Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

D.J.N. van Eijck

Determiners

Computer Science/Department of Software Technology   Report CS-R9039   August

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Determiners

Jan van Eijck
Centre for Mathematics and Computer Science
Kruislaan 413, 1098 SJ Amsterdam
The Netherlands

This paper gives an overview of the treatment of the semantics of determiner expressions in natural language.

Note: The paper will appear in the *Encyclopedia of Language and Linguistics*, Pergamon Press and Aberdeen University Press.

1985 Mathematics Subject Classification: 68S10.
CR Categories: H.2.3, I.2.4, I.2.7.
Keywords and Phrases: semantics of natural language, query languages, knowledge representation languages.

1 Introduction

Syntactically, determiners are operators that combine with nouns to form noun phrases. Semantically, determiners are functions that combine with noun denotations to form noun phrase denotations. Since noun phrase denotations can be viewed as sets of verb phrase denotations (namely, the denotations of those verb phrases that combine with the noun phrase to form a true sentence), determiner denotations can also be viewed as relations between noun denotations and verb phrase denotations. A determiner denotation relates noun denotations to the denotations of those verb phrases for which the sentence *Det N VP* is true.

This article first lists some general properties of determiner denotations. Next, a semantic account is given of the distinction between definite, indefinite and quantifier determiners. Subsequently, attention is paid to the internal structure of determiners and some special cases are discussed. The article ends with a sketch of a dynamic view on noun phrase and determiner interpretation, where unbounded anaphoric links are handled by interpreting the antecedent noun phrases as 'state changers'.

2 Determiners as Relations

In the simplest possible setup, disregarding the singular plural distinction and focussing on determiners combining with simple count nouns, determiner denotations are functions from sets of individuals (N denotations) to sets of sets of individuals (NP denotations), or, equivalently, determiner denotations are relations between sets of individuals (N denotations) and sets of individuals (VP denotations). Some examples will clarify this.

(1) $[_{\text{NP}}[_{\text{DET}}$ the $][_{\text{N}}$ men $]]$

(2) $[_{\text{NP}}[_{\text{DET}}$ a $][_{\text{N}}$ woman $]]$

(3) $[_{\text{NP}}[_{\text{DET}}$ at least three $][_{\text{N}}$ children $]]$

To facilitate talking about the denotations of the above example phrases, it is convenient to take a model firmly in mind and to use set–theoretic notation to

Report CS-R9039
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

talk about entities in that model: **men** is the set of men in the model, **women** the set of women, $|\textbf{children} \cap A|$ the number of elements of the intersection of the set of children and the set A, i.e. the number of entities that are both children and members of A. Using E for the universe of discourse, we then have the following:

(1) is interpreted as $\{A \subseteq E \mid \textbf{men} \subseteq A\}$ in case **men** has at least two members, it is undefined otherwise.

(2) is interpreted as $\{A \subseteq E \mid \textbf{women} \cap A \neq \emptyset\}$.

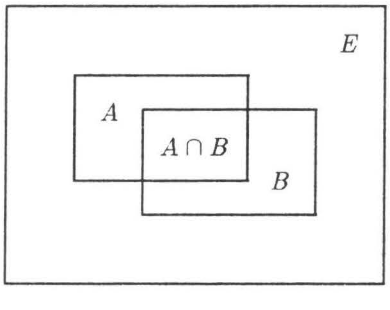
(3) is interpreted as $\{A \subseteq E \mid |\textbf{children} \cap A| \geq 3\}$.

Now, if for convenience we disregard the contextual factors that specify an appropriate subdomain for interpreting the definite description, we can say that *The men walked* is true in the model if and only if (i) **men**, the set of men in the model, has at least two elements, and (ii) the set of walkers is in $\{A \subseteq E \mid \textbf{men} \subseteq A\}$. Similarly, *At least three children played* is true if and only if the set of players is a member of $\{A \subseteq E \mid |\textbf{children} \cap A| \geq 3\}$.

If one views this uniform treatment of the semantics of subject predicate combinations from a slightly different angle, determiner denotations are two-place relations D between sets of individuals. Instead of $B \in DA$ we now write DAB. *The men walked* is true in a given model if and only if (i) there are more than two men in the model, and (ii) the relation of inclusion holds between **men** and **walked**. Thus, the determiner *the* is interpreted as the inclusion relation (modulo a uniqueness requirement for singular *the* and a semantic plurality requirement for plural *the*).

Abstracting from the domain of discourse, we can say that determiner denotations pick out binary relations on sets of individuals, on arbitrary universes E. Notation: $D_E AB$. See figure (1).

Figure 1: Determiners as Binary Relations



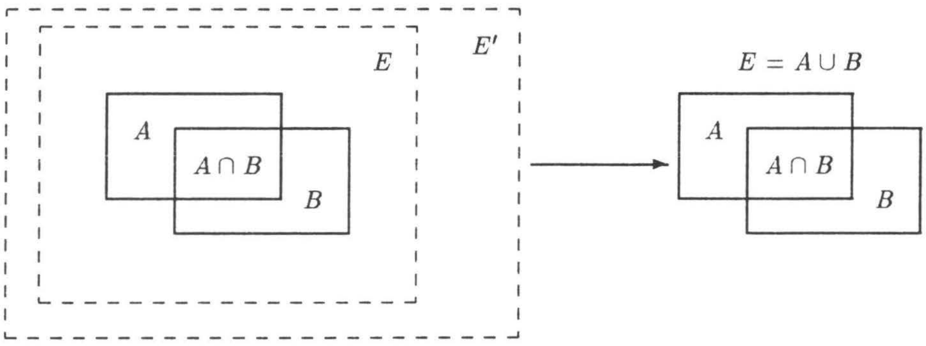
3 Global Conditions on Determiner Relations

Determiner relations satisfy certain requirements which depend on the semantic nature of the determiner. It is common to distinguish *definite*, *indefinite* and *quantifier* determiners. The boundary lines between these kinds can be drawn by semantic means, but first we must mention two semantic requirements that almost all determiners meet. A first requirement is *extension*:

EXT For all $A, B \subseteq E \subseteq E'$: $D_E AB \iff D_{E'} AB$.

A relation observing **EXT** is stable under growth of the universe. So, given sets A and B, only the objects in the minimal universe $A \cup B$ matter. See figure (2).

Figure 2: The Effect of EXT



Not all natural language determiners do satisfy **EXT**. An example of a determiner that does not is *many* in the sense of *relatively many.*

A second requirement for determiners is *conservativity*:

CONS For all $A, B \subseteq E$: $D_E AB \iff D_E A(A \cap B)$.

This property expresses that the first argument of a determiner relation (the interpretation of the noun) plays a crucial rôle: it sets the stage, in the sense that everything outside the extension of the first argument is irrelevant.

It is not difficult to think of noun phrase determiners that do not satisfy **CONS**. One example is *only* in example (4).

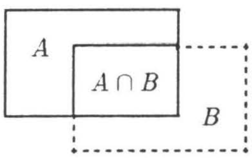
(4) *Only men came to the party.*

This example is true in a situation where all partygoers were men. Starting out from a situation like this, and adding some women to the partygoers will make (4) false. This shows non-conservativity. All is still well if it can be argued that noun phrases starting with *only*, *mostly* or *mainly* (two other sources of non-conservativity) are exceptional syntactically, in the sense that these noun phrase prefixes are not really determiners. In the case of *only*, it could be argued that *only men* has structure $[_{NP}[_{MOD}\text{only}][_{NP}\text{men}]]$, with *only* not a determiner but a noun phrase modifier, just as in (5).

(5) *Only John came to the party.*

However this may be, separating out the determiners satisfying **CONS** and **EXT** is important, for the two conditions taken together ensure that the truth of DAB depends only on $A - B$ and $A \cap B$. See figure (3).

Figure 3: The Combined Effect of EXT and CONS



4 Semantic Distinctions

Next, the relational perspective suggests a very natural way of semantically characterizing the three main kinds of determiner relations listed below.

Quantifier Determiners The determiners in *every child*, *no man*, *at least five donkeys*.

Definite Determiners The determiners in *the king of France*, *those books*, *John's girlfriend*.

Indefinite Determiners The determiners in *some woman*, *an unknown God*.

Determiners that are **quantifiers** satisfy the condition of *isomorphy*. **ISOM** (see the article QUANTIFIERS for a formulation) expresses that only the *cardinalities* (numbers of elements) of the sets A and B matter. If D satisfies **EXT**, **CONS** and **ISOM**, the truth of DAB depends only on the cardinal numbers $|A - B|$ and $|A \cap B|$.

The *definite* determiners are the determiners forming noun phrase denotations which are *principal filters* on some universe E.

A set $\mathcal{F} \subseteq \mathcal{P}(E)$ is a **principal filter** on E if and only if there is a set X such that $\mathcal{F} = \{A \subseteq E \mid X \subseteq A\}$. Here X is called the *generator* of the principal filter $\mathcal{F}$.

For example, the genitive determiner *John's* is definite, for if Sally is John's girlfriend (in some suitable domain of discourse E) then the noun phrase *John's girlfriend* is interpreted as a set of sets $\{A \subseteq E \mid s \in A\}$, which is the principal filter on E generated by $\{s\}$.

Also, proper names form definite noun phrases, for they are interpreted under the relational regime as the principal filters generated by the referents of the names. On a universe E, the noun phrase *John* is interpreted as the set $\{X \subseteq E \mid \mathbf{j} \in X\}$. In fact, this way of treating names is a slight overcomplication, for one could just as well have said that the proper name *John* is interpreted as an individual **j** in the domain under consideration, and *John smiled* is true just in case **j** is an element of the set of smiling entities in the domain. The bonus of treating proper names as principal filters is that one gets a uniform semantics for the move of combining noun phrases with verb phrases.

The *indefinite* determiners, finally, are the determiners forming noun phrase denotations which are filters, but not principal filters. Again we fix an appropriate universe E.

A set $\mathcal{F} \subseteq \mathcal{P}(E)$ is a **filter** on E if and only if the following hold:

1. if $A \in \mathcal{F}$ and $B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$.
2. if $A \in \mathcal{F}$ and $A \subseteq A' \subseteq E$ then $A' \in \mathcal{F}$.

Note that it follows directly from the definitions that all principal filters are filters but not vice versa. An example of a filter on E which is not principal is $\{X \subseteq E \mid X \cap A \neq \emptyset\}$, where A is some fixed set $\subseteq E$.

It is not difficult to see that the determiners *a* and *some* are indefinite, for *a man* is interpreted as the set of all sets containing at least one man, which is a filter, but not a principal filter, and similarly for *some women*. For a slightly different example, consider the compound determiner *somebody's*. This determiner is indefinite according to the definition given here, for *somebody's girlfriend* is interpreted as the set of all sets containing at least one girl who is somebody's girlfriend, and this is a non–principal filter. Similarly, the complex determiner *John's or Bill's* can be seen to be indefinite.

Note that it follows from the definitions that quantifiers and indefinites overlap. The interpretation of *some woman* observes **ISOM**, and therefore it is a quantifier, and it is also a non principal filter, and therefore it is an indefinite. On the other hand, *no* is an example of a quantifier determiner which is not an indefinite, for the interpretation of *no woman* observes **ISOM** but is not a filter. An example of the converse situation is *John's or Bill's*, which is an indefinite determiner, for the interpretation of *John's or Bill's girlfriend* is a set $\mathcal{F}$ which is a non principal filter, but the determiner is no quantifier, for $\mathcal{F}$ does not observe **ISOM**.

Definite and indefinite noun phrase denotations share the property of being *monotone increasing* in the following sense.

MON$\uparrow$ A set $\mathcal{A} \subseteq \mathcal{P}(E)$ is *monotone increasing* if and only if for all sets $A \subseteq A' \subseteq E$ it holds that $A \in \mathcal{A}$ implies $A' \in \mathcal{A}$.

Further information about **MON**$\uparrow$ and related properties is provided in the article QUANTIFIERS.

5 Internal Structure

The relational perspective makes it easy to interpret boolean compounds of determiners; see figure (4). Also, the treatment of possessives is relatively straightforward. In the figure, **N** is used for the interpretation of the noun *N*, **Det** for that of the determiner *Det*, and **NP** for that of the noun phrase *NP*. Also, in the entry for possessives, $F(Y)$ is everything which has something in Y that may count as its 'possessor'. The vagueness in the previous sentence is deliberate, for the nature of the possessive relation will generally depend on context.

Figure 4: Structure and Interpretation of Determiners

operation	syntax	interpretation in domain E
negation	[not D] N	$\{X \subseteq E \mid X \notin \mathbf{D(N)}\}$
conjunction	[D_1 and D_2] N	$\mathbf{D}_1\mathbf{(N)} \cap \mathbf{D}_2\mathbf{(N)}$
disjunction	[D_1 or D_2] N	$\mathbf{D}_1\mathbf{(N)} \cup \mathbf{D}_2\mathbf{(N)}$
possessive	[*NP's*] N	$\{X \subseteq E \mid \exists Y : Y \in \mathbf{NP}\ \&\ F(Y) \cap \mathbf{N} \subseteq X\}$
Adj restriction	[D *Adj*] N	$\{X \subseteq E \mid X \cap \mathbf{Adj} \in \mathbf{D(N)}\}$

One might also want to count adjectival restrictions as part of the determiner, giving *those blue-eyed girls* the structure (6).

(6) $[_{\text{NP}}[_{\text{DET}}$ those blue-eyed $][_{\text{N}}$ girls $]]$.

The semantics of adjectival restrictions are given in the final row of figure (4). Observe that this treatment only works for intersective adjectives such as *blonde*, *blue-eyed*, *long-haired* etcetera. Non-intersective adjectives such as *small* are in need of more subtle treatment (see the article ADJECTIVES), while intensional adjectives like *alleged*, *fake*, *would-be*, etcetera, call for a major shift of perspective (see the article INTENSION).

There may be a good syntactic case for grouping ordinary adjectives with nouns to form complex nouns instead of with determiners to form complex determiners, but the above shows that there are no semantic objections to the complex determiner option. For noun phrases containing superlative adjectives, on the other hand, there seems to be strong semantic evidence for considering the superlative as part of the determiner. Under the current regime, *the nicest girl* is interpreted as the principal filter generated by the person **g** who happens to be the nicest girl in the given context. This suggests taking *the Adj-est* as a complex determiner, to be interpreted as the function mapping any property **P** to the set $\{X \mid \mathbf{p} \in X\}$, where $\mathbf{p} \in \mathbf{P}$ is the individual satisfying **Adj** to the highest degree. The semantic specification assumes that 'grades of fulfilment', say on a continuous scale from 0 to 1, are available for all gradable adjectives.

To illustrate that semantics provides more leeway than may appear at first sight, here is sketch of a different approach. The semantic effect of the superlative morphology on a gradable adjective *Adj* (for convenience we restrict attention to intersective adjectives) is to change a function **Adj** mapping properties **P** to properties $\mathbf{P} \cap \mathbf{Adj}$ into a function **Adj-est** mapping properties **P** to properties with an associated **Adj** ordering, in other words to pairs of the form given in (7).

(7) $\langle \mathbf{P} \cap \mathbf{Adj}, \leq_{\text{adj}} \rangle$.

The definite determiner *the* is now interpreted as an expression constructing principal filters out of such pairs; the **Adj** ordering serves as the context providing the definiteness required. This account has a natural extension to complex determiners such as *the five*, as in *the five nicest girls*. The noun phrase gets interpreted as the principal filter generated by the set of girls occupying the first five places in the $\leq_{\text{nice}}$ ordering. See the article COMPARATIVES for further details on the semantics of comparatives and superlatives.

6 Determiners and Mass Nouns

In the above, attention was limited to determiners combining with count nouns to form noun phrases interpreted in a domain of run-of-the-mill individual entities. Extension to the case of determiners for non-count nouns is relatively straighforward: the interpretation of non-count nouns calls for a domain containing quantities of continuous stuff, with a 'part of' relation $\sqsubseteq$. The formulations of **EXT**, **CONS** and **ISOM** now need patching to take the switch from $\subseteq$ to $\sqsubseteq$ into account; these details are left to the reader.

Some wine is interpreted as the set of all quantities of stuff having some wine in it (to express this formally one needs the $\sqsubseteq$ relation), *little milk* as the set of all quantities of stuff not containing more than a certain small amount of milk, etcetera. Note that in these cases the denotations of the determiners remain basically the same. See the article MASS EXPRESSIONS for further details.

7 Determiners Involving Measure Phrases

Both count nouns and mass nouns can be preceded by complex determiner expressions involving measurement to form measured noun phrases.

(8) *I bought two pounds of apples/cheese.*

Basically, *Numeral F of*, where *F* is a measure function word such as *kilogrammes*, *inches*, *years*, is a complex determiner which is interpreted as a function mapping properties **P** (e.g., sets of apples) or quantities of stuff **S** (e.g., amounts of cheese) to (a characteristic function of) a set of sets (9) or a set of amounts (10).

(9) $\{X \mid F(X \cap \mathbf{P}) = n\}$.

(10) $\{X \mid F(X \sqcap \mathbf{S}) = n\}$.

Here F is the function which interprets the measure expression, and it is assumed that F measures along the right dimension for the sets of individuals or amounts of stuff under consideration (cheese is measured in grammes, pounds or kilogrammes, fabric in centimetres, inches or yards, detention in months, or under harsher regimes in years, and so on). $\sqcap$ is the intersection operation for amounts of stuff. This operation can be defined in terms of $\sqsubseteq$, but we will not bother to give the formal details. In fact, to make the whole setup work smoothly in a uniform framework, $\subseteq$ and $\sqsubseteq$ need to be subsumed under one relation, but the details are outside the scope of the current article.

Observe that the determiner *two pounds of* is indefinite, for the noun phrases that it forms are interpreted as non-principal filters (modulo the obvious patch to the definition to replace 'adding individuals' by 'adding stuff'). A similar patch to the definition of **ISOM** shows that it is also a quantifier determiner. *Less than two inches of*, on the other hand, is a quantifier determiner but not an indefinite determiner. *These three litres of*, finally, is a definite determiner. See the article MEASURE PHRASES for further details.

8 Generic Uses of Determiners

The English indefinite determiner *a(n)* has a special generic use, exemplified in (11), which has not been covered yet.

(11) *What this country needs is a good 5-cent cigar.*

It seems clear that whatever semantic treatment of generics one chooses, the same mechanism should be brought to bear on generic uses of the bare plural.

(12) *Good 5-cent cigars are hard to come by these days.*

One way of handling generics is to consider generic indefinites and generic bare plurals as *names* of members of a special class of generic entities. The assumption is that there is some entity **c** which is the generic good 5-cent cigar, and *a good 5-cent cigar* and *good 5-cent cigars* (in their generic uses) simply refer to that entity.

This seems straightforward enough, but there are some logical complications. If good 5-cent cigars form a natural class, what about bad 5-cent cigars? Or cigars costing less than five cents? These might form natural classes also, but note that if every subset of a given domain of regular individuals has a corresponding natural class with an associated generic individual, then there will be more generic individuals than regular individuals.

Next, how do generic individuals relate to regular individuals? First observe that they can share certain properties with regular individuals. Properties that the generic good 5-cent cigar must certainly have: costing only five cents, being good (according to the generic cigar smoker's standard, say), being approved of by at least one president of the United States (example (11) is a presidential quote), and so on. But there are also properties that no generic individual can have. To mention an example, the generic good 5-cents cigar does not share the property of regular good 5-cent cigars of exemplifying the generic good 5-cent cigar. If the example seems far-fetched, observe that it will not do to disregard the relation of 'exemplification' between regular individuals and generic individuals altogether, because natural language allows anaphoric linking between regular and generic uses of noun phrases.

(13) *They must sell a good 5-cent cigar in the US,*
for the president himself is smoking them.

In (13) the pronoun *them* refers to individual cigars, buth this pronoun is linked to a generic use of *a good 5-cent cigar* in the same sentence.

It should be clear that the above remarks only scratch the surface of the semantics of generic expressions. See the articles GENERICS and PLURALS for further information.

9 The Dynamics of Noun Phrase Interpretation

If one looks at noun phrase behaviour from a dynamic perspective, some noun phrases can be seen to prepare the ground for anaphoric links in a way that other noun phrases do not. In the following examples, the indices serve to indicate intended anaphoric links. Superscripts are used for antecedents, subscripts for anaphors.

(14) *Some^i man whistled. He$_i$ was happy.*

(15) *The^j man whistled. He$_j$ was happy.*

(16) *No^k man whistled. *He$_k$ was happy.*

Roughly, definite and indefinite noun phrases admit anaphora outside their scopes, other noun phrases (quantified noun phrases which are not also indefinites) do not. This is only an approximation because certain sentential operators—negation is an example—block anaphoric linking. A dynamic perspective borrowed from the semantics of imperative programming languages can account for the varieties of unbounded anaphoric behaviour.

The reason, by the way, for placing the antecedent indices on the determiners instead of the noun phrases they form is that noun phrases can have internal anaphors, i.e., noun phrases may contain pronouns anaphorically linked to their main determiner, as in (17).

(17) *Every^i man who thinks he$_i$ is a genius is conceited.*

Dynamic logic views the meaning of program statements as relations between machine states holding before the program statement was executed and machine states holding after the execution of the statement. Applying this perspective to natural language, the meaning of a noun phrase *[Det^i N]* (a determiner *Det* with index *i* and a noun *N*, the index being a device for indicating intended anaphoric links with the noun phrase as antecedent), given an appropriate verb phrase argument *VP*, is a relation between value assignments to pronouns holding before the processing of *[Det^i N] VP]* and value assignments to pronouns holding after the processing. The assignments that can hold before the processing of a sentence form the *input assignment set*, the assignments that can hold afterwards the *output assignment set*. Meaning is a relation between an input assignment set and an output assignment set.

Processing starts with an assignment set containing just the assignment mapping deictic pronouns to contextually given things, the *initial input assignment set*. A sentence or sequence of sentences is true if at the end of processing the set of output assignments is not empty.

To focus on an example: the interpretation for the noun phrase $some^i$ A will first take an appropriate second argument B and then relate a set of assignments G which do not have values for pro_i (the pronoun with index i) to a set of assignments G', with every g' in G' just like some g in G, except for the fact that pro_i now gets a value. Every member of G' maps the pronoun pro_i to an object d in $A \cap B$. The dynamism here reflects the fact that no pronouns preceding $some^i$ A can be anaphorically linked to the quantifier, while pronouns following the determiner plus its two arguments can be so linked by interpreting them as objects in the intersection of the noun interpretation and the verb phrase interpretation.

By contrast, the interpretation for the noun phrase no^i A will relate a set of assignments G which do not have values for pro_i to the same set G, provided the intersection $A \cap B$ is empty, where B is the interpretation of the second argument of the determiner, and to the empty set of assignments otherwise. This reflects the fact that neither pronouns preceding the noun phrase nor pronouns following it (and outside its scope) can be anaphorically linked to it.

Of course, the above account still leaves many semantic details unspecified; it does not, as it stands, do justice to the anaphoric possibilities inside an antecedent noun phrase and within its scope, and it also ignores the syntactic agreement constraints between noun phrase antecedents and pronouns. Nevertheless, if it has managed to illustrate the dynamic principle it has served its purpose. The article DYNAMIC INTERPRETATION provides further information.

References

[1] Carlson G 1982 'Generic Terms and Generic Sentences', *Journal of Philosophical Logic*, 11, 163–182.

[2] Van Benthem J 1983 'Determiners and Logic', *Linguistics and Philosophy* 6, 447–478.

[3] Zwarts F 1983 'Determiners: A Relational Perspective', in A. ter Meulen (ed.), *Studies in Model Theoretic Semantics*, Foris, Dordrecht, 37–62.