Notre Dame Journal of Formal Logic
Volume 30, Number 1, Winter 1989

The Theory of Descriptions Revisited

ALBERTO PERUZZI*

An *excursus* is carried out through the principal steps in the development of the theory of descriptions (TD) from B. Russell until now, and its most important advantages and disadvantages are sketched. TD is studied in the context of model theory (in A. Robinson's style), taking preservation and classification theorems based on normal forms into consideration. Finally, the categorical formulation of TD in topos theory, starting from M. Fourman and D. Scott, is presented with reference to sheaves.

1 Introduction A 'descriptive' operator is a function Δ which, given an open formula $\phi(x)$ of a language L as its input, gives an L-term $\Delta x\phi(x)$, called a 'description', as its output. (In the following, L is supposed to be a standard first-order language with identity.) Clearly, there are many such functions, and to account for at least some of them is a project not only of mathematical significance but also with extensive application to linguistics, because natural languages present a vast range of problems involving the articles "the" and "a"; the theory of descriptions (TD) concerns set theory, with the abstraction operator $\{\ |\ \ldots\}$, and recursion theory, from μ-operators to λ-terms. Here I shall limit myself to a general consideration of *definite* descriptions, i.e. singular terms generated by Δ's which can be read "the such and such", denoted by $\rceil x\phi(x)$.

The difficulties met within many contemporary attempts to formalize descriptions lead to approaching the problem in an unusual way, based on the intuition that the kernel of TD is the *presence (or absence) of symmetries* in a universe of discourse, and therefore in semantics. So, given a model $\mathfrak{M}$ for a theory T in L, we shall focus on the class $Aut_A(\mathfrak{M})$ of $\mathfrak{M}$-automorphisms pointwise (for simplicity) fixed on a set $A \subseteq \mathfrak{N}$, where $\mathfrak{M}$ is a substructure of $\mathfrak{N}$, in order to classify descriptions obtained through parameters from A. As A is varying, this classification proves to be strictly related both to philosophical and

*This article is an enlarged English version of a talk entitled "Federigo Enriques", held at the Dipartimento di Matematica, Università di Milano, on May 23, 1985.

Received December 12. 1986; revised May 4, 1987

mathematical questions: in such a context the spectrum from rigidity to complete freedom for the reference of a given description provides a criterion for testing the usefulness of ideas long discussed by logicians and linguists (e.g., the distinctions between *de dicto/de re* and attributive/referential). Moreover, it is striking how this usefulness depends on a program intended to generalize Galois theory from fields to arbitrary structures (even outside algebra). But if one wishes to study the phenomenology of symmetries of $\iota x\phi(x)$ one soon realizes that one must bear in mind the morphisms and functors associated with different concrete categories. Briefly, my slogan is: "once upon a time TD was a paradigm of philosophy, today it is a chapter of model-theoretic algebra in the categorial framework".

2 The landscape garden It is customary to say that (definite) description theory begins with Bertrand Russell's "On denoting", first published in 1905. I shall follow this custom. The characteristic feature of Russell's conception is that for any ϕ, $\iota x\phi$ is an L-term which can be eliminated by a *contextual* definition in the metalanguage:

$$\psi(\iota x\phi(x)) =_{df} \exists x[\forall y(\phi(y) \leftrightarrow x = y) \wedge \psi(x)].$$

So, to say that the x such that ϕ has the property ψ is the same as saying that there is one and only one x such that ϕ, and that one x has ψ. The aim of this choice is to allow one to speak about everything that may not exist (in the domain of interpretation) without having to widen the range of variables: ontological parsimony. The first disadvantage is that any sentence like $\psi(\iota x\phi)$ is simply *false* if $\neg\exists x\phi$. The same, however, would be true for definitions in the object language, replacing "$=_{df}$" with "$\leftrightarrow$", while if we substitute the equivalence for a weaker simple conditional, the critical case of a material implication with a false premise (when there is not a unique ϕ but there are ψ's) must be explained: a proposal is discussed in Section 6.5 of [12], but only has its value in criticism, pointing to the fact that what we are interested in is not the particular form of ι-elimination, but the conditions of persistence for the truth of "there is exactly one ϕ", formally $\exists!x\phi$, depending on the model-structure $\mathfrak{A}$ and the local/global character of terms in L. After all, to accept an implication from a false premise to a true consequence as conventionally true leads one nowhere, since knowledge requires *separation*. (Eliminability of descriptions would be much more revealing in a constructive logic.)

Another of Russell's problems originates from the interplay of ι with negation. If $\neg\exists!x\phi$ (that is, the description is "improper") then we have to specify the "scope" of the ι-term; otherwise, $\psi(\iota x\phi)$ is false, but $\neg\psi(\iota x\phi)$ is false too; therefore, under classical (bivalent) semantics, $\iota x\phi$ is a contradictory entity. So some qualification is needed. There are also problems concerning descriptions in the context of so-called propositional attitudes, but these problems are not so characteristic of descriptions as to be discussed here. Rather it should be stressed how in Russell's treatment there are no constraints on the structural dynamics which can lead, in the class of T-models, from $\mathfrak{A} \models \neg\exists x\phi$ to $\mathfrak{A}' \models \exists x\phi$ through a specified morphism $f\colon \mathfrak{A} \to \mathfrak{A}'$. Russell's approach, like most of its later variants, reveals a *pointlike* view of descriptions. In contrast, let us suppose that T-axioms codify the notion of a field and $\mathfrak{A}$ is $\mathbb{R}$, the field of real numbers, while ϕ means 'equal to $\sqrt{-1}$'; then any such $\mathfrak{A}'$, like the field $\mathbb{C}$ of complex numbers, will no longer be an ordered field. From this example one realizes that becoming "proper", for a $\rceil$-term, is strictly associated with structural properties of the models for L. Therefore TD is mathematically (and scientifically) meaningful, only with reference to a class K of structures, which behave as ontological background, and to a theory T (or a 'bunch' of theories) acting as a "Gestalt-selector" on K: T and K are parameters for the variation of a certain kind of definability.

With this project in mind, I shall also criticize some common alternatives to Russell's solution as merely formal and *ad hoc*, with respect to the real, profound, problems of TD. It is in this spirit that I consider puzzles such as the following: suppose $\forall x\phi$ holds, then by the axiom $\forall x\phi \to \phi(t)$, if there are no constraints on the formation of $\rceil$-terms, we get for $t = \rceil x\neg\phi$ the consequence $\phi(\rceil x\neg\phi)$, and by $\rceil$-definition $\exists!x(\neg\phi(x) \wedge \phi(x))$: a contradiction. Do logical rules no longer preserve truth? No, I can simply say that the premises were incompatible. Another example: given $\phi(y) = \exists x(y \neq x)$, let us form $\rceil y\phi$; while ϕ is consistent, any sentence containing such a term is not. One can understand how natural it might be to exclude such possibilities, constraining the formation of $\rceil$-terms on storaged information in T. It is just the proof-theoretic approach to TD which comes next, the first version of which was presented by Hilbert and Bernays in [7]. Their basic idea was the introduction of $\rceil$-terms with conditional axioms or rules such as

$$\frac{T \vdash \exists!x\phi(x)}{\rceil x\phi \text{ is a new term and } \phi(\rceil x\phi) \text{ a new axiom added to } T.}$$

There are many formal variants of such a style. The main advantage of this approach is the (constructive, if desired) control of the use of $\rceil$. The other side of the coin is, as Carnap stated in [3], that the notion of a term is no longer recursive, but at most recursively enumerable. To make virtue of necessity, one might reply: it is simply a dogma that the expressive resources of a language L cannot depend on results obtained by theories for which L is designed. I confess that I agree, in this case, with the acknowledgment of the factually inseparable interplay of definability and theoremhood as a *nonlinear* process, but what has again to be explained is our policy when $T \nvdash \exists!x\phi$ and $T \nvdash \neg\exists!x\phi$. Does the term $\rceil x\phi$ belong to a new Lymbus? If such an interplay as a heuristic schema is adopted, a positive theory of the constraints on the variability of reference for $\rceil$-terms in T-models must be worked out.

A different approach is exemplified in the proposal presented in [3]. It is the *conventional* solution: identify the value of $\rceil x\phi$ in a model $\mathfrak{M}$ with an arbitrary but fixed and uniformly assigned element in the domain of $\mathfrak{M}$, when $\neg\exists!x\phi$. It effectively relies on a certain practice in mathematical reasoning, especially when one gives a definition by the cases of a function. Sometimes, however, it is preferable to leave a function undefined; for example take $f\colon \mathbb{R} \to \mathbb{R}$ which for input x gives the inverse $1/x$; f is not defined at the argument 0, so in order to let f be total on $\mathbb{R}$ one might choose an arbitrary value for $1/0$, but this is not done. A symmetric situation occurs with the recursive definition of the factorial (we put $(0)! = 1$) or when the value ∞ is assigned as the limit of a raising function. So the general schema becomes:

$$|x\phi^{\mathfrak{M}} = \begin{cases} \text{the only } a \in M \text{ such that } \mathfrak{M} \vDash \phi[a] \\ 0 \text{ otherwise} \end{cases}$$

(where '0' names a particular individual in M).

Certainly we should explain why, after all, the conventions contained in high school textbooks are not *too* arbitrary. But even if we succeed, cf. Section 4.2 of [12], the problem lies elsewhere. It derives from the possibility that 0 is the value not only for any improper $|x\delta$, but also for some proper δ', so that $|x\delta = |x\delta'$ and therefore $\delta \leftrightarrow \delta'$. The latter, albeit counterintuitive, becomes conventionally true. When L has tools for codifying its own syntax, the story is the same for Frege's idea of giving the expression itself $\overline{|x\delta}$ as value to any improper δ. As is well known, there is another fregean option, but it falls just inside the next trend to be considered: the *external domain* solution.

This is an approach which postulates a plurality of domains attached to *each* L-model, in order to accomplish denotations for improper |-terms. The way out is simply to restrict the range of quantifiers to a subdomain of the range for free variables, considering the 'internal' subdomain D for quantifiers as consisting in the *actual*, existing, elements; and the 'external' one (ones: $D', D'', \ldots$) as consisting in the *possible*, nonexisting, elements, finally taking the union of all these D', D'', etc. as the ontological ambient. Such a step leads to new choices. For instance, a decision must be made between: (i) letting $|x\alpha = |x\beta$ be true on the grounds of $\forall x(\alpha \leftrightarrow \beta)$, by taking the internal domain alone seriously, or (ii) requiring the equivalence of α and β on the union of all the domains, $\alpha(x) \leftrightarrow \beta(x)$, in order to get $|x\alpha = |x\beta$.

A lot of variations can be made on such a theme even if their explanatory power is minimal, and which, moreover, complicate standard semantics in ways which have no connection with the practice of mathematicians, physical scientists, and researchers in (natural or artificial) language theory. At any rate, we could simplify the variety of D-'enlargements' to two coupled domains, containing actual and possible elements (relatively to the given model). With this modal intuition in mind one can imagine how many treatments come to hand, by explicit introduction of different types of quantifiers (as in Nino Cocchiarella's original proposal) $\forall^a, \exists^a$ and $\forall^p, \exists^p$, ranging over the two domains and giving rise to four classes of descriptions, respectively with prefixes $\exists^a\forall^a, \exists^p\forall^a, \exists^a\forall^p$, $\exists^p\forall^p$, where only the first prefix coincides with the usual $\exists!$.

At this point one is naturally led to consider a field of research where the simultaneous presence of multiple domains has been extensively studied: it is the so-called possible-world semantics for modal logic, often associated with Saul Kripke's contributions. Here the presence in L of propositional operators like $\Box$ (it is necessary that) and $\Diamond$ (it is possible that) leads to an investigation of several kinds of descriptions: $\exists x \forall y \Box \phi$, $\exists x \Box \forall y \phi, \ldots$. New problems soon arise with identity (essential to the uniqueness clause for |-terms). In the standard treatment, already in a 'poor' system like T one has $x = y \rightarrow \Box(x = y)$ as a theorem, while $x \neq y \rightarrow \Box(x \neq y)$ is obtained in the much stronger S5—see [8] for notation and proofs. This results in a hard life for descriptions in modal contexts. Moreover, let us think of the two paradigmatic cases, fusion (F) or bifurcation (B) of $|x\alpha$ and $|x\beta$ in the passage from a possible world m to an m' R-accessible to m in a given frame-model K:

(F) $m \models \rceil x\alpha \neq \rceil x\beta$, mRm', $m' \models \rceil x\alpha = \rceil x\beta$
(B) $m \models \rceil x\alpha = \rceil x\beta$, mRm', $m' \models \rceil x\alpha \neq \rceil x\beta$.

In the physical world we are acquainted with both (F) and (B) cases, while their *common* treatment is formally puzzling in modal logic. The problem is how to interpret this situation. As an indetermination of logico-linguistic practice or a fundamental inadequacy of modal logic? There is a highly debated semantic topic related to this problem: the *de dicto/de re* dichotomy. (Roughly, "*de dicto*" is an occurrence of a modal operator when it is applied to a closed formula, "*de re*" when applied to an open formula. So, to $\Box\exists!x\alpha$ there is a corresponding *de dicto* $\rceil$-term, to $\exists!x\Box\alpha$ a *de re* $\rceil$-term, resulting in distinct identifiability conditions, related to the difference between "the actual x such that in every alternative it continues to be 'the such and such'" and "the only x, possibly varying, which in each alternative is 'such and such'".) The foregoing implies a reidentification of individuals through possible worlds: a topic of high interest to psychology and artificial intelligence, welcome by many, on the contrary, as the horse of Troy for a metaphysics of a neo-medieval type. But here I must skip a detailed discussion of such aspects, with which I am presently dealing in another paper, "Lost modal horizon".

Thus I now leave the intrigue of modalities. However interesting the outcome of a modal framework is for analyzing (contingently improper) descriptions, one wonders whether the transition to a modal expansion L' of L could solve the original problem for L, unless one assumes that the same notion of existence requires explicit modal distinctions. Instead I wish just to point out two reinterpretations of modal expressions in classical model theory: first, one can assimilate K to $Mod(T)$, reading R as $\subset$ or $<$ or $<_n$, and properly translating the standard Kripke style conditions for $\Box$ and $\Diamond$; or one can assimilate K to $Aut_A(\mathfrak{M})$, $\mathfrak{M} \in Mod(T)$, reading R as any $\mathfrak{M}$-automorphism fixed on $A \subseteq \mathfrak{M}$—an 'internal' interpretation of modal logic. I think greater explanatory power can be obtained from these 'concrete' determinations of the modalities than if they were to be approached in a general (and generic) way.

More than in modal semantics, TD has received extensive attention in *free logic*. Here one can find a class of formal systems to which pertain both "inclusive" logics (for empty domains) and logics with terms which do not presuppose a denotation. As to inclusivity, in 1948 Andrej Mostowski showed that the implication from $x = x$ to $\exists x(x = x)$ proves false if one puts a simple restriction on variables: modus ponens must be applied only when free variables in the premises remain free in the conclusion. In this way we are able to have empty models. Finally, logic seems to have become the theory of really universal truths. As to existential presuppositions, many authors (Karel Lambert, Jaakko Hintikka, Bas Van Fraassen, Ermanno Bencivenga, among others) have tried to develop logic without necessarily denoting terms (see [2]). This aim has been pursued in many syntactical approaches originating from two main kinds of semantics, respectively grounded on the so-called *supervaluations* and the *substitutional* interpretation of quantifiers.

In fact, free logic embodies the option of constraining the use of quantificational axioms to terms which have been independently proven 'to exist', leaving the language to nondenoting terms open. Resulting models may give rise to truth-value gaps: if t does not refer to any element in A, then $\phi(t)$ has no value on $\mathfrak{A}$. This strategy plainly extends to $|$-terms. With the supervaluation method, one should recognize that classical bivalent models (without gaps) are related to an ideal knowing (in this case modeling) subject. However natural, this trend has encountered many problems. For instance, let us take an equation such as $t = t'$: it comes out not-true if t or t' does not refer, therefore one is led to adopt identities such as $t = t$ as *conventionally* true, which is certainly not in the spirit of a well-designed logic. With regard to the substitutional interpretation of $\forall$ and $\exists$ (especially advocated by Hughes Leblanc) as ranging over linguistic symbols, so that a set of sentences which includes $\exists!x\phi$ and $\neg\phi(a), \neg\phi(b), \ldots$, is not inconsistent in this framework even if it is not satisfiable (as in standard semantics), I confine myself to observe that such a change in the concepts of model and truth may be a step in the direction of an intralinguistic logic, but if logic has to be applied to scientific (or, at least, mathematical) theories, the problems of TD remain unexplained. That is, typically, the destiny of awkward analytical tricks. But a more general question persists: the prerequisite of existence can *never* be satisfied in pure free logic, so it is somewhat vacuous intralinguistically. If, on the other hand, we have a theory T and $T \vdash \phi(t)$ for a certain t, T-axioms have to play an essential role, so the prerequisite is rather superfluous. From this perspective, free logic seems to be just an appendix to conventionalism.

However, these arguments would not hold up if the existence clause for t were to be interpreted in another, mathematically relevant, fashion, as I am going to explain.

I have already spoken about the external domain approach, but I shall now discuss it in relation to the system presented in 1967 by Scott in [16]. I have left the discussion of Scott's system to the end of this section because its peculiar features give it the chance of becoming the basis of a much more general and useful tool, as may be seen in Section 4.

Scott has one type of quantifier ranging over a subset A of the domain B of interpretation (for free variables). The clauses for $\forall$ and $\exists$ are:

$$\mathfrak{B} \models \exists v\phi \text{ iff there is an } a \in A \text{ such that } \mathfrak{B} \models \phi[a]$$
$$\mathfrak{B} \models \forall v\phi \text{ iff for any } a \in A \; \mathfrak{B} \models \phi[a].$$

All terms which do not refer to 'actuals' in the standard subdomain A are now sent on $*$: $*$ is the totally undefined individual. So the global domain for $\mathfrak{B}$ becomes $A \cup \{*\}$. Scott admits $\neg\exists y(y = |x\phi) \to |x\phi = \bar{*}$ as a first $|$-axiom. Consequently, when neither ϕ nor $\neg\phi$ are uniquely satisfied, $\bar{*} = |z(z \neq z)$.

Now Scott expands the language with a new "existence" predicate E, which means "belongs to A". So $E(t) \leftrightarrow \exists v(v = t)$. Quantificational axioms become (as in 'free' style) $\forall v\phi \wedge E(t) \to \phi(t)$ and $\phi(t) \wedge E(t) \to \exists v\phi$. Finally, as a usual second $|$-axiom, we have $\forall y[y = |x\phi \leftrightarrow \forall x(x = y \leftrightarrow \phi(x))]$. The outcome is a very elegant treatment for descriptions, since changes within the classical framework of standard semantics are minimal, while it synthesizes penetrating intuitions of different origins (Frege, Russell, and Quine). What seems to be not completely successful in Scott's TD is the collapse on $*$ of all the improper descriptions relative to a given model.

3 Model-theoretic algebra The aim of the following project is to bring adepts of so-called philosophical logic to realize that we can get a lot of information on descriptions without explicitly naming them, and moreover without changing the standard framework, for applying a vast and conceptually rich stock of results in classical model theory to TD. Here I limit myself to elementary logic, but this is by no means compulsory: one may certainly have some pleasant surprises with descriptions in infinitary or higher-order languages. From the model-theoretic point of view, what is interesting in a description of something by ϕ is simply the character of existence and uniqueness for ϕ and the degree of invariance of 'the described'. So I shall deal first with the problems of *preserving* descriptions in the sense of preserving the associated $\exists!$-clause.

I have already observed the dependence of descriptions on "background knowledge", for simplicity's sake identified here with a pair $\langle T, \mathfrak{M} \rangle$ of a theory and one of its models; you can see that when T is complete the preservation problem is solved on the spot: however one treats $|$-terms, each of them is either always or never proper. This highly improbable case could be called the Hilbert–Bernays paradise. However, it does not provide information on another problem: the degree of *rigidity* of the $\mathfrak{M}$-element which is the denotation of $|x\phi$. When T is model-complete and $T \nvdash \exists!x\phi$, perhaps a T-model $\mathfrak{A}$ is such that $\mathfrak{A} \models \exists!x\phi$; then we are certain not only that, for any $\mathfrak{B} \supset \mathfrak{A}$, $\mathfrak{B} \models \exists!x\phi$, but also that the 'extension' of ϕ will be kept fixed: it is always the same $a \in A$ to be described by ϕ in all such $\mathfrak{B}$'s, just because they are elementary extensions of $\mathfrak{A}$. This kind of situation occurs more 'locally', when T is not model-complete, within the class G of generic T-structures—Robinson's infinitary model-theoretic forcing, where conditions are replaced by diagrams of T-models.

Even from so few observations, one can appreciate the possibility of a phenomenological study of the relation between the form of T-axioms and the form of the descriptions to be preserved: a typical *classification* problem to which the well-known prenex normal form theorem for classical logic offers the tools for a solution—obviously it is not a viable approach if one uses intuitionistic logic.

So one can start from existentially closed T-models and inductive T's, limiting one's attention to the basic complexity level of Σ_2-descriptions, see [12], chapter 7, and then proceed to investigate more particular and/or complex classes of formulas. One of the most direct cases occurs when ϕ is primitive positive, because $\exists!x\phi$ becomes a Horn-sentence, and, as such, it is preserved under reduced products of models and finite intersections of submodels.

But the most interesting aspect of the model-theoretic approach to TD reveals itself when, again following Robinson, we look more strictly at the generalization of concepts and constructions from algebra, in connection with bringing back the set of T-definable elements in $\mathfrak{M}$ to a particular case of the set of those $\mathfrak{M}$-members corresponding to n-valent formulas (i.e., formulas satisfied by a *finite* number n of elements in M), which Lolli [11] has called "quasi-descriptions". Their relevance is remarkable when ϕ is primitive, with parameters extracted from a fixed subset A of dom($\mathfrak{M}$) so that elements satisfying such formulas are rightly named "algebraic" for their direct analogy with roots for systems of equations (and disequations) taking coefficients from a given subset of the domain (take for example the use of integral coefficients).

Thus, it is not difficult to understand the meaning of the following result for TD: when restricted to universal theories with the amalgamation property, the notion of an algebraic element is equivalent to another provided by B. Jònsson: b is algebraic on A in $\mathfrak{M}$ iff for any $\mathfrak{N} \supset \mathfrak{M}$ such that $f_A\colon A(b) \rightarrowtail N$, we have $\bigvee_{i=1}^{n} (b = c_i)$, for $c_i \in \mathfrak{N}$.

Another line of research is related to T-polynomial models, a notion again defined by Robinson [14] and intended to generalize the algebraically closed structures for particular theories. His analysis of pre-polynomials (with a uniqueness condition instead of finiteness) is a starting point for the study of the A-definable subsets in $\mathfrak{M}$, and in particular of the descriptive closure of A, $dcl(A)$, as an extreme case of the algebraic closure, $acl(A)$ – note that A is d-closed when $A = dcl(A)$. These properties have been investigated and refined by Shelah [18], Lascar [10], and Poizat [13], emphasizing the *group-theoretic* aspects associated with $Aut_A(\mathfrak{M})$ and $Aut_A(\mathfrak{N})$, when $\mathfrak{M} < \mathfrak{N}$. By pursuing this theme one reaches a classification of M-elements, relative to A and $L(T)$, in rational ('describable'), algebraic, transcendent elements. This is a task which is also strictly related to the idea of recovering Galois-correspondences out of field theory, and in fact Poizat has proved the famous Galois Theorem on algebraically closed fields, with logical methods supplied by the approach under discussion. How could one say that this is not relevant to the parallel problem, taken from semantics for natural languages, which has to do with descriptions of ideal "fictitious entities"? Here the contrast can be stressed between the present approach and the more diffused one, for which the main purpose of TD ends up by giving a purely formal treatment of improper descriptions, for the simple reason that, when terms are 'properly' defined, there is no problem. Up to now, the analytic trend has led to a jungle of *ad hoc* systems based on axioms, rules, and semantics, constructed for managing a large range of (improper) descriptions in a uniform way. My idea is totally different: we are not searching for a universal melting pot in which to melt any possible context for a description, we are simply focusing on a typical scientific situation in which, given a theory T and an intended model $\mathfrak{M}$, we want to know if and why certain kinds of descriptions can remain proper through a change from $\mathfrak{M}$ to $\mathfrak{M}'$ in $Mod(T)$ or from T to T'. With this perspective in mind, we can appeal, and give methodological meaning, to such a vast and rich interplay between the classical tradition of algebra and contemporary model theory, in order to reach a 'soft' TD, ready to be applied as a predictive tool. It is a line of thought which does not prevent any previously formal attempts in TD from showing their heuristic value here and there.

In a more concrete manner, I advance the following schema as explanatory for the use of descriptions: one starts from $L = L(T)$ and L-terms such as $\iota x\phi$ which have been introduced on the basis of the ability to prove $\exists ! x\phi$ from T-axioms. Then one looks at a T-model $\mathfrak{A}$ (which may be its only relevant model) and enriches $L(T)$, taking into account those 'virtual' entities which work to solve problems in T relatively to $\mathfrak{A}$-parameters. Here is the creative function of descriptions, but at the same time a tension between the attitude to restrict or open the range of quantifiers can be perceived. Thus one proceeds to consider different structures $\mathfrak{A}'$ related through certain maps to $\mathfrak{A}$, and theories T' which possibly extend T, are individuated through information from $\mathfrak{A}$ and are such that $T' \vdash \exists ! x\psi$ while $T \nvdash \exists ! x\psi$. And so on, with $\mathfrak{A}'', T''$, in analogous relation to $\mathfrak{A}', T'$. But there are also inverse cases where one is led from consequences, derivable from $\exists ! x\phi$, to exclude certain 'slices' of previous models as 'unreal' parts. Combine the two roles of descriptions and you have an image of our ontological back and forth. There is no one moment at which we are, once and for all, given *all* possible terms: language develops with knowledge, and vice versa. Bootstrapping? Yes, but neither too much, nor conventionally. What breaks the vicious circle is the dynamic aspect of our constructive process and the stability we gain for model- and theory-variation.

4 Categories of sheaves Today, the most advanced framework for TD comes from category theory, and more specifically from topos theory. It is in this context that Scott further developed his proposal in connection with sheaves, together with Fourman (see [5] and [17]). This leads to the notion of "potential elements" as particular partial functions. The main reason for suggesting this radical shift from all other approaches cultivated within set-theoretic semantics is to be found in the combination of two features: (1) the abandonment of classical logic as the basis for TD and (2) the interference between the analysis of partially defined functions on a space, and the categorial interpretation of terms as arrows ranging in $Sub(1)$: the class of subobjects of the terminal object, viz., the class of monomorphisms (up to equivalence) with codomain the 'unique' object b such that, for any object a in the given category, there is exactly one arrow from a to b. In the case of sets, 1 is just any singleton.

In order to fix these ideas, it is useful to go back to *the* paradigmatic example. Suppose that f and g are two local sections of $A \xrightarrow{p} I$, considered as a fiber bundle, i.e., two partial injective functions from I to A such that $p \cdot f(i) = i$ and $p \cdot g(j) = j$, for i, j belonging to their respective domains included in I. As a measure of the degree at which "$f \approx g$" is true (where $\approx$ is the linguistic counterpart of $=$) one can take the subset of all those $i \in I$ where f and g receive the same values:

$$[\![f \approx g]\!] = \{ i : f(i) = g(i) \}.$$

Note that if one thinks of p as a bundle, in the topos Set/I the pair $\langle I, id_I \rangle$ (where id_I is the identity function on I; see [6] for details) acts as a terminal object, so $[\![f \approx g]\!]$ looks like a truth-value (see Figure 1).

In the case of $A \xrightarrow{p} X$, where A and X are topological spaces, and p is a local homeomorphism, $[\![f \approx g]\!]$ will have to be an *open* subset, so one takes the interior:

$$[\![f \approx g]\!] = \{ x \in X : f(x) = g(x) \}^0.$$

While in the universe of sets, such f's and g's form only a pre-sheaf, the local sections of p on a space X form a *sheaf*, i.e., they are in a precise sense "compatible". The class of all sheaves on X is a category, and more: a Grothendieck topos.

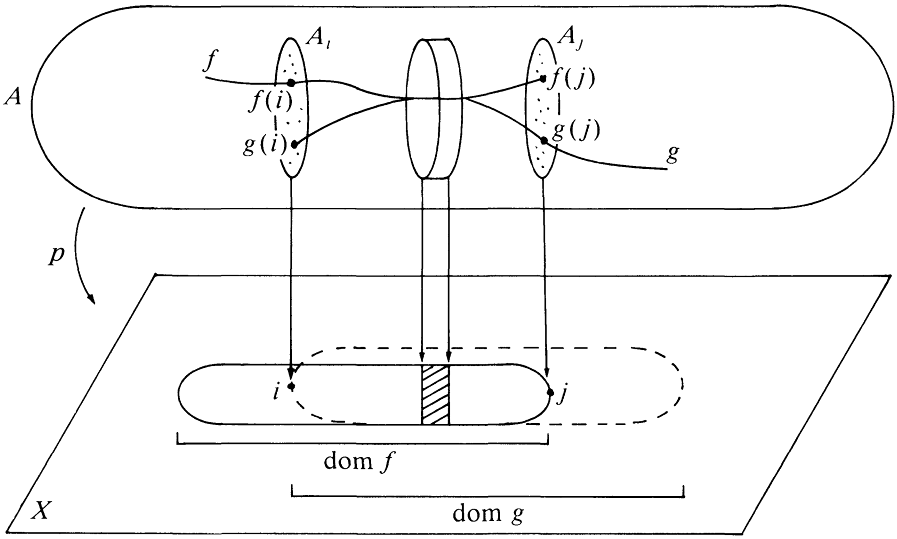


Figure 1.

The concept of existence, which remains mathematically vague in free logics, here receives a clear counterpart:

$$[\![E(f)]\!] = [\![f \approx f]\!] = \mathrm{dom}(f)$$

so the degree of existence for a partial f is measured by the size of its domain of definition. Observe the inverted perspective: you do not enlarge a basic domain, you separate in it and not arbitrarily.

One can now introduce an equivalence relation $\approxeq$ weaker than $\approx$, by putting $[\![f \approxeq g]\!] = -(\mathrm{dom}(f) \cup \mathrm{dom}(g)) \cup [\![f \approx g]\!]$. The definiens can be transformed in terms of the pseudocomplement $\Rightarrow$ in order to have a direct generalization to the intuitionistic-topological case. The idea is that $[\![f \approxeq g]\!]$ measures the degree at which both f and g are undefined or equal: $[\![E(f)]\!] \cup [\![E(g)]\!] \Rightarrow [\![f \approx g]\!]$. As a consequence "$f \approx f$" may be only *locally true* while $[\![f \approxeq f]\!]$ is always equal to I (or X) which here counts as 'True'. Quantification is analogously affected in quinean 'mood', and through glueing functions on overlapping domains you get $|x\phi(x)$ definable as $\bigcup_f f \restriction [\![\forall x(\phi(x) \leftrightarrow f \approx x)]\!]$.

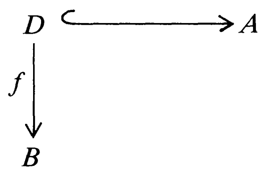
It is important, at this point, to remember the possibility of generalizing all this from a topology $\Theta(X)$ to any complete Heyting algebra Ω. The fruitfulness of such a shift will not emerge until a parallel generalization is reached from the set-theoretical background (the category Set of sets and functions) of our semantics and the related classical hypotheses, to which Scott's early TD is tied. Let us briefly present the essential steps in this new direction.

Any element x of a set A can be defined as an arrow from $1 = \{0\}$ to A, where "1" is our standard name for the terminal object of a category. So, when you leave Set for other topoi, the elements of an object a are arrows defined on 1, but: (i) 1 may now be very complex, (ii) we get the possibility of dealing directly with partial elements as arrows defined on subobjects of 1, or with even more generalized T-elements of a as arrows from a parameters-object T to a. Here I shall however confine myself to the previous option, i.e., f partial means $\mathrm{dom}(f) = u \rightarrowtail 1$.

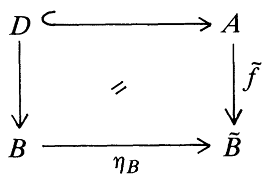
When Scott passed from A *to* $\tilde{A} = A \cup \{*\}$ he could also pass from the partial functions g associated with $\rceil$-terms to global ones, because now, once a formula $\gamma(x)$ has been specified, there is a $\tilde{g}: 1 \to \tilde{A}$ which corresponds to that element $x \in A$ such that $\{x\} = \gamma^{\mathfrak{A}}$ or to $*$. But this is only a particular case of a general procedure. Given $f: D \to B$ with $D \subset A$ (that is, f is partial from A to B, because $\exists x \in A$ such that $x \notin \mathrm{dom}(f)$) we could introduce a new entity $*$, with $* \notin$ B, letting $f(x) = *$ for any $x \notin \mathrm{dom}(f)$. To avoid $* \in B$, when B varies through sets, the trick consists simply in replacing B with an isomorphic copy (in Set, in bijective correspondence) $B' = \{\{y\} : y \in B\}$, taking $* = \varnothing$ in order to be certain that $* \notin B'$. Then $\tilde{B} = B' \cup \{\varnothing\}$. At this point we define $\tilde{f}: A \to B$ in the following way:

$$\tilde{f}(x) = \begin{cases} f(x), \text{ if } x \in D = \mathrm{dom}(f) \\ \varnothing, \text{ otherwise.} \end{cases}$$

So $\tilde{f}$ is derived in a unique way from f, and from a situation like

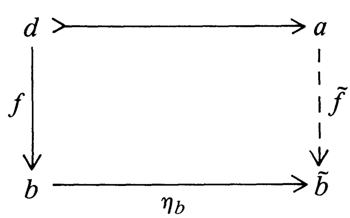


we get a (universally) commuting square, i.e.,



where $\eta_B(y) = \{y\}$, for any $y \in B$.

Once again, this procedure is a particular case within a more general one which pertains to topoi, and is summarized in the "Partial Arrows Classifier Theorem" (proved by William Lawvere and Michael Tierney) which states that in any topos **E**, given an **E**-object b there is a $\tilde{b}$ and an $\eta_b: b \to \tilde{b}$ such that for all arrows f from a subobject d of a there is exactly one $\tilde{f}$ which makes the diagram



a pullback.

When $a = 1$, we obtain an object $\tilde{b}$ of partial elements, in the sense that the partial (potential) elements of b have become total (actual) elements of $\tilde{b}$. Look at the heraclitean move: the opposition between potential and actual is *relative* to the topos, because if $u \rightarrowtail 1$ is partial in $\mathbf{E}$ it is global in $\mathbf{E}/\boldsymbol{u}$ (and vice versa).

Syntactically, L-constants are sent on partial elements. For a term t, an $\mathbf{E}$-model $\mathfrak{A}$ based on a verifies $E(t)$, equivalently $\exists v(v = t)$, when $t^{\mathfrak{A}}$ is a totally defined entity $1 \to a$, a global section in the topos of sheaves on a complete Heyting algebra Ω. Quantificational axioms are accordingly modified, adding the premise $E(t)$, and descriptions are introduced through the ι-axiom (Scott's format):

$$\forall y[(y \approx \iota x\phi(x)) \leftrightarrow \forall x(\phi(x) \leftrightarrow y \approx x)].$$

This strategy is ready to be applied to higher-order languages. Such an extension is due again to Fourman and Scott (see [5]). André Boileau and André Joyal propose a different conception but acknowledge that there is a direct translation of Fourman's system into their own. The way a semantical interpretation can be assigned to ι-terms in this categorial framework is sketched in Section 11.9 of [6].

I only wish to add some general remarks on those points which seem to me methodologically important in the categorial approach to TD, and to pose some open questions for the further development of this approach.

(1) In topos models one can appreciate the opportunities offered by the presence of nontrivial 'empty' domains (objects without global elements).

(2) Topos logic is generally intuitionistic, and from this point of view the property of a certain subdomain (that of the actual entities) to be inhabited or empty is no longer an alternative assertible in any case; and this also affects eliminability conditions (see Section 20 of [9]).

(3) The *local* character of descriptive activity, which emerges with sheaves, could receive further substantial support from the tradition of algebraic geometry (germs, 'points', etc.), especially when seen in the light of the problems encountered in semantics for natural language (Richard Montague's intensional logic, Jon Barwise and John Perry's "situations", etc.). The possibility of better fine tuning control of the case in which f and g are undefined at a 'point' i is still to be explored; in the spirit of Section 3, the collapse of f on g should be avoided, thus further reducing the import of conventional ingredients.

(4) A central feature of this new conception of TD is the variability in the status of ι-terms from topos to topos. For example, in M-Set (objects: actions of a monoid M on sets, arrows: equivariant maps) only rigid descriptions are possible, because all M-elements of A are fixed with respect to M-action.

(5) If the previous points are clearly related to traditional problems of TD from the viewpoint of external semantics on a topos, that quickly growing subject known as categorial logic is a rich source of new interesting problems for the model-theoretic approach presented in Section 3, when looked at 'internally'. Geometric morphisms between topoi do not guarantee the preservation of Σ_2-sentences of the form $\exists! x\phi$ even when ϕ is 'coherent', but they certainly do if we consider descriptions as geometric theories of the (sequential) form $\{\phi(x) \wedge \phi(y) \Rightarrow x = y,\ \top \Rightarrow \exists x\phi x\}$. There is much relevant work for TD to do in the direction of *internal* semantics.

Now, topoi allow one to deal uniformly with the problem of a generalized TD for extensions of (free) higher-order intuitionistic logic right up to the classical type theory (without ramification). But, at least up to the present time, category theory provides adequate semantic tools only for distributive logics, containing the law $\alpha \wedge (\beta \vee \gamma) \leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$, while a categorial treatment for $\rceil$-terms in the nondistributive case would also be desirable. As a major example of nondistributive logic still lacking such categorial treatment we have *quantum logic.* Dalla Chiara has made an interesting contribution to the development of a quantum TD in [4]: here one of the problems is that a proper formulation of clauses of existence and uniqueness for ϕ does not imply the individuation of a particular element d such that $\rceil x\phi$ is a 'name' for it.

In fact, we should not forget what is perhaps the basic aspect of 'describing': when you univocally identify something in an explicit way, you are recognizing a stable singularity in the universe of discourse. Can the logic-sensitive character of TD result from a variational process on a topological intuition? If so, then as in the case of the proximity-spaces in Section 2 of [1] it could be conjectured that only deeper investigation on the borderline between logic and (differential) topology will succeed in founding TD. This would also be an effective test for the real epistemological content of the 'dynamics' of descriptions I gave in Section 3 in the form of an explanatory schema. But now the same schema might become the object of research in functorial terms.

REFERENCES

[1] Bell, J. L., "A new approach to quantum logic," *British Journal for the Philosophy of Science*, vol. 37 (1986), pp. 83–99.

[2] Bencivenga, E., "Introduzione" to *Le Logiche Libere*, Boringhieri, Torino, 1975.

[3] Carnap, R., *Meaning and Necessity*, University of Chicago Press, Chicago, 1947.

[4] Dalla Chiara, M. L., "Nomi e descrizioni nella logica quantistica," pp. 435–445 in *Scienza e Filosofia*, ed., C. Mangione, Garzanti, Milano, 1985.

[5] Fourman, M., "The logic of topoi," pp. 1053–1090 in *Handbook of Mathematical Logic*, ed., J. Barwise, North-Holland, Amsterdam, 1977.

[6] Goldblatt, R., *Topoi. The Categorial Analysis of Logic*, 2nd revised ed., North-Holland, Amsterdam, 1984.

[7] Hilbert, D., and P. Bernays, *Die Grundlagen der Mathematik*, 2 vols. Springer, Berlin, 1934 and 1939.

[8] Hughes, G., and M. Cresswell, *An Introduction to Modal Logic*, Methuen, London, 1968.

[9] Lambek, J., and P. Scott, *Introduction to Higher-Order Categorial Logic*, Cambridge University Press, Cambridge, 1986.

[10] Lascar, D., "On the category of models of a complete theory," *The Journal of Symbolic Logic*, vol. 47 (1982), pp. 249–266.

[11] Lolli, G., "Sulle descrizioni: la radice quadrata di −1 e la prova ontologica," *Rivista di Filosofia*, vol. 3 (1975), pp. 409–417.

[12] Peruzzi, A., *Definizioni. La Cartografia dei Concetti*, Angeli, Milano, 1983.

[13] Poizat, B., "Une théorie de Galois imaginaire," *The Journal of Symbolic Logic*, vol. 48 (1983), pp. 1151–1170.

[14] Robinson, A., *Introduction to Model Theory and to the Metamathematics of Algebra*, North-Holland, Amsterdam, 1963.

[15] Russell, B., "On denoting," *Mind*, vol. 14 (1905), pp. 479–493.

[16] Scott, D., "Existence and descriptions in formal logic," pp. 181–200 in *Bertrand Russell, Philosopher of the Century*, ed., R. Schoenman, Allen & Unwin, London, 1967.

[17] Scott, D., "Identity and existence in intuitionistic logic," pp. 660–696 in *Applications of Sheaves*, eds., M. Fourman and D. Scott, Springer, Berlin, 1979.

[18] Shelah, S., *Classification Theory and the Number of Non-isomorphic Models*, North-Holland, Amsterdam, 1978.

Dipartimento di Filosofia
Università di Firenze
50139 Firenze, Italia