Jaakko Hintikka

CONTINUUM HYPOTHESIS AS A MODEL-THEORETICAL PROBLEM

1. *How to study set theory*

The continuum hypothesis (CH) is crucial in the core area of set theory, viz. in the theory of the hierarchies of infinite cardinal and infinite ordinal numbers. It is crucial in that it would, if true, relate the two hierarchies to each other. It says that the second infinite cardinal number, which is known to be the cardinality of the first uncountable ordinal, equals the cardinality $2^{\aleph_o}$ of the continuum. (Here $\aleph_o$ is the smallest infinite cardinal.)

At the same time CH offers an instructive case study of how to approach set theory. The question whether CH is true, known as the continuum problem, has in the last hundred years been usually approached as a deductive problem, viz. as the question whether CH is provable in a first-order axiomatic foundation of set theory, in the first place in the well-known Zermelo-Fraenkel set theory (ZF) with suitable new axioms added. Such approaches are seriously misguided, and should not be pursued. Current first-order axiomatizations of set theory are misuses of the axiomatic method. This method is calculated to help us to study a certain class of structures by capturing them as the models of an axiom system. Now problems like the continuum problem are questions concerning certain structures of sets. However, the models of any first-order theory are structures of particular objects (individuals), not of sets. How the study of the latter is supposed to help us with the study of the former has never been sufficiently clarified. It seems that set theorists initially assumed that one can simply consider sets as values of first-order variables. They should have seen the handwriting on the wall already at the time set theory was axiomatized by Zermelo. For the so-called Basic Assumption V in Frege's ill-fated

Grundgesetze essentially says just that one can do so. (Frege 1893-1903 and 1964.) And it was the failure of this very assumption that Frege himself quickly identified as the source of the paradoxes that affected his system.

Now CH is a statement about certain structures of sets. Hence it is wishful thinking that we could come to grips with it by first-order axiomatizations which pertain in the first place to structures of individuals. For the same reason, it should not have been any surprise when Kurt Gödel (1940) and Paul Cohen (1966) showed that CH is neither disprovable nor provable in ZF. Their results do not even indirectly show anything about the truth or falsity of CH. Nor is any help forthcoming from possible additional axioms. For already in ZF without any additions one can prove theorems that are false on a set-theoretical interpretation of variables of quantification. For instance, one can prove that there are true sentences whose Skolem functions do not all exist. (If this were not the case, one could define truth for a first-order set theory in the same set theory, which would violate Tarski's impossibility theorem. See here Tarski 1935, Hintikka 1998, 2004 and forthcoming.) Such false theorems cannot be eliminated by adding new axioms. Therefore the results of Gödel and Cohen should be considered not as contributions to the study of CH or any other set-theoretic problem, but warnings about the inadequacy of attempted first-order axiomatizations of set theory. It is thus ridiculous to present their results as "solutions" to the continuum problem, alias Hilbert's first problem.

In these circumstances, it is necessary to approach set-theoretical problems like the continuum problem by model-theoretical methods rather than by means of first-order axiomatizations. But if so, the logic that is used must be stronger than the principles of reasoning codified by ZF. For being a first-order theory, ZF cannot deal with its own model theory. For one thing, the basic concept of all model theory, the concept of truth, cannot be defined for a first-order theory in the theory itself according to Tarski's impossibility theorem (Tarski 1935).

Fortunately, a simple and straightforward extension of the usual first-order logic yields everything that is needed for the purposes of this paper. We can simply allow a quantifier (Q_2y) to be independent of another quantifier within whose scope it occurs. This change yields what is known as independence-friendly (IF) first-order logic. (For it, see e.g. Hintikka 1996 and 2000 (b).) It turns out that traditional mathematicians have ever since Weierstrass tacitly used independent quantifiers, for instance in using notions like uniform convergence.

Another simple way is to generalize the rule or existential instantiation. In its present form, it can be applied only to a sentence-initial existential quantifier $(\exists x)$, and it allows the variable x to be replaced by a new individual constant (say b) while the quantifier is omitted. In the generalized form, $(\exists x)$ can occur anywhere in a formula (in a negation normal form). The only difference is that the variable x is now replaced by a function term $f(y_1,y_2,\dots)$ where f is a new function constant and $(Q_1y_1),(Q_2y_2),\dots$ are all the quantifiers on which $(\exists x)$ depends.

In either case the resulting first-order logic will be strong enough to enable us to discuss the model theory of the relevant set-theoretical structures. For one thing, in such a first-order logic we can use a strong form of the axiom of choice which can be formulated by requiring the existence of the Skolem functions of each true sentence.

What must be given up is the idea of set theory as the study of a set theoretical universe. What can be done is to define certain set-theoretical structures and then to study them by means of suitable logic. The definitions need not be expressible by means of the logic in question, but can themselves be model theoretical. Such indefinability is no obstacle to studying the resulting structures by means of logic. For practically all purposes, a first-order logic extended along the lines just indicated is sufficient without any specifically set-theoretical assumptions. In many important cases, the usual first-order logic formulated so as to include the axiom of choice is enough.

Accordingly, CH is approached here as a model-theoretical proposition. It can be discussed by reference to a linear continuum, but its most accessible model-theoretical formulation is probably the one that states that the cardinality of the second number class (the sequence of all countable ordinals) equals the cardinality of the continuum or $2^{\aleph_0}$. This is the formulation (or equivalent form) of CH considered mostly in this paper.

2. *Finite constituents*

For the purpose, we need some unfamiliar model theory. In the usual approaches to model theory, the focus is in effect on what can be said of the elements of the model one by one. Here we are concentrating instead on sequences of elements that can be drawn from a given model. Accordingly, a fragment of model theory will be developed in the next few sections from this point of view.

Let us assume that a model M of a formal first-order language is given with the domain D(M). Without limiting the generality of the discussion it can be assumed that the vocabulary of the language consists of one two-place relation R(x,y) (which we will also express by $x \geq y$). We can then form the tree structure T(M) of all the ramified sequences of d individuals that can be drawn from D(M). Each node of T(M) is labeled by the (name of an) individual in D(M), together with a specification (in the form of a conjunction of negated or unnegated atomic formulas) of how it is related to the individuals lower in the same branch (and to itself).

From this ramified list of *all sequences* of d individuals in D(M) we form a related list (labelled tree) of *all* the different *kinds of sequences* that can be built from individuals in D(M).

The first step is to associate o variable to each node in L. The choice of the variables is irrelevant, as long as no confusion arises. For instance, for uniformity we could use the same variable for all nodes at the same height (level, aka depth) d, different at different levels.

Each node associated (with a variable , say) is associated with a label indicating its position in the tree. This label C will have to specify first of all how x is related to nodes y lower down in the same branch and to itself. Such a specification is accomplished by a nest of negated or unnegated atomic formulas. The conjunction of all the members of the list is said to be of the form

(2.1) $C^{(o)}[x, y_k, y_{k-1}, \ldots, y_1]$

We are in this paper using the exclusive interpretation of quantifiers. This means that (2.1) is always assumed in effect to include the conjuncts $(x \neq y_1), (x \neq y_2), \ldots, (x \neq y_k)$

Furthermore, the label associated with x must specify the nodes immediately above it. If these nodes are labeled $C_i^{(d-1)}[z, x, y_k, y_{k-1}, ..., y_1](i \in I)$,then we can use as the label for x a conjunction of the form

(2.2) $C^o[x, y_k, y_{k-1}, ..., y_1] \,\&\, \wedge_1 (\exists z) C_i^{(d-1)}[z, x, y_k, y_{k-1}, ..., y_1] \,\&\, (\forall z) \vee_i C_i^{(d-1)}[z, x, y_k, y_{(k-1)}, ..., y_1]$

Here $i \epsilon I$, the order of conjuncts and disjuncts is irrelevant, and s^o is the choice of the bound variables. Such a label (formula) is said to be a protoconstituent with the parameters $y_1, y_2, \ldots, y_k$. They are called

(2.3) $C_j^{(d)}[x, y_k, y_{k-1}, ..., y_1]$

Protoconstituents are not perfect structure descriptions, however, for they may contain repetitions of conjuncts and/or disjuncts. These repetitions can be easily removed one by one. The resulting structure, which is still of the form (2.3), is said to be a c-constituent.

Even though this elimination is trivial in the case of finite constituents (finite $\propto$), its structure is worth attention. If a structure C_1 of the form (2.3) results from another structure C_2 by omitting repetitions of conjuncts and/or disjuncts, then C_2 can be mapped marry-on onto C_1 with the preservation of the order (created by the relation R). In such a case, C_1 is said to be denser than C_2 . This is seen to be a partially ordering relation, and the C- constituents are minimal protoconstituents (with respect to the "denser than" relation).

Such minimal structures can be characterized formally. The conditions (2.1)-(2.2) (with no repetitions of conjuncts and/or disjuncts allowed) can be used as a recursive definition of a class of formulas of the form (2.3). (The choice of bound variables is again arbitrary.) These formulas are called constituents of depth d. (For somewhat obscure historical reasons, what intuitively is the height of a tree is here called its depth.) Such constituents have the structure of a tree with constituents of more parameters serving as labels of its nodes, as indicated by (2.2).

If repetitions of conjuncts and disjuncts are not allowed in (2.2), constituents are minimal in the partial order of "denser than" .

The basic relationships between the structures (formulas) so defined are easily seen from the way they are defined

Lemma 2.1 Every c-constituent is a protoconstituent

Lemma 2.2 Every protoconstituent is consistent (satisfiable)

Proof: By construction it is satisfied in the model from which it is formed.

Lemma 2.3 Every c-constituent is satisfiable

Proof. From lemmas (2.1)-(2.2)

Not all constituents are satisfiable, however. The question as to which of them are satisfiable is crucial in the theory of constituents.

Lemma 2.4 Every c-constituent is a constituent

Lemma 2.5 A constituent is satisfiable if and only if it is a c-constituent.

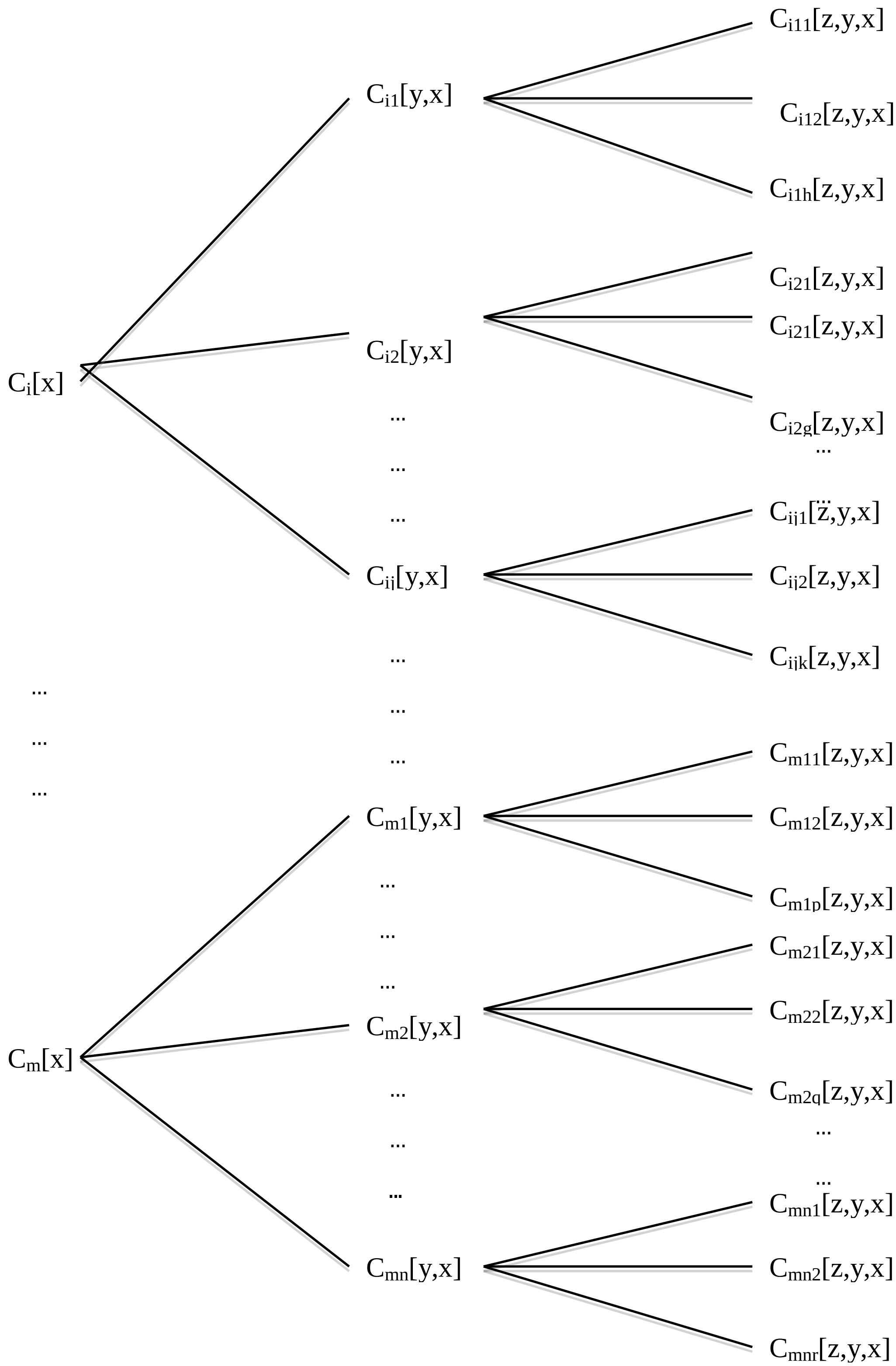
Proof: "If" part follows from *Lemma 2.3*. the "only if" part is seen by considering a model M of the constituent C in question. From M we can form C as one of its c-constituents.

Theorem (Exclusivity Theorem): If two instantiated constituents are satisfied in the same model, they are identical.

Proof: They can both be formed from the same model M by the same procedure.

The structure of a constituent is illustrated by Fig. 1. It is more complicated than is needed here because it will be used later to formulate more complex arguments.

[Fig. 1]


C_i[x]
C_m[x]
C_i1[y,x]
C_i2[y,x]
C_ij[y,x]
C_m1[y,x]
C_m2[y,x]
C_mn[y,x]
C_i11[z,y,x]
C_i12[z,y,x]
C_i1h[z,y,x]
C_i21[z,y,x]
C_i21[z,y,x]
C_i2g[z,y,x]
C_ij1[z,y,x]
C_ij2[z,y,x]
C_ijk[z,y,x]
C_m11[z,y,x]
C_m12[z,y,x]
C_m1p[z,y,x]
C_m21[z,y,x]
C_m22[z,y,x]
C_m2q[z,y,x]
C_mn1[z,y,x]
C_mn2[z,y,x]
C_mnr[z,y,x]

Theorem: Two different constituents of the same depth occurring as conjuncts of the same conjunction (in a consistent constituent) say

(2.4) $C_1^{(d)}[x, y_k, y_{k-1},..., y_1]$

(2.5) $C_z^{(d)}[z, y_k, y_{k-1},..., y_1]$

Cannot be satisfied by the same individual.

Proof: If they could, we consistently could add to the conjunction of (2.4) and (2.5) the identity (x=z). But then we could simply identify x and z and obtain a denser protoconsituent. This would contradict the definition of a c-constituent.

3. *Properties of constituents*

The role of constituents in model theory is illustrated by the following theorems and lemmas:

Theorem 3.0: An infinite sequence of deeper and deeper compatible constituents without parameters defines a complete theory.

Theorem 3.1: An infinite sequence of deeper and deeper compatible constituents with one free variable defines a type (in the model-theoretical sense).

Theorem 3.2: Any formula with ≤d layers of quantifiers and with the free variables $y_1, y_2, \ldots, y_k$ is equivalent to a disjunction of disjuncts of the form

(3.1) $C_i^{(d)},[y_k, y_{k-1},..., y_1]$

When the depth is increased to d+1, each disjunct (   ) is split into a disjunction of similar constituents but with depth d+1. Each constituent in (   ) is split into a disjunction of constituents, say

(3.2) $C^{(d-1)}[x, y_k, y_{k-1}, ..., y_1]$

is split into a disjunction of constituents of the form

(3.3) $C_j^{(d)}[x, y_k, y_{k-1}, ..., y_1]$

and so on.

Constituents have many interesting relations to each other and many symmetry properties.

Lemma 3.1 If we omit from a constituent (2.3) one layer of quantifiers, together with all atomic formulas (and identities) containing variables x_j bound to them plus all connectives that thereby become idle, we obtain a unique constituent of the form

(3.4) $C^{(d-1)}[x, y_k, y_{k-1}, ..., y_{j+1}, y_{j-1}, ..., y_1]$

This constituent is implied by (2.3).

Proof: follows from the construction of a constituent

Clearly (3.4) is the only constituent of depth d-1 that is compatible with (2.3).

In the direction of greater depth constituents split into disjunctions of deeper constituents. We can also omit from a constituent $C_i^{(d)}[x, y_k, y_{k-1}, ..., y_1]$ all atomic formulas (and identities) involving y, together with all connectives that are thereby made idle and obtain a constituent of the same depth d but without the parameter y_j. This constituent will be referred to as

(3.5) $C_i^{(d)}[x, y_k, y_{k-1}, ..., y_{j+1}, (y_j), y_{j-1}, ..., y_1]$

This notation can be generalized in an obvious way. The result (3.3) is implied by the original.

It follows that two constituents $C_m^{(d)}$ and $C_n^{(d+e)}$ are compatible (satisfiable in the same model) only if the latter is an extension of the former. For if $C_n^{(d+e)}$ is an extension of $C_m^{(d)}$, it obviously implies it.

One crucial question here is: When is a (formally defined) constituent consistent, that is, obtainable as a c-constituent from some model or other? The beginning of an answer is obtained by considering how two parallel constituents occurring as subconstituents of a larger one must be related to each other. A moment's reflection shows that the following two conditions must be satisfiable (see Fig. 1):

Compatibility condition: For any $C_i[x]$ and $C_m[x]$ there must be j and n such that

(3.6) $$C_{ij}[y,x] = C_{mn}[x,y]$$

Here all four constituents may depend on the same further parameter variables so that (3.2) can be generalized into

$$C_{ij}[y,x,z_k, z_{k-1}, \ldots , z_1] =$$

$$C_{mn}[x,y,z_k, z_{k-1},\ldots, z_1]$$

The structure of a constituent can be illustrated by Fig. 1, in more than one way. The branches of the tree are thought of as being all continued to some finite length (depth). The entire figure can be thought of as a displaying the structure of the constituent C_0 without variables

(3.6) $$(\exists x)C_1[x] \& (\exists x)C_2[x] \& \ldots \& (\exists x)C_m[x] \& \ldots \& (\forall x)(C_1[x] \vee C_2[x] \vee \ldots \vee C_m[x]).$$

The same figure can be thought of as illustrating the structure of the part of constituent that is above a given node $C_0[w_1, w_2, ...]$. Then we must think of all the constituents in (3.1) and in Fig. 1 as having suppressed additional arguments $w_1, w_2, ...$

Constituents have an important property which will be called exclusivity.

Theorem 3.3 (Exclusivity Theorem): If two constituents (i.e. minimal proto-constituents of the same depth) are satisfied in the same model, they are identical.

Proof: Let the proto-constituents C_1 and C_2 be satisfied in the same model M. When they are formed from M, the only reason why they might differ are repetitions of conjuncts or disjuncts. But if there are such repetitions, C_1 and C_2 cannot be both minimal.

Completability condition

For any $C_{ij}[x]$ as in Fig.1, there must be m and n such that (3.2) holds.

Here further variables may likewise be present. These two requirements are basic in the theory of constituents and distributive normal forms. (See Hintikka 1953.) The requirements do not alone guarantee consistency, however. For consistency, a stronger condition must be satisfied.

Consistency theorem. A formally defined constituent is consistent iff it has an infinite sequence of deeper and deeper extensions all satisfying the compatibility and completability conditions.

Constituents can serve many of the same purposes as the notions of the more familiar model theory. For instance, a sequence of increasingly deeper compatible consistent constituents (extensions of the previous ones) with no argument variables $C^{(d)}_{i(d)}(d = 1, 2...)$. A sequence of increasingly deep compatible constituents with one (and the same) free variable defines a type in the usual model-theoretic sense.

4. *Infinite constituents*

The notion of constituent and many of the properties can be extended to countably infinite constants. (The qualification “countable” will be omitted in the rest of this paper.) Such constituents are obtained from countably infinite sequences of individuals in the same way as finite constituents are obtained from sequences of individuals of a fixed finite length d. The main difference is that the nodes of the tree structures are now labeled by sequences of deeper compatible finite constituents rather than single finite constituents. If infinite conjunctions and disjunctions are allowed constituents will be infinite formulas analogous to finite constituents. No infinitary logic is nevertheless presupposed here.

Some of the properties of constituent are made simpler by the admission of infinite sequences. For instance, omitting a layer of nodes does not yield a compatible constituent of a lower depth, but the same (infinite) constituent. Likewise

Lemma : If $C[y,x]$ occurs in $C_1[x]$, then the constituent $C[x,y]$ occurs in $C_2[(x),y]$.

The structure of countably infinite constituents can be illustrated by the same Fig. 1 as in the finite case. The main difference, besides the infinity of the index sets, is that the C’s are now, not finite constituents, but countable sequences of deeper and deeper (finite) constituents. Furthermore, the branches of the tree now grow to (countable) infinity.

The most striking thing about infinite constituents is that the same structure, possibly with permutations, repeats itself everywhere. In the finite case, these repetitions were limited by the fact that a constituent with a smaller height d could reproduce only partially the structure of a constituent with a greater height e>d. This repetitive structure is manifested in different ways. Among the relevant properties of constituents there are the generalizations of the Compatibility and Completability Requirements to the infinite case. (Cf. Fig.1.)

Unlike what was found about finite constituents, these two requirements suffice, when applied to all the subconstituents of an infinite constituent, to guarantee its consistency

Again, consistent infinite constituents are precisely the ones that can be constructed as minimal proto-constituents from some model M. Intuitively, Compatibility and Completability Requirements together guarantee that the claims of the next individuals in forming the different ramified sequences of individuals all are made from the same store of individuals.

Thus, a consistent constituent is precisely one that can be formed as a minimal proto-constituent from a model M. Conversely, the models of C are the structures from which it can be so formed.

There are other basic relationships connecting different constituents with each other. From any constituent $C_0[y,x]$ with the arguments y, x we can form a constituent with x as its only argument by omitting all uses of y in $C_0[y,x]$. Formally speaking, we omit all atomic formulas containing y, all occurrences of & and $\vee$ that thereby become idle, and all quantifiers with y as its variable. The result is referred to as $C_0[(y),x]$. This notation can be extended to all sequences of free variables as arguments of constituents in the obvious way. Using this notation, other properties of constituents can be formulated.

Lemma 4.1. For any $C_{ij}[y,x]$, there is precisely one m such that $C_{ij}[y,(x)] = C_m[y]$.

This can be generalized to any sequence of arguments.

Lemma 4.2. For any $C_i[x]$ and $C_m[x]$, there is at least one j such that $C_{ij}[y,(x)] = C_m[y]$.

One useful property of infinite constituents is the following criterion of satisfaction:

A sequence $a_1, a_2, ...$ of individuals is compatible with the complete theory defined by an infinite constituent C_0 iff there is in C_0 a branch where each (4.1) is satisfied by $a_k, a_{k-1}, ..., a_2, a_1$.

Thus every sequence of individuals in a model of C_0 satisfies a branch in its infinite constituent and conversely every sequence of individuals satisfying such a branch is a part of a model of C_0.

Another pair of obvious facts about finite constituents can be extended to infinite constituents:

Lemma 4.3. A set of individuals $I = \{e_i\}$, with $i \in I$, constitutes a model of an infinite constituent C_0 iff the following conditions are satisfied:

Assume that Fig. 1 represents a constituent that is one of the approximations in C_0, with each constituent in Fig. 1 containing as parameters the same sequence *a,b,c,…* of members of *I*. Then

(i) For each individual $d \in I$ there is m such that d satisfies $C_m[d]$ (Universality requirement.)

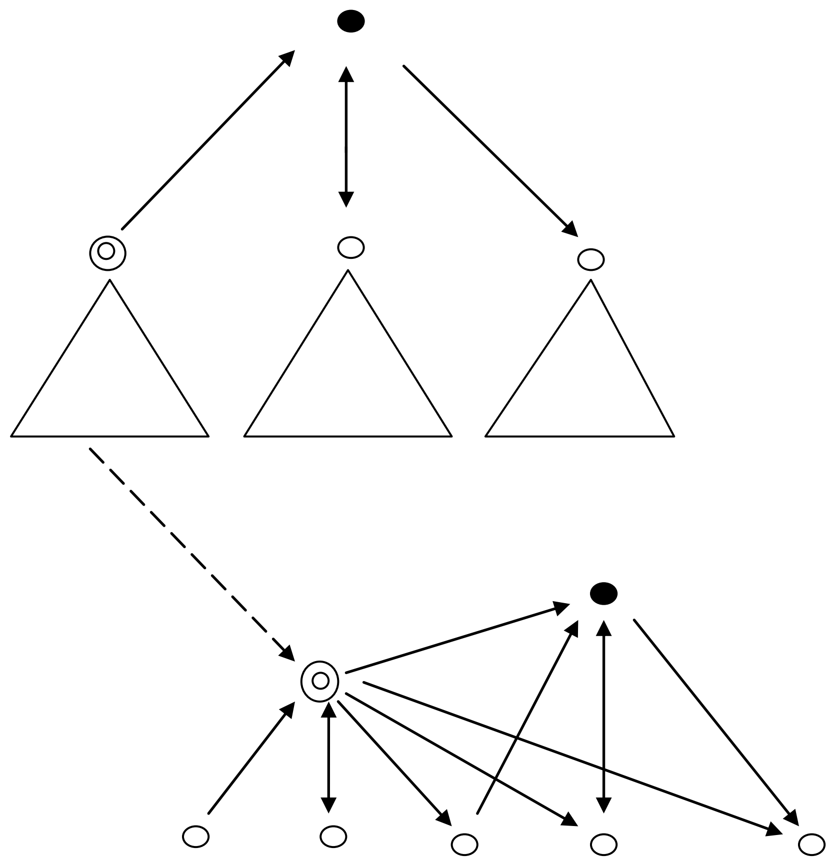
(ii) For each $C_m^m[x]$ there is an individual $d \in I$ that satisfies $C_m[d]$ (Existence requirement.)

It is important to appreciate what infinite constituents can and cannot express. Assume (as we have done) that our only nonlogical constant is an order relation. Then a consistent infinite constituent has an infinite descending sequence of individuals iff its models have the same property. Hence we can express well-ordering by means of an infinite constituent.

However, an infinite constituent does not require that each of its branches is instantiated in its models. It only requires that each initial segment of each branch is instantiated, which is tantamount to requiring that each node is reached by at least one instantiated branch. Hence an indefinitely branching branch need not be instantiated in a model of a consistent infinite constituent.

These observations throw some light on the nature of the continuum and the ways its structure can be specified logically. Suppose we have a linear continuum without end points defined by means of an ordering relation. The infinite constituent representing it is an indefinitely branching binary tree like that in Fig. 2. (Only a part is indicated.) But a model of such a constituent is a continuum only if all its branches are instantiated, in other words that the model in question is maximal. But this requirement cannot be enforced by means of a (countably) infinite constituent. In order for the model to form a continuum, all its branches must be instantiated.

Fig. 2



Similarly, consider the structure of the second number class Γ, and the infinite constituent C_γ true in it. This constituent imposes a linear ordering on each of its models M and also impose a well-ordering on M. Hence the models M of C_γ are well-orderings. However, C_γ is uniquely determined by the requirement that it is the totality of such well-ordering.

5. *Branches and Trees*

Hence the branching structure of constituents can be expected to repay further examination. By a branch of a given infinite constituent C_0 we mean a sequence of compatible infinite constituents in it:

(5.1) $C_1[x_1], C_2[x_2, x_1], C_3[x_3, x_2, x_1], \ldots$

Such branches have many interesting properties. Instantiated branches are in some sense building blocks of models. Not all branches of C_0 are instantiated in every model of C_0. However, each branch is compatible with C_0.

Lemma 5.1. Each initial segment of a branch of C_0 is instantiated in every model of C_0.

Lemma 5.2. Given a model M_1 of C_0 with the domain {a_1, a_2, …} and a branch (5.1) of C_0, there is a model M_2 of C_0 with the domain {a_1, a_2, … b_1, b_2, …} (it is not assumed that $a_i \neq b_j$ always) in which the branch is instantiated, i.e. the following are all there

(5.2) $C_1[b_1], C_2[b_2, b_1], \ldots$

The structure formed by b_1, b_2,… is a substructure of M_2. In brief, each branch of C_o can be satisfied by adding new individuals to any one of its models.

Intuitively, a branch is just that: a sequence of compatible constituents with variables (parameters) constituting a branch in the tree formed by the overall constituent C_o . Formally speaking, this means that

(5.3) $(\exists z)C_{i+1}[z, y_i, x_{i-2},..., x_1]$

occurs as a conjunct in

(5,4) $C_i[x_i, x_{i-1},..., x_1]$

We can form another sequence by as it were tracing an individual through in relation to a sequence of arbitrarily selected nodes of the constituent tree, for example from $C_i[x]$ in fig. 1 to $C_{mi}[y,x]$ to $C_{mni}[z,y,x]$ and so on.

If the resulting sequence is

(5.5) $$C_1[x_1], C_2[y_1, x_2], ..., C_j[x_1, x_2, ..., x_j]...$$

then

(5.6) $$(\exists z)C_{j+1}[x_1, x_2, ..., x_j, z]$$

must occur as a conjunct in

(5.7) $$C_j[x_1, x_2, ..., x_j]$$

But it is immediately seen that traces and branches are the same sequences, except for a different order of the parameter variables. Considering branches as traces nevertheless helps to see some of their characteristic properties. For instance, we can immediately see that the following holds

Lemma 5.1. A given individual a can satisfy only one trace (5.5)

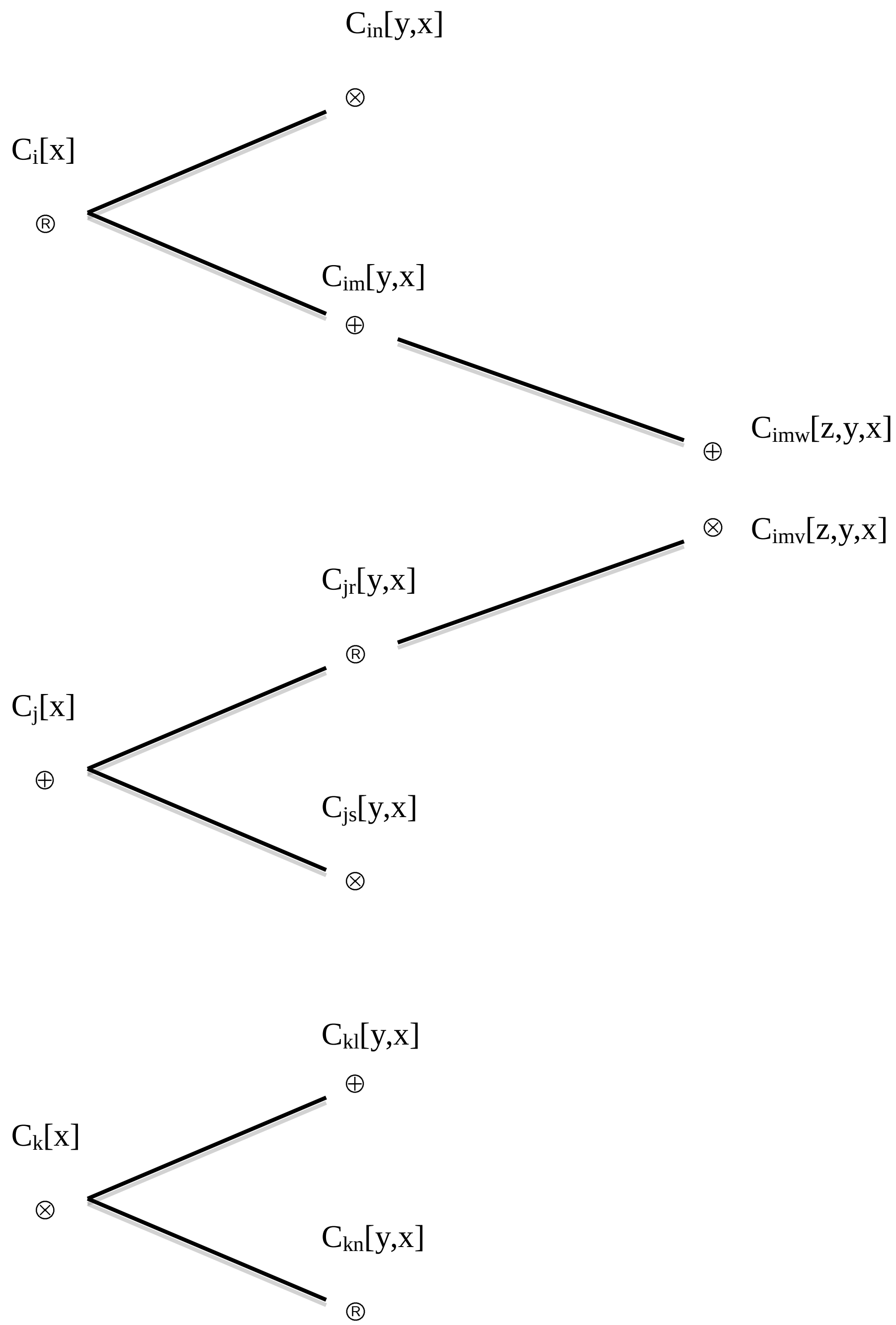
By satisfying (5.5), we mean there is a value of x_1 in it

6. *Perfect Branches*

One fundamental question here is: Can we impose maximality conditions on branches of constituents and not only on entire constituents? It can be shown that there are branches (5.1) such that the class of individuals $a_1, a_2, ...$ satisfying it satisfies the entire constituent. Such branches will be called perfect.

What is required of a branch to be perfect? An answer is provided by Lemma 4.3. It requires that at each depth two conditions are satisfied, the universality condition and the existence condition.

Fig. 3



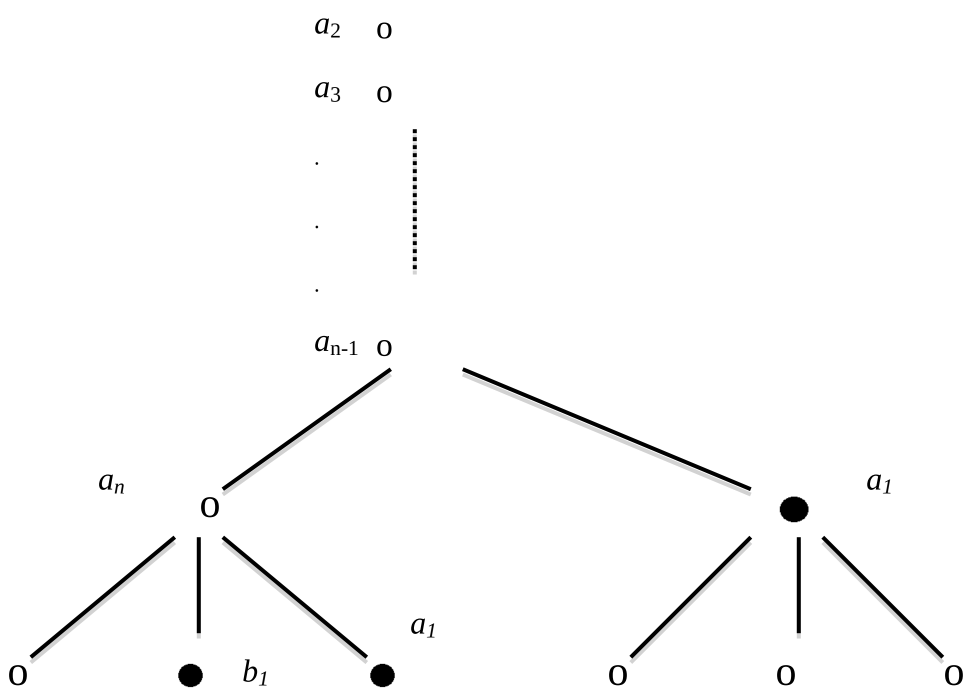
It is obvious that the successive formulas of any branch of any infinite constituent *C* are compatible with *C*. This means that all the universality requirements codified by *C* are satisfied. Hence it suffices, in order to show the existence of perfect branches, to see what happens when we try to construct one by one a branch whose successive members satisfy as many existence requirements as possible. For the purpose, consider Fig. 3. It can be taken to represent a constituent without free variables. However, only some subconstituents are indicated. An argument will be presented in terms of this figure.

Assume that $C_i\,[a_1]$ is the first member of some (instantiated) branch *C*. This means that $C_i\,[x]$ is instantiated in *C*. However, any other constituent $C[x]$ must also be instantiated in *C*. For a given $C_j[x]$, this can be guaranteed by choosing as the second member of *C* a constituent $C_{im}[y,x]$ such that $C_{im}[x,y]$ is identical with $C_{jm}[x,y]$. Such a constituent exists because the completability requirement is satisfied. Since the second member of *C* is $C_{im}[y,x]$, then some a_2 instantiates $C_{im}[a_2,a_1]$ and hence we have $C_{im}[a_2,(a_1)] = C_j[a_2]$. Hence $C_j[x]$ is also satisfied in the branch in question.

Likewise, by choosing the third member of the branch appropriately we can make sure that any third constituent $C_k[x]$ is instantiated in the branch *C*.

This can be generalized to any countable number of constituents $C[x]$ with one free variable. They can be satisfied by a countable number of choices of members of a branch. We cannot always satisfy all of them, for they may have an uncountable cardinality. But by a countable number of choices we can satisfy all the finite approximations to the infinite constituents $C_{\aleph}[x]$. This leaves a countable number of choices to be used for other purposes

Fig. 4



In the same way we can by a countable number of choices make sure that the individuals $b_1, b_2, ...$ substituted for $x_1, x_2, ...$ in (5.1) satisfy every approximation to each member of the sequence (5.1). This suffices to show that $a_1, a_2, ...$ satisfy all the formulas in the branch (5.1).

This proves the following:

Perfection Theorem: In every infinite constituent there is at least one perfect branch.

This can immediately be extended:

Corollary 6: Every initial segment of a branch is compatible with a perfect branch.

In the same way, we can speak of perfect traces. Since an individual can satisfy (as its initial member) only one trace, the same holds for branches.

The existence of perfect branches is thus virtually obvious on the basis of cardinality considerations. Yet it is a highly nontrivial conclusion. Their existence in the preceding line of thought is guaranteed by a sequence of complex applications of the axiom of choice. There does not seem to be any reason to expect that it could for instance be justified by the form of the axiom of choice that is built into the usual axiomatizations of set theory. A function that tells how far in a perfect branch we have to go in order to find the satisfying *a* given constituent number *m*, grows extremely rapidly.

Yet the existence of perfect branches can be proved by a seemingly much simpler argument. An infinite constituent C is equivalent to (specifies the same models as) a countable sequence of finite constituents. If it is satisfiable, by Skolem-Löwenheim theorem it is satisfiable in a countable model, say in a model with the domain $[a_1,a_2,\ldots]$. Then a_1, a_2, … satisfy a branch of C, which shows the existence of perfect branches.

An interesting result is obtained by noting that (5.1) can be transformed by a permutation of individuals into an equivalent sequence of the form

(6.1) $$C_1[x_1],\ C_2^*[x_1,x_2],\ C_3^*[x_1,x_3,x_2],\ \ldots$$

Hence a perfect branch of a constituent specifies how each of the individuals in the branch finds a slot in all the different constituents occurring in the branch.

It is seen immediately that the following results hold:

Lemma 6.1. An individual can satisfy at most one perfect branch as its first element.

Lemma 6.2. The cardinality of a model M_0 of an infinite constituent C_0 is the same or larger than the cardinality of perfect branches of C_0 instantiated in M_0.

7. *Second number class*

By means of these model-theoretical results, CH can be proved fairly easily. CH is equivalent with the statement that says that the cardinality of the second number class *Γ* is $2^{\aleph_0}$. The second number class is the structure formed by the set of all countable ordinals. For its properties, see any treatise of general set theory, for instance Kuratowski and Mostowski (1970). The structure of *Γ* and by implication the structure of *C*(*Γ*) are easily described. In *Γ* there is a countable initial segment μ_0 consisting of the definable elements of *Γ*. The rest is a well-ordered sequence $\mu_1 + \mu_2 + ... + \mu_\omega + ...$ (ω is an arbitrary countable ordinal) of segments of a countable number of elements. Within each μ_ω, each of its members is definable by reference to any one of them but not definable by reference to any one element or finite set of elements in other segments

The second number class has itself the structure of an ordinal, viz. the smallest uncountable ordinal. Its being uncountable means that the structure of each countable ordinal is a substructure of *Γ*. We will call this the Comprehensiveness Requirement. In order to prove CH, it thus suffices to show that every ordinal satisfying the Comprehensiveness Requirement has the cardinality of at least $2^{\aleph_0}$.

Consider now the infinite constituent *C*(*Γ*) true in *Γ*. It can be formed from *Γ* as described in sec. 4 above. One of its models is *Γ*, but it has other models *M* as well. What are these models like? There are no infinite descending chains in *Γ*, hence none in *C*(*Γ*) or in *M*, either. Moreover, *M* must be a linear ordering in which each element has an immediate successor. Hence the models *M* are simply infinite ordinals. All smaller ordinals are initial segments *G* of *M*, and each initial segment of *M* satisfying *C*(*Γ*) is an ordinal smaller than *G*. The infinite constituent *C*(*Γ*) has a number of perfect branches B. If B is not satisfied in a given model M of $C(\Gamma)$, we can according to Lemma 5.2 extend M so as to make it satisfied by adding to its domain $D(M)$ a countable number of individuals $A=\{a_1, a_2, \ldots\}$ in such a way that $a_1, a_2, \ldots$ satisfy B and together with the old elements also $C(\Gamma)$. Both $C(B)$ and M have the structure of ordinals smaller than M^*. If $D(M)$ is countable, then so is $D(M) \cup \{a_1, a_2, \ldots\}$. Since Γ is larger than any countable ordinal, a model M of $C(\Gamma)$ is Γ only if all perfect branches of $C(\Gamma)$ are satisfied in it. As is illustrated by Fig. 4, different perfect branches go together with different paths through the tree that is $C(\Gamma)$ in the sense indicated by the figure. If there is a subset of such paths each of which keeps splitting up indefinitely, there will be $2^{\aleph_0}$ perfect branches in $C(\Gamma)$. Each of the paths is satisfied by a different individual as its first member, according to Lemma 5.1. If all of them are satisfied in a model M, then the cardinality of $D(M)$ is at least the same as the number of different perfect branches. This number is determined by the way perfect branches split. If they keep on splitting indefinitely, this number will be $2^{\aleph_0}$. Hence, in order to show that the cardinality of Γ is $2^{\aleph_0}$ it suffices to show that the paths characterizing perfect branches with undefinable first members are always splitting, as in Fig. 4.

In order to see that such indefinite splitting happens with each perfect branch of $C(\Gamma)$, consider once again Fig. 1.

Let us assume that Fig. 1 depicts the beginning of $C(\Gamma)$. Consider two constituents $C_i[x]$ and $C_m[x]$ in $C(\Gamma)$ neither of which defines x in Γ. If $C_i[a]$ and $C_m[b]$, is the order $a \geq b$ or $b \geq a$ fixed? The answer is: never.

In order to prove this answer, note first that among the ordinals x in Γ, there are arbitrarily large a, b, c, in any order and arbitrarily far apart such that $C_i[a]$, $C_m[b]$, $C_m[c]$, a $\geq$b, $c \geq a$.

This can be proved by the following argument:

Let a satisfy $C_i[x]$. Then there is for each e an ordinal *o(e)* such that there are arbitrarily large ordinals b_e satisfying the approximation to $C_m[b]$ at depth c, that is, satisfying.

(7.1) $$C_m^{(c)}[b_e]$$

For if that is not the case, there is the least ordinal r beyond which there are no ordinals satisfying (7.1). This would define r, which is impossible because it would hope to be a member of the initial sequence of definable individuals and yet greater than some undefinable ordinal. But this is impossible because it was assumed that a, b, c are larger than all definable ordinals

According to the properties of the second number class, $b_0 = \lim_e(b_e)$ exists and is a member of Γ. Hence beyond b_0 there are arbitrarily large ordinals b such that $C_m[b]$, $b>a$.

By reversing the argument we can see that there are arbitrarily large a, b, c, such that $b>a$, $a>c$, $C_i[a]$, $C_m[b]$, $C_m[c]$.

In terms of C(Γ) this indeterminacy of the order of a and b means that in $C(\Gamma)$ there are two different perfect traces emanating from an x satisfying $C_i[x]$ and splitting from each other at the very first step. In the second member of one of them we have in the one case $C_{m,n}[x, y]$ = $C_{ij}[y,x]$ for suitable j, n and with $x>y$ and m,n in the other case with $y<x$ and hence $C_{ik}[y,x]$ =

$C_{mp}[x,y]$ for suitably different $k \neq j$ and $p \neq n$. Intuitively, in one perfect trace $C_i[x]$ is extended to $C_{mn}[x,y]$ and on the other one to $C_{mp}[x,y]$.

Essentially the same argument applies at any stage of a perfect trace. The argument can again be formulated by reference to Fig.1 except that we have to assume that each of the constituents depicted there contains a finite number of "hidden variables" $h_1, h_2, \ldots, h_d$ with specified relations to each other. Let us assume that a perfect trace has been continued all the way to $C_i[x] = C_i[x, h_1, h_2, \ldots, h_d]$. Here $C_i[x]$ determines between which of the $h_1, h_2, \ldots, h_d$ x is located. However, otherwise the order of an x with $C_i[x]$ and one with $C_m[x]$ is not determined.

From the properties of the second number class it follows that the ordinals satisfying $h_1, h_2, \ldots, h_d$ can be arbitrarily large and arbitrarily far apart from each other. The argument is similar to the argument in the case of the $h_1, h_2, \ldots, h_d$. Let these two constituents be $C_i[x] = C_i[x, h_1, h_2, \ldots, h_d]$ and $C_m[x] = C_m[x, h_1, h_2, \ldots, h_d]$. By the same argument as in the case d=0 it can be shown that the perfect trace can be continued from $C_i[x]$ to a constituent of the form $C_{m,n}[x,y]$ in two different ways. This shows that perfect traces [branches] always split, therefore the number of perfect traces [branches] is $2^{\aleph_0}$. According to Lemma 6.1 this suffices to show that the cardinality of Γ is $2^{\aleph_0}$. This completes the proof for CH. A different (but related) argument for the same conclusion is given in Hintikka (2009). The guiding ideas of both arguments go back to Hintikka (1993 and 2004).

8. *CH and maximality*

According to his biographers, Gödel conjectured that the key to a proof or disproof of CH would be some kind of maximality assumption not unlike Hilbert's Axiom of Completeness. (For this axiom, see Hilbert 1899, Baldus 1928 and Freudenthal 1957.) If this conjecture is correct, the difficulty of proving CH could perhaps be understandable. The way a usual logical language specifies a class of models (structures) is to impose different conditions on them one by one. In contrast, the maximality of a model turns on comparisons between different models.

In view of Gödel's conjecture, it is of some interest to see what role the notion of maximality plays in the argument of this paper. Various maximality properties are indeed the methodological key to the argument presented here. Constituents codify maximally fine distinctions between different models and between different kinds of individuals in them, given certain restrictions on their structure. The notion of perfect branch involves an interesting kind of maximality.

The crucial instances of maximality in the present argument nevertheless concerned the maximality in the sense of maximal richness of the intended models of a complete theory. A model of dense linear order is continuous if and only if it is maximal. Likewise, the second number class can be said to have the cardinality of the continuum because it contains all countable well-orderings as its substructures. Gödel's conjecture thus seems to have been an inspired one.

But notice how the relevant kind of maximality entered the argument of this paper. It was neither proved nor assumed as any kind of general set-theoretical principle. It was an attribute of the structures that were being studied. Maybe there is a methodological moral to our story. Maybe the business of set theory is not to study an elusive "set-theoretical universe", whatever it is or may be, but simply to study certain particular and particularly interesting mathematical structures that can be called set-theoretical. To adapt a phrase of Wittgenstein's to the present context, maybe there is no set theory, but there certainly are set-theoretical problems.

References

Baldus, Richard, 1928, "Zur Axiomatik der Geometrie I: Über Hilberts Vollständigkeitsaxiom", *Mathematische Annalen* vol. 100, pp. 321-333.

Cohen, Paul J., 1966, *Set Theory and the Continuum Hypothesis*, W.A. Benjamin, New York and Amsterdam.

Frege, Gottlob, 1893-1903, *Grundgesetze der Arithmetik, begriffschriftlich abgeleitet*, Herman Pohle, Jena.

Frege, Gottlob, 1964, *The Basic Laws of Arithmetic*, translated and edited by Montgomery Furth (partial translation of Frege 1893-1903), University of California Press, Berkeley and Los Angles.

Freudenthal, Hans, 1957, "Zur Geschichte der Grundlagen der Geometrie, *Nieuw Archief voor Wiskunde* vol. 5, pp. 105-142.

Gödel, Kurt, 1940, *The Consistency of the Axiom of Choice and the Generalized Continuum Hypothesis with the Axioms of Set Theory*, Princeton University Press, Princeton. Reprinted in Kurt Gödel, *Collected Papers*, ed. by Solomon Feferman et al, vol. 2, Oxford University Press, New York, 1990, pp. 33-101.

Hilbert, David, 1899, *Grundlagen der Geometrie*, Leipzig. Second edition, 1903. Seventh edition, Stuttgart, 1930. French translation by L. Laugel, 1900.

Hintikka, Jaakko, 1953, *Distributive Normal Forms in the Calculus of Predicates* (Acta Philosophiica Fennica vol.). Societas Philosophia Fennia, Helsinki

Hintikka, Jaakko, 1965, "Distributive normal forms in first-order logic", in J.N. Crossley and M.A.E. Dummett, editors, *Formal Systems and Recursive Functions*, North-Holland, Amsterdam, pp. 47-90.

Hintikka, Jaakko, 1993, "New foundations for mathematical theories", in Juha Oikkonen and Jouko Väänänen, editors, *Logic Colloquium '90* (Lecture Notes in Logic vol.2), Association for Symbolic Logic, Urbana, Illinois, pp. 122-144.

Hintikka, Jaakko, 1996, *The Principles of Mathematics Revisited*, Cambridge University Press, Cambridge.

Hintikka, Jaakko, 1998, "Truth-definitions, Skolem functions and axiomatic set theory", *Bulletin of Symbolic Logic* vol. 4, pp. 303-337.

Hintikka, Jaakko, 2004(a), "Independence friendly logic and axiomatic set theory", *Annals of Pure and Applied Logic*, vol. 126, pp. 313-333

Hintikka, Jaakko, 2004(b), "What is the true algebra of logic?" in Vincent Hendricks et al, editors, *First-Order Logic Revisited*, Berlin, Ontos Verlag, pp. 117-128

Hintikka, Jaakko, 2009, "Continuum problems and the structure of the second number class", *Synthese* (forthcoming).

Hintikka, Jaakko, forthcoming, "How to define (self-applied) truth: Why Tarski could not do it."

Hintikka, Jaakko, forthcoming, "Past, present and future of set theory"

Kuratowski, K., and A. Mostowski, 1976, *Set Theory*, North-Holland, Amsterdam.

Ramsey, Frank P., 1925, "The foundations of mathematics", *Proceedings of the London Mathematical Society*, Ser. II, vol. 25, pp. 338-384.

Rantala, Veikko, 1977, *Aspects of Definability* (Acta Philosophica Fennica vol. 29, nos. 2-3), North-Holland, Amsterdam.

Tarski, Alfred, 1935, "Der Wahrheitsbegriff in den formalisierten Sprachen", *Studia Philosophica* vol. 1, pp. 261-405. English translation in Tarski, *Logic, Semantics and Metamathematics*, Clarendon Press, Oxford, 1956, pp. 152-277.