From logical and linguistic generics to Hilbert’s tau and epsilon quantifiers
Stergios Chatzikyriakidis, Fabio Pasquali, Christian Retoré

To cite this version:

HAL Id: hal-01803717
https://hal.archives-ouvertes.fr/hal-01803717
Submitted on 12 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
From logical and linguistic generics to Hilbert’s tau and epsilon quantifiers

Stergios Chatzikyriakidis
CLASP, University of Gothenburg; Open University of Cyprus
stergios.chatzikyriakidis@gu.se

Fabio Pasquali
IRIF-CNRS, Université Paris Diderot
pasquali@dima.unige.it

Christian Retoré
LIRMM-CNRS, Université de Montpellier
christian.retore@lirmm.fr

Abstract

With our starting point being (universal) generics appearing in both natural language and mathematical proofs, and were further conceptualized in philosophy of language, we introduce the tau subnector that maps a formula \( F \) to an individual term \( \tau_x F \) such that \( F(\tau_x F) \) whenever \( \forall x F \). We then introduce the dual subnector \( \epsilon_x F \) which expresses the existential quantification \( F(\epsilon_x F) \equiv \exists x F \), and describe its use w.r.t the semantics of indefinite and definite noun phrases. Some logical and linguistic properties of this intriguing way to express quantification are discussed — but the reader is referred to the article by Abrusci in this volume for the impact of epsilon on Hilbert’s work as regards the logical foundations of mathematics.

Keywords: proof theory; quantifiers; generic objects; philosophy of language; formal linguistics

Thank the referees (or not), and anyone else for reasons relating to the paper as a whole, by adding \titlethanks{text} to the preamble (or \titlethanks{} to delete this reminder).

For acknowledgements specific to the author (e.g. thanking a particular author’s funding body), use \thanks{text} after the author name in the \addauthor command (or do not add the command at all for no such addition).

*The author gratefully acknowledges support from the Centre of Linguistic Theory and Studies in Probability (CLASP) in Gothenburg as well as the ANR project Polymnie in France.
1 Presentation

This introduction to the volume presents the epsilon and tau quantifiers that Hilbert came up with in the beginning of the XXth century. Although an introduction to these issues exists on the web, e.g. [21, 40], we believe that the current introduction, presenting a slightly different viewpoint, might be of interest to a number of researchers, especially because most of the literature on the topic [22] is in German, Russian and French — Ackermann’s seminal contribution [2] is also in German. We say that tau and epsilon are quantifiers because one can express quantification with them, but they map a formula to an individual term, Following Curry [11], these

---

1 The second one was written by Hartley Slater, who is also one of the authors in this volume. Hartley Slater unfortunately passed away before his paper was published. The reader will find an obituary in this volume as well.
quantifiers are called *subnectors* because each of them maps a formula $F$ into an individual term $\epsilon_x F$ and $\tau_x F$, in which $x$ is bound. $\epsilon_x F$ and $\tau_x F$ denote the existential and the universal generic objects w.r.t. a formula $F$ respectively. Basically $F(\epsilon_x F)$ means $\exists x.F$ and $F(\tau_x F)$ means $\forall x.F$ and because they belong to a classical, and not an intuitionistic, setting, one can be defined from the other: $\tau_x F = \epsilon_x \neg F$. The addition of epsilon and tau to the connectives may seem harmless, but it actually completely changes the logic, i.e. formulae of the epsilon calculus have no counterpart in first order or even higher order predicate calculus. In particular, because of *over binding* (also known as *in situ* binding), those quantifiers are closer to the syntactic behavior of quantifiers in natural language sentences. This is the reason why the epsilon subnectors have been used in the philosophy of language and formal linguistics to model quantifiers of natural language. [15, 16, 14, 42, 45, 46, 41, 38] Symmetrically, the subector tau could have been used to model universal quantification but it was not.

2 From the Ancient and Medieval view of universal quantification to tau

A long debated question in logic and metaphysics in the Ancient and the Medieval world (starting with Plato, Aristotle and all the way to Porphyry and the scholastics) is the relation between a universal or generic "dog" and the set of individuals "dogs". This is known as the *problem of universals*. [29, 24, 3, 13]

What is a concept like "horse"?

- a substance, that exists independently of the individuals falling under this concept (realism)
- a name without reality, i.e. a word that stands for the class of all individuals falling under the concept (nominalism)
- a concept, that is a mental construction having an empirical relation to the set of individuals (conceptualism)

In order to illustrate the debate between Abélard and Roscelin (see e.g. [13]), regarding the relation between the concept and the entities that fall under this concept, one could say: *If an illness causes the extinction of all tall dogs, would your concept of dog be altered?* Some have defended the position that each dog is as constitutive of the concept of dog as a wall or a roof is constitutive of the concept of a house.
The relation between concept and universal quantification is that, in case the generic object in a concept $P$ enjoys the property $P$, so do all $P$. This is of course also related to Aristotle’s proof rule of abstraction (alternatively called generalisation): if an integer enjoys a property $P$, then all integers enjoy $P$. Ancient and Medieval logic was not dealing with models. These only appeared with Frege at the end of the XIXth century. Philosophers ranging from Aristotle to scholastic logicians were mainly concerned with proofs and rules. Aristotle who was aiming at extracting the logical principles underlying mathematical reasoning, introduced the generalization rule for universal quantification (that he called abstraction).

Using this rule, to establish $F(x)$ for all elements $x$ of a class $C$ one proceeds as follows:

**Generalisation (a.k.a Abstraction)** Let $x$ be any element of $C$. A reasoning shows that $P(x)$ holds. The generalization rule asserts that the property holds for any element in $C$ — indeed $x$ does not possess anything special apart from being in $C$.

This is quite different from another technique used to establish $F(x)$ for all elements $x$ of a known, finite class $C$ which goes without saying (and that perfectly matches universal quantification in a given model of (many-sorted) first order logic, but.

**Conjunction** We can prove $P(c_1)$, then $P(c_2)$, $P(c_3)$, $P(c_4)$ and so on. Once we do so for all elements in $C$, we can form a conjunction out of all these formulae. This latter proof of a universally quantified statement is in fact, in modern terms, a reduction of universal quantification to conjunction: $\forall x \in C F(x) \equiv \land_{x \in C} F(x) \equiv P(c_1) \& P(c_2) \& P(c_3) \& P(c_4) \& \cdots$

This dual nature of universal quantification can be observed in the various wordings of universal quantifications, that some words rather refer to each individual in the collection (each), while some others rather refer to a (fictive) prototypical or average individual (any):

1. a. Each dog has four legs.
   b. All dogs have four legs.
   c. Every has four legs.
   d. A dog has four legs.
   e. Any dog has four legs.
f. Dogs have four legs.

The distributive reading, obliged when using *each*, cannot accept exceptions, may express a coincidence of properties that can be conjuncted, and it is not required that there is a reason (other than probability) to this coincidence and the domain can be complicated. The only good way to refute such a statement is to provide a counter example, i.e. an individual for which the property does not hold (one component of the conjunction fails).

(2) a. Each bird with both black and white feathers flies.
   b. Not this wound bird. (perfect)
   c. Not autruches. (not good refutation, since the relation between those two sets is not obvious)

Generic entities rather correspond to ideal and prototypical entities whose properties are derived by reasoning. Compared to noun phrases introduced by *each* can accept exceptions, and their domain cannot be complicated. The refutation of a sentence involving a generic is usually performed by another generic, or by a reasoning, i.e. one can go on at abstract level.

(3) a. Birds fly.
   b. Not this wound bird. (not a refutation, since generic readings admit exceptions)
   c. Not autruches. (perfect refutation)

The usual rule called generalisation or abstraction, is abbreviated by ∀i. i.e. ∀ introduction because it introduces the ∀ quantifier. This rule says that when a property has been established for an x which does not enjoy any particular property (i.e. is not free in any hypothesis), one can conclude that the property holds for all individuals.

In this paper, we use sequents: $H_1, \ldots, H_n \vdash C$ simply means that under hypotheses $H_1, \ldots, H_n$, conclusion $C$ holds. The ∀i (or generalisation) rule below simply means that from (1) $P(x)$ holds under hypotheses $H_1, \ldots, H_n$ without any free $x$ in any of the $H_i$ one may deduce that (2) $\forall x. P(x)$ holds under the hypotheses $H_1, \ldots, H_n$.

$$
\frac{H_1, \ldots, H_n \vdash P(x)}{H_1, \ldots, H_n \vdash \forall x. P(x)} \quad \forall_i \quad \text{when there is no free occurrence of } x \text{ in any } H_i
$$

2As we shall discuss later on, such a rule is enough to derive that all A are B, from the fact that B holds for an x satisfying A who has no specific property other than A.
The rule above can be formulated with a generic element, \( \tau_x P(x) \), a virtual element that has no specific relation to the property \( P \). If you think of \( P \) as being "to drink", \( \tau_x \text{drink}(x) \) is the most sober individual you can think about: \( \tau_x \text{drink}(x) \) drinks if and only if everyone else does:

\[
P(\tau_x P(x)) \text{ iff } \forall x. P(x)
\]

From this one easily defines the rules for quantification using the generic element w.r.t. \( P \), they simply are the usual rules for quantification:

\[
\frac{H_1, \ldots, H_n \vdash P(x)}{H_1, \ldots, H_n \vdash P(\tau_x P(x))} \tau_i - \text{when there is no free occurrence of } x \text{ in any } H_i
\]

There is another rule for quantifier, easier because without any restriction, called the specialisation or instantiation rule, which says that of something holds for any \( x \) then it holds for any constant or term.\(^3\)

\[
\frac{H_1, \ldots, H_n \vdash \forall x. P(x)}{H_1, \ldots, H_n \vdash P(a)} \forall_e
\]

This can be formulated with the same generic element \( \tau \) a subnector i.e. an operator that builds a term (of type individual) from a formula (Curry’s terminology).

\[
\frac{H_1, \ldots, H_n \vdash P(\tau_x P(x))}{H_1, \ldots, H_n \vdash P(a)} \tau_e
\]

If \( \tau_x P(x) \) enjoys the property \( P(_) \) then any individual does, and vice-versa. So \( \tau_x P(x) \) is an ideal entity which is absolutely independent from the property \( P(_) \) — that’s the reason why when it enjoys \( P(_) \) everything does.

3 Existential generics: from Russell’s iota to Hilbert’s epsilon

As opposed to the generic \( \tau_x P(x) \) that enjoys \( P \) if and only if every entity does, there is also "this P", "the P", i.e. the unique individual satisfying \( P \), if there is exactly one such individual, denoted as \( \iota_x P(x) \). Russell introduced \( \iota \) in [39] for

\[^3\]This rule can derive its relativised version that says that if all \( A \) are \( B \) then any particular \( A \) is \( B \).
definite descriptions (definite noun phrases like *the queen of England*). It is also, like \( \tau \), a subnector, given that it turns a formula into some individual term. It is the ancestor of Hilbert’s \( \varepsilon \).

However, as argued by von Heusinger, there is little difference between the logical form of definite descriptions and indefinite noun phrases. This is because uniqueness of the noun phrase with a definite article is not always observed:


(5) Taken in while very young by monks of the abbey of Reichenau on the island of the Constance lake, that fully took care of him, Hermann studied and became one of the most erudite monk of the XI\(^{th}\) century.

This example comes from a site dealing with first names and is really a pleasant coincidence\(^4\) indeed in the original paper by Egli and von Heusinger (1995, hence before the example) takes as a fictive example the case of the islands of the Constance lake to exemplify that someone seeing just one of the three islands of this lake could utter "the island of the Constance lake" while there are three of them.

Given that \( \iota \) is not expected to have good logical properties — its negation is "zero or more than two entities do not enjoy the property \( P(_{\_}) \), Hilbert introduced the existential \( \varepsilon \) subnector, building the existential generic \( \varepsilon_x F \):

\[
F(\varepsilon_x F) \equiv \exists x. \; F
\]

A term (of type individual) \( \varepsilon_x F \) associated with \( F \), usually contains occurrences of \( x \): as soon as an entity enjoys \( F \) the term \( \varepsilon_x F \) enjoys \( F(_{\_}) \). The operator \( \varepsilon \) binds in \( \varepsilon_x F \) the free occurrences of \( x \) in \( F \).

We do not give the introduction and elimination rules natural deduction style as above, because the elimination of the existential quantifier in natural deduction would deserve some explanation. Let us just say that they simply mimic the introduction and elimination rules for existential quantification.

If \( \tau_xdrink(x) \) is the most sober individual, \( \varepsilon_xdrink(x) \) is a soak: he drinks if and only if someone drinks.

### 4 Syntax of epsilon and tau first order calculus

Terms and formulae are defined by mutual recursion:

\(^4\)(http://www.prenoms.com/prenom/signification-prenom-HERMANT.html)
• Any constant in \( \mathcal{L} \) is a term.
• Any variable in \( \mathcal{L} \) is a term.
• \( f(t_1, \ldots, t_p) \) is a term provided each \( t_i \) is a term and \( f \) is a function symbol of \( \mathcal{L} \) of arity \( p \)
• \( \epsilon_x A \) is a term if \( A \) is a formula and \( x \) a variable — any free occurrence of \( x \) in \( A \) is bound by \( \epsilon_x \)
• \( \tau_x A \) is a term if \( A \) is a formula and \( x \) a variable — any free occurrence of \( x \) in \( A \) is bound by \( \tau_x \)
• \( s = t \) is a formula whenever \( s \) and \( t \) are terms.
• \( R(t_1, \ldots, t_n) \) is a formula provided each \( t_i \) is a term and \( R \) is a relation symbol of \( \mathcal{L} \) of arity \( n \)
• \( A \& B, A \lor B, A \Rightarrow B, \neg A \) when \( A \) and \( B \) are formulae.

There is no objection to simultaneously use the usual quantifiers \( \forall \) and \( \exists \). This superimposition is quite useful to have, since we can show that the epsilon/tau calculus restricted to formulas that are equivalent to usual first order logic, is a conservative extension of first order logic (first and second epsilon theorems, see e.g. Abrusci’s paper in the same issue). This means that we can have quantified formulas as well:

• \( \exists_x A \) is a formula if \( A \) is a formula and \( x \) a variable — any free occurrence of \( x \) in \( A \) is bound by \( \exists x \)
• \( \forall_x A \) is a formula if \( A \) is a formula and \( x \) a variable — any free occurrence of \( x \) in \( A \) is bound by \( \forall x \)

Quantification rules are the ones we already discussed above.

\( \tau \) rules

**introduction** The introduction rule of the tau universal quantifier is Aristotle’s rule of abstraction (also known as generalisation): from \( P(x) \) with \( x \) generic (i.e. not present in any hypothesis) infer \( P(\tau_x P(x)) \).

\[
\frac{H_1, \ldots, H_n \vdash P(x)}{H_1, \ldots, H_n \vdash P(\tau_x P(x))} \quad \tau_i \text{ — when there is no free occurrence of } x \text{ in any } H_i
\]
Hilbert’s epsilon and tau

The elimination rule of universal quantification (also known as instantiation or specialisation) is as usual: from \( P(\tau_x.P(x)) \) one may infer \( P(t) \) for any \( t \).

\[
\frac{H_1, \ldots, H_n \vdash P(\tau_x.P(x))}{H_1, \ldots, H_n \vdash P(t)} \quad \tau_e
\]

\( \epsilon \) rules

**introduction** The introduction rule of the epsilon existential quantifier is the usual one: from \( P(t) \) where \( t \) is any term, infer \( P(\epsilon_x.P(x)) \equiv \exists x \ P(x) \).

\[
\frac{H_1, \ldots, H_n \vdash P(t)}{H_1, \ldots, H_n \vdash P(\epsilon_x.P(x))} \quad \epsilon_e
\]

**elimination** The elimination of the epsilon universal quantifier is more tricky, as it is in natural deduction for first order logic format: assume that from \( P(x) \) and other hypotheses \( \Gamma \) not involving \( x \) a conclusion \( C \) without \( x \) is derivable and that \( P(\epsilon_x.P(x)) \) holds: then \( C \) holds under the hypotheses \( \Gamma \).

\[
\frac{K_1, \ldots, K_p \vdash P(\epsilon_x.P(x)) \quad H_1, \ldots, H_n, P(x) \vdash C}{K_1, \ldots, K_p, H_1, \ldots, H_n \vdash C} \quad \epsilon_e - \text{when there is no free occurrence of} \quad t
\]

From the rules for quantification, it is clear that epsilon/tau calculus is classical (as opposed to intuitionistic), because the following are easily derived:

\[
P(\epsilon_x.P(x)) \equiv \exists x P(x) \equiv \neg \forall x \neg P \equiv \neg \neg P(\tau_x.\neg P(x))
\]

\[
P(\tau_x.P(x)) \equiv \forall x P(x) \equiv \neg \exists x \neg P \equiv \neg \neg P(\epsilon_x.\neg P(x))
\]

Hence: \( \tau_x P(x) = \epsilon \neg P(x) \) and \( \epsilon_x P(x) = \tau \neg P(x) \)

Therefore one of the two subnectors/quantifiers \( \epsilon \) and \( \tau \) is enough, but most people have chosen \( \epsilon \), as in Bourbaki’s book on set theory for example.

The quantifier free epsilon calculus is a strict conservative extension of first order logic.

- Strict: there are formulas that are not equivalent to any formula of first order logic, e.g. \( P(\epsilon_x.Q(x)) \) with \( P, Q \) being distinct unary predicate symbols: if \( P \) and \( Q \) are predicates letters with no relation, this bound formula has no equivalent in first order logic.
• Conservative: With regards to first order formulas, the epsilon calculus derives the same formulas as first order logic, i.e. classical predicate calculus.

As by Claus-Peter Wirth in this issue, any first order formula can be turned into an equivalent epsilon formula. However, one should have in mind that such a translation yields formulas that are rather complex and hard to parse. For instance $\forall x\exists y P(x,y)$ can be written has: $\exists y P(\tau_x P(x,y),y)$ which itself can be written as: $P(\tau_x P(x,\epsilon_y P(\tau_x P(x,y),y)),\epsilon_y P(\tau_x P(x,y),y))$!

As mentioned in the article on epsilon and proof theory by Michele Abrusci in this issue, a major motivation of Hilbert’s, was to establish the consistency of arithmetic by elementary means, before Gödel’s incompleteness theorem. Following this objective that could not be met because of Gödel’s results, Hilbert nevertheless obtained the elimination of $\epsilon$ (1st & 2nd $\epsilon$-theorems), yielding the first correct proof of Herbrand’s theorem. These issues are also addressed in Abrusci’s paper.

The reader must be aware that epsilon calculus is very different from first order logic. As we said, many formulas of the epsilon calculus are not equivalent to first order formula. A number of published results on the epsilon calculus are known to be wrong, like cut-elimination and models among other things (see [41, 27] as explained in [7, 31]). As regards the first point, the problem is whether the complexity of the formulas in epsilon terms ought to be taken into account, and as regards the second point, the problem is that if there are sound and complete models for the whole epsilon calculus, they must substantially differ from usual models. However, all proposals so far try to provide usual models for a very unusual calculus. Nevertheless, intuitively, epsilon terms might be viewed as a kind of Henkin witnesses, or as choice functions that act simultaneously on all formulas, without any obligation to extend the first order language.

The relation between epsilon and choice functions introduced by Skolemisation is complicated. Choice functions are usually introduced when one considers a single formula (or a clause) in prenex form and introduces a function in the language, whose arguments are the universal variables, for each of the existential quantifiers. Epsilon introduces all choice functions at once but without reference to a particular formula (except the one defining the epsilon term), nor to universal variables, and the first order language (relations, functions) does not need to be extended. Regarding their syntax, there are no specific deduction rules for choice functions: the existential quantifiers are taken into account when interpreting the function — on the syntactic side they do not belong to first order logic.

---

5 We would like to thank Michele Abrusci, Mathias Baaz and Ulrich Kohlenbach for pointing out these problems to us.
5 Epsilon (and tau) in linguistics

5.1 Some critics of the usual treatment of natural-language quantification

Quantifiers, especially existential ones, are quite common in natural language. There are two sorts of quantifiers w.r.t natural language syntax. The first sort only needs the main predicate (something, someone, everyone, everything,...). The second sort (a, some, every, all,...) first applies to a common noun (a predicate) before being applied to the main predicate and all of its arguments but one — the verb plus its subject and complement but one.

(6) Something happened to me yesterday.
(7) A man comes on to tell me how white my shirts can be.
(8) Some girls give me money.
(9) Keith played a Beatles song.

The usual treatment of quantifiers initiated by Montague \[33\] takes place into the framework defined by Church \[9\] for writing first or higher order formulas of predicate calculi in a way that follows Frege’s compositionality principle.

This treatment usually assumes a base type for propositions \(t\) (\(o\) in Church writings) and another base type \(e\) (\(i\) in Church writings) for entities, individuals — usually there is just one base type for all individuals, the logic is single sorted as opposed to many sorted logic of, e.g., \[26\] used for semantics in \[37\].

In order to express linguistic semantics one needs logical constants for connectives and quantifiers:

- \(~\) of type \(t \rightarrow t\) (negation)
- \(\supset, \&, +\) of type \(t \rightarrow (t \rightarrow t)\) (implication, conjunction, disjunction)
- two constants \(\forall\) and \(\exists\) of type \((e \rightarrow t) \rightarrow t\)

Specific constants are needed to represent a first order language:

- \(R\) of type \(e \rightarrow (e \rightarrow (.... \rightarrow e \rightarrow t))\) (n-ary predicate — \(n\) times \(e\))
- \(f\) of type \(e \rightarrow (e \rightarrow (.... \rightarrow e \rightarrow e))\) (n-ary function symbol — \(n\) times \(e\))
As said above ∀ and ∃ are constants of type (e → t) → t to quantifiers that simply apply to the main predicate — they can interpret natural language quantifiers without a domain. This is the case with *everything*, *something* but other natural language quantifiers that apply to two predicates like *some*, *every* need more complex terms:

- **existential quantifier** (*some, a*):
  \[
  \lambda P^{e\to t} \lambda Q^{e\to t} (\exists x^t. (P x) (Q x)) : (e \to t) \to (e \to t) \to t
  \]

- **universal quantifier** (*every, all*):
  \[
  \lambda P^{e\to t} \lambda Q^{e\to t} (\forall x^t. (P x) (Q x)) : (e \to t) \to (e \to t) \to t
  \]

This modelling suffers from some inadequacies mainly due to a difference in the syntactic struture of the sentence and its logical form.

The standard analysis of the sentence 9 is: \((a (\text{Beatles song})) (\lambda z. \text{Keith played } z)\)

So:

\[
\underbrace{(\text{existential quantifier})}_{(\lambda P^{e\to t} \lambda Q^{e\to t} (\exists x^e. (P x) (Q x)))}
\]

is applied to

\[
\underbrace{(\text{being a Beatles song})}_{(\lambda u^e. \text{wrote}(\text{Beatles}, u) \& \text{song}(u))}
\]

and to

\[
\underbrace{(\text{being played by Keith})}_{(\lambda z. \text{Keith played } z)}
\]

As expected, these lambda terms reduce to

\[
(\exists x^e. \text{wrote}(\text{Beatles}, x) \& \text{song}(x) \& (\text{Keith played } x))
\]

i.e. to the expected meaning:

\[
(\exists x. \text{wrote}(\text{Beatles}, x) \& \text{song}(x) \& (\text{Keith played } x))
\]
5.1.1 Difference between the syntactic and semantic structures

The classical treatment of natural language quantifiers infringes the correspondence between syntax and semantics, i.e. the heart of compositionality. Indeed, observing a sentence, its syntactic structure, and its semantic structure:

(10) a. Keith played a Beatles song.
    b. syntax (Keith (played (a (Beatles song))))
    c. semantics: (a (Beatles song)) (λx. Keith played x)

it is clear that the underlined predicate does not correspond to any proper phrase (or subtree) of the sentence. A natural language quantifier is an in situ binder (as e.g. wh-interrogatives in s Chinese): the quantified noun phrase can also be nested deeply in the sentence parse tree, may apply to the whole parse tree.

5.1.2 Asymmetry between the domain of quantification and the main predicate

It is easily observed in the following examples that, as opposed to the usual logical formulas representing meaning, one cannot swap the two predicates, for instance in Aristotle I sentences. However, even when these can be swapped, as in the last example, the meaning is not the same, because the focus is different, and depending on the context (in a university or in a company) only one of the two can be said, the other one begin unnatural.

(11) a. Some politicians are crooks.
    b. ?? Some crooks are politicians.
    (12) a. Some students are employees.
        b. Some employees are students.

5.1.3 Semantic nature of the quantified noun phrase

According to the usual treatment, a quantified noun phrase is a function that maps a predicate to a proposition. However, intuitively, the type of a quantified noun phrase, and especially the type of an existentially quantified noun phrase should rather be an individual. This is confirmed by the cognitive process: when "an A" is uttered, one actually imagines such an "A", so a quantified noun phrase may have a reference as an individual before uttering the main predicate (if any):

(13) Cars, cars, cars,... (Blog)
    (14) What a thrill — My thumb instead of an onion. (S. Plath)
5.2 Using epsilon for indefinite noun phrases

Coming back to the linguistic motivation, some scientists have been modelling existential quantification (rather than universal quantification) by means of epsilon, since the pioneering work of von Heusinger and Egli [15, 16] that has been further pursued in a number of papers [43, 44, 45, 46]. The leitmotiv in these papers, is to model existentially quantified noun phrases like "an $A$ is $B$" (Aristotle's I sentences) by $B(\epsilon x. A(x))$. Nothing in a formula like this says that $\epsilon x. A(x)$ has the property $A$, i.e. that there are some $As$. Thus, the presupposition $A(\epsilon x. A(x))$ could (or should) be added.

Here we should make an important remark on those sentences In general, for instance if $B$ and $A$ have $B(\epsilon x. A(x))$ is not equivalent to any ordinary formula. In particular it is not equivalent to $\exists x. (B(x) \land A(x))$, but the two formulas are related as follows:

**In general $B(\epsilon x. A(x))$ does not entail $\exists x. (B(x) \land A(x))$** Indeed, it is possible that $\exists x. (B(x) \land A(x))$ is false while $B(\epsilon x. A(x))$ is true. Indeed, let $B(x)$ be ($x = x$) and let $A(x)$ be ($x \neq x$) i.e. $\neg B(x)$. Then $B(\epsilon x. A(x)) \equiv B(\epsilon x. \neg B(x)) \equiv B(\tau x. B(x)) \equiv \forall x. B(x) \equiv \forall x. x = x$ which is clearly true. But $\exists x. (B(x) \land A(x)) \equiv \exists x. (B(x) \land \neg B(x)) \equiv \exists x. (x = x \land x \neq x)$ which is clearly false. The argument works with any formula of one variable that is universally true like here $B(x) \equiv (x = x)$.

\[ B(\epsilon x. A(x)) \land A(\epsilon x. A(x)) \text{ entails } \exists x. B \land A(x) \text{ Indeed } B(\epsilon x. A(x)) \land A(\epsilon x. A(x)) \text{ entails } B(\epsilon x. B \land A(x)) \land A(\epsilon x. B \land A(x)) \text{ that is } B \land A(\epsilon x. (B \land A(x))) \text{ which means } \exists x. B \land A(x). \]

\[ \exists x. A(x) \land \forall y (A(y) \Rightarrow B(y)) \text{ entails } B(\epsilon x. A(x)) \text{ Indeed, } \epsilon\text{-terms are usual terms, a universal quantifier can be instantiated to an epsilon term.} \]

In accordance with the small difference between definite and indefinite descriptions, von Heusinger proposes to interpret both “a” (introducing indefinite noun phrases) and “the” (introducing definite description) by an epsilon term. Only the interpretation differentiates them: the “a” always refers to a new individual in the class, while “the” refers to the most salient one in the context. This context dependent interpretation lead to the indexed epsilon calculus [32] and further studied by Hans Leiß in this volume.

(15) A student entered the lecture hall. He sat down. A student another student left the lecture hall.
A student arrived lately. The professor looked upset. The student the same student left.

The epsilon modeling avoids all the three aforementioned drawbacks of the standard interpretation, because $\epsilon_x F(x)$ is an individual term (as natural language quantifiers) which allows in situ binding (as natural language quantifiers).

5.1.3 A quantified noun phrase can be interpreted as an individual without the main predicate: indeed $\epsilon_x F(x)$ is an individual term.

5.1.1 The semantics $P(\epsilon_x F(x))$ of an existential sentence e.g. of an I sentence "some F is P." follows the syntactical structure: $\epsilon_x F(x)$ is an individual term, which is the usual semantics of a noun phrase, that is inserted into the main predicate $P$ to obtain the sentence semantics — which is comparable to there exists some $x$ satisfying $F(x)$ such that $P(x)$.

5.1.2 The asymmetry between subject and predicate is restored $P(\epsilon Q) \neq Q(\epsilon P)$.

The interpretation of noun phrases with $\epsilon$ solves the so-called E-type pronouns interpretation of Gareth Evans [17] where the semantic of the pronoun is the copy of the semantic of its antecedent:

(17) A man came in. He sat down.

(18) [$He$] = [$A$ man] = ($\epsilon_x$ Man($x$)).

Slater pointed out WHERE? AT LEAST DURING THE MEETING POSSIBLY IN HIS PAPER a possible problem with the modelling of indefinite noun phrases by epsilon terms. The two sentences A man enters. A man left. are respectively modeled as entered($\epsilon_x$ man($x$)) and left($\epsilon_x$ man($x$)) and if you consider them simultaneously, you may infer that entered&left($\epsilon_x$ man($x$)) and consequently entered&left($\epsilon_x$ entered&left($x$)) so $\exists x.$ entered($x$)&left($x$)! The problem is that nothing says that the two propositions should be conjoined by a usual & and so nothing tells us that the epsilon term that appears in both propositions should be interpreted in the same way twice.

So Slater proposes to interpret a man entered by entered($\epsilon_x$ man($x$)&entered($x$)) a man left by left(man($x$)&left($x$)). This avoids the unpleasant consequence above. But if you allow yourself to interpret in the same way the epsilon terms that appear in different propositions, then with Slater’s approach you end up with another problem. Consider: The professor presented first order logic. A student left. He introduced sequent calculus. A student left. Then, the second and fourth sentences yields exactly the same semantic representation, left($\epsilon_x$ student($x$)&left($x$)), we
have just at least one student that left, while we should have at least two. So the semantic representation if just one step, and there after one has to be careful with the interpretation of the epsilon terms.

This is the reason why Mints, in [32] introduced and studied indexed epsilon calculus, where the interpretation is relative to a given context, and contexts are indices for the calculus (see also the paper by Hans Leiß in this volume).

5.3 Tau and universally quantified noun phrases

Coming back to the subnector tau discussed at the beginning of the paper, it is fairly natural and coherent with the treatment of existential quantification to model universally quantified noun phrases introduced by "all", "every" and bare plurals by tau terms, as follows:

\[
\begin{align*}
\text{(19) a. Every man is mortal.} \\
\text{b. mortal}(\tau_x \text{man}(x))
\end{align*}
\]

This proposal has been explored in [?]. It enables a distinction between each and every the first one being a conjunction $\&_{x \in D} P(x)$ over a known domain $D$ and the later one being the usual universal quantifier with the generic object formalised as $\tau_x M(x)$. The analysis is a proof theoretical one, where statements are interpreted as the set of their possible justification as in [1].

6 Aristotle square of opposition revisited with epsilon and tau

With such an interpretation of quantified noun phrases, existential noun phrases as epsilon terms and universals noun phrase as tau terms, one may reformulate Aristotle’s I A E O sentences and revisit Aristotle’s square of opposition.

<table>
<thead>
<tr>
<th></th>
<th>Every S is P</th>
<th>Some S is P</th>
<th>No S is P</th>
<th>Some S are not P</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$P(\tau_x S(x))$</td>
<td>$P(\epsilon_x S(x))$</td>
<td>$\neg P(\epsilon_x S(x))$</td>
<td>$\neg P(\tau_x S(x))$</td>
</tr>
<tr>
<td>I</td>
<td>Universal Affirmative</td>
<td>Particular Affirmative</td>
<td>Universal Negative</td>
<td>Particular negative</td>
</tr>
</tbody>
</table>

Aristotle showed that those formulas define a square of opposition, i.e.: $A \equiv \neg O$, $E \equiv \neg I$ $A$ and $E$ are contradictory (they cannot both hold), $I$ and $O$ cannot both fail and are thus subcontraries, $A$ entails $I$, which is said to be a subaltern of $A$, and $E$ entails $O$ which is said to be a subaltern of $E$.

Using standard diagrams, in the Hilbert’s $\epsilon$-calculus, at least one of the following two squares is a square of opposition provided $P$ is bivalent with respect to $S$ this
condition being defined as $P(\varepsilon_x S(x)) \vdash P(\tau_x S(x))$ or $P(\tau_x S(x)) \vdash P(\varepsilon S(x))$ (see [35] for more details.

\[
\begin{array}{ccc}
P(\tau_x S(x)) & \rightarrow & \neg P(\varepsilon_x S(x)) \\
\downarrow & & \downarrow \\
P(\varepsilon_x S(x)) & \rightarrow & \neg P(\tau_x S(x))
\end{array}
\quad
\begin{array}{ccc}
P(\tau_x S(x)) & \rightarrow & \neg P(\varepsilon_x S(x)) \\
\downarrow & & \downarrow \\
P(\varepsilon_x \neg S(x)) & \rightarrow & \neg P(\tau_x \neg S(x))
\end{array}
\]

\section{Typed Hilbert operators}

\subsection{Some critics of the unique universe of Frege}

As discussed in [?] universal quantification there are two ways to describe universal quantification (each of these ways is expressed using a different word in most languages, e.g. each vs. every in English):

- universal quantification as a conjunction over the domain (model theoretical view) $\&_{x \in D} P(x)$
- universal quantification as a property of the generic member of its class (proof theoretical view) $\forall x P(x)$ or better $P(\tau_x P(x))$

In the model theoretic view, the domain of quantification is clear, but in the proof theoretic view, the domain is not specified. This means that one cannot write $\&_{x \in D} P(x)$ and a fortiori cannot derive it. The completeness theorem for first order logic [20] makes sure that both notions of quantification agree.

As argued by some medieval philosophers, e.g. Abu’l-Barakāt al-Baghdādī, a property is always asserted from an entity, as being a member of some particular class and not as an entity of the class of all entities [12]. For quantifying (universally or existentially) a notion of class is needed as well: in proof theory the class is simply absent, and in standard model theory, only the largest class of all entities is known. Nevertheless quantifying over a part of the domain is quite common:

(20) All students passed.
(21) All students who passed logic passed algebra.
(22) Numbers ending in 3 have at least one multiple having all 1.

\textsuperscript{6}This avoids problems like my daughter being both tall (as a girl of her age) and not tall (as a member of the family for taking a group picture). [10]
S. Chatzikyriakidis, F. Pasquali, Ch. Retoré

(23) Some student passed algebra.

(24) Some student who passed logic passed algebra.

(25) Some prime numbers are one less than a power of two.

In defining predicate calculus Frege [18, 19] used a “trick” to formalise the quantification on classes. For instance: Aristotle’s A and I sentences $\forall x : S. P(x)$ or $\exists x : S. P(x)$ can be respectively written as $\forall x. S(x) \Rightarrow P(x)$ and $\exists x. S(x) \& P(x)$. Given the classical symmetries $\&/\lor, \forall/\exists$, because of the classical definition of $A \Rightarrow B$ as $(\neg A) \lor B$, one obtains that $\neg (\forall x : S. P(x))$ is $\exists x : S. \neg P(x)$ and $\neg (\exists x : S. P(x))$ is $\forall x : S. \neg P(x)$

However it is clear that there is no similar trick for generalised quantifiers:

(26) a. for 1/3 of the $x : S P(x)$ $\not\equiv$ for 1/3 of the $x (S(x) \Rightarrow P(x))$

b. for few $x \in S P(x)$ $\not\equiv$ for few $x (S(x) \& P(x))$

In these cases, specific generics could be used also introduced by subnector. This issue has recently started to get explored. [36]

7.2 Quantifiying over a sort and typed epsilon and tau

Sorts and classes with specific quantifiers may be a good direction. It should observed that one can defined a many sorted variant of first order logic, where predicates and functions use sorts fo objects that possibly follow a hierarchy (the power-set of a sort can be a sort, etc.). Such a many sorted logic has been precisely defined and studied in the chapter 5 of [26]

Within the usual Montagovian semantic analysis, the epsilon and tau operators for first order individuals should be of type $\epsilon : (e \rightarrow t) \rightarrow e$. If we have many sorts or types, there are two possibilities: $\tau^*, \epsilon^* : \Pi \alpha. \alpha$ and $\tau, \epsilon : \Pi \alpha. (\alpha \rightarrow t) \rightarrow \alpha$, where $\pi$ is basically quantification over types. The first one maps any type to an element in the type (to be understood as the existential or universal generic element in the type). The second one maps a predicate on $\alpha$-object o and $\alpha$-object.

The types for $\epsilon$ and $\tau$ yield more general subnectors than the ones of $\epsilon^*$ and $\tau^*$. This is because types can be mirrored as predicates, but not the converse. It might be useful to apply $\epsilon$ to a complicated predicate like scooter with two front wheels for interpreting a sentence like A scooter with two front wheels just passed in the

\footnote{x : A: is in the class Aand x is of type A. We do not want to be very precise on this, since we have a priori no type theory, no set theory.}

\footnote{No problem of consistency arises with such constants whose type in unprovable (like fix point Y). Any of those type entails the other: $\epsilon^* = \epsilon\{\Lambda \alpha. \alpha\}(\lambda x^{\Pi \alpha. \alpha}. x(t)) : \Lambda \alpha. \alpha \epsilon = \epsilon^*\{\Pi \alpha. (\alpha \rightarrow t) \rightarrow \alpha\}$.}
street. While it would be counterintuitive to consider scooter with two front wheels as a type\(^9\) When there is a type \(T\) one can introduce a constant \(\hat{T}\) of type \(e \to t\) where \(e\) is the largest type which includes any other type.

If the predicate \(S\), to whom the subnector is applied, is of type \(U \to t\), the subnector yields an individual term of type \(U\) and nothing asserts that this terms enjoys the property \(P\). As natural language quantifiers only appear in a proposition which holds (at least locally), one has to add the presupposition \(S(\epsilon S)\), and in case \(S\) is a type, this can be written as \(\epsilon S : S\).

These questions, in particular w.r.t the semantics of determiners, are further discussed in \[38\]

### 7.3 Categorical approaches in the study of typed epsilon and tau subnectors

The fact that the entire apparatus of logic can be recast in purely category-theoretic terms is well known and goes back to the pioneering work of F. W. Lawvere. A description of the History of category theory and of its branch of categorical logic goes well beyond the purposes of this brief section and we refer curious readers to \[30\] and references therein.

One of the most relevant concepts in category theory is that of elementary topos. The definition of elementary topos as a category with finite limits and power objects was introduced by Lawvere and Tierney as a generalization of the previous notion of Grothendieck toposes. It was immediately clear that the concept was of fundamental importance and, since its introduction, applications of elementary toposes have been proposed for an extremely large variety of fields \[23\]. Focusing on logic, we can associate a calculus, often called *internal language of a category*, to every category with finite limits, and therefore to every elementary topos.

The internal language of a category with finite limits is a multi-typed language. By multi-typed language we mean a 5-tuple \((\mathcal{T}, \mathcal{O}, \mathcal{R}, \omega, \rho)\) where \(\mathcal{T}, \mathcal{O}\) and \(\mathcal{R}\) are collections whose elements are called types, operation symbols and relation symbols respectively, while \(\omega\) and \(\rho\) are arity functions. More specifically \(\rho\) sends each element of \(\mathcal{R}\) to a finite, possibly empty, list of types \(A_1, A_2, \ldots, A_n\) and \(\omega\) sends each element of \(\mathcal{O}\) to a non-empty finite list of types \(A_1, A_2, \ldots, A_n, A_{n+1}\). A context \(\Gamma\) is a finite list of typed variables, i.e. \(\Gamma\) is of the form \(x_1 : A_1, x_2 : A_2, \ldots, x_n : A_n\). Terms and

\(^{9}\)However, take a look at \[28, 8\] for an alternative view.
formulas are written in context and the expressions
\[ \Gamma \vdash t : B \quad \Gamma \vdash \phi \]
are used to denote that \( t \) is a well formed term of type \( B \) in the context \( \Gamma \) and \( \phi \) is a well formed formula in the context \( \Gamma \).

The definition of Hilbert’s epsilon calculus in the framework of multi-typed languages does not present many difficulties. What one wants is to require that for every well formed formula \( \Gamma, x : A \mid \phi \) there exists a term \( \Gamma \vdash \epsilon_A^\phi : A \) such that
\[
\frac{\Gamma \vdash H_1, \ldots, H_n \vdash P(t)}{\Gamma \vdash H_1, \ldots, H_n \vdash P(\epsilon_A^\phi)} \quad \epsilon_i
\]
where \( t \) is a term of type \( A \) in the context \( \Gamma \), i.e. \( \Gamma \vdash t : A \), and
\[
\frac{\Gamma \vdash K_1, \ldots, K_p \vdash P(\epsilon_A^\phi) \quad \Gamma, x : A \mid H_1, \ldots, H_n, P(x) \vdash C}{\Gamma, x : A \mid K_1, \ldots, K_p, H_1, \ldots, H_n \vdash C} \quad \epsilon_e
\]
If the underlying collection of types \( T \) has some special types constructor, it might be convenient, or even necessary, to restrict the existence of terms of the form \( \Gamma \vdash \epsilon_A^\phi : A \) to those \( A \)'s that belong to a specific class of types. This caution becomes necessary whenever \( T \) contains the zero type \( 0 \), which can be thought of as an abstract notion of the empty set. In this case the existence of a term \( \Gamma \vdash \epsilon_0^\phi : 0 \) makes the calculus collapse to an inconsistent calculus. Hence, the typed Hilbert’s epsilon calculus is a multi-typed calculus where for every well formed formula \( \Gamma, x : A \mid \phi \), where \( A \) is not the zero type, there exists a term \( \Gamma \vdash \epsilon_A^\phi \) satisfying the rules \( \epsilon_i \) and \( \epsilon_e \) defined above.

The typed Hilbert’s tau calculus is defined analogously. In particular, the formation rules of tau terms must obey the constraint that tau terms has not to be of type \( 0 \). Of course, and within the typed epsilon calculus, the law of excluded middle implies the existence of tau terms and vice versa.

The existence of a model for the classical typed epsilon calculus is not an issue. A standard set-based model, where a non-zero type \( A \) is interpreted as a non-empty set \( \llbracket A \rrbracket \), works. In this model a context \( \Gamma \) is interpreted as the cartesian product of the interpretations of the types in it, a term \( \Gamma \vdash t : A \) as a function \( [t] : [\Gamma] \to [A] \) and a well formed formula in the context \( \Gamma \) as a subset of \( [\Gamma] \). Entailment is inclusion of subsets. Now suppose that \( \phi \) is a well formed formula in the context \( \Gamma, x : A \)
and suppose that $\llbracket \phi \rrbracket$ is its interpretation. For every $g$ in $\llbracket \Gamma \rrbracket$ let $E(g)$ be the set 
\{ $a \in [A] \mid (g, a) \in [\phi]$ \} and define a set $G = \{ g \in [\Gamma] \mid E(g) \text{ is not empty} \}$. By the axiom of choice there is a function which sends each $g$ in $G$ to an element $s(g) \in E(g)$. The interpretation of the function $\llbracket \epsilon^A_\phi \rrbracket : [\Gamma] \to [A]$ is given by the following assignment 

$$
\llbracket \epsilon^A_\phi \rrbracket (g) = \begin{cases} 
  s(g) & \text{if } g \in G \\
  x & \text{if } g \notin G
\end{cases}
$$

where $x$ is any element of $[A]$, which certainly exists as $[A]$ is not empty.

In the previous interpretation we made use of the axiom of choice as well as the law of excluded middle and since toposes can be seen as universes of sets, it is not surprising that, under some mild hypotheses, the interpretation above can be rephrased in any elementary topos validating the axiom of choice and the law of excluded middle. Note that in toposes, by Diaconescu’s argument [14], the axiom of choice implies the excluded middle, so the the second requirement is redundant.

Thus, a more interesting problem is to find an interpretation of the intuitionist typed epsilon calculus. The problem is even more interesting if one considers that in the intuitionist epsilon calculus the axiom of choice is derivable. Therefore any category in which one can carry out the argument of Diaconescu does not provide the desired interpretation.

To the best of our knowledge, the first study of the typed intuitionist epsilon calculus in categories is due to J. L. Bell [6]. In his work Bell used a special class of elementary toposes as a model of the calculus and he called such toposes Hilbertian. To avoid the excluded middle hold, Bell considers a calculus in which epsilon terms exists only for those well formed formulas with at most one free variables. This is a fragment of the full intuitionist epsilon calculus in which the axiom of choice is no longer derivable, therefore the argument of Diaconescu is no longer valid, and Hilbertian toposes need not collapse to classic ones.

An approach that is in some respects complementary to that of Bell’s, is to consider categories which are weaker than toposes. In fact the full power of an elementary topos is required to carry out Diaconescu’s argument. Therefore, if one considers other weaker categories, such as Heyting categories, the validity of the axiom of choice does not imply the validity of excluded middle. A study of those categories that provides an interpretation of the intuitionistic epsilon calculus with all epsilon terms can be found in [34].
8 Conclusion

This partial introduction describes quantification with epsilon and tau, in particular the most frequent cases, i.e. existentials. This avoids the drawbacks of the usual interpretation: epsilon terms follow the syntactic structure, they refer to individuals and they further avoid the unpleasant symmetry between CN and VP in existential statements.

The use of epsilon suggest intriguing connections between types and properties, type theory and first order logic. In this setting, underspecified scope is natural: which quantifier, existential or universal, comes first in \( P(\tau_x. A(x), \varepsilon x. B(x)) \). It also sheds new light on generalised quantifiers: these could be defined syntactically, using a generic as a Hilbert subnector tau and epsilon. Some first insights for “most” are given in [36].

If all roads lead to Rome, most segments of the transportation system lead to Roma Termini! (Blog: Ron in Rome)

References

Hilbert’s epsilon and tau


