Routledge Encyclopedia of PHILOSOPHY

General Editor
EDWARD CRAIG

ROUTLEDGE
London and New York

paische Verlagsanstalt. (A historical account of the development of Bernstein's ideas.)

Meyer, T. (1977) *Bernsteins konstruktiver Sozialismus* (Bernstein's Constructive Socialism), Berlin and Bonn: Dietz. (On Bernstein's contribution to socialist theory.)

* Tudor, H. and Tudor, J.M. (eds and trans) (1988) *Marxism and Social Democracy: The Revisionist Debate 1896–1898*, with intro. by H. Tudor, Cambridge: Cambridge University Press. (Contains all but two of the articles entitled 'The Problems of Socialism'.)

H. TUDOR

BETH'S THEOREM AND CRAIG'S THEOREM

Beth's theorem is a central result about definability of non-logical symbols in classical first-order theories. It states that a symbol P is implicitly defined by a theory T if and only if an explicit definition of P in terms of some other expressions of the theory T can be deduced from the theory T. Intuitively, the symbol P is implicitly defined by T if, given the extension of these other symbols, T fixes the extension of the symbol P uniquely. In a precise statement of Beth's theorem this will be replaced by a condition on the models of T. An explicit definition of a predicate symbol states necessary and sufficient conditions: for example, if P is a one-place predicate symbol, an explicit definition is a sentence of the form $(x)(Px \equiv \phi(x))$, *where* $\phi(x)$ *is a formula with free variable x in which P does not occur. Thus, Beth's theorem says something about the expressive power of first-order logic: there is a balance between the syntax (the deducibility of an explicit definition) and the semantics (across models of T the extension of P is uniquely determined by the extension of other symbols).*

Beth's definability theorem follows immediately from Craig's interpolation theorem. For first-order logic with identity, Craig's theorem says that if ϕ *is deducible from* ψ, *there is an interpolant* θ, *a sentence whose non-logical symbols are common to* ϕ *and* ψ, *such that* θ *is deducible from* ψ, *while* ϕ *is deducible from* θ. *Craig's theorem and Beth's theorem also hold for a number of non-classical logics, such as intuitionistic first-order logic and classical second-order logic, but fail for other logics, such as logics with expressions of infinite length.*

1 The axiomatic method

Questions concerning the definability of concepts arose within the development of the formal axiomatic method by late nineteenth- and twentieth-century mathematicians such as Moritz Pasch, Giuseppe Peano, David Hilbert and Alfred Tarski. What sets apart the formal axiomatic method from the earlier axiomatic thinking of, say, Euclid's geometry is that the primitive terms occurring in the axioms are uninterpreted and the axioms devoid of any meaning; and that the rules of reasoning have been made completely explicit and formal. The axioms state purely formal relationships between the terms. This suffices for the purpose of deducing theorems from the axioms by rigorous reasoning. The development of the formal axiomatic method is the culmination of the movement to rigorize mathematics started in the early nineteenth century by mathematicians such as Cauchy and Bolzano. But these earlier efforts were directed chiefly towards ontology, for example, cleansing the language of analysis of visual images and reference to movement – and a mathematical theory still had subject matter, albeit abstract. In the formal axiomatic method the subject matter is provided from the outside by an interpretation of the primitive terms of the theory. One and the same theory may be open to radically diverse interpretations.

A perspicuous system of axioms requires under the formal axiomatic method the independence of each axiom from the others: a dependent axiom can be dropped without loss of content. Peano developed a method to prove independence: give an interpretation of the primitive terms which makes the one axiom false and the others true. This now common method is the formalization of an earlier idea of Eugenio Beltrami who felt that the so-called non-Euclidean geometry of Lobachevskii lacked a 'real foundation', that is a foundation in actual physical space. In 1868 Beltrami offered an interpretation of this geometry in terms of the acceptable Euclidean geometry: Lobachevskian 'geometry' could be understood as being about a special kind of line (a geodesic) in a special kind of plane (a surface with constant negative curvature) in Euclidean space. But Beltrami's project was one of meaning and he did not have the consistency of non-Euclidean geometry in mind nor was he concerned with the independence of Euclid's parallel axiom. In fact, only two years later, the French mathematician Guillaume Hoüel pointed out that Beltrami's construction showed the independence of the parallel axiom: while the other axioms are true for all planes, the parallel axiom holds only in planes of zero curvature (see Scanlan 1988).

Another idea from the pre-formal period was absorbed by the formal axiomatic method through the work of the Italian mathematician Alessandro Padoa, a close collaborator of Peano. Padoa was concerned with the definability of concepts. He pointed out that there exists a parallel between, on the one hand, the methodological concepts of being an axiom and being derivable as a theorem and, on the other hand, concepts from the theory of definition: the notion of primitive term corresponds to the notion of being an axiom, and the notion of definability corresponds to the notion of deducibility. A perspicuous system of axioms therefore also requires the independence of the primitive concepts used in the axioms. Padoa was thus led to ask in 1900 whether there could be a method to prove independence of concepts, just as Peano's method had shown independence of axioms.

Early in the nineteenth century the French mathematician Jose Diez Gergonne had suggested the contrasting terms 'explicit' and 'implicit' as regards definition (Gergonne 1918–19). Gergonne's distinction was suggested by the difference, in algebra, between a set of solved equations, which gives as it were explicit definitions of the unknowns, and a set of unsolved equations which is strong enough to determine a unique solution for the unknowns. Gergonne characterizes implicit definitions as 'phrases that make us understand one of the words that occur in it through the known meaning of the other words'. Analogously to the algebraic case, Gergonne requires that the number of unknown words should be equal to the number of phrases that together implicitly define them. Completely forgotten by the end of the nineteenth century it was Giovanni Vacca, an assistant of Peano, who around 1896 gave a short account of Gergonne's paper on definitions. Padoa's paper (1901), read at the First International Congress of Philosophy in Paris, is clearly motivated by Gergonne's work. To prove the independence of a concept from the other concepts occurring in a theory Padoa proposed a new method: find a true interpretation of the theory, considered as an abstract system, that remains a true interpretation when solely the meaning of that concept is changed. Thus, though Padoa did not refer to Gergonne or use the term 'implicit definition', his so-called 'two-model method' establishes that a term is not implicitly defined by a theory in the sense of Gergonne.

Since Padoa did not indicate how to construct the two models it may be better to speak of the 'Padoa criterion' for undefinability. Padoa claimed without further proof that his criterion was both necessary and sufficient for the explicit undefinability of a given concept by means of the other concepts in a given theory. Sufficiency is clear: if an explicit definition were to be implied by the theory two such models could not exist since the truth of the explicit definition would force the uniqueness of the interpretation of the explicitly defined term given an interpretation of the other terms. But necessity is not obvious. Does the absence of two such models guarantee that an explicit definition exists and is derivable? With hindsight this question could not have been answered at that time for it requires a more careful specification of the underlying logical system than was available to Padoa.

Alfred Tarski (1935) answered the question affirmatively for a modification of the ramified theory of types of Whitehead and Russell's *Principia Mathematica* (see THEORY OF TYPES). His proof was a rather straightforward derivation within the system, since the meta-claim that Padoa's two models do not exist can be expressed in the language of type theory.

2 Beth's theorem and Craig's interpolation theorem

In 1953 the Dutch philosopher and logician Evert Beth proved the necessity of Padoa's criterion for first-order or elementary logic. Beth showed that if no explicit definition of a term can be deduced from a theory, two models of the theory exist that differ only in the interpretation of the term in question. Moreover, in his so-called semantic tableau method, Beth found the means to construct systematically, in the absence of definability, the two models required by Padoa's criterion, albeit often through an infinite process, while a closed tableau makes it possible to find an explicit definition of the term in question (see NATURAL DEDUCTION, TABLEAU AND SEQUENT SYSTEMS §4). Beth thus took away some of the concerns of the American mathematician Oswald Veblen who had remarked in 1902 that what Padoa proposed 'seems hardly adequate' when the issue was to replace an axiomatic system by one with independent axioms and independent terms, since he gave no method to find the two models and thus prove independence, or to construct the explicit definition, in the case of dependency.

Let L be a first-order language and P an arbitrary non-logical constant not in L. Let $L(P)$ denote the language obtained by adding P to L. To simplify our notation we will assume that P is a one-place predicate symbol. If T is an arbitrary theory in the language L, then $T(P)$ will be a theory in the language $L(P)$. Deducibility in first-order logic will be denoted by '$\vdash$'. An interpretation or model M for L specifies extensions in a domain D for all the non-logical constants of L (see MODEL THEORY). If M is a model of L, a model of $L(P)$ will be denoted by (M, X), where X is a subset of the domain D of M. Thus P is here interpreted as the subset X.

We will now define the notions of explicit and implicit definability. $T(P)$ is said to *define P explicitly* if there is a formula $\phi(x)$ of L such that $T(P) \vdash (x)(Px \equiv \phi(x))$. If P is not a one-place predicate symbol, this definition can be modified in the obvious way. Furthermore, let T be the set of first-order consequences of $T(P)$ in the language L, in which P does nor occur. Then $T(P)$ is said to *define P implicitly* when, for every model M of T, there is exactly one expansion (M, X) of M which is a model of $T(P)$.

'Beth's definability theorem' for first-order logic states that a theory $T(P)$ defines a term P implicitly if and only if $T(P)$ defines P explicitly. Beth's original proof of 1956 uses a modification of Gentzen's 'extended *Hauptsatz*', which shows that, in first-order logic, every proof can be carried out without any detours. Nowadays Beth's theorem is usually proved to be a direct implication of Craig's interpolation theorem.

'Craig's interpolation theorem' for first-order logic with identity says that if a sentence ψ of first-order logic entails a sentence θ there is an 'interpolant', a sentence ϕ in the vocabulary common to θ and ψ, that entails θ and is entailed by ψ. William Craig originally proved this theorem as a lemma to be used in obtaining a simpler proof of Beth's theorem (Craig 1957). Since then, however, the result has come to stand on its own.

We will now sketch a proof that Craig's theorem implies Beth's theorem. Since first-order logic is complete, implicit definability, a model-theoretic condition, is equivalent to the following deducibility condition, which is in fact Beth's original definition of (implicit) definability. Let P' be a one-place predicate not in L and distinct from P, and let $T(P')$ be the theory in $L(P')$ obtained by replacing P by P' in $T(P)$ wherever it occurs. Then P is implicitly defined by $T(P)$ if and only if $T(P) \cup T(P') \vdash (x)(Px \equiv P'x)$. Assume now that this condition holds, and that $T(P)$ is a finite set of sentences, or, rather, one big conjunction of axioms, and similarly for $T(P')$. So we can write $T(P) \,\&\, T(P') \vdash (x)(Px \equiv P'x)$. But then also $T(P) \,\&\, Pc \vdash (T(P') \rightarrow P'c)$, where c is a new individual constant not in L. By Craig's theorem there is an interpolant $\phi(c)$ such that $T(P) \,\&\, Pc \vdash \phi(c)$ and $\phi(c) \vdash (T(P') \rightarrow P'c)$. Since P' does not occur in $\phi(c)$ it is also true that $\phi(c) \vdash (T(P) \rightarrow Pc)$. Thus $T(P) \vdash (Pc \equiv \phi(c))$. Since c does not occur in $T(P)$, we also have $T(P) \vdash (x)(Px \equiv \phi(x))$. This completes the proof that if $T(P)$ defines P implicitly, then $T(P)$ defines P explicitly. The other direction of Beth's theorem follows independently of Craig's theorem.

3 Further developments and applications

Again, let T be the set of first-order consequences of $T(P)$ in the language L. We said that $T(P)$ defines P implicitly when for every model M of T there is exactly one expansion (M, X) which is a model of $T(P)$. In Padoa's method two models (M, X) and (M, X') of $T(P)$ are exhibited. Another way in which implicit definability can be violated is if there is a model of T that cannot in any way be expanded to a model of the full $T(P)$. Karel de Bouvère (1959) studied this so-called one-model method to show undefinability of addition and multiplication in number theory. In the philosophy of science literature this is called a failure of Ramsey eliminability of the term.

Not to be confused with the above concept of definability of a term in a theory is the concept of definability of a set in a model which, for Tarski, belongs to semantic definability rather than to the formal definability involved in Padoa's question since now we have a fixed model for an interpreted language. Given a model M of L, a subset X of its domain D is *definable in the model M* if there is a formula of L with one free variable $\phi(x)$ such that $(x)(Px \equiv \phi(x))$ is true in (M, X), where P is interpreted as X. Obviously, if P is explicitly definable in $T(P)$ and if (M, X) is a model of $T(P)$, then X is definable in M. Moreover, the concept of definability in a model can be iterated, whereas definability of a term in a theory cannot since a set of terms is not itself a term of the language.

We say that the predicate P is definable in a model (M, X) of $T(P)$ if an explicit definition of P holds in (M, X). Different models of $T(P)$ may satisfy different, non-equivalent definitions. But a theorem proven by Lars Svenonius in 1959 shows that if P is definable in every model (M, X) of $T(P)$ then each model (M, X) of $T(P)$ satisfies one of a finite list of definitions. That is, $T(P)$ implies a (finite) disjunction of explicit definitions of P. This property is called 'explicit definability up to disjunction' or 'piecewise definability'.

In model theory the concept of 'a logic' is defined and logics for which Craig's interpolation theorem hold are said to have the Craig or interpolation property; similarly for the Beth property. Any usual logic with the Craig property has the Beth property, but the latter has been shown to be weaker.

See also: DEFINITION; GEOMETRY, PHILOSOPHICAL ISSUES IN; LOGICAL AND MATHEMATICAL TERMS, GLOSSARY OF

References and further reading

Barwise, J. and Feferman, S. (eds) (1985) *Model-Theoretic Logics*, New York: Springer. (The papers by Ebbinghaus and by Makowsky give details about the relation between the Beth property and the Craig property in abstract model theory.)

Benthem, J.F. van (1978) 'Ramsey Eliminability', *Studia Logica* 37: 321–36. (A detailed exposition of Ramsey eliminability with a discussion of its relevance for the philosophy of the empirical sciences.)

* Beth, E.W. (1953) 'On Padoa's Method in the Theory of Definition', *Indagationes Mathematica* 15: 330–9. (Gives Beth's result that for first-order logic implicit definability implies explicit definability in Beth's original proof-theoretic formulation.)

—— (1956) *L'Existence en Mathématiques*, Paris: Gauthier-Villars. (Includes an example of the one-model method to show undefinability. This involves constructing a model of the consequences of the theory without the terms whose undefinability has to be demonstrated, a model which cannot be extended to a model of the full theory.)

—— (1962) *Formal Methods: An Introduction to Symbolic Logic and to the Study of Effective Operations in Arithmetic and Logic*, Dordrecht: Reidel. (Shows how Beth's semantic tableau gives a method for constructing a Craig interpolant, and how this leads to the construction of an explicit definition in case Padoa's method fails.)

* Bouvère, K.L. de (1959) *A Method in Proofs of Undefinability, With Applications to Functions in the Arithmetic of Natural Numbers*, Amsterdam: North Holland. (De Bouvère's dissertation with the one-model method, which shows that, under certain conditions, a term is essentially undefinable, that is, undefinable in any consistent extension.)

Chang, C.C. and Keisler, H.J. (1973) *Model Theory*, Amsterdam: North Holland. (Gives a complete account of interpolation and definition results for classical logic.)

* Craig, W. (1957) 'Three Uses of the Herbrand–Gentzen Theorem in Relating Model Theory and Proof Theory', *Journal of Symbolic Logic* 22: 269–85. (Derives Craig's theorem about the existence of an 'intermediate' formula or interpolant from a consequence of Gentzen's extended *Hauptsatz*, and applies it in a proof of Beth's definability theorem, formulating it as a result relating model theory and proof theory.)

Ebbinghaus, H.-D. (ed.) (1987) *Bibliography of Mathematical Logic*, vol. 3, *Model Theory*, Berlin: Springer. (§C40 gives an extensive bibliography on interpolation and definability.)

* Gergonne, J.D. (1918–19) 'Essai sur la théorie des definitions', *Annales de Mathématiques Pures et Appliquées* 9: 1–35. (Referred to in §1. Defines the concept of implicit definition before the advent of the formal axiomatic method.)

Kleene, S.C. (1967) *Mathematical Logic*, New York: Wiley. (A good account of the formal axiomatic method.)

* Padoa, A. (1901) 'Essai d'une théorie algébrique des nombres entiers, précédé d'une introduction logique à une théorie déductive quelconque', in *Premier Congrès International de Philosophie*, vol. 3, *Logique et Histoire des Sciences*, Paris: Armand Colin, 309–65; partial trans. 'Logical Introduction to any Deductive Theory', in J. van Heijenoort (ed.) *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge, MA: Harvard University Press, 1967, 118–23. (Referred to in §1. The 'Logical Introduction on Deductive Theories in General' has the first statement of Padoa's criterion for the undefinability of a symbol on the basis of a theory.)

* Scanlan, M.J. (1988) 'Beltrami's Model and the Independence of the Parallel Postulate', *History and Philosophy of Logic* 9: 13–34. (Referred to in §1. Compares the work of Beltrami and Hoüel with the treatment of Euclidean and non-Euclidean geometry after the development of the formal axiomatic method by Pasch, Peano and others.)

* Tarski, A. (1935) 'Einige methodologische Untersuchungen über die Definierbarkeit der Begriffe', *Erkenntnis* 5: 80–100; trans. 'Some Methodological Investigations on the Definability of Concepts', in *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, trans. and ed. J.H. Woodger, Oxford: Clarendon Press, 1956, 296–319. (Referred to in §1. Includes the proof that in the logical system of *Principia Mathematica* the two extensions required by Padoa's method do exist if a term cannot be explicitly defined.)

Tarski, A. and Kuratowski, C. (1931) 'Les opérations logiques et les ensembles projectifs' (Logical Operations and Projective Sets), *Fundamenta Mathematicae* 17: 240–8. (The first occurrence of the expression 'implicit definition' in the sense of Gergonne in the modern literature. The authors point out that definitions by induction are a type of implicit definition.)

Tuomela, R. (1973) *Theoretical Concepts*, Vienna: Springer. (An accessible discussion of definability and Ramsey eliminability in the empirical sciences.)

* Vacca, G. (1896–9) 'Sui precursori della logica mathematica. II. J.D. Gergonne' (On Precursors of Mathematical Logic. II. J.D. Gergonne), *Rivista di matematica* 6: 183–6. (Referred to in §1. Brings