

IDENTITY CRITERIA OF COMMON NOUNS AND DOT-TYPES FOR COPREDICATION

STERGIOS CHATZIKYRIAKIDIS¹ & ZHAOHUI LUO²

¹ *Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg*

² *Royal Holloway, University of London*

ABSTRACT

Copredication, especially when combined with quantification, provides interesting examples to support the idea that common nouns have their own identity criteria, as once argued for by Geach and subsequently studied by others. In this paper, revisiting the use of dot-types in modern type theories to model copredication, we show that, when both copredication and quantification are involved, CNs are not just types but should better be interpreted as types associated with their own identity criteria. In other words, formally, CNs are setoids – pairs whose first component is a type that interprets the domain of a CN and whose second component gives the identity criterion for that CN. For copredication with quantification, identity criteria play an essential role in giving a proper treatment of individuation and counting and hence constructing appropriate semantics to facilitate reasoning correctly. With CNs being setoids, the dot-type approach provides an adequate theory for copredication in general and for copredication with quantification in particular. It is further explained that the CNs-as-types approach is still the appropriate characterisation of our approach to interpreting CNs since, in phenomena that do not involve the interaction of copredication with quantification, the identity criteria of related CNs are essentially the same and can be safely ignored.

[1] INTRODUCTION

Copredication (Pustejovsky 1995) is the phenomenon in which more than one predicate (verb or adjective) that require different types of arguments are used in coordination and applied to the “same” CN argument. When combined with quantification, copredication provides interesting examples to support the idea that common nouns have their own identity criteria, an idea first discussed by Geach (1962) and further studied by others including, for example, that by the second author (Luo 2012a) in the semantic framework based on modern type theories. In this paper, revisiting the use of dot-types in modern type theories to model copredication (Luo 2010, 2012b), we show that, when both copredication and quantification are involved, CNs are not just types but should better be in-

terpreted as types associated with their own identity criteria, formally called setoids.¹ For copredication with quantification, identity criteria play an essential role in giving a proper treatment of individuation and counting and hence constructing appropriate semantics to facilitate reasoning correctly. With CNs being setoids, the dot-type approach provides an adequate theory for copredication in general and for copredication with quantification in particular. Consider the following example of copredication:

- (1) John picked up and mastered the book.

In the above sentence, the predicates *pick up* and *master* require physical and informational objects as their respective arguments: formally, their domains are *Phy* and *Info*, the types of physical and informational objects, respectively. However, they manage to be applied in coordination to the argument *the book*. Thus, *book* in this case is used in its physical sense with respect to the predicate *picked up* and in its informational sense with respect to the predicate *mastered*.

The term copredication was coined by Pustejovsky (1995), according to whom lexical items like *book* in (1) are complex and have more than one sense, to be chosen appropriately according to the context. The topic has since been studied by many researchers and different accounts of copredication have been proposed in various semantic frameworks.² For instance, the second author has proposed dot-types in modern type theories for the semantic study of copredication (Luo 2010, 2012b). The idea was that, the phenomenon that a single occurrence of *book* may be used in both physical and informational senses can be captured by stipulating that the type *Book*, that interprets *book*, be a subtype of the dot-type of those of physical and informational objects:³

$$Book \leq Phy \bullet Info.$$

The copredication phenomenon as exhibited in (1) can then be dealt with straightforwardly: the predicates *pick up* and *master* can be applied in coordination to the argument *the book* because, with the above stipulation in the setting of dot-types, they are both of type $Book \rightarrow Prop$ (see later for a more detailed description of dot-types and their use for copredication.)

-
- [1] The current authors have considered the issue in (Chatzikiyriakidis & Luo 2015) where, however, the necessity of considering identity criteria was not recognised – the current paper studies identity criteria in this context, based on the abstract presented at the Workshop on Approaches to Coercion and Polysemy in Oslo (Chatzikiyriakidis & Luo 2017b).
- [2] Besides Pustejovsky’s initial proposal (Pustejovsky 1995), these include, to mention a few among many: Asher (2012) using his Type Composition Logic, Luo (2010, 2012b) using modern type theories, Bassac et al. (2010) using second-order λ -calculus, and Gotham (2014, 2017) using a mereological account.
- [3] In formal semantics based on modern type theories (MTT-semantics) (Luo 2012b; Chatzikiyriakidis & Luo 2018), CNs are interpreted as types (not predicates), an idea originally coming from Ranta’s work (Ranta 1994).

When giving an account of copredication, one needs to take care of a more advanced issue – the involvement of quantification. (2) is an example, taken from (Asher 2012), where copredication and quantification interact:

(2) John picked up and mastered three books.

Because of the presence of the quantifier *three* in (2), individuation and counting come into play in the semantic analysis of this sentence with copredication. In other words, an adequate account of copredication must take into consideration these parameters – but, what is needed for an adequate account? To explain, let's start by considering the following simpler sentences that do not involve copredication:

(3) John picked up three books.

(4) John mastered three books.

The first example, i.e. (3), is true in case John picked up three distinct physical objects. Thus, it is compatible with a situation where John picked up three copies of a book that are informationally identical as long as three distinct physical copies are picked up. Similarly, example (4) is true in case three distinct informational objects are mastered but it does not impose any restrictions on whether these three informational objects should be different physical objects or not.⁴ A more explicit way of explaining this is that, intuitively, the following entailments should be the case:

(5) John picked up three books \Rightarrow John picked up three physical objects

(6) John mastered three books \Rightarrow John mastered three informational objects

At first appearance, the above examples may be puzzling. The books John picked up and the books he mastered must have different individuation criteria – being physically identical is different from being informationally identical. Are they the same books? Actually, they are not. The collection of books with *being physically the same* as its identity criterion (call it $=_p$) is different from the collection of books with *being informationally the same* as its identity criterion (call it $=_i$). In other words, identity criteria play a crucial role in fixing collections of objects represented by CNs, although they are not made explicit in NL sentences. Books

[4] We adopt the view that two physically identical books can be informationally different, as assumed by (Asher 2012) and (Gotham 2014). However, the authors actually incline to believe that, in reality, if two books are physically identical, they cannot be informationally different. For example, Asher considers the example of several books in one volume, thinking that the same volume is involved. But we do not think that *volume* and the books in concern are anything comparable in such an example. But, as we said, we still adopt Asher/Gotham's view since this is not important as far as giving the examples is concerned and it is easier for comparisons as well.

in (3) and (5) refer to those with identity criterion $=_p$ while *books* in (4) and (6) refer to those with identity criterion $=_i$. Note that it is the predicates (*pick up* and *master*) that determine which identity criterion should be: in (3), the domain of *picked up* is the type *Phy* of physical objects, and the identity criterion for books is $=_p$ (being physically the same); and in (4), the domain of *master* is the type *Info* of informational objects, and the identity criterion for books is $=_i$ (being informationally the same).

Now, a further complication arises in cases like (2), where copredication and quantification interact and, therefore, in its semantic analysis, both identity criteria have to be used, i.e. we have to consider both collections of books: one with *being physically the same* as its identity criterion and the other with *being informationally the same* as its identity criterion. Giving the semantics properly, we should be able to get the following entailment:

- (7) John picked up and mastered three books \Rightarrow
 John picked up three physical objects and mastered three informational objects

Here, there is some sort of double-distinctness that should be accounted for, involving both identity criteria. Any adequate theory of copredication should not only account for (5) and (6), but also for (7), dealing with counting and individuation correctly.

In order to deal with copredication with quantification, this paper further develops the idea on identity criteria for CNs, studied for MTT-semantics by the second author in (Luo 2012a). We propose that, in general, CNs are interpreted as types associated with their identity criteria, i.e., they are setoids. A setoid is a pair $(A, =)$, where A is a type (domain of the setoid) and $=$ is an equivalence relation on A (equality of the setoid). If a setoid interprets a CN, its domain interprets that of the CN and its equality gives the identity criterion for the CN. For example, *books* in (3) refer to those in the collection of books whose identity criterion is $=_p$ (being physically the same), while *books* in (4) refer to those in the collection of books whose identity criterion is $=_i$ (being informationally the same). They are books with different identity criteria; or put in a better way, they belong to different collections of books whose identity criteria are different. In the CNs-as-setoids setting, it is shown that dot-types offer a proper treatment of copredication, not just the usual cases but also the sophisticated cases involving both copredication and quantification.

At this point, one might want to ask: why do people usually say that, in MTT-semantics, CNs are interpreted as types and do not pay attention to their identity criteria? This is because most cases are not as sophisticated as copredication with quantification and, in these cases, related CNs usually have the ‘same’ identity criteria. For example, *Man* inherits its identity criterion from *Human*: two

men are the same if and only if they are the same as humans. This explains why we usually consider the entailment below induced by the subtyping relationship $Man \leq Human$ as ‘straightforward’ without even mentioning their identity criteria:

(8) Three men talk \Rightarrow Three humans talk

In other words, when interpreting a CN, its identity criterion can usually be safely ignored. Of course, this is not always the case: *books* in (3) and (4) are referring to different collections of books which have different identity criteria and, in (2) where both copredication and quantification are involved, more than one identity criterion will be involved in constructing appropriate semantics. In other words, in general, CNs are setoids but, in the usual cases, their identity criteria can be safely ignored and they are just types.

In §[2], we shall give a brief introduction to dot-types in MTT-semantics. Then, in §[3], after an informal introduction to identity criteria for CNs, we shall study the generic semantics of numerical quantifiers. Copredication with quantification is studied in §[4], where we show how dot-type and the generic quantifier can be used together to deal with such situations. Related and future work is discussed in the last section.

[2] DOT-TYPES: A BRIEF INTRODUCTION

Dot-types in MTT-semantics were introduced by the second author (Luo 2010, 2012b) to model copredication in MTT-semantics. Here is a brief introduction.

Consider the copredication example (1), repeated here:

(9) John picked up and mastered the book.

Let *Phy* and *Info* be the types of physical and informational objects, respectively, and the interpretations of *pick up* and *master* have the following types:

$$\begin{aligned} \textit{pick up} & : Human \rightarrow Phy \rightarrow Prop \\ \textit{master} & : Human \rightarrow Info \rightarrow Prop \end{aligned}$$

In (9), *pick up* and *master* are applied in coordination to *the book*. In order for such coordination to happen, the two verbs must be of the same type.⁵ Our question is: given the above typing, can both of them be of the same type? The introduction of dot-types make this happen naturally.

Informally, $Phy \bullet Info$ is the dot-type that satisfies the following property: it is a subtype of both *Phy* and *Info*. Then, the phenomenon according to which

[5] In MTT-semantics, the conjunction *and* can be given a polymorphic type $\Pi A : LType.A \rightarrow A \rightarrow A$, where *LType* is the universe of linguistic types. See (Chatzikyriakidis & Luo 2012) for details.

a book has both a physical and an informational aspect can be captured by stipulating that *Book* is a subtype of *Phy • Info*: formally, we have

$$\begin{aligned} Book &\leq Phy \bullet Info \leq Phy \\ Book &\leq Phy \bullet Info \leq Info \end{aligned}$$

Therefore, by contravariance of subtyping for function types, we have

$$\begin{aligned} pick\ up &: Human \rightarrow Phy \rightarrow Prop \\ &\leq Human \rightarrow Phy \bullet Info \rightarrow Prop \\ &\leq Human \rightarrow Book \rightarrow Prop \\ \\ master &: Human \rightarrow Info \rightarrow Prop \\ &\leq Human \rightarrow Phy \bullet Info \rightarrow Prop \\ &\leq Human \rightarrow Book \rightarrow Prop \end{aligned}$$

In other words, *pick up* and *master* are both of type $Human \rightarrow Book \rightarrow Prop$ and, therefore, the coordination in (9) and its interpretation can proceed straightforwardly as intended.

In general, given types A and B , we can form the dot-type $A \bullet B$ in the case that A and B do not share common parts (formally called components). For instance, $Phy \bullet Info$ is a legitimate dot-type, while $Phy \bullet (Phy \bullet Info)$ is not, because in the latter, the constituent types Phy and $Phy \bullet Info$ share the common part Phy . Furthermore, as exemplified above, a dot-type $A \bullet B$ is a subtype of both of its constituent types A and B .

The formal definition of component and the rules for dot-types are given in Appendix A. Note that, because of the non-sharing requirement for components and the subtyping relationships with its component types, a dot-type is not an ordinary type (like an inductive type) already available in MTTs. One has to introduce dot-types by means of the specific rules in Appendix A.

As mentioned in Footnote 3, in MTT-semantics CNs are interpreted as types (rather than predicates): for instance, the CNs *table* and *man* can be interpreted as types *Table* and *Man*, respectively. Every CN corresponds to a type, but not vice versa: not every type represents a CN. Dot-types are examples: dot-types such as $Table \bullet Man$ do not represent any CNs in natural language.

[3] IDENTITY CRITERIA: COMMON NOUNS AS SETOIDS

[3.1] *Individuation and Setoids*

In simple terms, individuation is the process by which objects in a particular collection are distinguished from one another. Individuation provides us with the

means that enables one to count individual cats/dogs or any other type of objects and differentiate among them. At the same time, individuation also provides one with a sameness criterion, i.e. a way to decide whether two members of a particular collection are the same or not. The discussion on individuation goes back to at least Aristotle and has been the subject of enquiry of great philosophers in both the continental as well as the analytical tradition. However, it is outside the scope of the present chapter to present an overview of individuation from a philosophical point of view. Our point of departure will be individuation as discussed in the philosophy of language and linguistic semantics tradition.

In linguistic semantics, individuation is very much related to the idea that a CN may have its own identity criterion, as discussed by Geach (1962). In mathematical terms, the idea amounts to the association of an equivalence relation (the identity criterion) to each CN. In fact, in the tradition of constructive mathematics, a set or a type is indeed a collection of objects together with an equivalence relation that serves as identity criterion of that collection.⁶ For CNs in MTT-semantics, this is the same as saying that a CN should in general be interpreted as a *setoid* – a type associated with an equivalence relation over the type.

In our discussions of MTT-semantics, the current authors have been extensively discussing and endorsing the view that CNs are better treated as *types* rather than predicates. In doing so, we have been skipping some detail for the sake of simplicity. These details will now become significant, as well as handy, when discussing the issue of individuation concerning CNs. Put it in another way, these details we have been skipping are now crucial in giving an account of individuation and correct predictions in complex cases like (5) and (6) of copredication with quantification, where individuation, and the derived issue of counting, becomes significant and cannot be ignored anymore.

The crucial detail that has been mostly ignored is embodied in the proposal put forward by the second author in (Luo 2012a), according to which the interpretation of a CN is not just a type, but rather a type associated with an identity criterion for that specific CN. In other words, a common noun is in general interpreted as a setoid, i.e. a pair

$$(10) \quad (A, =)$$

where A is a type and $=: A \rightarrow A \rightarrow Prop$ is an equivalence relation over A . The notion of setoids is not new in type theory and reflects the view that a type is basically a setoid, meaning that it is comprised of a type plus an equivalence relation on this type. We apply this to linguistic semantics and this view makes

[6] For this, the interested readers may consult writings in constructive mathematics including, just to mention two of them, (Bishop 1967; Beeson 1985). The idea that sets/types may have different identity criteria is fundamentally different from that in classical mathematics where there is a universal equality relation between all objects of the formal theory (say, set theory).

meanings of CNs more nuanced with respect to identity criteria: two CNs may have the same base type, but their identity criteria can be different and, if so, they are different CNs.

To see how this proposal works, let's consider some simple examples for illustration. In a simplified view, one would only interpret a CN as a type: for instance, *human* would be interpreted as a type *Human*, as in (11). However, in the elaborate view, CNs are interpreted as setoids, i.e., pairs of the form (10) and, therefore the CN *human* is now interpreted as in (12).

(11) $[human] = Human : Type$ (CNs-as-types view)

(12) $[human] = (Human, =_h)$ (CNs-as-setoids view)

where $=_h : Human \rightarrow Human \rightarrow Prop$ is the equivalence relation that represents the identity criterion for humans.

Interpreting CNs as setoids makes explicit the individuation criteria (or identity criteria) and, based on this, we shall be able to use dot-types as introduced in the previous section to deal with copredication properly, even for the sophisticated cases when quantification is involved (see §[4]). However, before that, we shall first deal with the usual situations where the identity criteria of related CNs are essentially the same – inherited from supertypes.

[3.2] *Inheritance of Identity Criteria: Usual Cases of Individuation*

Consider the following sentences (13) and (14) and their formal interpretations (15) and (16), respectively:

(13) A man talks.

(14) A human talks.

(15) $\exists m : Man.talk(m)$

(16) $\exists h : Human.talk(h)$

where *talk* is interpreted as a predicate of type $Human \rightarrow Prop$.⁷ Note that $talk(m)$ in (15) is only well-typed because of the following subtyping relationship:

(17) $Man \leq Human$ ⁸

Then, the following is expected to be the case:

[7] When CNs are interpreted as setoids, the interpretations of verbs/adjectives should be *IC-respecting* predicates: for example, for $talk : Human \rightarrow Prop$, $talk(h_1) \Leftrightarrow talk(h_2)$ if $h_1 =_h h_2$.

[8] Such relations are intuitively true. Formally, it can be realised by more than one way: for example, *Man* may be defined as $\Sigma x : Human.male(x)$ with $male : Human \rightarrow Prop$ and then $Man \leq Human$ via the first projection as coercion.

(18) (15) \Rightarrow (16)

and, in fact, it can be easily proven because of (17). It is worth noting that, analysing the above example, we have not needed to consider the identity criteria of *man* and *human* at all, simply because that, for the simple sentences like (13) and (14), identity criteria are irrelevant.

However, this is not the case anymore if we consider the following sentences (19) and (20) and their semantics (21) and (22), respectively:

(19) Three men talk.

(20) Three humans talk.

(21) $\exists x, y, z : Man. x \neq_M y \ \& \ y \neq_M z \ \& \ x \neq_M z \ \& \ talk(x) \ \& \ talk(y) \ \& \ talk(z)$

(22) $\exists x, y, z : Human. x \neq_H y \ \& \ y \neq_H z \ \& \ x \neq_H z \ \& \ talk(x) \ \& \ talk(y) \ \& \ talk(z)$

where $Man = (Man, =_M)$ and $Human = (Human, =_H)$ are setoids and the identity criterion for men and that for humans are used in (21) and (22), respectively, to express that x, y and z are distinct from each other.

The difference of (19)/(20) with the earlier examples (13)/(14) is the presence of quantifier *three* (a numerical quantifier bigger than one), which makes it necessary to consider the individuation criteria explicitly by using the identity criteria $=_M$ and $=_H$. The relationship between the setoids *Man* and *Human* is not just that the domain of the former is a subtype of that of the latter ($Man \leq Human$), but their identity criteria are also essentially the same: the identity criterion for men is the restriction of the identity criterion for humans to the domain of men. Put in another way, the identity criterion for men is inherited from that for humans: two men are the same if, and only if, they are the same as human beings. In symbols, we have:

(23) $(=_M) = (=H)|_{Man}$

One may wonder how $=_M$ and $=_H$ may be specified so that (23) is true. For example, if *Man* and *Human* are both base types with assumed coercion c such that $Man \leq_c Human$, then given $=_H$, we can define $m =_M m'$ as $c(m) =_H c(m')$. Then, (23) is true. Another possibility is that, in the case that the type of men is defined to be the type of humans who are male – formally, $Man = \Sigma x : Human. male(x)$ (cf., Footnote 8), we can then define $m =_M m'$ as $\pi_1(m) =_H \pi_1(m')$, where π_1 is the operator for first projection. With this, (23) is true as well.

Because of both (17) and (23), we have the following result as expected, which would not be provable if we only have (17) but not (23):

(24) (21) \Rightarrow (22).

It is important to notice that, in usual cases, the relationships between related CNs are like that between *man* and *human*: one of the domains is the subtype of the other (like $Man \leq Human$) and one of the identity criteria inherits the other (like $=_M$ inheriting $=_H$ as shown above). Such an inheritance relationship occurs in many cases, in fact, in all of the usual cases: examples of such pairs include *man* and *human*, *red table* and *table*, and many others. It may be useful to provide a name for this relation and this what the following definition does.

Definition (sub-setoid) We say that a setoid $A = (A, =_A)$ is a sub-setoid of $B = (B, =_B)$, notation $A \sqsubseteq B$, if $A \leq B$ and $=_A$ is the same as $(=_B)|_A$ (the restriction of $=_B$ over A). We often write $=_B$ for $(=_B)|_A$, omitting the restriction operator.

For example, besides $Man \sqsubseteq Human$, we also have $(RTable, =_t) \sqsubseteq (Table, =_t)$, where *RTable* is the type $\Sigma x : Table.red(x)$ representing the domain of red tables and $=_t$ is the equivalence relation representing the identity criterion for tables (and inherited for red tables). Note that, in the restricted domain like *Man* or *RTable*, the identity criteria coincide with those in *Human* and *Table* and, in such a case, they are essentially the same and we can safely ignore them in semantic studies.

Of course, there are more sophisticated cases where identity criteria are not inherited – they are in fact rather different. Copredication with quantification provides interesting examples, to be studied in §[4].

[3.3] Generic Semantics of Numerical Quantifiers

Another way to consider the semantic interpretations of (19)/(20) is to define a generic semantics of the numeral quantifiers such as *three* and then define the semantics using the generic operator. In the following, we shall use *three* as an example of a numerical quantifier to show how to give generic semantics to them.

A first attempt to define the semantics of *three* is to consider the following definition (25): for any setoid $asetoidB = (B, =_B)$, and any predicate $P : B \rightarrow Prop$,

$$(25) \quad Three_0(B, P) = \exists x, y, z : B. D[B](x, y, z) \ \& \ P(x) \ \& \ P(y) \ \& \ P(z)$$

where $D[B](x, y, z) = x \neq_B y \ \& \ y \neq_B z \ \& \ x \neq_B z$. Using $Three_0$, the semantics of (21) and (22) can be rewritten as (26) and (27), respectively:

$$(26) \quad Three_0(Man, talk)$$

$$(27) \quad Three_0(Human, talk)$$

And, similarly, we can express the semantics of (28) as (29), which is (30) after

$Three_0$ is expanded, where $Phy = (Phy, =_p)$ is the setoid for the collection of physical objects and $pick\ up : Phy \rightarrow Prop$:

(28) John picked up three physical objects.

(29) $Three_0(Phy, pick\ up)$

(30) $\exists x, y, z : Phy. D[Phy](x, y, z) \ \& \ pick\ up(x) \ \& \ pick\ up(y) \ \& \ pick\ up(z)$

A common feature of the examples (19), (20) and (28) is that the verb's domain is the same as that of the object CN: for example, in (28), the domain of *pick up* is the same as *Phy*, the type of physical objects. Because of this restriction, the operator $Three_0$ is not generic enough: it does not cover semantics of sentences with reference to more general situations, including (31):

(31) John picked up three pens.

(31) is an example where, first, the verb is applied to a CN whose domain is more restricted (but not the same): *pick up* can be applied to any physical object, not just pens; secondly, the identity criterion of the object CN is inherited from that for the domain of the verb (two pens are the same if, and only if, they are the same as physical objects). The identity criterion for pens inherits that for physical objects and can be determined from the contextual information in the sentence, more specifically, from the verb *pick up*: only physical objects can be picked up (the semantic type of *pick up* is $Phy \rightarrow Prop$). The semantics of the CN *pen* is the setoid $(Pen, =_p)$, whose identity criterion comes (is inherited) from the setoid of physical objects. In general, the identity criteria are determined by those of the verbs or adjectives applied to them.

This has led us to the following generic definition.

Definition (*Three*) Let A be a type and $B = (B, =_B)$ a setoid such that $A \leq B$, and $P : B \rightarrow Prop$ a predicate over B . Then, we define the generic semantics of *three* as follows:

(32) $Three(A, B, P) = \exists x, y, z : A. D[B](x, y, z) \ \& \ P(x) \ \& \ P(y) \ \& \ P(z)$.

where $D[B](x, y, z) = x \neq_B y \ \& \ y \neq_B z \ \& \ x \neq_B z$. Among *Three*'s arguments, the type A and the domain B of setoid B are related in one of the following manners (all of them satisfy that A is a subtype of B):

(i) $B = A$: in this case, $Three(A, B, P)$ is just $Three_0(B, P)$. In other words, $Three_0$ is a special case of *Three*. For example, the semantics of (28) can be re-written as $Three(Phy, pick\ up)$.

(ii) B is different from A , but it is not a dot-type. An example of this is (31)

whose semantics can be given as (33), which is (34) when *Three* is expanded:

(33) $Three(Pen, Phy, pick\ up)$

(34) $\exists x, y, z : Pen. D[Phy](x, y, z) \& pick\ up(x) \& pick\ up(y) \& pick\ up(z)$

(iii) *B* is different from *A*, but it is a dot-type – we shall discuss this in §[4].

When the setoid *B* interprets a CN, the predicate *P* is usually the interpretation of a verb phrase or an adjective. We usually require that such a predicate respect the identity criteria in the sense that, if $x =_B y$, $P(x) \Leftrightarrow P(y)$ (cf., Footnote 7.) Sometimes, *B* does not interpret a CN; an example is when *B* is a dot-type, which represents a typical case of copredication, to be discussed in the next section.

[4] COPREDICATION WITH QUANTIFICATION

When both copredication and quantification are involved, the situation becomes more sophisticated and requires special treatment w.r.t the setoids concerned in order to give proper semantics.

[4.1] *Dot-types and Quantification*

Dot-types in MTTs have been developed for copredication in linguistic semantics (Luo 2010, 2012b). The basic idea and formal rules are sketched in §[2], with the simple example of (9), repeated below as (35), which involves copredication, but no quantification. This is different in (2), repeated here as (36), where copredication interacts with quantification manifested in the numerical quantifier *three*.

(35) John picked up and mastered the book.

(36) John picked up and mastered three books.

In such cases as (36) where both copredication and quantification are involved, a proper semantic treatment becomes more sophisticated and require more careful considerations.

Let's start by considering the simpler subcases (37) and (38), which do not involve copredication.

(37) John picked up three books.

(38) John mastered three books.

First, it is not difficult to realise that (37) and (38) are not much different from (31), if we replace pens by books, and they should have similar semantics. Actually, they do: the semantics of (37) and (38) are (39) and (40), respectively:

(39) $Three(Book, Phy, pick\ up).$

(40) $Three(Book, Info, master).$

where $Phy = (Phy, =_p)$ and $Info = (Info, =_i)$.

However, it is important to note that the CN *book* in (37) refers to a different collection from that referred to by *book* in (38): for the collection of books in (37), two books are the same if they are physically the same, while for that in (38), two books are the same if they are informationally the same. Put in another way, although they share the same domain *Book*, their identity criteria are different and, therefore, they are different collections. This is reflected in their semantics: to compare books, $=_p$ is used in (39) and $=_i$ in (40). So, the CN *book* in (37) stand for a collection of books that is different from that in (38), represented by the setoids $Book_1$ and $Book_2$ as follows:

(41) $Book_1 = (Book, =_p)$

(42) $Book_2 = (Book, =_i)$

At this point, a question that naturally arises is how the identity criterion for books is determined: why do we use $=_p$ in (39) and $=_i$ in (40)? The answer is that it is the verb (and its semantics) that determines the identity criterion of the object CN. In (37), the verb is *pick up*: $Phy \rightarrow Prop$, its domain *Phy* has determined that it is $=_p$, the identity between physical objects, that should be used for the books; and in (38), the verb is *master*: $Info \rightarrow Prop$, its domain *Info* has determined that it is $=_i$, the identity between informational objects, that should be used for the books. In general, one may express this as follows: given a predicate $V : Dom(V) \rightarrow Prop$ that interprets the verb, the identity criterion of the object CN *N* (*book* in the above examples) to which the verb is applied is determined by the domain of the predicate: in the case $Dom(V) \in \{Phy, Info\}$,

$$(43) \quad IC^{N,V} = \begin{cases} =_p & \text{if } Dom(V) = Phy \\ =_i & \text{if } Dom(V) = Info \end{cases}$$

What the above does not cover is the crucial case of conjunction – for example, coordination of verbs *pick up* and *master* in (36). In order for such a coordination to happen, the two verbs must be of the same type, but the originally given types to *picked up* and *mastered* are different: $Human \rightarrow Phy \rightarrow Prop$ and $Human \rightarrow Info \rightarrow Prop$, respectively. As explained in §[2], the introduction of dot-type $Phy \bullet Info$ and the subtyping in (44) make both verbs be of the same type and, hence, the two verbs can be coordinated.

(44) $Book \leq Phy \bullet Info$

However, the treatment illustrated in §[2] is only for the situations when CNs are interpreted as types. As explained, (36) also involves quantification, due to the additional quantifier *three*, as well as copredication. This makes it necessary to consider CNs as setoids, with explicit identity criteria. As mentioned in the Introduction, the correct semantics of sentences like (36) makes use of both identity criteria, $=_p$ for physical objects and $=_i$ for informational objects. In other words, one can update the above (43) into (45):

$$(45) \quad IC^{N,V} \Rightarrow \begin{cases} =_p & \text{if } \text{Dom}(V) = \text{Phy} \\ =_i & \text{if } \text{Dom}(V) = \text{Info} \\ ??? & \text{if } \text{Dom}(V) = \text{Phy} \bullet \text{Info} \end{cases}$$

The question is: what is the equivalence relation if the domain is a dot-type? Put in another way, assuming that we have the setoid $\text{Phy} \bullet \text{Info}$, its domain is $\text{Phy} \bullet \text{Info}$, what is $=_{\text{Phy} \bullet \text{Info}}$? This is to be defined below in the following subsection.

[4.2] Setoids for Dot-types

We shall first give the definition of setoids for dot-types and then explain the related issues for clarification.

Definition (setoids for dot-types) Let $A = (A, =_A)$ and $B = (B, =_B)$ be setoids. Then the dot-setoid $A \bullet B$ is defined as follows:

$$A \bullet B = (A \bullet B, =_{A \bullet B}),$$

where $\langle a_1, b_1 \rangle =_{A \bullet B} \langle a_2, b_2 \rangle$ if, and only if, $(a_1 =_A a_2) \vee (b_1 =_B b_2)$.

Note that disjunction is used in the above definition. This is mainly because that, with this definition, the semantics by means of the generic numerical quantifier (say *Three*) is exactly what we want. For instance, the semantics for (36) is (46), which is (47) when *Three* is expanded, where $pm : \text{Human} \rightarrow \text{Phy} \bullet \text{Info} \rightarrow \text{Prop}$ is the interpretation of *pick up and master*:

$$(46) \quad \text{Three}(\text{Book}, \text{Phy} \bullet \text{Info}, pm(j))$$

$$(47) \quad \exists x, y, z : \text{Book}. \\ D[\text{Phy}](x, y, z) \ \& \ D[\text{Info}](x, y, z) \ \& \ pm(j, x) \ \& \ pm(j, y) \ \& \ pm(j, z)$$

The semantics captures the ‘double distinctness’ as expected: that is, there are books x , y and z which are different physically *and* informationally. Note that this is achieved through defining the equivalence relation for dot-types by means of disjunction of both identity criteria and, then, we obtain double distinctness by negating the disjunction.

At this point, one might want to question the equivalence relation for dot-types, as given in the definition above. The first question would be: Why is it

defined by disjunction of the constituent equalities? The answer to this is that, besides other reasons, the definition delivers the correct semantics by means of the generic quantifier, as shown above.

Another subtler and possibly deeper question is: is the definition appropriate? Instead of answering this question directly, let us discuss two related issues which we hope would clear away some misunderstandings that may be behind asking the question and clarify the issues at hand. The first issue concerns the relationship between dot-types and linguistic entities like CNs. Although CNs can be interpreted as types, there is no CN that can be interpreted as a dot-type. In particular, the equivalence relation does not represent any identity criterion of a CN. That's why we have been careful in not calling the equivalence relation $=_{A \bullet B}$ an identity criterion – it is simply not.

The other related issue is more general than the first: we think that, unlike other data types in MTTs (say, the type of natural numbers), a dot-type $A \bullet B$ is not a representation of a collection of objects, although its constituent types A and B may be (this concerns the detailed definition of dot-types – see, for example, (Luo 2012b) for some discussions.) Therefore, in a setoid whose domain is a dot-type, its equivalence relation is not supposed to be the equality for a collection. This explains the flexibility we have in defining the equivalence relation for dot-types. In fact, as far as correctness is concerned, as long as it is an equivalence relation, it would do. Our definition does result in an equivalence relation.

[4.3] *Verbs Plus Adjectives: More Examples of Copredication with Quantification*

Consider the following example:

(48) John mastered three heavy books.

Interpreting the above sentence, one needs to capture that John mastered three informational objects that are also heavy as physical objects. In effect, both the verb and the adjective have a word on the decision concerning the identity criteria involved in interpreting this sentence. Formally, we have:

(49) $heavy : Phy \rightarrow Prop$

(50) $master(j) : Info \rightarrow Prop$

So, both identity criteria for physical/informational objects are in play.

This is a more sophisticated copredication phenomenon. Let us break the problem into two parts: first we look at the adjectival modification and, then, the result of *three* quantifying over *heavy books* and, then, the application of *mastered(j)* to the whole quantified NP *three heavy books*. The interpretation of the adjectival

modification is taken to be a Σ -type:⁹ *heavy book* will be interpreted as (51), which is a subtype of *Book* as shown in (52):¹⁰

$$(51) \quad HBook = \Sigma(Book, heavy)$$

$$(52) \quad HBook \leq Book \leq Phy \bullet Info$$

The interpretation we get for the whole sentence (48) is (53), which is (54) when *Three* is expanded:

$$(53) \quad Three(HBook, Phy \bullet Info, master(j))$$

$$(54) \quad \exists x, y, z : HBook. \\ D[Phy](x, y, z) \& D[Info](x, y, z) \& master(j, x) \& master(j, y) \& master(j, z)$$

The above account can also be extended to the case when a noun is modified by more than one adjective that may induce different identity criteria, such as (55):

$$(55) \quad \text{John mastered three heavy informative books.}$$

In such a case, what we need to capture is the fact that John mastered three informational objects that are heavy as physical objects but informative as informational objects. Let us see how this works. We first form the Σ -type in a nested manner,¹¹ and we have similar subtyping relationships (58):

$$(56) \quad IBook = \Sigma(Book, informative), \text{ where } informative : Info \rightarrow Prop.$$

$$(57) \quad HIBook = \Sigma(IBook, heavy), \text{ where } heavy : Phy \rightarrow Prop.$$

$$(58) \quad HIBook \leq IBook \leq Book \leq Phy \bullet Info.$$

The interpretation we get for the whole sentence (55) is the following:

$$(59) \quad Three(HIBook, Phy \bullet Info, master(j))$$

Shown by the above examples, what is obvious is that the way the identity criteria are decided are a little bit more complicated, given that we might have cases where both a verb and an adjective play a role in this decision.

Another related issue is how many identity criteria can be used. In all our examples, the maximum number is two. This is not an accident as it basically represents the ability that a common noun to be associated with identity criteria. This is true for common nouns like *book* that are associated with two aspects: in

[9] For more information on adjectival modification and modification in general in MTTs, please consult (Chatzikyriakidis & Luo 2013, 2017a)

[10] (51) is just another notation for $\Sigma x : Book. heavy(x)$, the type of books which are heavy.

[11] We note that there is another interpretation of the adjective modification: heavy informative books is seen as heavy and informative books. Then similar but slightly different analysis would follow.

the case of book, a physical and an informational one. Whether we have cases where more than three identity criteria are involved for the same common noun boils down to the question of whether common nouns with that many aspects can be found. One such case is *newspaper*: it is associated with three senses: (a) physical, (b) informational, and (c) institutional. However, rather interestingly, only two of the three senses can appear together. More specifically, [Antunes & Chaves \(2003\)](#) argue that, whereas senses (a) and (b) can appear together in a coordinated structure, sense (c) cannot appear with any of the other two:

- (60) # That newspaper is owned by a trust and is covered with coffee.
- (61) # The newspaper fired the reporter and fell off the table.
- (62) # John sued and ripped the newspaper.

In ([Chatzikyriakidis & Luo 2015](#)), we have argued that one can find cases where two senses are actually coordinated:

- (63) The newspaper you are reading is being sued by Mia.

Whatever the data are, however, there are no cases where all three senses are coordinated nor we know of any other common noun that allows this kind of situation. This means that an account of individuation can be reduced to dot-types with two senses, not more than that, even though in principle the account presented here, as well as other accounts like the one proposed by ([Gotham 2017](#)), could easily be generalized for n senses if there is a need to do so.

[5] RELATED WORK AND CONCLUSION

The dot-type approach offers an adequate account of copredication in general, and a proper treatment of individuation and counting in copredication with quantification in particular. It is shown that, in this latter sophisticated case, identity criteria should be taken into account – in another word, CNs should be interpreted as setoids. Whence the identity criteria of CNs are taken into account, correct predictions and expected reasoning results can be obtained.

We have only considered the interpretation of CNs as setoids in the specific context of copredication with quantification. However, the general paradigm of CNs-as-setoids has not been developed in depth: these include, for example, its relationship with the simpler CNs-as-types paradigm and its other applications. It should be explored further in future work.

Related work. The origin of the notion of criteria of identity can be traced back to [Frege \(1884\)](#) when he considered abstract mathematical objects such as numbers or lines. As far as the authors know, it was [Geach \(1962\)](#) who first connected iden-

tity criteria with CNs in linguistic semantics. Subsequent studies include Baker (2003), Barker (2008), Gupta (1980) and Luo (2012a), among others. For the first time, this paper has studied the issue of identity criteria in the context of copredication with quantification. As mentioned in Footnote 1, we have studied the issues of copredication with quantification in (Chatzikyriakidis & Luo 2015) but did not realise the necessity of considering identity criteria – the current paper does the job.

The individuation problem was studied by Asher (2008, 2012), among others. Gotham (2017), starting from his PhD thesis (Gotham 2014), has successfully looked at the individuation problem in copredication with quantification in a mereological setting, as related to (Link 1983) among others. It is the first account to deal in depth with the problems discussed in this paper in which Gotham manages to provide a compositional account of co-predication that derives the correct identity criteria in the cases where double distinctness is needed. One of the interesting connections with our work is the assumption in the first paragraph of Section 2 of (Gotham 2017) – it is related to our disjunction-based equivalence relation for dot-types (see Definition in §[4.2]). Deeper reflections and analysis of these different approaches are called for.

ACKNOWLEDGEMENTS

The first author supported by by grant 2014-39 from the Swedish Research Council, which funds the Centre for Linguistic Theory and Studies in Probability in the Department of Philosophy, Linguistics, and Theory of Science at the University of Gothenburg. We are grateful to N. Asher, M. Gotham, R. Cooper and C. Retoré for providing useful discussion and comments throughout the years that we have been working on this issue.

REFERENCES

- Antunes, S. & R.P Chaves. 2003. On the licensing conditions of co-predication. In *Proc of the 2nd Inter. Workshop on Generative Approaches to the Lexicon (GL 2007)*, .
- Asher, N. 2008. A type driven theory of predication with complex types. *Fundamenta Informaticae* 84(2). 151–183.
- Asher, N. 2012. *Lexical Meaning in Context: a Web of Words*. Cambridge University Press.
- Baker, Mark C. 2003. *Lexical categories: Verbs, nouns and adjectives*, vol. 102. Cambridge University Press.
- Barker, C. 2008. Nominals don't provide criteria of identity. In A. Alexiadou &

- M. Rathert (eds.), *Nominalizations across languages and frameworks* Interface Explorations, .
- Bassac, C., B. Mery & C. Retoré. 2010. Towards a type-theoretical account of lexical semantics. *Journal of Logic, Language and Information* 19(2).
- Beeson, M.J. 1985. *Foundations of constructive mathematics*. Springer-Verlag.
- Bishop, E. 1967. *Foundations of constructive analysis*. McGraw-Hill.
- Chatzikiyriakidis, S. & Z. Luo. 2012. An Account of Natural Language Coordination in Type Theory with Coercive Subtyping. In Y. Parmentier & D. Duchier (eds.), *Proc. of Constraint Solving and Language Processing (CSLP12)*. LNCS 8114, 31–51. Orleans.
- Chatzikiyriakidis, S. & Z. Luo. 2013. Adjectives in a modern type-theoretical setting. In G. Morrill & J.M Nederhof (eds.), *Proceedings of Formal Grammar 2013*. LNCS 8036, 159–174.
- Chatzikiyriakidis, S. & Z. Luo. 2015. Individuation criteria, dot-types and copredication: A view from modern type theories. In *Proceedings of the mathematics of language 2015, acl anthology*, .
- Chatzikiyriakidis, S. & Z. Luo. 2017a. Adjectival and adverbial modification: The view from modern type theories. *Journal of Logic, Language and Information* 26(1). 45–88.
- Chatzikiyriakidis, S. & Z. Luo. 2017b. Identity Criteria of CNs: Quantification and Copredication. *Workshop on Approaches to Coercion and Polysemy*. Oslo .
- Chatzikiyriakidis, S. & Z. Luo. 2018. *Formal Semantics in Modern Type Theories*. Wiley & ISTE Science Publishing Ltd. (to appear).
- Frege, G. 1884. *Grundlagen der arithmetik*. Basil Blackwell. (Translation by J. Austin in 1950: *The Foundations of Arithmetic*).
- Geach, P. 1962. *Reference and generality: An examination of some medieval and modern theories*. Cornell University Press.
- Gotham, M. 2014. *Copredication, quantification and individuation*. University College London PhD dissertation.
- Gotham, M. 2017. Composing criteria of individuation in copredication. *Journal of Semantics* 34(2). 333–371.
- Gupta, A. 1980. *The logic of common nouns*. Yale University Press.

- Link, G. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Schwarze C. Bauerle R. & von Stechow A (eds.), *Meaning, use an interpretation of language*, 302–323. Mouton De Gruyter, Berlin.
- Luo, Z. 1999. Coercive subtyping. *Journal of Logic and Computation* 9(1). 105–130.
- Luo, Z. 2010. Type-theoretical semantics with coercive subtyping. *Semantics and Linguistic Theory 20 (SALT20)*, Vancouver .
- Luo, Z. 2012a. Common Nouns as Types. In D. Bechet & A. Dikovsky (eds.), *Logical Aspects of Computational Linguistics (LACL'2012)*. LNCS 7351, .
- Luo, Z. 2012b. Formal semantics in modern type theories with coercive subtyping. *Linguistics and Philosophy* 35(6). 491–513.
- Luo, Z., S. Soloviev & T. Xue. 2012. Coercive subtyping: theory and implementation. *Information and Computation* 223. 18–42.
- Pustejovsky, J. 1995. *The generative lexicon*. MIT.
- Ranta, A. 1994. *Type-theoretical grammar*. Oxford University Press.

A RULES FOR DOT-TYPES

The following formal rules for dot-types are given in (Luo 2012b). The notion of component with notation $\mathcal{C}(_)$ is used in the Formation Rule below, as defined as follows.

Definition A.1 (components) Let $T : Type$ be a type in the empty context. Then, $\mathcal{C}(T)$, the set of components of T , is defined as, where $\text{SUP}(T) = \{T' \mid T \leq T'\}$:

$$\mathcal{C}(T) =_{df} \begin{cases} \text{SUP}(T) & \text{if the normal form of } T \text{ is not of the form } X \bullet Y \\ \mathcal{C}(T_1) \cup \mathcal{C}(T_2) & \text{if the normal form of } T \text{ is } T_1 \bullet T_2 \end{cases}$$

Presenting the rules, we shall assume the knowledge of coercive subtyping (Luo 1999; Luo et al. 2012). Also, we shall use \leq instead of $<$ for subtyping and, for this, it may be useful to know that the coherence condition for coercive subtyping implies that, if $A \leq_c A$, then $c = [x : A]x$, the identity function over A .

Formation Rule

$$\frac{\Gamma \text{ valid } \langle \rangle \vdash A : Type \quad \langle \rangle \vdash B : Type \quad \mathcal{C}(A) \cap \mathcal{C}(B) = \emptyset}{\Gamma \vdash A \bullet B : Type}$$

Introduction Rule

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B \quad \Gamma \vdash A \bullet B : Type}{\Gamma \vdash \langle a, b \rangle : A \bullet B}$$

Elimination Rules

$$\frac{\Gamma \vdash c : A \bullet B}{\Gamma \vdash p_1(c) : A} \quad \frac{\Gamma \vdash c : A \bullet B}{\Gamma \vdash p_2(c) : B}$$

Computation Rules

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B \quad \Gamma \vdash A \bullet B : Type}{\Gamma \vdash p_1(\langle a, b \rangle) = a : A} \quad \frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B \quad \Gamma \vdash A \bullet B : Type}{\Gamma \vdash p_2(\langle a, b \rangle) = b : B}$$

Projections as Coercions

$$\frac{\Gamma \vdash A \bullet B : Type}{\Gamma \vdash A \bullet B \leq_{p_1} A : Type} \quad \frac{\Gamma \vdash A \bullet B : Type}{\Gamma \vdash A \bullet B \leq_{p_2} B : Type}$$

Coercion Propagation

$$\frac{\Gamma \vdash A \bullet B : Type \quad \Gamma \vdash A' \bullet B' : Type \quad \Gamma \vdash A \leq_{c_1} A' : Type \quad \Gamma \vdash B = B' : Type}{\Gamma \vdash A \bullet B \leq_{d_1[c_1]} A' \bullet B' : Type}$$

where $d_1[c_1](x) = \langle c_1(p_1(x)), p_2(x) \rangle$.

$$\frac{\Gamma \vdash A \bullet B : Type \quad \Gamma \vdash A' \bullet B' : Type \quad \Gamma \vdash A = A' : Type \quad \Gamma \vdash B \leq_{c_2} B' : Type}{\Gamma \vdash A \bullet B \leq_{d_2[c_2]} A' \bullet B' : Type}$$

where $d_2[c_2](x) = \langle p_1(x), c_2(p_2(x)) \rangle$.

$$\frac{\Gamma \vdash A \bullet B : Type \quad \Gamma \vdash A' \bullet B' : Type \quad \Gamma \vdash A \leq_{c_1} A' : Type \quad \Gamma \vdash B \leq_{c_2} B' : Type}{\Gamma \vdash A \bullet B \leq_{d[c_1, c_2]} A' \bullet B' : Type}$$

where $d[c_1, c_2](x) = \langle c_1(p_1(x)), c_2(p_2(x)) \rangle$.

AUTHOR CONTACT INFORMATION

Stergios Chatzikyriakidis

Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg

stergios.chatzikyriakidis@gu.se

Zhaohui Luo

Royal Holloway, University of London

zhaohui.luo@hotmail.co.uk