Minds and Machines (2005) 15: 229–257
DOI 10.1007/s11023-004-6590-1 © Springer 2005

Canonical Naming Systems

LEON HORSTEN
Center for Logic, Philosophy of Science & Philosophy at Language, Kardinaal Mercierplein 2, 3000 Leuven, Belgium; E-mail: Leon.Horsten@hiw.kuleuven.ac.be

Abstract. This paper outlines a framework for the abstract investigation of the concept of canonicity of names and of naming systems. Degrees of canonicity of names and of naming systems are distinguished. The structure of the degrees is investigated, and a notion of relative canonicity is defined. The notions of canonicity are formally expressed within a Carnapian system of second-order modal logic.

Key words: canonical names, Carnap, modal logic

1. Introduction

There are many standard or canonical naming systems for the natural numbers. Here are a few:

- the binary notation system: 0, 1, 10, 11,100,101, ...
- the standard arabic numerals: 0, 1, 2, 3, ...
- the numerals of the language of Peano Arithmetic: 0, $S0$, $SS0$, $SSS0$, ...
- the roman numerals: [0], I, II, III, IV, ...

These notation systems share a striking property, which explains to a large extent why they are so useful to us: for any two names of such a standard notation system, we can at least in principle find out whether they denote the same natural number.

Consider the class of names of Turing machines that code countable recursive ordinal numbers. Recursion theory proves that there is no systematic way of telling for any two names of Turing machines whether or not they name the same ordinal number. For all we know, there may even be Turing machines T_1 and T_2, in fact denoting the same ordinal, but for which we will never come to know that they denote the same ordinal. In that sense, our best naming system for the recursive ordinals is less canonical and more opaque than our standard naming systems for the natural numbers.

There are naming systems for finite structures which may have an even lower degree of canonicity:

Example 1: Consider the naming system $S = \{1, a, b\}$, where a and b are defined as follows:

– $a =: 1$ if the Twin Prime Conjecture is true; $a =: 0$ otherwise;
– $b =: 1$ if $a = 0$; $b =: 0$ otherwise.

We do not presently know, and may never know, whether the Twin Prime Conjecture is true. Therefore we may never have any way of telling whether '*a*' and '1', or '*b*' and '1', name the same natural number. In this sense, *S* is not a very good naming system.

About the name *a* in Example 1, we know at least that it denotes either 0 or 1. It is somewhat harder to think of a putative example of an 'absolutely random' natural number presentation, i.e., a name *n* such that for *every* standard numeral *k,* we may never be able to find out that $n \neq k$ (the reader is invited to try before looking at the following example). Here is the best I have been able to come up with:

Example 2: Let α be the smallest ordinal number (>0) such that $2^{\aleph_{\alpha-1}} > \aleph_\alpha$ if there is such an ordinal number, and let $\alpha =: 0$ otherwise. Then define *n* as follows:

– $n =: 0$ if either the Generalized Continuum Hypothesis is true or $\alpha \geq \omega$.
– $n =: \alpha$ otherwise.

If, as some maintain, we will forever have to live with the fact that 'the Continuum Hypothesis may fail anywhere', then for every standard numeral *k*, we will never be able to tell whether $n = k$. If, as others say, we will (or have?) come to know that $2^{\aleph_0} = \aleph_2$, then *n* will not serve as an example of an absolutely arbitrary number.

It will be harder to come up with a putative example of an infinite naming system *S* such that each $a \in S$ uniquely names a natural number and each natural number has a unique name in *S*, and such that for every $a \in S$ and every standard numeral *k,* we will never be able to know whether $a = k$. In sum, progressively non-canonical concrete notation systems for the natural numbers appear to be increasingly hard to find.

In this paper, a systematic investigation into the logical properties of degrees of canonicity of naming systems is undertaken. Naming systems for sets of natural numbers serve as prime illustrations, and play a considerable heuristic role in the development of the theory. But in principle, the investigation is intended to be more general. Naming systems for all sorts of structures of objects, concrete or abstract, are intended to be in its scope.

The paper is structured as follows. In Section 2, the need for a theory of canonicity of naming systems is explained, and crucial and difficult conceptual questions that have to be answered in order to develop such a theory are discussed. In Section 3, an abstract logical framework of Carnapian second-order modal logic is sketched. Within this framework, the theory of canonical notation systems is then expressed (Section 4). Degrees of canonicity are isolated and proved to be distinct, and a concept of relative canonicity is defined. The class of standard Peano-numerals is investigated as an illustration of a highly canonical notation system. In Section 5, a notion of extreme opaqueness of presentations is described, and the relation with degrees of canonicity is to a small extent mapped.

The reader will notice that the present paper has a programmatic character. Many issues are touched upon only to be dropped long before they have been properly addressed or even before the central problems that are involved have been adequately formulated. Detailed elaboration of much that is brought up here is deferred to future work. At best, therefore, this paper can claim to be a sketch of a future theory of canonical presentations and degrees of canonicity.

2. The Problem of Quantifying into Provability Contexts

In mathematics, one often speaks of proving something *of* a natural number. For instance, one hears mathematicians say that so-and-so has proved of the number 29 that it is the only prime number with a certain property. But what does it mean to prove *of* a given number *that* it has a certain mathematical property? What does *de re* provability mean?

Problems of quantifying into intensional contexts have been discussed extensively in the philosophy of language.[1] In that context it has been argued that sense can be made of *de re necessity* (using a metaphysical notion of necessity). It has also been argued that sense can be made of *de re* knowledge and belief (using the notion of acquaintance). But it is harder to see how, in the context of mathematics, sense can be made of *de re* provability, unless one is willing to posit a notion of acquaintance with abstract objects such as natural numbers, and even then.

Strictly speaking, *de re* provability seems to be unintelligible. The things that can be proved are theorems, and theorems are complex interpreted linguistic expressions. This suggests that one can only prove things *of* interpreted expressions. For instance, one can prove *of* an expression that when it is appended to a certain other expression, it yields a provable sentence. That, in the end, seems all that proving something *of* a mathematical object that it has a property can possibly mean.

Mathematical logicians, in contrast to most philosophers and philosophical logicians, have recognized this fact ever since the work of Gödel on the incompleteness theorems. In mathematical logic, particularly in proof theory, the following sort of notation is ubiquitous. Let $\varphi(x)$ be a formula of the language of Peano Arithmetic, and let Bew_{PA} be the standard provability predicate for PA. Then (existentially) quantifying into the provability predicate is represented in terms of the *dot-notation* as

$$\exists x Bew_{PA}(\ulcorner \varphi(\dot{x}) \urcorner),$$

which is shorthand for

$$\exists x Bew_{PA}(\underset{.}{\text{sub}}(\underset{.}{\text{num}}(\text{x}), \ulcorner \text{x} \urcorner, \ulcorner \varphi(\text{x}) \urcorner)),$$

where sub̩ is a name in the language of Peano Arithmetic of the substitution function, and num̩ is a name in the language of Peano Arithmetic of the function that assigns to every number *n* the (code of the) Peano Arithmetic-numeral

$$S \ldots S0(n \text{ times}).$$

So in words, what the above expression means is: there is a number *n*, such that when its *standard presentation* $S \ldots S0$ replaces all free occurrences of the variable x in $\varphi(x)$, a theorem of Peano Arithmetic results. This, then, is a *loose* sense in which we can say that quantification into provability contexts *does* make sense. But observe that this only works because we have a 'standard' naming system for the natural numbers. In general, extensions to other domains are not obvious.

From the 1980s onward, so-called *epistemic formalizations of mathematical theories* were proposed and investigated.[2] These systems aim at giving an axiomatic description of the notion of *provability in principle* against the background of formalized mathematical theories. Such epistemic theories allow quantification into the scope of the sentential *operator* that expresses the notion of provability in principle. Therefore they are confronted with the problems of quantifying into intensional contexts. In this context too, the philosophical point made above must be conceded: *It is only possible to prove things of interpreted expressions.*

For the language of the standard theories of first-order epistemic *arithmetic,* an unobjectionable interpretation of quantifying in *can* indeed be given, along the gödelian lines sketched above. So in a sense, for these languages there is no real problem. Nevertheless, the philosophical point was not sufficiently *explicitly* recognized.[3] This has not been without consequences for the research program of epistemic formalizations of mathematical theories. The epistemic systems of set theory and of *higher-order* arithmetic that have been proposed are much less natural than Shapiro's system of epistemic first-order arithmetic.

From one point of view, the present paper can be seen as an attempt at carrying out a logical investigation of the notion of provability in principle which explicitly recognizes the philosophical point. In view of what was said above, a theory of provability in principle must necessarily involve a theory of (linguistic) *presentations of (mathematical) objects*. From another point of view – and this is the point of view that we will focus on in this paper – we will see that a theory of presentations of objects can be naturally formulated in a framework in which a notion of provability plays a central role.

In the following section, an abstract, general, but manageable framework for investigating *canonical naming systems* for structures of objects is described. Its expressive power will be seen to be large enough to allow the expression of many of the distinctions that are theoretically important. At the same time, it is not too cumbersome to work with and reason in.

3. Carnapian Modal Logic

> "I believe, however, that there is a simpler way to achieve [a solution to Quine's problem about quantification into intensional contexts]. It is similar to that of Church but avoids the use of two kinds of variables for the same type. This use is [...] an unnecessary duplication. It is sufficient to use variables of one kind which are neutral in the sense that they have classes as value-extensions and properties as value-intensions..."
>
> [Carnap, 1956, p. 195]

We will work in the framework of Carnap's systems of modal logic.[4] Carnap's systems of modal logic differ from the now common systems of modal logic in their interpretation. In Carnapian modal logic, *all* expressions (including the variables, as we will see) have different interpretations in extensional contexts and in intensional contexts (i.e. in the context of the modal operator). In extensional contexts, expressions have their usual denotations. In intensional contexts, they denote their usual intension or meaning.[5] This can be called, using an expression of Kripke, *Carnapian double talk*.

It is known that even first-order Carnapian modal languages have a high degree of expressive power.[6] We will use a second-order Carnapian language to formulate the highly intensional distinctions that are needed in the theory of canonical presentations. At the same time, Carnapian logics are natural to reason in. We will see that the quantifier laws retain their usual form, and we have full second-order comprehension. Some care is needed in the formulation of the laws of identity.

3.1. CARNAPIAN LANGUAGES

3.1.1. *Syntax*

We opt for *second-order* Carnapian languages of modal logic. Such languages contain individual variables $x, y, \ldots$, second-order variables $X, Y, \ldots$, individual constants $a, b, \ldots$, function symbols $f, h, \ldots$, predicate constants $P, Q, \ldots$. Beside the usual second-order logical vocabulary, in which we include the identity symbol $=$ as a primitive predicate, these languages also have an intensional sentential operator $\Box$. The operator $\Box$ will be interpreted as "It can be established in the actual world by person K that", "It is verifiable in the actual world by person K that", or "It can be shown in the actual world by person K that". Here the person K, our nondescript epistemic agent, will be kept fixed throughout.

3.1.2. *Semantics*

The distinguishing feature of Carnapian modal languages is that in extensional contexts, expressions denote their usual extensions, whereas in intensional contexts (i.e., in the scope of $\Box$), expressions denote their intensions. This holds even for individual constants and variables. For example, in the expression:

$$a = b \rightarrow \Box(a = b),$$

the first occurrence of a denotes an object (the reference of a), whereas the second occurrence of a denotes an intension (the 'meaning' of a). Semi-formally, we will let an expression φ *always* denote an ordered pair $\langle o, p_o \rangle$, consisting of an object o and a presentation p_o of o. In extensional contexts, the presentations do not matter for determining truth-values; in intensional contexts, the presentations do matter.

This naturally leads to a notion of *models* for Carnapian languages. A model $\mathscr{M}$ for a Carnapian language $\mathscr{L}_C$ is determined by a universe of objects D, a collection of presentations (of objects, sets, relations) taken from a language $\mathscr{L}$, a collection of sentences S of this language $\mathscr{L}$, and assignments to expressions of $\mathscr{L}_C$. The elements of S make up the extension of $\Box$ in $\mathscr{M}$: the elements of S are the sentences the truth of which our epistemic agent K is able to establish. Truth in a model, validity, consequence are then defined on the basis of such a notion of models for Carnapian languages.

This semi-formal way of explicating the semantics of Carnapian languages of course begs for mathematical precision. We will not provide it here.[7] We will, in the sequel, semi-formally sketch some models for Carnapian languages. We trust that the reader is able to semi-formally interpret sentences of Carnapian languages in such models. Here are a few simple examples of how such readings go:

Example 3: Consider the formula $x = a$. In a given model, the variables x and a denote ordered pairs $\langle o, p_o \rangle$ and $\langle o', a \rangle$ respectively. Since in this formula, both x and a occur in extensional contexts, it is true (in the model) if and only if $o = o'$.

Example 4: Consider the formula $\Box(x = y)$. In a given model, the variables x and y denote ordered pairs $\langle o, p_o \rangle$ and $\langle o', p_{o'} \rangle$, respectively. x and y occur within the scope of $\Box$. The formula $\Box(x = y)$ is therefore true (in the model) if the second member of the denotation of x (in the model), concatenated with $=$, concatenated with the second member of the denotation of y, which results in the identity statement $p_o = p_{o'}$, is verifiable, i.e., if this sentence is in the extension of $\Box$.

Example 5: The formula $\exists x \Box (x = a)$ is true if there is a presentation p_o of some object o such that $p_o = a$ is a verifiable sentence.

Example 6: The formula $\exists x \Box \exists y (x = y)$ is true if there is a presentation p_o of some object o such that $\exists y (p_o = y)$ is verifiable.

3.1.3. *Extensionality*

> "Propositions in which a function ϕ occurs may depend, for their truth-value, upon the particular function ϕ or may depend only upon the *extension* of ϕ. In the former case, we will call the proposition concerned an intensional function of ϕ; in the latter case, an extensional function of ϕ [...] the mark of an extensional function f of a function $\phi ! \hat{z}$ is
>
> $\phi ! x . \equiv_x \cdot \psi ! x : \supset_{\phi, \psi} : f(\phi ! \hat{z}) . \equiv \cdot f(\psi ! \hat{z})$."
>
> [Russell and Whitehead, 1980 [1910], p.187]

It is possible to to give a *definition* of the notion of *extensionality* of predicates in the language of Carnapian logic:

Definition 1: $EXT(F) \equiv \forall x [Fx \leftrightarrow \forall y (y = x \rightarrow Fy)]$.

Note that for this definition to express a *nontrivial* condition on predicates F, it is crucial that the principle

$$x = y \rightarrow \Box (x = y)$$

is *not* generally valid. In most standard systems of modal logic, which can be broadly called *Kripkean,* this principle does come out valid. In these systems, one can define a distinction between extensionality and intensionality of *higher-order* predicates (in the Russell–Whitehead way), but not of first-order predicates.[8] In Carnapian languages, in contrast, the substitution principle is not in general valid. So here the extra expressive power of Carnapian languages becomes visible.

We can also define the notion of the *extensional completion* $Q(x)$ of a (one-place) predicate $P(x)$ in the following way:

Definition 2: $EXTC(P) \equiv \exists y (y = x \wedge Py)$.

In a similar fashion, the extensional completion of an n-place predicate can be defined. For instance, the extensional completion of a *two-place* predicate $P(x, y)$ is defined as

$$\exists z \exists u(z = x \wedge u = y \wedge P(z, u)).$$

It is important to realize, in this context, that extensionality is *not* the same as transparency. Take, for instance, the predicate P_x, defined as: $x = 1$ if Goldbach's conjecture is true; $x = 0$ otherwise. Then $EXTC(P)$ is *by definition* extensional, but it is by no means transparent.

In the sequel, $x \in S$ will often be used instead of $S(x)$, and $S \subseteq T$ as an abbreviation of $\forall x[S(x) \to T(x)]$. Note that in general, these are *intensional* relations. It may well be, for instance, that x and y refer to the same object, while $x \in S$ but $y \notin S$. We can also define notions of *extensional subset, extensional coincidence* and *extensional elementhood*:

Definition 3: $S \subseteq_e U \equiv [\forall x \in S \exists y(y \in U \wedge y = x)]$.

We can define what it means for two predicates to be *co-extensional*:

Definition 4: $S =_e U \equiv [S \subseteq_e U \wedge U \subseteq_e S]$.

Definition 5: $x \in_e S \equiv \exists y(y \in S \wedge y = x)$.

More use will be made of this terminology later on (Section 5).

3.2. CARNAPIAN MODAL LOGIC

Let us now look at the valid logical principles of such languages. We formulate a basic system **C** of *basic Carnapian logic*. Most of the usual laws of classical second-order logic hold in Carnapian logic. The propositional laws and the laws of quantification are formulated in the usual way, and the usual second-order comprehension axiom is valid. Only the principle of substitution of identicals has to be weakened.

Concretely, the system **C** of *Carnapian Logic* is just like classical $S4$ second-order logic, except for the following modifications and additions:

3.2.1. *Comprehension*

A moment's reflection reveals that we may validly *strengthen* the usual second-order comprehension axiom

$$\exists Y \forall x(Yx \leftrightarrow \Phi(x)),$$

with Y not occurring free in Φ, to

$$\exists Y \Box \forall x(Yx \leftrightarrow \Phi(x))(\text{with } Y \text{ not occurring free in } \Phi).$$

Indeed, the formula Φ furnishes the required presentation for Y to make

$$\Box \forall x(Yx \leftrightarrow \Phi(x))$$

true.

Note that unrestricted comprehension allows the definition of *intensional* classes:

Example 7: Consider the following instance of comprehension:

$$\exists Y \forall x(Yx \leftrightarrow \Box x = a).$$

The class Y defined by this sentence can be denoted as $\{x | \Box x = a\}$. This is clearly an intensional class.

For this reason, second-order variables cannot be uniformly regarded as constituting *extensional* contexts.

3.2.2. *Substitutivity of Identicals*

We postulate of course that identity is an equivalence relation. But the usual principle of substitutivity of identicals

$$(x = y) \rightarrow (\Phi \leftrightarrow \Phi[y/x])$$

will have to be restricted somewhat. It is valid in Carnapian logic only for *first-order* formulas Φ that contains *no* occurrences of $\Box$. For all other formulas Φ we only postulate the weaker principle of substitutivity of *provable* identicals:

$$\Box(x = y) \rightarrow (\Phi \leftrightarrow \Phi[y/x]).$$

Similarly, Carnapian logic contains the second-order analogue

$$\Box \forall x(Xx = Yx) \rightarrow (\Phi \leftrightarrow \Phi[Y/X])$$

of the principle of substitution of provably identical terms.

3.2.3. *Propositional Modal Logic*

The natural choice appears to be to let the propositional logic of $\Box$ be given by the $S4$ laws of modal logic.[9]

3.2.4. *Barcan and Choice*

At first sight, there appears no reason to think that the Barcan formulas

$$\forall x \Box \varphi(x) \rightarrow \Box \forall x \varphi(x),$$
$$\forall X \Box \varphi(X) \rightarrow \Box \forall X \varphi(X)$$

are valid. Already the first of these Barcan formulas looks like a formalized ω-rule. So we will not include them in the logic at this point.

The full second-order *axiom of choice* will also not be included in the system **C** of Carnapian logic. In Horsten, (1998), an argument is developed which is intended to cast doubt on its validity in a Carnapian intensional setting. Here I merely remark that in the literature, scepticism about the validity of intensional versions of the axiom of choice has been voiced.[10]

This completes the description of the basic logical principles of the Carnapian logic **C**. The particular system, which is formulated in the language of second-order arithmetic plus the provability operator $\Box$, and which consists of the axioms of *Peano Arithmetic* (**PA**) plus the principles of the Carnapian modal logic **C**, will be called **CPA**.[11]

3.2.5. *Digression: Descriptions in Carnapian Modal Logic*

In ordinary classical logic, the following rule describes the logical behavior of the description operator ι:

$R\iota$ From $\exists! x \varphi(x)$, infer $\forall y(y = \iota x \theta(x) \leftrightarrow \theta(y))$.

The justification for this rule is roughly the following. If one has proved that exactly one object satisfies some condition θ, then one may call this object "*the* θ".

In our Carnapian modal logic, the rule $R\iota$ is not admissible. Reinhardt, though working in a different setting[12] and discussing rules for introducing new individual constants by definition, has given an argument which shows why. Here follows an adaptation to the present setting of Reinhardt's argument:[13]

Consider the theory $\mathbf{CPA}+\exists x(\theta(x) \wedge \neg\Box\theta(x))$, for some arithmetical formula $\theta(x)$. Call this theory $\mathbf{CPA}^+$. Let $\theta^*(x)$ abbreviate:

$$\theta(x) \wedge \neg\Box\theta(x) \wedge \forall y < x : \neg[\theta(y) \wedge \neg\Box\theta(y)].$$

Then

$$\mathbf{CPA}^+ \vdash \exists x \theta^*(x) \wedge \forall y(\theta^*(y) \rightarrow y = x),$$

i.e., $\mathbf{CPA}^+ \vdash \exists! x \theta^*(x)$. Now suppose we are allowed to infer

$$\mathbf{CPA}^+ \vdash \forall y(y = \iota x \theta^*(x) \leftrightarrow \theta^*(y)).$$

Then, since of course $\iota x\theta^*(x) = \iota x\theta^*(x)$, we have $\mathbf{CPA}^+ \vdash \theta^*[\iota x\theta^*(x)]$. On the one hand, by the necessitation rule, this implies $\mathbf{CPA}^+ \vdash \Box\theta^*[\iota x\theta^*(x)]$, whereby $\mathbf{CPA}^+ \vdash \Box\theta[\iota x\theta^*(x)]$. On the other hand, $\mathbf{CPA}^+ \vdash \theta^*[\iota x\theta^*(x)]$ directly implies $\mathbf{CPA}^+ \vdash \neg\Box\theta[\iota x\theta^*(x)]$. Contradiction.

The formula

$$\exists x\Theta^*(x) \wedge \forall y(\Theta^*(y) \rightarrow y = x)$$

says that there exist one or more pairs $\langle o, p_o^1\rangle, \langle o, p_o^2\rangle, \ldots$ satisfying $\theta^*(x)$, but they must all have the same first member o. Why can we not stipulatively introduce a name, $\iota x\theta^*(x)$, referring to this unique object o? The answer to this question is that we *can* stipulatively introduce a name referring to the object o. But this name will *not* satisfy, on pain of contradiction, the intensional description $\theta^*(x)$. In other words, the description rule $R\iota$ cannot be valid for this name.

For this reason,[14] descriptions do not function in the classical way in Carnapian logic. In the sequel, we will take Carnapian languages not to contain a description operator.

3.2.6. *Consistency*

Proposition 1: C is consistent.

Proof: The translation function τ which from any second-order modal formula removes all occurrences of $\Box$ (the so-called eraser-translation), translates all proofs in ***C*** into proofs in classical second-order logic. ■

4. Notation Systems and Canonicity

"At first sight, we could assume that a set is defined by prescribing how its elements are formed. This we do when we say that the set of natural numbers $\mathbb{N}$ is defined by giving the rules:

$$0 \in \mathbb{N}, \quad \frac{a \in \mathbb{N}}{a' \in \mathbb{N}}$$

by which its elements are constructed. However, the weakness of this definition is clear: 10^{10}, for instance, though not obtainable with the given rules, is clearly an element of $\mathbb{N}$, since we know that we can bring it to the form a' for some $a \in \mathbb{N}$. We thus have to distinguish the elements which have a form by which we directly see that they are the result of one of the rules, and call them canonical, from all the other elements, which we will call non-canonical". [Martin-Löf, 1984, p. 7]

Quite in general, a collection of interpreted names, or, equivalently, of ordered pairs the first element of which is an object and the second of which is a presentation of this object, can be considered as a *notation system* or *naming system* for a set of objects. In this section, we take a closer look at the phenomenon that some notation systems are more 'canonical' than others. We distinguish *degrees of canonicity,* which can be defined in the language of Carnapian logic. We show that these degrees are distinct by constructing *examples* that separate them. We also define and illustrate a notion of *relative canonicity*. At the end of this section, we look at one example of a highly canonical notation system in more detail: the system of *standard numerals of Peano Arithmetic*.

4.1. NOTATION SYSTEMS

Suppose we have a class of objects O, and a collection S of primitive and/or complex expressions which name these objects. Then this constitutes a *naming system* for this class of objects. Let the members of S be called *names* for elements of O. Suppose that in addition it is determined for which names the namer can ascertain that they denote identical c.q. different objects. Then these verifiable identity- and difference-relations determine a degree of *canonicity* that this notation system has for the namer.

From now on it is assumed that all notation systems are *denumerable*. This restriction is motivated by the fact that all languages which humans could ever speak and out of which they can form notation systems for classes of objects, are denumerable.

Let us call a notation system *recursively defined* if consists of *all* the names that can be constructed from the basic building blocks out of which its names are constructed. The class of recursive naming systems then forms an important subclass of the class of all notation systems. The system of the standard numerals of Peano Arithmetic $(0, S0, SS0, \ldots)$ is a recursive naming system: it is *recursively generated* from the name 0 and the function symbol S. Below, it will be shown how it is possible to define, at least 'up to provable identity', recursively defined notation systems such as the system of the standard Peano-numerals in Carnapian languages.

Related to this, there is a notion of the (recursively defined) system of names *associated with a language* $\mathscr{L}$. The associated notation system of a language $\mathscr{L}$ consists of all the names that can be formed in $\mathscr{L}$. Derivatively, one can even speak of the (recursively defined) notation system associated with a *theory* T: it consists of all the names that can be formed in the language in which T is expressed.

Even though it goes without saying that the notion of being recursively defined is important, we will not be concerned with it much in the remainder of the paper. In this paper, *canonicity* properties of systems of names will be investigated while abstracting from the recursive nature of these systems.

4.2. GRADES OF CANONICITY

First, we define four notions of canonicity for names. Two of these notions concern verifiable identity of denotation (strong and weak positive canonicity), and two concern verifiable difference of denotation (strong and weak negative canonicity). Let S be any formula of our Carnapian language. We define

Definition 7: $SC_S^-(x) \equiv \forall y \in S : x \neq y \rightarrow \Box x \neq y.$

Definition 8: $WC_S^-(x) \equiv \forall y \in S : x \neq y \rightarrow \exists u \in S(u = y \wedge \Box x \neq u).$

Definition 9: $SC_S^+(x) \equiv \forall y \in S : x = y \rightarrow \Box x = y.$

Definition 10: $WC_S^+(x) \equiv \forall y \in S : x = y \rightarrow \exists u \in S(u = y \wedge \Box x = u).$

In such canonicity conditions, a formula of the form:

$\forall x, y \in S : \ldots$

is of course a concise way of denoting:

$\forall x \forall y((Sx \wedge Sy) \rightarrow (\ldots)),$

etc.

Being WC_S^+ is a trivial property. For if the antecedent of $WC_S^+(x)$ holds, then x itself serves as a witness for the existential quantification in the consequent. In other words, we have:

Proposition 2: For every $x \in S, WC_S^+(x)$ holds.

So the notion WC_S^+ is useless. It will be abandoned.

The following property is immediate:

Proposition 3: $\forall x \in S : SC_S^-(x) \rightarrow WC_S^-(x).$

In terms of these notions of canonicity for names, degrees of canonicity of notation systems can be expressed. The guiding intuition is that a notation system S is canonical if for each of the objects which S names, S contains a canonical name. The following degrees can then be distinguished:

1. $SC^+_\forall(S) \equiv \forall x \in S : SC^+_S(x)$
2. $WC^-_\forall(S) \equiv \forall x \in S : WC^-_S(x)$
3. $SC^-_\forall(S) \equiv \forall x \in S : SC^-_S(x)$
4. $SC^+_\exists(S) \equiv \forall x \in S \exists y \in S : y = x \wedge SC^+_S(y)$
5. $WC^-_\exists(S) \equiv \forall x \in S \exists y \in S : y = x \wedge WC^-_S(y)$
6. $SC^-_\exists(S) \equiv \forall x \in S \exists y \in S : y = x \wedge SC^-_S(y)$

Note that these canonicity conditions are *intensional*. It is perfectly possible for one notation system S to satisfy one of these canonicity conditions whereas another notation system S' fails to satisfy this condition, even though S and S' are co-extensional.[15]

These degrees of canonicity conditions can be glossed informally:

1. $SC^+_\forall$: Every name is strongly positively canonical. For any two names of an $SC^+_\forall$ notation system, if they denote the same object, then this can be ascertained.
2. $WC^-_\forall$: Every name is weakly negatively canonical: any object which differs from the object named by a given name, has at least one name which verifiably separates it from the object named by this weakly negatively canonical name.
3. $SC^-_\forall$: Every name is strongly negatively canonical. For any two names of an $SC^-_\forall$ notation system, if they denote different objects, then this can be ascertained.
4. $SC^+_\exists$: Every object named by the naming system is named by at least one strongly positively canonical name: every presentation naming the same object as this strongly positively canonical name can be ascertained to do so.
5. $WC^-_\exists$: Every object named is named by at least one weakly negatively canonical name.
6. $SC^-_\exists$: Every object named by the naming system is named by a strongly negatively canonical name: every presentation naming an object different from the object named by this strongly negatively canonical name can be ascertained to do so.

The following property is easily established for all S:

Proposition 5: $SC^+_\forall(S) \leftrightarrow SC^+_\exists(S)$.

So $SC^+_\exists$ is redundant: we can (and will) drop this notion.

Note that each of the five remaining degrees of canonicity of S indeed requires that each object named by S is named by a presentation which is in some sense (WC^-, SC^-, SC^+) canonical. Therefore they are rightly called *degrees* of canonicity. Note also that more complex initial quantifier strings than the ones used ($\forall x \in S$ and $\forall x \in S \exists y \in S : y = x \wedge \ldots$), such as

$$\forall x \in S \exists y \in S : y = x \wedge \forall z \in S(z = y \rightarrow \ldots)$$

yield no new degrees of canonicity.

Nevertheless, more notions of canonicity of notation systems can be distinguished. First, a notation system can of course be a combination of some of the five degrees of canonicity that were defined. In the sequel, of special importance in this respect will be the combination of being both $SC^{+}_{\forall}$ and $SC^{-}_{\forall}$. We will abbreviate this combination as $SC^{\pm}_{\forall}$. We will see how it can be shown that the class of standard Peano-numerals constitutes such a $SC^{\pm}_{\forall}$ notation system. Secondly, there are notions of canonicity of notation systems which cannot be naturally expressed in terms of weak and strong canonicity of names. Here is a *very weak* such notion, which will play a role in what follows:

Definition 11 (very weakly canonical):

$$VWC(S) \equiv \forall x, y \in S : x \neq y \rightarrow \exists z, u \in S(x = z \wedge y = u \wedge \Box u \neq z).$$

In words: Every two objects named are verifiably separated by names in the notation system. Thirdly, one can express notions of being *almost canonical*. For instance, one can also define a notation system S to be *almost* $WC^{-}_{\forall}$ if all $x \in S$ are such that for *almost* all $y \in S$, i.e. for all y except a *finite* number of elements of S,

$$x \neq y \rightarrow \Box x \neq y$$

.

There are undoubtedly even more notions of canonicity of notation systems that can be thought of. But we focus exclusively on the notions defined here, which appear to be natural ones.

The notion of weak canonicity of names can be generalized:

Definition 12: For all x, S:

$$WC^{0}_{S}(x) \equiv x \in S;$$
$$WC^{n+1}_{S}(x) \equiv \forall y[y \neq x \rightarrow \exists z(z = y \wedge WC^{n}_{S}(z) \wedge \Box z \neq x)];$$
$$WC^{\infty}_{S}(x) \equiv \forall n \in \mathbb{N} : WC^{n}_{S}(x).$$

A name for an object o is therefore WC^{2}_{S}, for instance, if it is not only WC^{-}_{S}, but also related by the provable difference relation to WC^{-}_{S} names for all objects o' such that $o' = o$.

The same can be done for strong canonicity:

Definition 13: For all x, S:

$$SC_S^0(x) \equiv x \in S;$$
$$SC_S^{n+1}(x) \equiv \forall y[y \neq x \rightarrow (SC_S^n(y) \wedge \Box y \neq x)];$$
$$SC_S^{\infty}(x) \equiv \forall n \in \mathbb{N} : SC_S^n(x).$$

The following shows that for strong canonicity, this generalization is not very useful:

Proposition 6: For all S:

$$\exists x \in S : SC_S^2(x) \Rightarrow \forall x SC_S^{\infty}(x).$$

In other words, if S contains a SC_S^2 presentation, then S is an $SC_{\forall}^{-}$ system.

The generalized notion of weak canonicity displays somewhat more interesting structure. We first introduce some more terminology:

Definition 14: For all $S, x \in S$:

$$R_S^0(x) \equiv \{x\};$$
$$R_S^{n+1}(x) \equiv R_S^n(x) \cup \{y \in S | \exists x : x \in R_S^n(x) \wedge \Box y \neq x\};$$
$$R_S^{\infty}(x) \equiv \bigcup_{n \in \omega} R_S^n(x).$$

So $R_S^{\infty}(x)$ is, as it were, the collection of names of S *generated* from x by the provable difference relation. Observe that if $x \in S$ and $SC_S^2(x)$, then $R_S^{\infty}(x) = S$; this is not always so if $WC_S^2(x)$.

Then the following proposition is easy to establish:

Proposition 7: For all S:

Suppose $x \in S$ and $WC_S^{\infty}(x)$.

Then $R_S^{\infty}(x)$ is a $WC_{\exists}^{-}$ subsystem of S naming the same collection of objects as S does.

Actually, already $WC_S^2(x), R_S^2(x)$ (and also $R_S^{\infty}(x)$) will be a $WC_{\exists}^{-}$ subset of S naming the same collection of objects as S does.

4.3. ORDERING OF THE DEGREES

The notions $SC_{\forall}^{+}, \ldots, VWC$ are intended to capture extents of *transparency* or *canonicity* of a notation system. It is not claimed that the degrees of canonicity that we distinguish are the only or even the most important factors in determining the naturalness or 'niceness' of a notation system.[16] But we do think that the degree of canonicity of a naming system is highly significant for determining its practical usefulness.

We will show that the partial ordering relation between the degrees of canonicity is given by the following table:

$$\begin{array}{ccccccc} & & & & SC^{-}_{\exists}(S) & & \\ SC^{\pm}_{\forall}(S) & \Rightarrow & SC^{-}_{\forall}(S) & \Rightarrow & \not\Updownarrow & \Rightarrow & WC^{-}_{\exists}(S) \Rightarrow VWC(S) \\ \Downarrow & & & & WC^{-}_{\forall}(S) & & \\ SC^{+}_{\forall}(S) & & & & & & \end{array}$$

It is left to the reader to verify that the $\Rightarrow$-directions in this table indeed hold (it is not hard to see). We will show that none of the converse arrows hold. Also, the degree $SC^{+}_{\forall}(S)$ does not imply any of the other degrees. We will not be satisfied with constructing arbitrary models in the sense of Section 3.1. In order to develop intuitions, we want to construct somehow 'realistic imaginary situations' which separate the notions $SC^{\pm}_{\forall}, \ldots, VWC$.

These imaginary situations or models will take the following form. They will contain one namer, who lives in a universe, and who assigns (at most denumerably many) names to the objects in the universe. Moreover, the namer develops a correct but generally incomplete theory about the identity and difference of the objects denoted by the names that he has assigned to them. The collections of names for objects in the universe which the namer constructs in this way will be a notation system associated with the model or situation. One can then inquire into the degree of canonicity of this notation system.

4.3.1. *Example of a VWC Notation System Which is not* $WC^{-}_{\exists}$

Suppose a universe $\mathcal{U}$ consisting of a countably infinite number of objects, and a namer who cannot move about in this universe and cannot see an infinite number of objects. All he can do is to extrapolate on the basis of his experiences with a finite number of objects in his neighborhood.

For any finite collection F of objects of $\mathcal{U}$, let there be a property P such that:

- all objects not in F lack property P;
- all objects in F have property P to some measurable degree;
- no two objects in F have property P to exactly the same degree.

Suppose that the namer, on the basis of his local experiments, comes to correctly conjecture that this is the case. Then he can use the infinite supply of properties $P_1, \ldots, P_i, \ldots$ to construct a naming system S for $\mathcal{U}$ which contains infinitely many names for every object in $\mathcal{U}$: one name for each finite collection to which the object belongs. This naming system contains all names of the form: 'the second-least but nonzero degree P_{25}', 'the 34th smallest but nonzero degree P_{100}',... Here of course a default stipulation has to be made to ensure reference of the term if no second-least-degree P_{25}, 34th smallest P_{100}, ... exists. For instance, the namer could stipulate that terms not meeting this assumption refer to some fixed local object o_l in his neighborhood.

S is a VWC notation system. Any two objects o_1, o_2 of $\mathcal{U}$ are verifiably separated by at least one pair of names. Take, for instance, the names determined by the property P_k which singles out $\{o_1, o_2\}$. Then 'the least-degree but nonzero (and some default object if zero) P_k' and 'the highest-degree P_k' separate o_1 and o_2.

But S is not a $SC^-_\exists$ notation system. Take any name, e.g. 'the second-smallest P_{25}'. Suppose that the object named by this term does not have the property P_2. Then 'the second-smallest P_{25}' is not verifiably separated from 'the smallest P_2'.

Also, S is not a $WC^-_\forall$ notation system. For take any two names of S: 'the lth-degree P_i' and 'the mth-degree P_j' referring to o_a, o_b respectively. Now suppose o_b lacks property P_i. Then there is no name in S under which o_b can be verifiably separated from 'the lth-degree P_i'. S is not even a $WC^-_\exists$ system. For take any object o_a. Every name n_a of o_a is of the form 'the kth degree P' for some property P which is true of only a *finite* set S_P of objects. We must show that n_a is not weakly canonical. For this purpose, take any $o_b \notin S_P$. No name n_b for o_b can be verifiably separated from n_a by the namer. So n_a is not weakly canonical.

4.3.2. *Two Examples of Notation Systems Which are $WC^-_\forall$ but not $SC^-_\exists$*

A finite example: Let a naming system $S = \{n^1_1, n^1_2, n^2_1, n^2_2\}$ be defined such that:

- the extension of S is a set of objects $\{o_1, o_2\}$ with $o_1 \neq o_2$,
- n^1_1, n^1_2 are names of o_1,
- n^2_1, n^2_2 are names of o_2;
- the (correct) theory of the namer entails that:
  - n^1_1 and n^2_1 name distinct objects,
  - n^1_2 and n^2_2 name distinct objects;
- the theory of the namer does *not* entail that:
  - n^1_1, n^1_2 name the same object,
  - n^2_1, n^2_2 name the same object,
  - n^1_1 and n^2_2 name distinct objects,
  - n^1_2 and n^2_1 name distinct objects.

It is not hard to verify that S is $WC_{\forall}^{-}$. But S is not $SC_{\exists}^{-}$: S contains no sufficiently canonical names for o_1,o_2.

The discrimination and identification abilities of the namer can be represented in a table:

$$\begin{array}{ccccc} n_2^1 & -- & -- & -- & n_2^2 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ n_1^1 & -- & -- & -- & n_1^2 \end{array}$$

Here dotted lines indicate nondiscernibleness and nonidentifiability. Striped lines indicate discernibleness.

More generally, notation systems can be represented as simple, undirected graphs with labelled nodes. The nodes stand for objects; any two nodes with the same label stand for the same object. Edges stand for *verifiable* difference or identity relations (depending on whether the nodes they connect have the same label). Let us call such graphs *naming system graphs*. Notation system graphs G are in general reflexive, and what one may call *weakly transitive,* i.e. if (v_1, v_2), $(v_2, v_3) \in G$, and v_1 and v_2 have the same label, then $(v_2, v_3) \in G$.

An infinite example: Consider a universe $\mathcal{U}$, in which a namer names objects. The namer can interact only with a forever determined finite number of objects. On the basis of this interaction, he develops the following correct theory. There is a denumerably infinite number of objects and a denumerably infinite collection of properties. Each object possesses each property to some degree. For any given property, the set of degrees in which objects have the property form a well-ordering of order type ω. No two objects have any given property to the same degree.

On the basis of this theory, the namer gives many names to each object. For example: 'the object which has property 28 to the 5th least degree', which can be represented as $\langle 28, 5\rangle$. Call the resulting naming system S.

Now let the following infinite matrix represent the denotation relation for S:

...	...	...	...	...	...
$\langle 5,1\rangle$	$\langle 5,2\rangle$	$\langle 5,3\rangle$	$\langle 5,4\rangle$	$\langle 5,5\rangle$	...
$\langle 4,1\rangle$	$\langle 4,2\rangle$	$\langle 4,3\rangle$	$\langle 4,4\rangle$	$\langle 4,5\rangle$	...
$\langle 3,1\rangle$	$\langle 3,2\rangle$	$\langle 3,3\rangle$	$\langle 3,4\rangle$	$\langle 3,5\rangle$	...
$\langle 2,1\rangle$	$\langle 2,2\rangle$	$\langle 2,3\rangle$	$\langle 2,4\rangle$	$\langle 2,5\rangle$	...
$\langle 1,1\rangle$	$\langle 1,2\rangle$	$\langle 1,3\rangle$	$\langle 1,4\rangle$	$\langle 1,5\rangle$	...

Each column of this matrix represents (again) the set of names of a single object. So the object which has property 1 to the lowest degree is in fact also the object which has property 2 to the lowest degree, and so on. Now let the verifiable difference relations be as follows:

$$\begin{array}{cccccc}
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\langle 5,1\rangle & \langle 5,2\rangle & \langle 5,3\rangle & \langle 5,4\rangle & \underline{\langle 5,5\rangle} & \cdots \\
\langle 4,1\rangle & \langle 4,2\rangle & \langle 4,3\rangle & \underline{\langle 4,4\rangle} & \overline{|\langle 4,5\rangle} & \cdots \\
\langle 3,1\rangle| & \langle 3,2\rangle & \underline{\langle 3,3\rangle} & \overline{|\langle 3,4\rangle} & \langle 3,5\rangle| & \cdots \\
|\langle 2,1\rangle & \underline{\langle 2,2\rangle}| & \overline{|\langle 2,3\rangle} & \langle 2,4\rangle| & \langle 2,5\rangle & \cdots \\
\underline{\langle 1,1\rangle} & \overline{|\langle 1,2\rangle} & \langle 1,3\rangle| & \langle 1,4\rangle & \langle 1,5\rangle & \cdots
\end{array}$$

Each column of this matrix represent the set of names of a single object. And the verifiable differences follow the diagonals drawn from objects on the lower row – no other differences in denotation are assumed verifiable by the namer.

A few of these verifiable difference relations are indicated by underlining and side-lining, e.g., the theory of the namer entails that $\langle 1,3\rangle$ and $\langle 3,5\rangle$ denote different objects; also, it entails that $\langle 1,2\rangle$ and $\langle 2,1\rangle$ denote different objects. But the namer's theory does not prove that $\langle 3,5\rangle$ and $\langle 1,4\rangle$ denote different objects.

This notation system S then is $WC_{\forall}^{-}$. But S is not $SC_{\forall}^{-}$. For instance, as noted above, the namer's theory does not prove that $\langle 3,5\rangle$ and $\langle 1,4\rangle$ denote different objects.

4.3.3. *Example of a $WC_{\exists}^{-}$ Notation System Which is not $SC_{\exists}^{-}$*

Consider the following finite notation system:

$$\begin{array}{ccc}
\underbrace{n_2^1}| & |\bar{n}_2^2 & | \overbrace{n_2^3} \\
\overline{\underline{n_1^1}} & \underbrace{\overbrace{n_1^2}} & \underline{n_1^3}|
\end{array}$$

In this table, each column again contains names denoting the same object. The underlinings and overlinings again express verifiable difference relations. For example n_1^2 and n_2^3 are verfiably different, but n_1^2 and n_1^1 are not verifiably different. It is easy to verify that this system is $WC_{\exists}^{-}$. But it is not $SC_{\exists}^{-}$.

This system is not $WC_{\forall}^{-}$, for n_2^3 is not a weakly canonical name. If we would add $\langle n_1^1, n_2^3\rangle, \langle n_2^3, n_1^3\rangle$ as extra verifiable differences, then the resulting system would be $WC_{\forall}^{-}$. But still it would not be $SC_{\exists}^{-}$. Therefore

$$WC_{\forall}^{-} \not\Rightarrow SC_{\exists}^{-}.$$

Since the finite example in Section 4.3.2 already showed us that

$$SC_{\exists}^{-} \not\Rightarrow WC_{\forall}^{-},$$

we indeed have

$SC^{-}_{\exists} \not\Leftrightarrow WC^{-}_{\forall}$.

4.3.4. *Example of a $SC^{-}_{\exists}$ Notation System Which is not $SC^{-}_{\forall}$*

Left to the reader (easy).

4.3.5. *Example of a Notation System Which is $Sc^{-}_{\forall}$;* but not $SC^{\pm}_{\forall}$

Consider Kleene's notation system O for the constructive ordinals.[17] There exists a Turing machine T which, when started on any two $a, b \in O$, eventually prints 0 if a and b name different ordinal numbers. But there exists no Turing machine which, in addition to this behavior, also always eventually prints 1 if a and b are ordinal notations for the same ordinal number. For in order to do that, the Turing machine would effectively have to be able to tell whether two Turing machines enumerate the same set of numbers; it is well known that no such Turing machine exists.

Now suppose that the capacities of a human namer for identifying and discriminating ordinal notations do not exceed those of Turing machines. Then for a human namer, the system O can never be $SC^{\pm}_{\forall}$. At best, O will for him be a $SC^{-}_{\forall}$ notation system.

4.4. THE STANDARD NUMERALS

As an illustration, we will now investigate the system of the standard numerals 0, *S*0, *SS*0, *SSS*0, ... of the language of Peano Arithmetic as a concrete example of a $SC^{\pm}_{\forall}$ notation system. First, it is shown how the class of standard numerals can be defined, up to provable identity, by a formula $SN(x)$ ('Standard Numeral') of our second-order Carnapian language. Second, it is outlined how in basic Carnapian logic, augmented with Peano Arithmetic (**PA**), it can be *proved* that *SN* is a $SC^{\pm}_{\forall}$ notation system.

4.4.1. *Defining the Standard Numerals*

Let $SN\ (x)$ be the following condition:

$$\forall X \forall z\{[(\Box(z = 0) \rightarrow Xz) \wedge (\Box Xz \rightarrow \Box X(Sz))] \rightarrow Xx\}.$$

SN is then of course an intensional set. We claim that *SN* defines, up to provable identity, the class of the standard numerals.

Proposition 8: For all x:

$SN\ (x) \Leftrightarrow x$ is provably identical with a standard Peano numeral.

Proof$\Rightarrow$: It suffices to show that there exists a class containing the standard numerals. But that is straightforward. Take the class $\{x \mid x = x\}$.

$\Leftarrow$: By the minimal closure condition. ■

4.4.2. *Proving in CPA that SN is* $SC^{\pm}_{\forall}$

Lemma 1: $\forall x \in SN : \Box(x = 0) \vee \exists z[\Box(x = Sz) \wedge z \in SN]$

Proof (Sketch): Essentially by the minimal closure condition in the definition of SN. For suppose not. Then consider $SN \setminus \{x\}$. This would still satisfy the defining condition, thus contradicting the fact that SN is the smallest such.■

Theorem 1: $\forall x, y \in SN : x = y \rightarrow \Box(x = y)$.

Proof: We proceed by induction on x $(= y)$.

(*a*) *Basis*. $x = 0$. Then by the previous lemma and the reflexivity principle $\Box\phi \rightarrow \phi$, we must have $\Box$ $(x = 0)$. Likewise, we must have $\Box$ $(y = 0)$. So, by the substitutivity of provable identicals, we have $\Box$ $(x = y)$.
(*b*) *Induction step*. By the previous lemma, we must have $\Box$ $(x = Sz)$ for some $z \in SN$, and $\Box$ $(y = Sz')$ for some $z' \in SN$, where $z = z'$. By the inductive hypothesis, we then have $\Box$ $(z = z')$. Therefore, by the substitution of provable identicals, we get $\Box$ $(Sz = Sz')$. Again by the substitutivity of provable identicals, we then obtain $\Box$ $(x = y)$. ■

Theorem 2: $\forall x, y \in SN : x \neq y \rightarrow \Box(x \neq y)$.

Proof: We proceed by induction on x $(= y)$.

(*a*) *Basis*. $x = 0$. Then by the lemma, we must have $\Box$ $(x = 0)$, and $\Box$ $(y = Sz)$ for some z. Then from the supposition that $x = y$, we can by substitution of provable identicals derive a contradiction. So by the necessitation rule we obtain $\Box(x \neq y)$.
(*b*) *Induction step*.
There are two possibilities for y:
(*b1*) $\Box(y = 0)$.
But then, by the same reasoning as in *a*, we obtain $\Box(x \neq y)$.
(*b2*) $\Box(y = Sz')$ for some $z' \in SN$.

From the supposition that $x \neq y$, i. *e*. that $Sz \neq Sz'$, we derive in *PA* that $z \neq z'$. Then by the induction hypothesis we get $\Box(z \neq z')$, whereby in a few steps we obtain $\Box(Sz \neq Sz')$, using the Necessitation rule and distributivity of $\Box$ over $\rightarrow$. Then substitution of provable identicals yields the desired result. ■

So it can be proved in **CPA** that the standard numerals form a particularly nice notation system. It seems likely that along similar lines. one can prove in second-order Carnapian Peano Arithmetic that the recursive ordinals form a $SC_{\forall}^{-}$ notation system – although I have not checked the details.

In classical **PA**, the class of Peano-numerals can already be defined – Gödel has shown how. Also, it is of course true that it can be shown in **PA** that the equality-relation on the Peano-numerals is decidable (in **PA**). But all this proceeds via coding. The Carnapian results in this section do not rely on a coding machinery.

4.5. RELATIVE CANONICITY

Let us now try to express what it means to *effectively*, *finitely* reduce questions of identity and difference concerning a notation system S to identity and difference questions about another system T, i.e., let us try to express a notion of *relative canonicity*.

First, we need to express that notation system X contains identity and/or difference information about a *finite* subset of notation system Y. To do this, we need to be able to express, a.o., what it means for a set S of *presentations* (relation X on *presentations*) to be *finite*. Up to provable identity, this can be done in Carnapian modal logic. Let the formula $FIN(S)$ ($FIN(X)$) express this.[18] Moreover, let $FUN(X)$ express that X is a *function* on presentations, and (for X a function), let $DOM(X)$ be a formula true of all elements in the *domain* of X, and $RAN(X)$ be true of all elements in the *range* of X. Using these notions we can indeed express, in Carnapian logic, the characteristic function of a finite part of the identity relation on Y:[19]

Definition 15: $FID_Y(X) \equiv$

$$FIN(X) \wedge FUN(X) \wedge$$
$$DOM(X) \subseteq Y \times Y \wedge RAN(X) = \{\underline{0}, \underline{1}\} \wedge$$
$$\Box \forall x, y \in Y \times Y \exists z : \Box X(x, y, z).$$

Now we are finally able to express what it means for the identity/difference relation on S to be *effectively, finitely* reducible to that on T:

Definition 16: $S \preceq T \equiv$

$$\begin{aligned}&T \subseteq S\wedge\\&\forall x, y \in S : \exists F, G : FID_S(F) \wedge FID_S(G)\wedge\\&\qquad\Box\left\{\forall u, v \in T\begin{bmatrix}(F(u,v) = \underline{1} \leftrightarrow u = v)\wedge\\(F(u,v) = \underline{0} \leftrightarrow u \neq v)\end{bmatrix} \rightarrow x = y\right\}\wedge\\&\qquad\Box\left\{\forall u, v \in T\begin{bmatrix}(G(u,v) = \underline{1} \leftrightarrow u = v)\wedge\\(G(u,v) = \underline{0} \leftrightarrow u \neq v)\end{bmatrix} \rightarrow x \neq y\right\}.\end{aligned}$$

Evidently variations on this definition are possible.

As an illustration, we give a simple example of concrete, *finite* notation systems S, T such that $S \preceq T$.

Definition 17:

$m \equiv 1$ if Goldbach's conjecture is true;
$m \equiv 0$ otherwise.

Definition 18 (the systems T and S):

$$T \equiv \{m, \underline{0}\};$$
$$S \equiv \{m, \underline{0}, \underline{1}\}.$$

We may suppose here (although this is not essential for what follows) that Goldbach's conjecture is absolutely undecidable, i.e., that:

$$\neg\Box GOLD \wedge \neg\Box\neg GOLD.$$

We claim that $S \preceq T$.

Proof (*Informal sketch*). There are 3 cases for x, y that need to be considered:

(1) $x = m$ and $y = 0$.

(1a) Set $F \equiv \{\langle m, \underline{0}, \underline{1}\rangle\}$. Suppose $F(m, \underline{0}) = \underline{1} \rightarrow m = \underline{0}$. We know that by definition: F $(m, 0) = 1$. So we deduce m $= 0$. So indeed:

$$\Box\{\forall u, v \in T[(F(u,v) = \underline{1} \leftrightarrow u = v) \wedge (F(u,v) = \underline{0} \leftrightarrow u \neq v)] \rightarrow m = \underline{0}\}.$$

(1b) Set $G \equiv \{\langle m, \underline{0}, \underline{0}\rangle\}$. Suppose $G(m, \underline{0}) = \underline{0} \rightarrow u \neq v$. We know that by definition G $(m, 0) = 0$. So indeed

$$\Box\{\forall u, v \in T[(G(u,v) = \underline{1} \leftrightarrow u = v) \wedge (G(u,v) = \underline{0} \leftrightarrow u \neq v)] \rightarrow m \neq \underline{0}\}.$$

(2) $x =$ m and $y = 1$.

(2a) Set $F \equiv \{\langle m, \underline{0}, \underline{0}\rangle\}$. Suppose $F(m, \underline{0}) = \underline{0} \rightarrow m \neq 0$. Then indeed $m \neq 0$. Therefore by definition of m: $m = 1$. So indeed

$\Box\{\forall u, v \in T[(F(u, v) = \underline{1} \leftrightarrow u = v) \wedge (F(u, v) = \underline{0} \leftrightarrow u \neq v)] \rightarrow m = \underline{1}\}$.

(2b) Set $G \equiv \{\langle m, \underline{0}, \underline{1}\rangle\}$.

(3) $x = 0$ and $y = 1$.

(3a) Set $F \equiv \{\langle \underline{0}, \underline{1}, \underline{1}\rangle\}$. Suppose $F(\underline{0}, \underline{1}) = 1 \rightarrow 0 = 1$. We know that $F(\underline{0}, \underline{1}) \equiv \underline{1}$. So we deduce $0 = 1$.

(3b) Set $G \equiv \{\langle \underline{0}, \underline{1}, \underline{0}\rangle\}$. ■

5. Indiscernible Names

It is impossible to prove in second-order Carnapian Peano Arithmetic for any notation system that it is *not* $SC^{\pm}_{\forall}, \ldots, VWC$. The reason is that basic Carnapian is purely *positive,* in the sense that it cannot prove that for some objects or properties absolutely undecidable properties exist:

Proposition 9: For all $A \in \mathscr{L}_C$: **C** does not prove

$$\exists X_1 \ldots \exists X_k \exists x_1 \ldots \exists x_l (\neg\Box A \wedge \neg\Box\neg A).$$

Proof. The translation τ (see the proof of proposition 1) translates

$$\exists X_1 \ldots \exists X_k \exists x_1 \ldots \exists x_l (\neg\Box A \wedge \neg\Box\neg A)$$

into a contradiction. ■

This entails as a special consequence that **C** cannot prove that some identity statements are absolutely undecidable. If we want to establish that some nonempty notation system are not $SC^{\pm}_{\forall}, \ldots, VWC$, then we must explicitly assume or *postulate* that some identity statements are absolutely undecidable.

For a given notation system *S,* we define a notion of *S*-randomness, or better perhaps, *S -indiscernibleness* as follows:

Definition 19: $Ind_S(x) \equiv \forall y \in S : \neg\Box y \neq x$.

The name *n* that was defined in Example 2 is a putative concrete example of an *SN*-indiscernible name.

We can also distinguish a notion of *weak S-indiscernibleness:*

Definition 20: $WInd_S(x) \equiv \forall y \in S \exists z : z = y \wedge \neg\Box(x \neq z)$.

Let us now make some simple observations about indiscernibleness. For convenience, we first introduce some terminology:

Definition 21: For any notation system S, and for any $a \in S$, let $d\,(a)$ be the object named by a.

Definition 22: For any notation system S, let $Obj\,(S)$ be the set of objects that have names in S.

From now on, we assume that the naming systems that we will consider are *nontrivial,* i.e. that for every naming system S, $Obj\,(S)$ contains at least two objects.

Proposition 10: Let S be any (*nontrivial*) $WC^{-}_{\exists}$ system, and let $x \in_e S$ and $Ind_S\,(x)$. Then $S \cup \{x\}$ is exactly $WC^{-}_{\exists}$, i.e., it is $C^{-}_{\exists}$ but not $SC^{-}_{\exists}$ and not $WC^{-}_{\forall}$.

Proof.

(1) $S \cup \{x\}$ is not $SC^{-}_{\exists}$. For consider any $o \in Obj(S)$ such that $o \neq d(x)$. Then o cannot have any strongly canonical name in S, for (by $Ind_S\,(x)$) no name is distinguishable from x. So $S \cup \{x\}$ is not $SC^{-}_{\exists}$.
(2) $S \cup \{x\}$ is not $WC^{-}_{\forall}$. Reason: x cannot be distinguished from any name $a \in S$ such that $d(a) \neq d(x)$.
(3) $S \cup \{x\}$ is $WC^{-}_{\exists}$. For this, it suffices to show that $d(x)$ has a weakly canonical name. But since $x \in_e S$, there must be at least one name $a \in S$ such that $d(a) = d(x)$. Any such a will serve as a weakly canonical name for $d(x)$. ■

Adding a weakly indiscernible to a canonical notation system does not lower the degree of canonicity as much as adding an indiscernible. Using an argument analogous to that of the proposition above, it is easy to see that $SC^{-}_{\forall} + WInd$ is not a $SC^{-}_{\forall}$ naming system. But it can happen that the resulting notation system is $SC^{-}_{\exists}$ or $WC^{-}_{\forall}$.

Recall from the previous section that in Carnapian second-order Peano Arithmetic we can prove the existence of a $SC^{\pm}_{\forall}$ notation system (the system *SN* of standard numerals). Since the proofs of the propositions above can be carried out inside the Carnapian logic **C**, Carnapian second-order Peano Arithmetic can prove that *if* there are *SN*-indiscernibles, *then* there are notation systems for the natural numbers that are no more canonical than $WC^{-}_{\forall}$.

There are more technical notions to be distinguished here, and more theorems to be proved, of course. But we will not do this here: the foregoing should give some indication of the importance of the general notion of indiscernibleness for the investigation of canonical notation systems.

6. Concluding Remarks

In this paper, a formal framework for the systematic investigation of canonical notation systems was proposed, and first steps were taken in the investigation of its logical properties. Let us now, by way of conclusion, summarize our findings. The usefulness of the framework of Carnapian logic appears from its ability of:

1. Defining *degrees of canonicity*. This was shown in Section 4.2.
2. Proving *general* theorems about the structure of degrees of canonicity and about indiscernible objects. This was to some extent done in Section 4.3, and in Section 5.
3. *Defining* notation systems for systems of objects. For this purpose, $\mathscr{L}_C$ must contain the basic expressions out of which these notation systems are built. For this reason, for instance, we defined in Section 4.4.1 the class *SN* in second-order Carnapian *arithmetic*.
4. *Proving, in* a Carnapian system, that *particular* naming systems have a *precise* degree of transparency *n*. To this effect, one has to prove first that the system in question has a degree of canonicity which is *at least n,* and second, that its canonicity degree is *no greater than n*. The first part can often be achieved in the relevant Carnapian system. As an illustration, this was shown for *SN* in Section 4.4.2. The second part, however, requires explicitly assuming the existence of indiscernible names, as was shown in Section 5.
5. *Defining* in Carnapian modal logic, notions of *relative canonicity* (Section 4.5).

In addition to this, it may be interesting to *apply* the framework of Carnapian logic to the investigation, within a classical framework, of epistemic notions and problems, some of which have been studied before in the context of intuitionistic arithmetic. One might try to investigate the intuitionistic notion of apartness in the present framework. One might try to formulate epistemic choice principles, principles concerning lawless sequences, and analogues of Church's Thesis in Carnapian Peano Arithmetic, and to investigate their logical properties.

Acknowledgements

Research for this paper was supported in part by a research scholarship from the Fund for Scientific Research – Flanders, which is gratefully acknowledged. I am indebted to John Burgess, Tony Anderson and Jan Heylen for discussions on the subject matter of this paper. Thanks also to an anonymous referee for helpful suggestions for improvement.

Notes

[1] For a 'rational history' of the discussion of quantifying into intensional contexts, see Neale, (2000).
[2] The classical reference is the papers in Shapiro (1985). But see also the papers by Reinhardt on this topic (e.g., Reinhardt, 1986).
[3] An exception is Reinhardt (1986).
[4] The *locus classicus* is Carnap (1956, chapter V). The distinctive flavor of the Carnapian approach is already foreshadowed in Carnap (1946, Section 12).
[5] This idea of course goes back to Frege (1960).
[6] See Kripke (1992).
[7] The formal notion of a model for Carnapian languages is spelled out in more detail in Horsten (1998).
[8] The intensional systems in the literature concerning Epistemic Arithmetic are mostly of the Kripkean kind.
[9] One could contemplate strengthening this modal logic, roughly in the direction of $S4.1$. See Horsten (1997), for a discussion of this line of reasoning.
[10] See Anderson (1989, p. 100) and (Gödel (1990, p. 139).
[11] It should be noted that the full second-order principle of mathematical induction is not intuitively valid in the Carnapian setting. The antecedent of the induction axiom is satisfied by a property Φ if Φ holds of the numerals $0, S0, SS0, \ldots$ But for the consequent to be valid, Φ must hold not only for these standard numerals, but for *all* presentations of natural numbers. In response to this difficulty, one may want to restrict universal quantifier in the consequent of the induction axiom to the *standard Peano numerals* (see Section 4.4.1).
[12] Reinhardt is working in a *Churchian* setting in which two types of individual variables are distinguished: intensional and extensional ones.
[13] This argument is adapted from Reinhardt (1986, pp. 452–453).
[14] I do not claim that Reinhardt himself would have agreed with my diagnosis of the version of Reinhardt's argument described here.
[15] The notion of coextensiveness was defined in Section 3.1.3.
[16] For a brief catalogue of properties for natural notation systems *for ordinal numbers* (See Weiermann, 2004, p. 3).
[17] For a definition of the notation system O see e.g. Rogers (1987[1967], pp. 208–209).
[18] The reader might want to write out this formula in the language of Carnapian modal logic. It is not *completely* obvious how this is done: note that it is not the same as saying, in Carnapian logic, that the *extension* of S (X) is finite.
[19] In the sequel, the underlining in 0, e.g., means that, up to provable identity, the *numeral* and not the number 0 is intended.

References

Anderson, C.A. (1989), 'Russellian Intensional Logic', in J. Almog et al., eds., *Themes From Kaplan*, Oxford University Press, pp. 67–103.

Carnap, R. (1946), 'Modalities and Quantification', *Journal of Symblic Logic* 11, pp. 33–64.

Carnap, R. (1956), *Meaning and Necessity. A Study in Semantics and Modal Logic*, 3rd edition, University of Chicago Press.

Church, A. (1943), 'Review of: Notes on Existence and Necessity, by W.V. Quine (Journal of Philosophy **40**(1943))', *Journal of Symbolic Logic 8*, pp. 45–47.

Frege, G. (1966), 'Über Sinn und Bedeutung'. (1892) Reprinted in G. Frege, ed., *Funktion, Begriff, Bedeutung. Fünf logische Studien*. Herausgegeben und eingeleited von Günther Patzig. Vandenhoeck & Ruprecht, pp. 40–65.

Gödel, K. (1990), 'Russell's Mathematical Logic [1944]'. Reprinted, in S. Feferman et al. eds., *Kurt Gödel. Collected Works*. Part II, Oxford University Press, pp. 119–141.

Horsten, L. (1997), 'Provability in Principle and Controversial Constructivistic Principles', *Journal of Philosophical Logic* 26, pp. 635–660.

Horsten, L. (1998), 'In defense of Epistemic Arithmetic', *Synthese* 116, pp. 1–25.

Kripke, S. (1992), 'Individual Concepts: Their Logic, Philosophy and Some of Their Uses', *Proceedings and Addresses of the APA* 66, pp. 70–73.

Martin-Löf, P. (1984), *Intuitionistic Type Theory*, Bibliopolis.

Neale, S. (2000), 'On a Milestone of Empiricism', in A. Orenstein and P. Kotatko eds., *Knowledge, Language and Logic*, Kluwer, pp. 237–346.

Reinhardt, W. (1986), 'Epistemic Theories and the Interpretation of Gödel's Incompleteness Theorems', *Journal of Philosophical Logic* 15, pp. 427–474.

Rogers, H. (1987[1967]), *Theory of Recursive Functions and Effective Computability*, MIT Press.

Russell, B. Whitehead, A.N. (1980[1910]), *Principia Mathematica to *56*, Cambridge University Press.

Shapiro, S. ed., (1985), *Intensional Mathematics*, North-Holland.

Shapiro, S. (1991), *Foundations Without Foundationalism. A Case for Second-Order Logic*, Clarendon-Press.

Weiermann, A. (2004), 'Analytic combinatorics and transfinite ordinals', Manuscript.