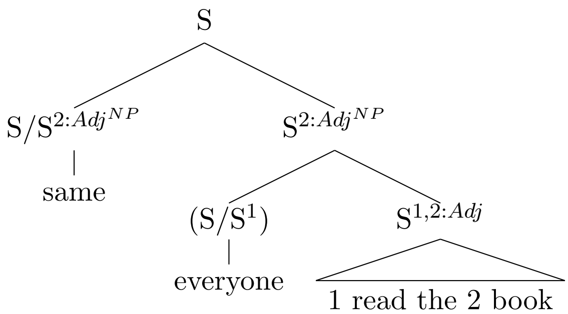
First, *everyone* and *same* are replaced with the corresponding variables and the tree nodes are marked with the variables' names and types: (As usual, we elide NP in marking):



We can adjoin the raised $everyone : (S/S^1)$ as usual obtaining



but then we are stuck. The raised *same* has the higher type $S/S^{n:Adj^{\Gamma}}$ rather than the simple $S/S^{n:Adj}$: it binds choice functions over adjectives rather than just adjectives. Therefore, we have to use the dependent quantification when adjoining *everyone*, which leads to the successful analysis:



The truth conditions are read off as usual (keeping in mind the dependent quantification):

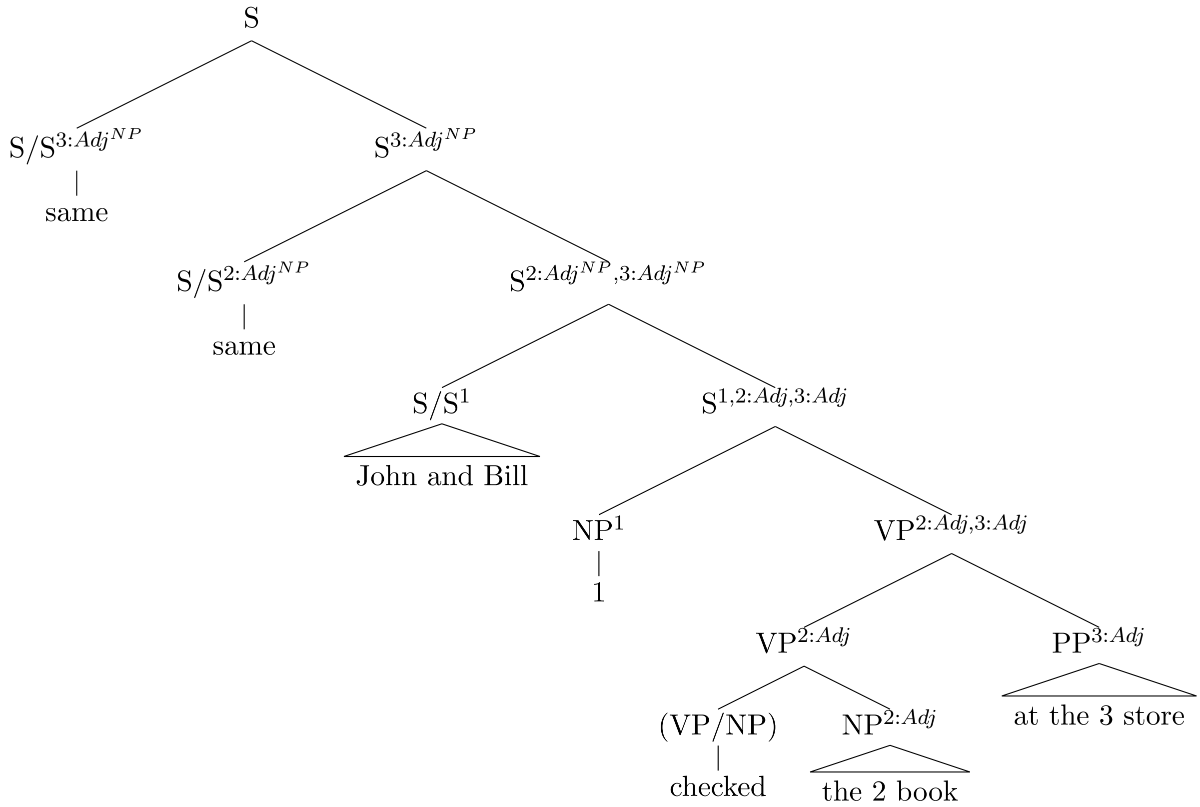
$$same\ (\lambda f_2.\, everyone\ (\lambda y_1.\, (\lambda y_1.\, \lambda p_2.\, \iota z.\, read(z, y_1) \wedge p_2(z) \wedge book(z))\ y_1\ (f\ y_1)))$$
$$\rightsquigarrow\quad \exists i.\, \forall y.\, \iota z.\, read(z, y) \wedge z = i \wedge book(z)$$

If we replace "everyone" with "John" in the sample sentence, the analysis does not go through: there is no longer a universal quantifier that could do the dependent quantification and introduce the choice function variable for the raised *same* to bind. We thus predict that the internal reading is only possible when some element in a sentence can do dependent quantification.

Like Barker (2007), we relate *same* with the universal quantification and use choice functions (of different types though). However, we existentially quantify the choice function wider than the universal and hence we avoid the parasitic scope. We also avoid the unusual postulate of (Barker, 2007) that the universal quantifier quantifies over groups of entities rather than atomic entities.

Our analysis also easily deals with multiple occurrence of *same*, for example

(2a) John and Bill checked the same book at the same store.



Unlike before, the new analysis of *same* extends to *different*:

Category	$S/(Adj\backslash S)$
Trace named n	$Adj^{n:Adj}$
Category of raised *different*	$S/S^{n:Adj^{\Gamma}}$
Its semantic type	$((\Gamma((et)(et)))t)t$
Denotation	$\lambda C : (\Gamma((et)(et)))t.$ $\exists p : (e(et)).\, (\forall u : \Gamma.\, \forall v : \Gamma.\, u \neq v \Rightarrow p(u) \cap p(v) = \emptyset) \wedge$ $C\ (\lambda y : \Gamma.\, \lambda P : (et).\, \lambda x : e.\, P(x) \wedge p(y)(x))$

Thus *different* finds such property p indexed by $y : \Gamma$ that different indices correspond to non-intersecting properties. Since the types and categories are those of *same*, *same* and *different* are freely interchangeable. For example, it is trivial to adjust the above analysis for (2b) and even (2c):

(2b) John and Bill checked the same book at different stores.

(2c) John and Bill checked the same book at the same store on the same day for the same topic.

The final piece of the puzzle are the summative phrases like (5b).

(5b) The two men lost the total of 10,000.

Although *the total of s* is an NP rather than *Adj*, its analysis is similar:

Category	$S/(NP\backslash S)$
Trace named n	$NP^{n:NP}$
Category of raised *the total of s*	$S/S^{n:NP^{\Gamma}}$
Its semantic type	$((\Gamma e)t)t$
Denotation	$\lambda C:(\Gamma e)t.$ $\exists f:(ee).\,(s = \sum_{v:\Gamma} f(v)) \wedge C\ (\lambda y:\Gamma.\, f\ y)$

6 Conclusions and Future work

To summarize, we propose that on the internal reading, *same*, *different*, as well as *the total of s*, are widest-scope generalized quantifiers. Unusually however, whereas *same* or *different* look like an adjective in situ, what they quantify is a choice function over adjectives rather than merely a variable of the *Adj* category. The conversion from an *Adj* to a choice function is due to the dependent quantification. A sentence therefore must have some universal quantificational element, capable of dependent quantification, for *same* and *different* to have internal readings.

Our analysis is strictly compositional and accommodates multiple occurrences of symmetric and summative phrases within the same sentence. It also predicts the ambiguity in

Different waiters served everyone lunch and dinner.

with two quantifiers.

The quantification over choice functions has independent motivations, discussed by Szabolcsi (2000, §3.2.1) (in the context of works by Reinhart and Winter), for example, to explain both readings of

(11) Three relatives of mine inherited a house.

(i) 'there are three relatives of mine who together inherited a house'

(ii) 'there are three relatives of mine who each inherited a house'

especially the first.

The immediate future work is analyzing respective readings of plural and conjoined expressions.

Acknowledgments.

I am indebted to Yusuke Kubota for his encouragement, many insightful conversations and introducing a wealth of the empirical material. I thank anonymous reviewers for the helpful comments and the recommendation to explain the analysis in simple and familiar framework.

This work was partially supported by JSPS KAKENHI Grant Numbers 22300005, 25540001, 15H02681.

References

Barker, Chris. 2007. Parasitic scope. *Linguistics and Philosophy* 30(4):407–444.

Heim, Irene and Angelika Kratzer. 1997. *Semantics in Generative Grammar*. Oxford: Blackwell Publishers.

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

Kubota, Yusuke and Robert Levine. 2013. Empirical foundations for hybrid type-logical categorial grammar. the domain of phenomena.

Kubota, Yusuke and Robert Levine. 2014. The syntax-semantics interface of ‘respective’ predication: A unified analysis in hybrid type-logical categorial grammar. in print.

Szabolcsi, Anna. 2000. The syntax of scope. In *Handbook of Contemporary Syntactic Theory*, 607–634. Blackwell.