Sets and Their Sizes

by

Fredric M. Katz

S.B., Massachusetts Institute of Technology
(1971)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

April 1981

Signature of Author ______________________________
Department of Philosophy
April 7, 1981

Certified by ______________________________
George Boolos
Thesis Supervisor

Accepted by ______________________________
George Boolos
Chairperson

Abstract

Cantor's theory of cardinality violates common sense. It says, for example, that all infinite sets of integers are the same size. This thesis criticizes the arguments for Cantor's theory and presents an alternative.

The alternative is based on a general theory, called "CS" (for "Class Size"). CS consists of all sentences in the first-order language with a subset predicate and a less-than predicate which are true in all interpretations of that language whose domain is a finite power set. Thus, CS says that less-than is a linear ordering with highest and lowest members and that every set is larger than any of its proper subsets. Because the language of CS is so restricted, CS will have infinite interpretations. In particular, the notion of one-one correspondence cannot be expressed in this language, so Cantor's definition of similarity will not be in CS, even though it is true for all finite sets.

We show that CS is decidable but not finitely axiomatizable by characterizing the complete extensions of CS. CS has "finite completions", which are true only in finite models, and "infinite completions", which are true only in infinite models. An infinite completion is determined by a set of "remainder principles", which say, for each natural number n, how many atoms remain when the universe is partitioned into n disjoints subsets of the same size.

We show that any infinite completion of CS has a model over the power set of the natural numbers which satisfies an additional axiom, OUTPACING:

> If initial segments of A eventually become smaller than the corresponding initial segments of B, then A is smaller than B.

Models which satisfy OUTPACING seem to accord with common intuitions about set size. In particular, they agree with the ordering suggested by the notion of asymptotic density.

Acknowledgements

This dissertation has been made possible by Cantor--who invented set theory, Tarski--who discovered model theory, Abraham Robinson--who introduced model-completeness, and George Boolos--who made all of these subjects accessible to me. Prof. Boolos has read several handwritten versions and as many typed versions of this thesis, with care and patience each time around. I owe him thanks for finding (and sometimes correcting) my errors, for directing me to some useful literature, and simply for being interested. The reader should join me in thanking him for his help with style organization and for his insistence on filling in "mere details". Nevertheless, whatever errors, gaps, or infelicities remain are solely the responsibility of the author.

The final typescript was prepared with KAMAL, a formatting program for the IBM 8100 written by Dr. Douglas Kalish of Computer System Architects. Dr. Kalish has, in aid of this dissertation, stretched KAMAL far beyond its initial design specifications, for which I am indebted to him. Whatever typographical problems remain are solely the responsibility of the IBM 8100 and its operating system.

Finally, I would like to thank my employer, John Thron of Computer System Architects, for the use of the computer and for his considerable patience while I have completed this dissertation.

Contents

A Note to the Reader

1. The notations used in this thesis are summarized in the appendix.

2. The appendix also lists some model theoretic notions and facts used in the text.

3. There is an index in the back which covers models, theories, statements, predicates, and languages.

4. The "findings" in this thesis are mostly contained in the first three chapters and the last; chapters 4 and 5 are devoted to the proof of a completeness theorem. The last section of chapter 5 also contains some interesting corollaries of the completeness theorem and its lemmas.

1 INTRODUCTION

1.1 THE PROBLEM

This paper proposes a theory of set size which is based on intuitions, naive and otherwise. The theory goes beyond intuitions, as theories will, so it needs both justification and defense. I spend very little time justifying the theory; it is so clearly true that anyone who comes to the matter without prejudice will accept it. I spend a lot of time defending the theory because no one who comes to the matter comes without prejudice.

The prejudice stems from Cantor's theory of set size, which is as old as sets themselves and so widely held as to be worthy of the name, "the standard theory". Cantor's theory consists of just two principles:

ONE-ONE: Two sets are the same size just in case there is a one-one correspondence between them.

CANTOR<: A set, x, is smaller than a set, y, just in case x is the same size as some subset of y, but not the same size as y itself.

A "one-one correspondence" between two sets is a relation which pairs each member of either set with exactly one member of the other. For example, the upper-case letters of the alphabet can be paired with the lower-case letters:

ABCDEFGHIJKLMNOPQRSTUVWXYZ

abcdefghijklmnopqrstuvwxyz

so, the standard theory says, the set of upper-case letters is the same size as the set of lower-case letters. Fine and good.

The standard theory also says that the set of even numbers is the same size as the set of integers, since these two sets can also be paired off one-to-one:

..., -n, ..., -2, -1, 0, 1, 2, ..., n, ...

..., -2n, ..., -4, -2, 0, 2, 4, ..., 2n, ...

Similarly, the standard theory says that the set of positive even integers is the same size as the set of prime numbers: pair the n-th prime with the n-th positive even integer. In both of these cases, common sense chokes on the standard theory.

In the first case, common sense holds that the set of integers is larger than the set of even integers. The integers contain all of the even integers and then some. So it's just good common sense to believe there are more of the former than the latter. This is just to say that common sense seems to follow

> SUBSET: If one set properly includes another, then
> the first is larger than the second.

even into the infinite, where it comes up against the standard theory.

Common sense can make decisions without help from SUBSET. Though the set of primes is not contained in the set of even integers, it is still clear to common sense that the former is smaller than the latter. One out of every two integers is even, while prime numbers are few and far between. No doubt, to use this reasoning, you need a little number theory in addition to common sense; but, given the number theory, it's the only conclusion common sense allows.

The theory proposed here accommodates these bits of common sense reasoning. It maintains SUBSET and a "few and far between" principle and much else besides. To state this theory, I use three two-place predicates: '<', '=', and '>'. If 'A' and 'B' name sets, then

'A < B' is read as 'A is smaller than B',
'A = B' is read as 'A is the same size as B',
'A > B' is read as 'A is larger than B'.

Incidentally, we will assume throughout this thesis that the following schemata are equivalent, item by item, to the readings of the three predicates given above, assuming that A is the set of ALPHAS and B is the set of BETAS:

A) There are fewer ALPHAS than BETAS.
There are just as many ALPHAS as BETAS.
There are more ALPHAS than BETAS.

B) The number of ALPHAS is less than the number of BETAS.
The number of ALPHAS is the same as the number of BETAS.
The number of ALPHAS is greater than the number of BETAS.

and, finally,

C) The size of A is smaller than the size of B.
A and B are (or "have") the same size.
The size of A is larger than the size of B.

Regarding this last group, we emphasize that we are not arguing that there "really are" such things as set sizes, nor that there "aren't really" such things. Statements about "sizes" can be translated in familiar and long-winded ways into statements about sets, though we won't bother to do so.

I have identified the standard theory, Cantor's, with two principles about set size. The term 'size', however, is rarely used in connection with Cantor's theory; so it might be wondered whether the standard theory is really so standard. In stating ONE-ONE and CANTOR<, Cantor used the terms 'power' and 'cardinal number' rather than 'size'. In the literature the term 'cardinal number' (sometimes just 'number') is used most frequently. If someone introduces 'cardinal number' as a defined predicate or as part of a contextual definition (e.g. "We say two sets have 'the same cardinal number iff ..."), there is no point in discussing whether that person is right about size.

Though Cantor's theory is usually taken as a theory of set size, it can also be taken as "just" a theory of one-one correspondences. More specifically, saying that two sets are "similar" iff they are in one-one correspondence can either be taken as a claim about size or be regarded as a mere definition. Whether or not "similarity" is coextensive with "being the same size", the definition is worth making. The relation picked out is well studied and well worth the study. The technical brilliance of the theory attests to this: it has given us the transfinite hierarchy, the continuum problem, and much else. In addition, the theory has consequences which don't, *prima facie*, seem to have anything to do with size or similarity; the existence of transcendental numbers comes to mind. All of this is to say that the interest in one-one correspondence has not been sustained solely by its identification with the notion of size. Hence, denying that they are the same does not endanger the theory of one-one correspondences, *per se*.

But most mathematicians and philosophers don't use 'cardinal number' as a mere abbreviation. They use the term in just the way that we use 'size' and slide freely among (A), (B), and (C). This is true, in particular, of Cantor, who offered ONE-ONE as a theory; indeed, he offered an argument for this theory.

1.2 CANTOR'S ARGUMENT

Cantor bases his argument for ONE-ONE on the idea that the size of a set, its cardinal number, depends on neither the particular elements it contains nor on how those elements are ordered:

> (1):
>
> The cardinal number, CARD(M), of a set M (is the general concept which) arises from M when we make abstraction of the nature of its elements and of the order in which they are given. (Cantor, p.86).

But to say that the cardinal number of a set doesn't depend on certain things is not to say what the cardinal number is. Neither does it insure that two sets have the same cardinal number just in case they are in one-to-one correspondence. To flesh out this notion of "double abstraction", Cantor reduces it to a second abstraction operator, one which works on the elements of sets rather than the sets themselves:

> Every element, m, if we abstract from its nature becomes a "unit" ... (Cantor, p.86).

And so, concludes Cantor:

> The cardinal number, CARD(M) is a set composed of units

> (which has existence in our minds as an intellectual image or projection of M.) (Cantor, p.86)

According to Jourdain, Cantor

> distinguished very sharply between an aggregate and a cardinal number that belongs to it: "Is not an aggregate an object <u>outside</u> us, whereas its cardinal number is an abstract picture of it <u>in our</u> mind." (Cantor, p.80)

I have parenthesized the expressions above where Cantor describes cardinal numbers as mental entities. Nevertheless, I can only make sense of his arguments insofar as he treats cardinal numbers as sets: he refers to them as "definite aggregates", supposes that they have elements, and employs mappings between cardinal numbers and other sets.

The following three statements seem to express Cantor's intent:

1.1 CARD(M) = {y: Ex(x ∈ M & y = ABST(x)) }

1.2 (x)(Ey)(y = ABST(x) & UNIT(y))

1.3 (M)(y)(y ∈ card(M) -> UNIT(y))

"ABST(x)" is to be read as "the result of abstracting from the element x", "UNIT(x)" as "x is a unit", and "CARD(M)" as "the cardinal number of M".

So (1.1) gives a definition of cardinal number, in terms of the operation of abstraction, from which Cantor proves both ONE-ONEa and ONE-ONEb.

ONE-ONEa: CORR(M,N) -> CARD(M) = CARD(N)

ONE-ONEb: CARD(M) = CARD(M) -> CORR(M,N)

ONE-ONEa is true, says Cantor, because

> the cardinal number CARD(M) remains unaltered if in the place of one or many or even all elements m of M other things are substituted. (p. 88)

and so, if f is a one-one mapping from M onto N, then in replacing each element, m, of M with f(m)

> M transforms into N without change of cardinal number. (p.88)

In its weakest form, the principle Cantor cites says that if a single element of M is replaced by an arbitrary element not in M then the cardinal number of the set will remain the same. That is, (1.4) (a $\in$ M & $\neg$(b $\in$ M)

& N = {x: (x $\in$ M & x $\neq$ a) ; (x = b)})

-> CARD(M) = CARD(N)

The reasoning here is clear: so far as the cardinal number of a set is concerned, one element is much the same as another. It is not the elements of a set, but only their abstractions, that enter into the cardinal number of a set. But abstractions of elements are just units; so one is much the same as another.

ONE-ONEb is true, Cantor says, because

> ... CARD(M) grows, so to speak, out of M in such a way that from every element m of M a special unit of CARD(M) arises. Thus we can say that CORR(M,CARD(M)). (p.88)

So, since a set is similar to its cardinal number, and similarity is an equivalence relation, two sets with the same cardinal number are similar. Unless each element of a set abstracts to a "special", i.e. distinct, unit, the correspondence from M to its cardinal number will be many-one and not one-one. A weak version of this principle is:

(1.5) If M = {a,b} and a $\neq$ b, then

CARD(M) = {ABST(a),ABST(b)} and ABST(a) ≠ ABST(b).

These two arguments do one another in. (1.4) says that replacing an element of a set with any element not in the set does not affect the cardinality. But, by the definition of CARD(M) (1.1), this means that

(1.6) (x)(y)(ABST(x) = ABST(y))

For, consider an arbitrary pair of elements, a and b. Let M = {a} and let N = {b}. So, the conditions of (1.4) are met and CARD(M) = CARD(N). But CARD(M) = {ABST(a)} and CARD(N) = {ABST(b)}, by (1.1). Generalizing this argument yields (1.6).

So Cantor's argument for ONE-ONEa only works by assigning all non-empty sets the same, one-membered, cardinal number. But this contradicts ONE-ONEb.

Conversely, the argument that a set is similar to its cardinal number relies on (1.5), which entails

(1.7) (x)(y)(x ≠ y -> ABST(x) ≠ ABST(y))

assuming only that any two objects can constitute a set. But if the abstractions of any two elements are distinct, then no two sets have the same cardinal number as defined by (1.1), _contra_ ONE-ONEa.

There is no way to repair Cantor's argument. Rather than leading to a justification of ONE-ONE, Cantor's definition of cardinal number is sufficient to refute the principle. The negation of (1.6) is:

(1.8) ExEy(ABST(x) ≠ ABST(y))

So one of (1.6) and (1.8) must be true. We have shown that (1.6) contradicts ONE-ONEb. Similarly, (1.8) contradicts ONE-ONEa: if a and b have distinct abstractions, then {a} and {b} have distinct cardinal numbers, {ABST(a)} and {ABST(b)}, despite the fact that they are in one-one correspondence. So, ONE-ONE is false whether (1.6) or its negation, (1.8), is true.

1.3 CANTOR AND THE LOGICISTS.

Though both Frege and Russell accepted Cantor's theory of cardinality, neither accepted Cantor's argument. Frege spends an entire chapter of the Grundlagen mocking mathematicians from Euclid to Schroder for defining numbers as sets of "units". He neatly summarizes the difficulty with such views:

> If we try to produce the number by putting together different distinct objects, the result is an agglomeration in which the objects remain still in possession of precisely those properties which serve to distinguish them from one another, and that is not the number. But if we try to do it in the other way, by putting together identicals, the result runs perpetually together into one and we never reach a plurality...
>
> The word 'unit' is admirably adapted to conceal the difficulty... We start by calling the things to be numbered "units" without detracting from their diversity; then subsequently the concept of putting together (or collecting, or uniting, or annexing, or whatever we choose to call it) transforms itself into arithmetical addition, while the concept word 'unit' changes unperceived into the proper name 'one'.(pp. 50-51)

These misgivings about units don't prevent Frege from basing his definition of 'number' and his entire reduction of arithmetic on Cantor's notion of one-one correspondence. "This opinion," says Frege, "that numerical equality or identity must be defined in terms of one-one correlation, seems in recent years to have gained widespread acceptance among mathematicians" (pp.73-74). Frege cites Schroder, Kossak, and Cantor.

Russell displays similar caution about Cantor's argument (see Principles of Mathematics, p.305) and similar enthusiasm for his theory (see the quote at the beginning of chapter 2, for example.)

Of course, Frege and Russell "cleaned up" Cantor's presentation of the theory. Russell, for example, notes that Cantor's statement (1) is not a "true definition" and

> merely presupposes that every collection has some such property as that indicated -- a property, that is to say, independent of its terms and of their order; depending, we might feel tempted to add, only upon their number. (Ibid, pp. 304-305)

So Russell, and similarly Frege, relied upon the principle of abstraction to obtain a "formal definition" of cardinal numbers, in contrast to Cantor, who had "taken" number "to be a primitive idea" and had to rely on "the primitive proposition that every collection has a number." (Ibid)

So, while some people regard Cantor's ONE-ONE as "just" a definition and others embrace it as a theory, the logicists have it both ways: adding ONE-ONE as a "formal definition" to set theory (or, as they would have it, logic) they have no obligation to defend it and can steer clear of peculiar arguments about "units"; at the same time, they can advance it as a great lesson for simple common sense.

The logicists' adoption of Cantor's theory of cardinality needs no great explanation: it "came with" set theory and, to a large extent, motivated set theory and determined its research problems. But there are two specific reasons that they should have seized upon ONE-ONE and CANTOR<. First, they both have the form of definitions, no matter how they are intended. So the notion of cardinality is "born reduced".

Second, Cantor's theory clears the way for other reductions. Suppose, for example, you wish to reduce ordered pairs to sets. Well, you have to identify each ordered pair with a set and define the "relevant" properties of and relations among ordered pairs in terms of properties and relations among sets. One of the relations that has to be maintained is, of course, identity: so each ordered pair must be identified with a distinct set. In addition, the relative sizes of sets of ordered pairs should be preserved under translation. But, if ONE-ONE is the correct theory of size, then this second condition follows from the first, since the existence of one-one correspondences will be preserved under a one to one mapping.

1.4 AIMS AND OUTLINE

It would be naive to suppose that people's faith in Cantor's theory would be shaken either by refuting specific arguments for ONE-ONE or by associating the acceptance of that theory with a discredited philosophy of mathematics. Such points may be interesting, but in the absence of an alternative theory of size, they are less than convincing.

This dissertation presents such an alternative. Chapter 2 canvasses common sense intuitions for some basic principles about set size. Chapter 3 reorganizes those principles into a tidy set of axioms, offers an account of where the intuitions come from (viz. known facts about finite sets), and mines this source for additional principles to incorporate in our theory. Chapters 4 and 5 prove that the theory so obtained is "complete", in the sense that it embraces all facts about finite sets of a certain kind (i.e. expressible in a particular language). Finally, chapter 6 elaborates additional principles that concern only sets of natural numbers and demonstrates that these additional principles, together with the theory in chapter 3 are satisfiable in the domain of sets of natural numbers.

> "The possibility that whole and part may have the same number of terms is, it must be confessed, shocking to common sense ... Common sense, therefore, is here in a very sorry plight; it must choose between the paradox of Zeno and the paradox of Cantor. I do not propose to help it, since I consider that, in the face of the proofs, it ought to commit suicide in despair." (Russell, B. Principles of Mathematics, p.358)

Is common sense confused about set size, as Russell says, or is there a way of elaborating on common sense to get a plausible and reasonably adequate theory of cardinality? To be "plausible", a theory should at least avoid principles and consequences which violate common sense. To be "reasonably adequate", a theory has to go beyond bare intuitions: it should not rest with trivialities and it should answer as many questions about set size as possible, though it needn't be complete. Plausibility and adequacy are conflicting demands: the first says that there should not be too many principles (no false ones, consistency), the second that there should not be too few principles.

In the Introduction, I argued that a coherent theory of cardinality has to contain some principles that refer to the kinds of objects in sets, *pace* Cantor. In this chapter, however, I want to see how far one can go without such principles, i.e. how much you can say about the *smaller-than relation* without using predicates (other than the identity predicate) which relate the members of the sets being compared. I shall begin by stating a number of principles and explaining why they are included in the general theory.

2.1 THE THEORY CORE

First, there is the SUBSET principle:

> SUBSET: If x is a proper subset of y, then x is smaller than y.

The reason for including SUBSET should be obvious. What has prompted the search for an alternative to the standard theory of cardinality is the conflict between ONE-ONE and SUBSET. Now, it is often said that common sense supports both of these principles and that it is in doing so that common sense is confused. From this it is supposed to follow that common sense cannot be relied upon, so we should opt for ONE-ONE, with the technical attractions that it provides.

But there's a difference in the way that common sense supports these two principles. There is no doubt that you can lead an unsuspecting person to agree to ONE-ONE by focusing their attention on forks and knives, husbands and wives, and so forth: i.e. on finite sets. With carefully chosen examples, say the odds and the evens, you might even convince someone that ONE-ONE is true for infinite sets, too. Now, I don't think that such guile is need to lead someone to agree to SUBSET, but that's not what my argument depends on. The argument hinges on a suggestion about how to resolve cases where mathematical intuitions seem to conflict. The suggestion is to see what happens with particular cases on which the principles conflict _before_ you've lead someone to agree to either of the general statements.

So, if you want to find out what common sense really thinks about SUBSET and ONE-ONE, you would present people with pairs of infinite sets, where one was a proper subset of the other. I've actually tried this, in an unscientific way, and what I've gotten, by and large, is what I expected: support for the subset principle. ("By and large" because many people think that all infinite sets have the same size: "Infinity".)

Naturally, I wouldn't venture that this sort of technique, asking people, is any way to find out which of SUBSET and ONE-ONE is _true_. People's intuitions about mathematics are notoriously unreliable, not to mention inconstant. Of course harping on this fact might engender some unwarranted skepticism about mathematics. What I am suggesting is that there might be a rational way of studying mathematical intuitions and that we should at least explore this possibility before proclaiming common sense to be hopelessly confused on mathematical matters.

So, SUBSET, all by itself, seems to be a plausible alternative to Cantor's theory, though it surely isn't enough. Given just this principle, it's possible that one set is smaller than another just in case it is a proper subset of the other. It's clear that we need some additional principles. All of the following seem worthy (where "$x < y$" is to be read as "x is smaller than y", "$x = y$" as "x is the same size as y", and "$x > y$" as "x is larger than y". The Appendix gives a full account of the notations used throughout this thesis.)

Theory 2.1.1 QUASI-LOGICAL

ASYM<: $x < y \rightarrow \neg\, y < x$

ASYM>: $x > y \rightarrow \neg\, y > x$

TRANS<: $(x < y \,\&\, y < z) \rightarrow x < z$

TRANS>: $(x > y \,\&\, y > z) \rightarrow x > z$

INDISC=: $x = y \rightarrow \mathrm{INDISC}(x,y)$

SYM=: $x = y \rightarrow y = x$

REF=: $x = x$

TRANS=: $(x = y \,\&\, y = z) \rightarrow x = z$

DEF>: $x > y \leftrightarrow y < x$

where

INDISC(x,y) abbreviates

$$(z)((z < x \leftrightarrow z < y) \,\&\, (x < z \leftrightarrow y < z))$$

We shall call the principles listed above "quasi-logical principles" because it is tempting to defend them as logical truths. Consider the first principle, for example, in unregimented English:

ASYM<: If x is smaller than y, then y is not smaller than x.

This sentence can be regarded as an instance of the schema:

ASYM{F}: If x is F-er than y, then y is not F-er than x.

where 'F' is to be replaced by an adjective from which comparatives can be formed, e.g. 'tall', 'short', 'happy', 'pretty', but not 'unique', or 'brick'. It appears that every instance of this schema is true, so it could be maintained that each instance is true in virtue of its form, that each is a logical truth.

The other principles might be defended in the same way, though the schema for INDISC= would have to be restricted to triples of corresponding comparatives, for example: 'is smaller than', 'is larger than', 'is the same size as'.

But using such observations to support these principles would be problematic for two reasons. First, it would require taking positions on many questions about logical form and grammatical form which would take us far afield and, possibly, antagonize first-order logicians. Second, there are some instances of the schemata that make for embarrassing counterexamples: "further east than" (in a round world) and "earlier than" (in, I'm told, a possible world).

So, it might be that casting the principles above as instances of the appropriate schemata would only explain why they are part of common sense. What remains clear is that a theory of cardinality which openly denied any of these principles would be implausible: it would be ridiculed by common sense and mathematical sophisticates alike. I can just barely imagine proposing a theory which, for fear of inconsistency, withheld judgment on on or more of these principles. But to do so without good reason would be counterproductive. It seems that if a case could be made that these statements, taken together, are inconsistent with SUBSET, that would be good reason to say that there is no reasonably adequate alternative to Cantor's theory. Since my goal is to counter such a conclusion, it seems that the proper strategy is to include such seemingly obvious truths and to show that the resulting theory is consistent. So the strategy here is not to adduce principles and argue for the truth of each. This would be impossible, given that most of the principles are logically contingent. Instead, the idea is to canonize what common sense holds to be true about cardinality and to show that the result is consistent and reasonably adequate.

So long as we restrict ourselves to SUBSET and the quasi-logical principles, consistency is no problem. After all, what do the quasi-logical principles say? Only that *smaller-than* is a partial ordering, that *larger-than* is the converse partial ordering, that *the same size as* is an equivalence relation, and that sets of the same size are indiscernible under the partial orderings. So, if we are given any domain of sets, finite or infinite, we get a model for our theory by assigning to '<' the relation of *being a proper subset of*, assigning to '>' the relation of *properly including*, and assigning to '=' the identity relation. Since common sense knows that different sets can have the same size, there must be some additional principles to be extracted from common sense.

We shall now consider some principles which cannot be regarded as quasi-logical.

First, there is TRICHOTOMY,

TRICH: x < y ¦ x = y ¦ y < x

which says that any two sets are comparable in size. While a theory of set size which excluded TRICH might escape ridicule, it would surely be regarded with suspicion. Indeed, if the principles of common sense were incompatible with TRICH, this would undoubtedly be used to discredit them.

Second, there is the "representation principle",

$$\text{REP<}: x < y \rightarrow (Ex')(x' \approx x \;\&\; x' \subset y)$$.

which says that if a set, x, is smaller than another, y, then x is the same size as some proper subset of y. Now this is a principle which common sense has no particular feelings about. Analogous statements about physical objects are neither intuitive nor very clearly true. For example, (1) doesn't stand a chance of being regarded as true

> (1) If one table is smaller than another, then the first is the same size as some proper part of the second.

if 'part' is taken to mean 'leg or top or rim or ...'. Even if common sense can be persuaded to take particles of tables and arbitrary fusions of such as parts of tables, no one should condemn its residual caution about (1). If REP< is true, then I think that that's an interesting and special fact about sets.

REP< was originally included in this theory for "technical" reasons which will emerge; it makes it easier to reduce the set of axioms already presented and it provides a basis for several principles not yet presented. REP< may be open to doubt, but it is not a principle that Cantorians could complain about, for it is entailed by Cantor's definition of '<':

$$\text{CANTOR<}: x < y \leftrightarrow \neg(x = y) \,\&\, (Ex')(x' = x \,\&\, x' \subset y)$$

Note that if CANTOR< is regarded as a principle instead of a definition, then it is entailed by the principles we have already mentioned:

> If $x < y$, then $\neg(x = y)$, by INDISC= and ASYM<. By REP<, some proper subset of y, say x', must be the same size as x. But $x' < y$, by SUBSET; so $x < y$, by INDISC-.

There are more principles to come, but before proceeding, I'd like to take stock of what we already have. First, I want to reduce all of the principles mentioned above to a tidy set of axioms. Second, I want to estimate how far we've gone.

The entire set of principle already adopted are equivalent to the following, which will be referred to as "the core theory".

Theory 2.1.2. CORE

SUBSET: $x \subset y \rightarrow x < y$

DEF>: $x > y \leftrightarrow y < x$

DEF≈: $x \approx y \leftrightarrow \mathrm{INDISC}(x,y)$

REP<: $x < y \rightarrow (Ex')(x' \approx x \;\&\; x' \subset y)$

IRREF<: $\neg (x < x)$

TRICH: $x < y \mid x \approx y \mid y < x$

The only axiom in CORE that has not already been introduced is DEF≈, which is logically equivalent to the conjunction of INDISC≈ and its converse, ≈INDISC:

INDISC≈: $x \approx y \rightarrow \mathrm{INDISC}(x,y)$

≈INDISC: $\mathrm{INDISC}(x,y) \rightarrow x \approx y$

≈INDISC says that if two sets fail to be the same size, then their being different in size is attributable to the existence of some set which is either smaller than one but not smaller than the other or larger than one but not larger than the other.

Theorem 2.1.3.

Let T = QUASI-LOGICAL;SUBSET;TRICH;REP<.

Then CORE ≡ T.

Proof:

(a) T |- CORE

We only need to show that ≈INDISC is entailed by T. Suppose ¬(x ≈ y). So x < y or y < x, by TRICH. But ¬(x < x) and ¬(y < y), by IRREF.

(b) CORE |- T

(i) TRANS<. If y < z, there is a y' such that y' ≈ y and y' c z, by REP<. If x < y, then x < y' because y' ≈ y, by DEF≈. So there is an x' such that x' ≈ x and x' c y'. So x' c z, and, by SUBSET, x' < z. But then x < z by DEF≈.

(ii) ASYM<. If x < y and y < x, then x < x, by TRANS<, contra IRREF<.

(iii) TRANS>, ASYM>, and IRREF> follow from the corresponding principles for '<' and DEF>.

(iv) INDISC≈, SYM≈, TRANS≈, and REF≈ are logical consequences of DEF≈.

CORE is consistent. In fact, there are two kinds of models which satisfy CORE.

Definition 2.1.4

(a) $\underline{A}$ is a *finite class model with basis* x iff

(i) $A = P(x)$ for some finite set x.

(ii) $\underline{A} \models a \subset b$ iff a is a proper subset of b.

(iii) $\underline{A} \models a < b$ iff $cd(a) < cd(b)$.

(b) $\underline{A}$ is a *finite set model* iff

(i) $A = \{x : x$ is a finite subset of $Y\}$ for some infinite Y.

(ii) $\underline{A} \models a \subset b$ iff a is a proper subset of b.

(iii) $\underline{A} \models a < b$ iff $cd(a) < cd(b)$.

Note that models can be specified by stipulating the *smaller-than* relation since *larger-than* and *the same size as* are defined in terms of *smaller-than*.

Fact 2.1.5

(a) If $\underline{A}$ is a finite class model, $\underline{A} \models$ CORE.

(b) If $\underline{A}$ is a finite set model, $\underline{A} \models$ CORE.

Proof: In both cases, the finite cardinalities determine a quasi-linear ordering of the sets in which any set is higher than any of its proper subsets.

The normal ordering of finite sets by size is, in fact, the only one that satisfies CORE. By adding TRICH we have ruled out all non-standard interpretations of '<'.

Theorem 2.1.6. Suppose that $\underline{A}$ is a model such that

(i) If $x \in A$, x is finite,

(ii) If $x \in A$ and $y \subset x$, then $y \in A$, and

(iii) $\underline{A} \models$ CORE.

Then,

(iv) $\underline{A} \models x < y$ iff $cd(x) < cd(y)$.

Proof: We prove (*) for every n by induction.

(*) If $cd(a) = n$, then $\underline{A} \models (a \approx b)$ iff $cd(b) = n$.

First suppose $n = 0$.

If $cd(a) = 0$, $a = \emptyset$.

So if $cd(b) = 0$, $\underline{A} \models (a \approx b)$ by REF≈.

If $cd(b) \neq 0$, then $b \neq \emptyset$.

So $a \subset b$ and, by SUBSET, $\underline{A} \models (a < b)$.

So, by INDISC≈, $\underline{A} \not\models (a \approx b)$.

Now suppose (*) is true for all $i \leq n$.

If $cd(a) = cd(b) = n + 1$, then $\underline{A} \models (a \approx b)$:

Suppose $\underline{A} \models (a < b)$.

So, $\underline{A} \models (a' \approx a)$, for some $a' \subset b$, by REP<

But if $a' \subset b$, then $cd(a') \leq n$;

so, $cd(a) \leq n$, by (*), <u>contra</u> our hypothesis

Now suppose $cd(a) = (n+1)$ and $\underline{A} \models (a \approx b)$.

We claim: $cd(b) = (n+1)$.

Because $cd(b) \geq (n+1)$, by induction.

But if $cd(b) > (n+1)$,

choose $b' \subset b$, with $cd(b') = (n+1)$.

$b' \in A$ by condition (ii).

So $\underline{A} \models (b' < b)$, by SUBSET

and $\underline{A} \models (b' \approx a)$, (see above).

Thus, $\underline{A} \models a < b$, <u>contra</u> our supposition.

So $cd(b) = (n+1)$.

2.2 ADDITION OF SET SIZES

We shall now extend CORE to get an account of addition of set sizes. Since the domains of our intended models contain only sets and not sizes of sets, we have to formulate our principles in terms of a three-place predicate true of triples of sets: 'SUM(x,y,z)' is to be read as "the size of z is the sum of the sizes of x and y".

The following principles are sufficient for a theory of addition.

Theory 2.2.1 Addition

FUNC+: Functionality of addition

(a) $SUM(x,y,z) \rightarrow (x \approx x' \leftrightarrow SUM(x',y,z))$

(b) $SUM(x,y,z) \rightarrow (y \approx y' \leftrightarrow SUM(x,y',z))$

(c) $SUM(x,y,z) \rightarrow (z \approx z' \leftrightarrow SUM(x,y,z'))$

DISJ+: Law of addition for disjoint sets.

$$x \wedge y = \emptyset \rightarrow SUM(x,y,x \vee y)$$

MONOT+: Monotonicity of addition

$$SUM(x,y,z) \rightarrow x < z \mid x \approx z$$

The functionality of addition says that sets bear SUM relations to one another by virtue of their sizes alone. This condition must clearly be met if 'SUM' is to be read as specified above.

DISJ+ tries to say what function on sizes the SUM relation captures by fixing the function on paradigm cases -- disjoint sets. But FUNC+ and DISJ+ leave open the possibility that addition is "cyclic": suppose we begin with a finite class model whose basis has n elements and assign to SUM those triples <x,y,z> where

cd(z) = (cd(x) + cd(y)) modulo n+1.

Both FUNC+ and DISJ+ will be satisfied, though the interpretation of SUM does not agree with its intended reading. MONOT+ rules out such interpretations.

Given an interpretation of '<' over a power set there is at most one way of interpreting SUM which satisfies ADDITION. We shall show this by proving that ADDITION and CORE entail DEF+:

DEF+: SUM(x,y,z) <->
Ex'Ey'(x = x' & y = y' & x' /\ y' = ∅ & x' \/ y' = z)

So, if DEF+ is true, the extension of 'SUM' is determined by the extension of '<'.

A model of CORE must satisfy an additional principle, DISJU, if SUM is to be interpreted in a way compatible with ADDITION.

DISJU: $(x \approx x' \ \& \ y \approx y' \ \& \ x \wedge y = \emptyset \ \& \ x' \wedge y' = \emptyset) \rightarrow$

$(x \vee y) \approx (x' \vee y')$.

(Note: The proofs in this chapter will use boolean principles freely, despite the fact that we haven't yet introduced them.)

Theorem 2.2.2. CORE + ADDITION |- DISJU Proof:

Suppose $x \approx x'$,

$y \approx y'$,

$x \wedge y = \emptyset$,

and $x' \wedge y' = \emptyset$.

Then SUM$(x, y, x \vee y)$

and SUM$(x', y', x' \vee y')$, by DISJ+

so SUM$(x', y, x \vee y)$, by FUNC+(a)

so SUM$(x', y', x \vee y)$, by FUNC+(b)

so $(x \vee y) \approx (x' \vee y')$, by FUNC+(c).

We shall now use this fact to show that if the minimal conditions on addition are to be satisfied in a model of CORE, then SUM has to be definable by DEF+.

Theorem 2.2.3. CORE + ADDITION |- DEF+.

Proof:

<-. Assuming that the matrix of the right hand side of DEF+ is satisfied,

then SUM(x',y',z), by DISJ+

so SUM(x, y',z), by FUNC+(a)

so SUM(x, y, z), by FUNC+(b).

->. Suppose SUM(x,y,z)

so x < z or x ≈ z, by MONOT+.

But if x ≈ z,

let x' = z

and y' = ∅;

then SUM(x',y',z), by DISJ+

so SUM(x, y',z), by FUNC+(a)

so y' ≈ y, by FUNC+(b).

and if x < z

pick x' c z

with x' ≈ x, by REP<.

let y' = z - x'

but y' ≈ y, as before.

Theory 2.2.4. EXCORE, "the extended core", is

CORE \/ ADDITION

Fact 2.2.5. The following are consequences of EXCORE:

(a) (x /\ y1 = x /\ y2 = ∅ & y1 ≃ y2)

-> x \/ y1 ≃ x \/ y2

(b) (x1 /\ y1 = x2 /\ y2 = ∅ & x1 ≃ x2 & y1 < y2)

-> x1 \/ y1 < x2 \/ y2

(c) (x /\ y1 = x /\ y2 = ∅ & y1 < y2)

-> x \/ y1 < x \/ y2

(d) (x c z & y c z & x < y) (RC<)

-> (z - y) < (z - x)

(e) (x c z & y c z & x ≃ y)) (RC≃)

-> (z - y) ≃ (z - x)

(f) x < y -> (Ey')(y' ≃ y & x c y')

(g) x ≃ y -> (x - (x /\ y)) ≃ (y - (x /\ y))

(h) x < y -> (x - (x /\ y)) < (y - (x /\ y))

The proofs are elementary.

3 A FORMAL THEORY OF CLASS SIZE

In the preceding chapter, we searched for principles that accord with pre-Cantorian ideas about sizes of sets. We produced several such principles, constituting EXCORE, and found two kinds of interpretations which satisfied these principles. In one case, the domains of the interpretations were power sets of finite sets; in the other case the domains consisted of all finite subsets of a given infinite set. Insofar as both kinds of interpretation had the quantifiers ranging over finite sets, we may say that they demonstrate that EXCORE is true when it is construed as being about finite sets.

Our goal, of course, is to show that the general theory of set size we present can be maintained for infinite sets as well as for finite sets. We shall show this by constructing a model for the general theory whose domain is the power set of the natural numbers. But the ability to construct such a a model is interesting only to the extent that the general theory which it satisfies is reasonably adequate. Suppose, for example, that we offered as a general theory of size the axioms of CORE other than REP<. Call this theory 'T'. So T just says that <u>smaller-than</u> is a quasi-linear ordering which extends the partial ordering given by the proper subset relation. Since any partial ordering can be extended to a quasi-linear ordering, T has a model over P(N). But unless we have a guarantee that the model constructed will satisfy, say, DISJU, the existence of the model does not rule out the possibility that T is incompatible with DISJU. For a particular principle, ø, in this case DISJU, we may take one of three tacks: (1) add ø to T, obtaining T', and show that T' has a model over P(N); (2) show that ø _is_ inconsistent with T and argue that T is somehow more fundamental or more intuitive than ø; (3) acquiesce in ignorance of whether T and ø are compatible and argue that _if_ they are incompatible then T should be maintained anyway.

Below, we deal with DISJU as in (1), since DISJU is in EXCORE. Cantor's principle ONE-ONE is dealt with as in case (2). It seems futile to try to rule out the need to resort to the third approach for any cases at all, but we can reduce this need to the extent that we include in our general theory, T, as many plausible statements as possible.

Of course, we can't construct T by taking _all_ statements which are true for finite sets; not only is ONE-ONE such a statement, but using the notion of _all statements true for finite sets_ presupposes that we have some idea of the range of "all statements". To avoid the problems involved in speaking of "all statements", we might instead settle for "all statements in L", where L is some judiciously chosen language. To avoid ONE-ONE, L must fall short of the full expressive power of the language of set theory.

Consider, now, the axioms in EXCORE. Other than size relations, these axioms involve only boolean operations and inclusion relations among sets. They do not use the notions of "ordered pair", "relation", or "function". In short, the only set theory implicit in these axioms is boolean algebra, or a sort of "Venn diagram set theory". This is not to say that the axioms do not apply to relations, functions, or other sets of ordered pairs, but only that they do not refer to these sorts of objects *as such*.

In the next section, we define a language just strong enough to express EXCORE. We then construct a theory by taking all statements *in this language* which are true over all finite power sets. By drawing statements only from this relatively weak language, we arrive at a theory which can be satisfied over infinite power sets. But since we include in the theory *all* statements of the language which are true over any finite power set, we know that no statement in that language can arise as something which ought to be true over infinite power sets but might be incompatible with our theory.

There remains the possibility that we could follow the same strategy with a more expressive language, though it would have to remain less expressive than the full language of set theory. In fact, we shall indicate in the final chapter that such a language can be obtained by including a notion of the product of set sizes. This in turn opens the possibility of a succession of richer languages and a corresponding succession of stronger theories of size. At this point, the possible existence of any such hierarchy is sheer speculation; we mention it only to emphasize that no claim is made here that we have the strongest possible general theory of set size.

3.1 CS - THE THEORY OF CLASS SIZE

The theories discussed in this paper will be formulated within first order predicate logic with identity. To specify the language in which a theory is expressed, then, we need only list the individual constants, predicates, and operation symbols of the language and stipulate the _rank_, or number of argument places, for each predicate and each operation symbol.

Definition 3.1.1

(a) L(C), _the language of classes_, is the first order language with individual constants ∅ and I, the one-place predicate ATOM, the two place predicate c, and the two-place operation symbols -, /\ , and \/ .

(b) L(<), *the language of size* is the first order language with the one-place predicate UNIT, the two-place predicates < and =, and the three place predicate SUM.

(c) L(C<), *the language of class size*, is the first order language containing all and only the non-logical constants in L(C) and L(<).

Following the strategy outlined above, we define the theory of class size in terms of interpretations of L(C<) over finite power sets.

Definition 3.1.2

(a) If L(C) $\subseteq$ L, then $\underline{A}$ *is a standard interpretation of* L iff

(i) A = P(x) for some set x, and

(ii) $\underline{A}$ assigns the usual interpretations to all constants of L(C), thus:

$\underline{A}$(I) = x,

$\underline{A}$(Ø) = Ø,

$\underline{A}$ |= (a c b) iff a c b.

(b) If $\underline{A}$ is a standard interpretation, and A = P(x), then x is the *basis* of $\underline{A}$, B($\underline{A}$).

Definition 3.1.3 $\underline{A}$ is a <u>standard finite interpretation</u> of L(C<) iff

(i) $\underline{A}$ is a standard interpretation of L(C<),

(ii) $\underline{A}$ has a finite basis, and

(iii) $\underline{A} \models (a < b)$ iff $cd(a) < cd(b)$

$\underline{A} \models (a \approx b)$ iff $cd(a) = cd(b)$

$\underline{A} \models UNIT(a)$ iff $cd(a) = 1$

Definition 3.1.4 CS, <u>the theory of class size,</u> is the set of all sentences of L(C<) which are true in all finite standard interpretations of L(C<).

By drawing only on principles which can be stated in L(C<) we have at least ruled out the most obvious danger of paradox. That is to say, since the notion of one-to-one correspondence cannot be expressed in this language, Cantor's principle ONE-ONE will not be included in the theory CS, even though it is true over any finite power set.

Since CS has arbitrarily large finite models, it has infinite models. It isn't obvious that CS has <u>standard</u> infinite models, in which the universe is an infinite power set. In chapter 6 we show that such models do exist.

The present chapter is devoted to getting a clearer picture of the theory CS. Section 2 develops a set of axioms, CA, for CS. Section 3 outlines the proof that CA does indeed axiomatize CS. This proof is presented in chapter 5, after a slight detour in chapter 4.

3.2 CA - AXIOMS FOR CS

Here we shall develop a set of axioms, CA, for the theory CS. This will be done in several stages.

3.2.1 BA - Axioms for atomic boolean algebra.

We'll begin with the obvious. Since all of the universes of the interpretations mentioned in the definition of CS are power sets, they must be atomic boolean algebras and, so, must satisfy BA:

Definition 3.2.1. BA is the theory consisting of the following axioms:

$$x \vee y = y \vee x$$

$$x \wedge y = y \wedge x$$

$$x \vee (y \vee z) = (x \vee y) \vee z$$

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z$$

$$x \wedge (y \vee z) = (x \wedge y) \vee (y \wedge z)$$

$$x \vee (y \wedge z) = (x \vee y) \wedge (y \vee z)$$

$$x \wedge (I - x) = \emptyset$$

$$x \vee (I - x) = I$$

$$x \subset y \leftrightarrow (x \vee y) = y \ \& \ x \neq y$$

$$ATOM(x) \leftrightarrow (y)(y \subset x \leftrightarrow y = \emptyset)$$

$$x \neq \emptyset \rightarrow (Ey)(ATOM(y) \ \& \ (y \subset x \mid y = x))$$

Remark. These axioms are adapted from (Monk, Def. 9.3, p. 141 and Def. 9.28, p.151). Though we refer to this theory simply as BA, note that atomicity is included.

BA is clearly not a complete axiomatization of CS, since BA doesn't involve any size notions. But BA does entail all sentences in CS which do not themselves involve size notions. To show this we need to draw on some established facts about the complete extensions of BA (in the language L(C)). The key idea here is that complete extensions of BA can be obtained either by stipulating the finite number of atoms in a model or by saying that there are infinitely many atoms.

Definition 3.2.2. For $n \geq 1$,

(a) ATLEAST{n} is a sentence which says that there are at least n atoms:

Ex1...Ex.n(ATOM(x1) & ... & ATOM(x.n)
& (& {(x.i /\ x.j = Ø) : 0 < i < j ≤ n}))

(b) EXACTLY{n} is a sentence which says that there are exactly n atoms:

ATLEAST{n} & ¬ATLEAST{n+1}

(c) INF is a set of sentences which is satisfied in all and only infinite models of BA:

INF = {ATLEAST{n} : n > 1}

(d) BA{n} = BA; EXACTLY{n}

(e) BAI = BA \/ INF

Fact 3.2.3

(a) For $n \geq 1$, BA{n} is categorical (Monk, Cor 9.32, p.152)

(b) For $n \geq 1$, BA{n} is complete. (Immediate from (a)).

(c) For $n \geq 1$, any n-atom [illegible]ic boolean algebra is isomorphic to any finite standard interpretation of L(C) with an n-element basis. (Monk, Prop. 9.30, p.151).

(d) BAI is complete. (Monk, Theorem 21.34, p. 360).

(e) BAI and the theories BA{n}, $n \geq 1$, are the only complete, consistent extensions of BA.

Fact 3.2.4. If BAI $\vdash \phi$, then ϕ is true in some finite model of BA.

Proof:

BA $\cup$ INF $\vdash \phi$.

By compactness, then, there is a k such that

$$\text{BA} \cup \{\text{ATLEAST}\{n\} : 1 \leq n \leq k\} \vdash \phi$$

So ϕ is true in any atomic boolean algebra with more than k atoms.

Theorem 3.2.5. If CS $\vdash \phi$, and $\phi \in L(C)$, then BA $\vdash \phi$.

Proof:

If BA $\nvdash \phi$, then $\neg\phi$ is true in some atomic boolean algebra, $\underline{A}$. If $\underline{A}$ is finite, then $\underline{A}$ is isomorphic to some finite standard interpretation $\underline{A}'$ of L(C). But then $\neg\phi$ is true in $\underline{A}'$, so ϕ is not true in $\underline{A}'$ and ϕ is not in CS.

If $\underline{A}$ is infinite, then $\neg\phi$ is consistent with BAI. But BAI is complete, so BAI $\vdash \neg\phi$. By 3.2.4, $\neg\phi$ is true in some finite model $\underline{A}$ of BA. Hence, $\neg\phi$ is true in some finite standard interpretation of L(C) and, again, ϕ is not in CS.

3.2.2 Size principles

Here we just gather the principles presented above as EXCORE.

Definition 3.2.6. SIZE consists of the following axioms:

SUBSET: $x \subset y \rightarrow x < y$

DEF>: $x > y \leftrightarrow y < x$

DEF≈: $x \approx y \leftrightarrow \mathrm{INDISC}(x, y)$

REP<: $x < y \rightarrow (Ex')(x' \approx x \ \& \ x' \subset y)$

IRREF<: $\neg (x < x)$

TRICH: $x < y \mid x \approx y \mid y < x$

DEF+: $\mathrm{SUM}(x, y, z) \leftrightarrow$

$$Ex'Ey'(\ x \approx x' \ \& \ y \approx y' \ \& \ x' \cap y' = \emptyset \ \& \ x' \cup y' = z)$$

DISJU: $(x \approx x' \ \& \ y \approx y' \ \& \ x \cap y = \emptyset \ \& \ x' \cap y' = \emptyset)$

$$\rightarrow \ (x \cup y) \approx (x' \cup y').$$

DEF1: $\mathrm{UNIT}(x) \leftrightarrow \mathrm{ATOM}(x)$

Combining the principles of boolean algebra and the size principles, we obtain our first serious attempt at a general theory of set size.

Definition 3.2.7. BASIC = BA $\cup$ SIZE

3.2.3 Division principles

Basic is not a complete axiomatization of CS. In this section we shall exhibit an infinite number of principles which need to be added to BASIC in order to axiomatize CS. When we are done we will have an effective set of sentences, CA (for Class-Size, Axiomatic), though we will not prove that CA $\equiv$ CS until chapter 5.

To show that the new principles really do need to be added, we'll need some non-standard models of BASIC. These models will be similar in that (1) their universes will be subsets of P(N), (2) their atoms will be the singletons in P(N), and (3) all boolean symbols will receive their usual interpretations. The models will, however, include different subsets of N and will also assign different size orderings to the these sets.

In chapter 6, these models of BASIC will reappear as submodels of various standard models of CS over P(N). So, in addition to the immediate purpose of establishing independence results, these models provide a glimpse of how sets of natural numbers are ordered by size.

Every finite standard interpretation, $\underline{A}$, of L(C<) satisfies exactly one of the following:

$$\text{EVEN: } (Ex)(Ey)(x \approx y \ \& \ x \wedge y = \emptyset \ \& \ x \vee y = I)$$

$$\text{ODD: } ExEyEz(x \approx y \ \& \ x \wedge y = x \wedge z = y \wedge z = \emptyset$$
$$\& \ \text{ATOM}(z) \ \& \ x \vee y \vee z = I)$$

$\underline{A} \models$ EVEN if cd(B($\underline{A}$)) is even and $\underline{A} \models$ ODD if cd(B($\underline{A}$)) is odd. So CS $\vdash$ (EVEN ¦ ODD). But BASIC $\nvdash$ (EVEN ¦ ODD). Consider the model $\underline{F}$ whose universe consists of all and only the finite and cofinite subsets of N, where (for a and b in F):

$\underline{F} \models (\underline{a} < \underline{b})$ iff

(i) a and b are finite and cd(a) < cd(b), or

(ii) a and b are cofinite and cd(N−a) > cd(N−b), or

(iii) a is finite and b is infinite.

$\underline{F}$ is a model of BASIC. But neither EVEN nor ODD is true in F, for any two sets that are the same size are either both finite, in which case their union is also finite, or both cofinite, in which case they cannot be disjoint.

So (EVEN ¦ ODD) is in CS, but not entailed by BASIC. As you might suspect, this is just the tip of the iceberg of principles missing from an axiomatization of CS. Informally, we can extend $\underline{F}$ to a model that satisfies EVEN by including the set of even numbers and the set of odd numbers and making them the same size. To round out the result to a model of BASIC, we also need to include all sets which are "near" the set of evens or the set of odds, i.e. those that differ from the evens or odds by a finite set. With these additions made, the new model will be closed under boolean operations and will satisfy BA. There is a (unique) way of ordering these added sets by less-than that will satisfy BASIC: rank them according to the size and direction of their finite difference from the odds or the evens. So, we can construct an infinite model of BASIC;(EVEN ¦ ODD).

But this model will still not satisfy CS, as we can see by generalizing the argument above. (EVEN ¦ ODD) says that the universe is "roughly divisible" by two: EVEN says that the universe is divisible by two without remainder; ODD says that there is a remainder of a single atom. We can construct a similar statement that says the universe is "roughly divisible" by three -- with remainder 0, 1, or 2. As with (EVEN ¦ ODD), this statement will be in CS; but it will be satisfied by neither our original model nor the model as amended. Again, we can extend the model and again we can produce a statement of CS which is false in the resulting model.

We shall now formalize this line of reasoning.

Definition 3.2.8. If 0 ≤ m < n,

(a) MOD{n,m} is the sentence

Ex1...Ex.nEy1...y.m (

(ATOM(x1) & ... & ATOM(x.n)

& (& {x.i /\ x.j = ∅ : 1 ≤ i < j ≤ n})

& (& {y.i /\ y.j = ∅ : 1 ≤ i < j ≤ m})

& (ATOM(y1) & ... & ATOM(y.m)

& (x1 \/ ... \/ x.n) /\ (y1 \/ ... y.m) ≠ ∅

& (x1 \/ ... \/ x.n) \/ (y1 \/ ... y.m) = I)

MOD{n,m} says that the universe can be partitioned into n sets of the same size and m remaining atoms.

(b) DIV{n} is the sentence

MOD{n,0} ; ... ; MOD{n,n-1},

which says that the universe can be roughly divided into n sets of the same size with fewer than n atoms remaining.

Fact 3.2.9. If 0 ≤ m < n and A is a finite standard interpretation of L(C<), then

(a) A |= MOD{n,m} iff cd(B(A)) ≡ m mod n.

(b) A |= DIV{n}.

(c) CS |- DIV{n}.

Theorem 3.2.10. BASIC $\not\vdash$ CS.

Proof:

If $n > 1$, BASIC $\not\vdash$ DIV{n}. The model $\underline{F}$ defined above satisfies BASIC but not DIV{n}, for any n finite sets have a finite union and any two cofinite sets overlap. So, BASIC $\not\vdash$ CS, by 3.2.9c.

We could consider adding all DIV{n} sentences to BASIC in the hope that this would yield a complete set of axioms for CS. I did consider this, but it doesn't work. To demonstrate this, we need some independence results for sets of DIV sentences.

Definition 3.2.11

(a) $DIV(J) = \{ DIV\{n\} : n \in J\}$

(b) $BDIV(J) = BASIC \cup DIV(J)$

(c) $BDIV(j) = BDIV(\{j\})$

Remark: "DIV" appears on the right hand side of (a) as the name of the schematic function defined in 3.2.8b. (a) defines a function, whose name is "DIV", from sets of natural numbers to theories.

Our independence results will be obtained by constructing models of BASIC which satisfy specific sets of DIV sentences. To build such models from subsets of N, we shall include sets which can be regarded as fractional portions of N.

Definition 3.2.12. For $n > 0$,

(a) x is an <u>n-congruence class</u> iff $x = \{n*\underline{k} + m\}$ for some m, $0 \leq m < n$.

(b) x is an <u>n-quasi-congruence class</u> iff x is the union of finitely many n-congruence classes.

(c) x is a <u>congruence class</u> iff x is an n-congruence class for some n.

(d) x is a <u>quasi-congruence class</u>) iff x is an n-quasi-congruence class for some n.

(e) $QC(n) = \{x : x \text{ is an n-quasi-congruence class}\}$

(f) $QC = \bigvee \{QC(n) : 0 < n\}$

Examples:

(a) The set of evens, $\{2k\}$, and the set of odds, $\{2k+1\}$, are both 2-congruence classes.

(b) $\{3k + 2\}$ is a 3-congruence class.

(c) N is a 1-congruence class;

(d) N is an n-quasi-congruence class for every $n > 0$:

$$N = \{n\underline{k} + 0\} \vee \cdots \vee \{n\underline{k} + (n-1)\}$$

Fact 3.2.13

(a) If $x \in QC(n)$ and $y \in QC(n)$, then $(x \vee y) \in QC(n)$.

(b) If $x \in QC(n)$, then $(N - x) \in QC(n)$.

Proof:

(a) Suppose

$x = a1 \vee \ldots \vee a.k$ and

$y = b1 \vee \ldots \vee b.j$

Then

$(x \vee y) = a1 \vee \ldots \vee a.k \vee b1 \vee \ldots b.j$

(b) Note that N itself is the union of n n-congruence classes. If x is the union of m of these classes, then (N - x) union of the remaining (n-m) n-congruence classes.

Definition 3.2.14. x <u>is near</u> y, NEAR(x,y), iff (x - y) and (y - x) are both finite.

Fact 3.2.15. x is near y iff there exist finite sets w1 and w2 such that $x = (y \vee w1) - w2$.

Proof:

(->) Let $w1 = (x - y)$ and let $w2 = (y - x)$

(<-) Suppose $x = (y \vee w1) - w2$. Then $(x - y) \subseteq w1$ and $(y - x) \subseteq w2$, so (x - y) and (y - x) are finite.

Fact 3.2.16. If $x1 \subseteq x \subseteq x2$, x1 is near y, and x2 is near y, then x is near y.

Proof:

Since $x \subseteq x2$, $(x - y) \subseteq (x2 - y)$.

But $(x2 - y)$ is finite, so $(x - y)$ is finite.

Since $x1 \subseteq x$, $(y - x) \subseteq (y - x1)$.

But $(y - x1)$ is finite, so $(y - x)$ is finite.

Fact 3.2.17. NEAR is an equivalence relation.

Proof:

(a) x is near x, since $x - x = \emptyset$, which is finite.

(b) If x is near y, then y is near x. Immediate.

(c) Suppose x is near y and y is near z. Note that

$$(z - x) = ((z \wedge y) - x) \vee ((z - y) - x).$$

But $(z - y) - x$ is finite because $(z - y)$ is finite, and $((z \wedge y) - x)$ is finite because

$$((z \wedge y) - x) \subseteq (y - x),$$ which is finite.

So the union, $(z - x)$ is finite.

Similarly,

$$(x - z) = ((x \wedge y) - z) \vee ((x - y) - z),$$

where $((x \wedge y) - z) \subseteq (y - z)$

and $((x - y) - z) \subseteq (x - y)$.

So $(x - z)$ is also finite.

Hence x is near z.

Fact 3.2.18.

(a) If x1 is near x2, then (x1 \/ y) is near (x2 \/ y).

(b) If x1 is near x2 and y1 is near y2, then (x1 \/ y1) is near (x2 \/ y2).

(c) If x is near y, then (N - x) is near (N - y).

Proof:

(a) (x1 \/ y) - (x2 \/ y) $\subseteq$ x1 - x2

and (x2 \/ y) - (x1 \/ y) $\subseteq$ x2 - x1.

Since x1 is near x2, (x1 - x2) and (x2 - x1) are finite. Hence, any subsets of these sets are also finite.

(b) By (a), (x1 \/ y1) is near (x2 \/ y1)

and (x2 \/ y1) is near (x2 \/ y2).

So (x1 \/ y1) is near (x2 \/ y2), by trans.

(c) (N - x) - (N - y) = (y - x) and

(N - y) - (N - x) = (x - y).

So if x and y are near each other, so are their complements.

We can now define the domains of the models we'll be using.

Definition 3.2.19.

(a) Q(n) = {y : y is near an n-quasi-congruence class}.

(b) Q = \/ {Q(n) : n > 0}

Examples

(a) $Q(1) = \{y : y \subseteq N$ and y is finite or cofinite$\}$

(b) $Q(2) = \{y : y \subseteq N$ and y is finite, cofinite, near $\{2n\}$, or near $\{2n + 1\}\}$. So $Q(2)$ is the domain of the model constructed above to satisfy DIV$\{2\}$.

Before we construct models of SIZE over these domains we will show that if $n > 0$, then $Q(n)$ forms an atomic boolean algebra under the usual set-theoretic operations.

Fact 3.2.20. If A is a class of sets such that

(i) $\bigcup A \in A$,

(ii) if $x \in A$, then $(\bigcup A - x) \in A$, and

(iii) if $x \in A$ and $y \in A$, then $(x \cup y) \in A$,

then A forms a boolean algebra under the usual set-theoretic operations, where 'I' is interpreted as A. (See Monk, Def. 9.1, p.141 and Corr 9.4, p.142)

Theorem 3.2.21. For any n, $Q(n)$ is an atomic boolean algebra under the usual set-theoretic operations.

Proof:

First we show that $Q(n)$ is a boolean algebra by verifying each of the conditions in Fact 3.2.20.

(a) $\bigvee Q(n) = N$, since $N \in Q(n)$ and if $x \in Q(n)$, $x \subseteq N$.

(b) If $x \in Q(n)$, then $(N-x) \in Q(n)$:

Suppose $y \in QC(n)$

and x is near y.

So $(N - y) \in QC(n)$, by 3.2.13b

and $(N - x)$ is near $(N - y)$, by 3.2.18b.

So $(N - x) \in Q(n)$.

(c) Suppose $x \in Q(n)$

and $y \in Q(n)$.

Then x is near x'

and $x' \in QC(n)$, for some x'.

And y is near y'

and $y' \in QC(n)$, for some y'.

But then $(x' \vee y') \in QC(n)$, by 3.2.13a

and $(x \vee y)$ is near $(x' \vee y')$ by 3.2.18b.

So $(x \vee y) \in QC(n)$.

Thus, $Q(n)$ is a boolean algebra. Moreover, every singleton is in $Q(n)$, since all singletons are near $\emptyset$. So $Q(n)$ is an atomic boolean algebra.

We shall now define a size function on all sets which are near quasi-congruence classes. The sizes assigned to sets are ordered pairs. The first member is a rational between 0 and 1 which represents the density of the set. The second member is an integer which represents the finite (possibly negative) deviation of a set from "average" sets of the same density. First, we define the ordering and arithmetic for these sizes with the intention of inducing the size ordering and SUM relation for sets from the assignment of sizes to sets.

Definition 3.2.22.

(a) A <u>size</u> is an ordered pair $\langle \bar{p}, \bar{d} \rangle$, where $\bar{p}$ is a rational and $\bar{d}$ is an integer.

(b) If $\theta 1 = \langle \bar{p}1, \bar{d}1 \rangle$ and $\theta 2 = \langle \bar{p}2, \bar{d}2 \rangle$ are sizes, then

(i) $\theta 1 < \theta 2$ iff $\bar{p}1 < \bar{p}2$ or ($\bar{p}1 = \bar{p}2$ and $\bar{d}1 < \bar{d}2$).

(ii) $\theta 1 + \theta 2 = \langle \bar{p}1 + \bar{p}2, \bar{d}1 + \bar{d}2 \rangle$.

Only some of these sizes will actually be assigned to sets. Specifically, a size will be assigned to a set only if $0 \leq \bar{p} \leq 1$. Moreover, if $\bar{p} = 0$, the $\bar{d} \geq 0$ and $\bar{p} = 1$, then $\bar{d} \leq 0$.

Our intention in assigning sizes to sets is as follows: Suppose x is near an n-quasi-congruence class x', so x' is the union of $k \leq n$ n-congruence classes. The set x' has density k/n and this is the value, β, assigned to x. The σ value assigned to x is the finite number of elements added to or removed from x' to obtain x. The definitions and facts below formalize this intention and demonstrate that the assignment of sizes to sets is well-defined.

Fact 3.2.23

(a) If x ∈ QC, y ∈ QC, and x ≠ y, then x is not near y (i.e. no two quasi-congruence classes are near each other.)

(b) Any set is near at most one quasi-congruence class.

Proof:

(a) Let n be the least number such that

x ∈ QC(n) and y ∈ QC(n).

So each is a union of n-congruence classes:

x = x.1 \/ ... \/ x.k

and y = y.1 \/ ... \/ y.l.

Suppose a ∈ (x - y)

Then a ∈ x.i, for some i

but ¬(a ∈ y.j), for any y.

So x.i ≠ y.j, for any j.

So x.i ⊆ (x - y).

so (x - y) is infinite, since x.i is infinite.

Similarly, if a ∈ (y - x), then (y - x) is infinite.

(b) If x were near two quasi-congruence classes, the two would have to be near each other, since NEAR is an is transitive. But this is impossible by (a).

Notice, however, that a set can be in $Q(n)$ and in $Q(m)$ even if $n \neq m$, because a given quasi-congruence class may be in $QC(n)$ and $QC(m)$.

Definition 3.2.24. If x is near a quasi-congruence class, then:

(a) $C(x)$ = the quasi-congruence class near x.

(b) $D1(x) = x - C(x)$.

(c) $D2(x) = C(x) - x$.

Note that $D1(x)$ and $D2(x)$ are finite and that

$$x = (C(x) \cup D1(x)) - D2(x).$$

Definition 3.2.25. If $x \in QC$, then

(a) $\bar{a}(x)$ = the least n such that $x \in QC(n)$.

(b) $\bar{b}(x)$ = the least k such that x is the union of k $\bar{a}(x)$-congruence classes (i.e., the unique k such that x is the disjoint union of k $\bar{a}(x)$-congruence classes).

Examples:

(a) If $x = \{2n + 1\}$, $\bar{a}(x) = 2$ and $\bar{b}(x) = 1$.

(b) If $x = \{4n + 1\} \cup \{4n + 2\}$, $\bar{a}(x) = 4$, $\bar{b}(x) = 2$.

(c) If $x = \{4n + 1\} \cup \{4n + 3\}$, $\bar{a}(x) = 2$, $\bar{b}(x) = 1$, since $x = \{2n + 1\}$.

Definition 3.2.26. If $x \in Q$, then

(a) $\beta(x) = \delta(C(x))/\eth(C(x))$

(b) $\eth(x) = cd(D1(x)) - cd(D2(x))$

(c) $\theta(x) = \langle\beta(x),\eth(x)\rangle$

We can, at last, define the models to be used in our independence proof.

Definition 3.2.27. For $n > 0$, $\underline{Q}(n)$ is the interpretation $\underline{A}$ of L(C<) such that

(i) $A = Q(n)$,

(ii) Boolean symbols receive their usual interpretation,

(iii) $\underline{A} \models x < y$ iff $\theta(x) < \theta(y)$,

(iv) $\underline{A} \models x \cong y$ iff $\theta(x) = \theta(y)$,

(v) $\underline{A} \models$ UNIT(x) iff $\theta(x) = \langle 0,1\rangle$,

(vi) $\underline{A} \models$ SUM(x,y,z) iff $\theta(x) + \theta(y) = \theta(x)$.

To show that the models $\underline{Q}(n)$ satisfy BASIC we will need the following facts about congruence classes.

Fact 3.2.28

(a) If $x = \{a*\underline{n} + b\}$ and $a2 = a*c$, then

$$x = \bigvee \{ \{a2*\underline{n} + (i*a + b)\} : 0 \leq i < c\}$$

(b) If $x \in QC(n)$, and $m = k*n$, then $x \in QC(m)$.

(c) If $x \in QC$ and $y \in QC$, there is an n such that $x \in QC(n)$ and $y \in QC(n)$.

Proof:

(a) If $k \in \{a2*\underline{n} + (i*a + b)\}$ for some i,

$0 \leq i < c$, then, for some n1,

$k = (a2*n1 + i*a + b)$

$= a*c*n1 + i*a + b$

$= a*(c*n1 + i) + b$

So $k \in x$.

If $k \in x$, there is an n1 such that $k = a*n1 + b$.

Let n2 be the greatest n such that

$a2*n \leq k$.

Let $k' = k - a2*n2$.

Since $k \equiv b \bmod a$ and

$a2*n2 \equiv 0 \bmod a$,

$k' \equiv b \bmod a$.

So $k' = a*i + b$, where $0 \leq i < c$.

But $k = a2*n2 + k'$

$= a2*n2 + a*i + b$.

So $k \in \{a2*\underline{n} + (i*a + b)\}$.

(b) By (a), each n-congruence class is a finite union of disjoint m-congruence classes.

(c) Suppose $x \in QC(n1)$ and $y \in QC(n2)$.

Then, by (b), both x and y are $\in QC(n1*n2)$.

Theorem 3.2.29. For any $n > 0$, $\underline{Q}(n) \models$ BASIC. Proof: By 3.2.21, $\underline{Q}(n)$ is an atomic boolean algebra; so $\underline{Q}(n) \models$ BASIC. The $<$ - relation of $\underline{Q}(n)$ is induced from the linear ordering of sizes; so it is a quasi-linear ordering and IRREF, TRICH, and DEF= are satisfied. As for the remaining axioms:

(a) SUBSET:

Suppose $x \subset y$:

If $C(x) = C(y)$

then $\bar{\beta}(x) = \bar{\beta}(y)$

and $D1(x) \subseteq D1(y)$ and $D2(y) \subseteq D2(x)$

where at least one of these inclusions is proper.

So $\bar{\partial}(x) < \bar{\partial}(y)$.

But if $C(x) \neq C(y)$

then $C(x) \subset C(y)$

So $\bar{\beta}(x) < \bar{\beta}(y)$.

In either case, $\theta(x) < \theta(y)$ so $\underline{Q}(n) \models x < y$.

(b) REP<:

Suppose $\underline{Q}(n) \models x < y$

so $\theta(x) = \langle k1/n, d1 \rangle$

and $\theta(y) = \langle k2/n, d2 \rangle$

and either $k1 < k2$

or $k1 = k2$ and $d1 < d2$.

We want to find some x' such that

$\underline{Q}(n) \models x \approx x'$

and $x' \subset y$:

If $k1 = k2 > 0$, then y must be infinite;

so x' can be obtained by removing (d2 − d1)
atoms from y.

If $k1 = k2 = 0$, then $0 \leq d1 < d2$;
so, again, x' can be obtained by removing
(d2 − d1) atoms from y.

If $k2 > k1 > 0$, then let y1 be the union of
k1 n-congruence classes contained in C(y).
So (y − y1) is finite and (y1 − y) is
infinite. Let $y2 = y1 - (y1 - y) = y1 \wedge y$.
So $\theta(y2) = \langle k1/n, -d3 \rangle$ where $d3 = cd(y1-y)$
Finally, let $d4 = d3 + d1$ and $x' = y2 \vee y3$,
where $y3 \subseteq (y1 - y)$ and $cd(y3) = d4$.

If $k1 = 0$, then x is finite. So if $k2 > 0$ (y is
infinite), there is no problem. If $k2 = 0$,
then $d1 < d2$; so y is finite but has more
members than x. So let x' be some proper
subset of y with d1 members.

(c) DISJU: It is sufficient to show that if x and y are
disjoint, then $\theta(x \vee y) = \theta(x) + \theta(y)$. We
need the following three facts:

(i) $C(x \vee y) = C(x) \vee C(y)$ (see Fact 3.2.18b)

(ii) $D1(x \vee y) = (D1(x) \vee D1(y)) - (C(x) \vee C(y))$
(If $a \in x \vee y$ but a not $\in C(x \vee y)$, then
$a \in D1(x)$ or $a \in D1(y)$; any element of D1(x)
is also in $D1(x \vee y)$, unless it is in C(y);
any element of D1(y) is in $D1(x \vee y)$, unless
it is in C(x).)

(iii) $D2(x \cup y) = (D2(x) \cup D2(y)) - (D1(x) \cup D1(y))$

(Note that if $x \in Q$, $y \in Q$ and $x \cap y = \emptyset$,
then $C(x) \cap C(y) = \emptyset$; otherwise $C(x)$ and $C(y)$
have an infinite intersection.

Hence, if $a \in C(x \cup y) - (x \cup y)$
then $a \in C(x) - x$, i.e. $a \in D2(x)$
or $a \in C(y) - y$, i.e. $a \in D2(y)$.

And if $a \in D2(x)$, then $a \in D2(x \cup y)$
unless $a \in D1(y)$

And if $a \in D2(y)$, then $a \in D2(x \cup y)$
unless $a \in D1(x)$.

From (ii) we obtain (iia):

(iia) $cd(D1(x \cup y)) = cd(D1(x) \cup D1(y))$
$\quad - cd((D1(x) \cup D1(y) \cap (C(x) \cup C(y)))$

and from (iii) we obtain (iiia):

(iiia) $cd(D2(x \cup y)) = cd(D2(x) \cup D2(y))$
$\quad - cd((D1(x) \cup D1(y) \cap (D2(x) \cup D2(y)))$

But $D1(x)$ and $D1(y)$ are disjoint, since x and y are disjoint, so

(iv) $cd(D1(x) \cup D1(y)) = cd(D1(x)) + cd(D1(y))$
$\quad = \eth1(x) + \eth1(y)$

And since $D2(x) \subseteq C(x)$,
and $D2(y) \subseteq C(y)$,
and $C(x) \cap C(y) = \emptyset$,
and $D2(x) \cap D2(y) = \emptyset$,

we may conclude:

(v) $cd(D2(x) \vee D2(y)) = \bar{\delta}2(x) + \bar{\delta}2(y)$

So,

$\bar{\delta}1(x \vee y) - \bar{\delta}2(x \vee y)$

$= (\bar{\delta}1(x) + \bar{\delta}1(y)) - (\bar{\delta}2(x) + \bar{\delta}2(y))$

$= (\bar{\delta}1(x) - \bar{\delta}2(x)) + (\bar{\delta}1(y) - \bar{\delta}2(y))$

$= \bar{\delta}(x) + \bar{\delta}(y)$

And $\bar{\beta}(x \vee y) = \bar{\beta}(x) + \bar{\beta}(y)$, by (i)

So $\theta(x \vee y) = \theta(x) + \theta(y)$

(d) DEF+:

Suppose $\underline{Q}(n) \models SUM(x,y,z)$

So $\theta(z) = \theta(x) + \theta(y)$.

Clearly $\theta(x) \leq \theta(z)$

Assume $\theta(x) = \theta(z)$

then $\theta(y) = \langle 0,0 \rangle$

so $y = \emptyset$ and the consequent of DEF+ is

satisfied (let $x' = z$, $y' = \emptyset$.)

Assume $\theta(x) < \theta(z)$

Then $\underline{Q}(n) \models x' \simeq x$

and $x' \subset z$, for some x', since $\underline{Q}(n) \models$ REP<.

Let $y' = z - x'$.

So x' and y' are disjoint sets whose union is z

We claim $\theta(y) = \theta(y')$ (so $\underline{Q}(n) \models y \simeq y'$):

For $\theta(x') + \theta(y') = \theta(z)$, by DISJU.

But $\theta(x') = \theta(x)$

So $\theta(y') = \theta(z) - \theta(x)$

$= \theta(y)$ ("Cancelling" is valid for

sizes because it is valid for

rationals and integers.)

Conversely,

$$\text{If } \theta(z) = \theta(x') + \theta(y')$$

$$\text{and } \theta(x') = \theta(x)$$

$$\text{and } \theta(y') = \theta(y)$$

$$\text{then } \theta(x) = \theta(x) + \theta(y)$$

Theorem 3.2.30. For $n > 0$, $\underline{Q}(n) \models DIV\{m\}$ iff $m|n$.

Proof:

(<-) For each i, $0 \leq i < n$, let $A.i = \{n*\underline{k} + i\}$.

So $N = \bigvee \{A.i : 0 \leq i < n\}$.

If $i \neq j$, $A.i \wedge A.j = \emptyset$ and

$\underline{Q}(n) \models A.i \approx A.j$,

since $\theta(A.i) = \theta(A.j) = \langle 1/n, 0\rangle$.

Letting $p = n/m$, group the n sets A.i into m collections with p members in each:

$B1, \ldots, B.m$.

Letting $b.j = \bigvee B.j$ for $1 \leq j \leq m$,

$b.j \in Q(n)$ and

$\theta(b.j) = \langle p/n, 0\rangle = \langle 1/m, 0\rangle$.

Furthermore, $b.1 \vee \ldots \vee b.m = N$.

(->) $\theta(N) = \langle 1, 0\rangle$.

Hence, if m disjoint sets of the same size exhaust N, they must each have size $\langle 1/m, 0\rangle$.

But if $x \in Q(n)$, then $\theta(x) = \langle a/n, b\rangle$, for integral a and b. So $b = 0$ and $a = m/n$.

Definition 3.2.31. If $J \neq \emptyset$ and J is finite, then <u>the least common multiple of</u> J, $\nabla(J)$, is the least k which is divisible by every member of J.

Remark. $\nabla(J)$ always exists since the product of all members of J is divisible by each member of J. Usually, the product is greater than $\nabla(J)$.

Corollary 3.2.32. If J is finite, then

(a) If BDIV(n) |- DIV{m}, then m | n.

(b) If BDIV(J) |- DIV{m}, then m | ∇(J).

(c) There are only finitely many m for which BDIV(J) |- DIV{m}.

Proof.

(a) If m ∤ n, then Q(n) |= BDIV(n);¬DIV{m}.

(b) Q(∇(J)) |= BDIV(J) since it satisfies DIV{j} for each j ∈ J, by (a). But if m ∤ ∇(J), then Q(∇(J)) |≠ DIV{m}.

(c) Only finitely many m divide ∇(J), so by (b), BDIV(J) entails only finitely many DIV{m}.

We are now ready to show that BDIV{n} |≠ CS by finding a sentence in CS which entails infinitely many DIV sentences. Such sentences can be produced by generalizing the notion of divisibility to all sets instead of applying it only to the universe.

Definition 3.2.33

(a) If 0 ≤ N, then Times{n}(x,y) is the formula:

Ex0...x.n (x0 = ∅ & x.n = y &

& {SUM(x.(i-1),x,x.i) : 1 ≤ i ≤ n})

So Times{n}(x,y) says that y is the same size as the disjoint union of n sets, each the same size as x.

(b) If 0 ≤ m < n, then Mod{n,m}(z) is the formula:

ExEyEvEw(Times{n}(x,v)

& UNIT(y)

& Times{m}(y,w)

& SUM(v,w,z))

So Mod{n,m}(z) says that z can be partitioned into n sets of the same size and m atoms.

(c) Div{n}(z) is the formula

Mod{n,0}(z) ¦ ... ¦ MOD{n,n-1}(z)

(d) Adiv{n} is the sentence

(x)Div{n}(x)

Remark. We have taken this opportunity to formulate the divisibility predicates purely in terms of size predicates. Notice that in the presence of BASIC,

MOD{n,m} ≡ Mod{n,m}(I)

Fact 3.2.34. CS |- ADIV{n}, for every n.

Proof: Every set in every finite standard interpretation is a finite set, and all finite sets are roughly divisible by every n.

Fact 3.2.35. BASIC;ADIV{n} |-:

(a) ADIV{n**m}, for all m,

(b) DIV{n}, and

(c) DIV{n**m}, for all m.

Proof:

(a) By induction on m: if m = 1, then n**m = n, so ADIV{n} |- ADIV{n**m}. If T |- ADIV{n**k}, A |= T, and x ∈ A, then x can be partitioned into n**k sets of the same size and i atoms, where i < n**k. Each non-atomic set in the partition can be further partitioned into n sets of the same size and j atoms, where j < n. Thus, we have partitioned x into (n**k)*n sets of the same size and ((n**k)*j + i) atoms. But (n**k)*n = n**(k+1) and, since i < n**k and j < n, (n**k)*j + i < n**(k+1). Hence A |= ADIV{n**(k+1)}.

(b) Obvious.

(c) Immediate from (a) and (b).

Theorem 3.2.36. BDIV(N) |/- ADIV{n} for any n > 1.

Proof:

If BDIV(N) |- ADIV{n}, then by compactness there is a finite set J such that BDIV(J) |- ADIV{n}. But then BDIV(J) |- DIV{n**k} for every k, by fact 3.2.35c. But this contradicts the fact that BDIV(J) entails only finitely many DIV{n} sentences (Fact 3.2.32).

So, even if we add all of the DIV-sentences to BASIC, we are left with a theory weaker than CS. Since this weakness has arisen in the case of ADIV sentences, it is reasonable to attempt an axiomatization of CS as follows:

Definition 3.2.37. CA = BASIC \/ {ADIV{n} : n > 0}

The remainder of this chapter and the next two are devoted to showing that CA is, indeed, a complete set of axioms for CS.

3.3 REMARKS ON SHOWING THAT CA = CS.

We know that CS $\vdash$ CA and we want to show that CA $\equiv$ CS, i.e. that CA $\vdash$ CS. To do so, it will be sufficient to show that every consistent, complete extention of CA is consistent with CS:

Fact 3.3.1.

(a) (Lindenbaum's lemma) Every consistent theory has a consistent, complete extension. (See Monk, Theorem 11.13, p.200).

(b) If every consistent, complete extension of T2 is consistent with T1, then T2 $\vdash$ T1.

Proof of b:

Suppose T1 $\vdash \emptyset$, T2 $\nvdash \emptyset$. Then T = T2;$\neg\emptyset$ is consistent. T has a consistent, complete extension, T', by Lindenbaum's lemma. Since T2 $\subseteq$ T, T' is also a consistent complete extension of T2. But T' is not consistent with T1.

Definition 3.3.2. T' is a <u>completion</u> of T iff T' is a complete, consistent extension of T.

To prove that every completion of CA is consistent with CS, we define two kinds of completions of a theory.

Definition 3.3.3.

(a) T' is a <u>finite completion</u> of T iff T' is true in some finite model of T.

(b) T' is an <u>infinite completion</u> of T iff T' is true in some infinite model of T.

Fact 3.3.4. If T' is a completion of T, and BA $\subseteq$ T, then (T' is a finite completion of T iff T' is not an infinite completion of T).

Proof:

(->). Suppose $\underline{A}$ is a finite model and $\underline{A}$ |= T'. Since T' is complete,

T' |- EXACTLY{n}, where n is the number of atoms in $\underline{A}$. So T' |- ¬ATLEAST{n+1} and has no infinite models.

(<-). T' |- ATLEAST{n} for every n, so T' has no finite models.

Fact 3.3.5

(a) Every finite completion of CA is equivalent to CA;EXACTLY{n}, for some n.

(b) Every finite completion of CA is consistent with CS.

Proof:

(a) CA |- BASIC and, by theorem 2.1.6, BASIC is categorical in every finite power.

(b) The model, $\underline{A}$, of CA;EXACTLY{n} is a standard finite interpretation. So $\underline{A}$ |= CS.

Definition 3.3.6

(a) CAI = CA \/ INF

(b) CSI = CS \/ INF

So, to show that every completion of CA is consistent with CS, we may now concentrate on showing that every completion of CAI is consistent with CSI.

What, then, are the completions of CAI? Recall that CA entails DIV{n} for every n > 0, where DIV{n} is

MOD{n,0} ¦ ... ¦ MOD{n,n-1}

So any completion, T, of CAI has to "solve" the disjunction DIV{n} for each n -- that is, T has to entail one of the disjuncts. The main result of this chapter is that we can complete CAI by specifying, for each n, the number of atoms remaining when the universe is divided into n disjoint subsets of the same size.

Definition 3.3.7

(a) f:N+ ==> N is a <u>remainder function</u> iff $0 \leq f(n) < n$ for all n ∈ Dom(f). (Henceforth, 'f' ranges over remainder functions.)

(b) f is <u>total</u> iff Dom(f) = N+; otherwise f is <u>partial</u>.

(c) f is <u>finite</u> iff Dom(f) is finite.

(d) n <u>is a solution for</u> f iff for any i ∈ Dom(f), $n \equiv f(i) \bmod i$.

(e) f is <u>congruous</u> iff for any i and j ∈ Dom(f), gcd(i,j) | (f(i)-f(j)); otherwise f is inconqruous.

(f) The remainder theory specified by f, RT(f), is {MOD(n,m} : f(n) = m}.

(g) If T is a theory, T(f) = T \/ RT(f).

It will be shown in chapter 5 that if f is total, CAI(f) is complete and that these are the only complete extensions of CAI. In this section, we will show that CAI(f) is consistent just in case CSI(f) is consistent.

Fact 3.3.8

(a) If f is finite, then f has a solution iff f is congruous iff f has infinitely many solutions. (See Griffin, Theorem 5-11, p. 80.)

(b) f is congruous iff every finite restriction of f is congruous.

(c) There are congruous f without any solutions. (Let f(p) = p-1 for all primes p. Any solution would have to be larger than every prime.)

Theorem 3.3.9

(a) If f is finite, then CS(f) is consistent iff f is congruous.

(b) CS(f) is consistent iff f is congruous.

(c) CSI(f) is consistent iff f is congruous.

Proof:

(a) (->) Let $\emptyset$ = & {RT(f)}. Since CS;$\emptyset$ is consistent, there is some n such that $\underline{F}$.n |= $\emptyset$. So n is a solution of f and, hence, f is congruous by 3.3.8a.

(<-) If f is congruous, f has a solution, n. So $\underline{F}$.n |= CS;$\emptyset$.

(b) (->) For every finite restriction, g, of f, CS(g) is consistent. By (a), each such g is congruous. Hence f is congruous by 3.3.8b.

(<-) Every finite restriction, g, of f is congruous. So CS(g) is consistent, by (a). By compactness, then, CS(f) is consistent.

(c) (->) If CSI(f) is consistent, so is CS(f). So, by (b), f is congruous.

(<-) By compactness, it is sufficient to show that every finite subtheory, T, of CSI(f) is consistent. But if T is such a theory, then

$$T \subsetneq CS(g) \cup \{ATLEAST(i): i < n\}$$

for some n and some finite restriction, g, of f. Since f is congruous, g is as well, by 3.3.8b. So g has arbitrarily large solutions and CS(g) has finite models large enough to satisfy T. Hence, T is consistent.

We now want to prove a similar theorem for CA, our proposed axiomatization of CS. To do this, we must first establish that certain sentences are theorems of CA.

Lemma 3.3.10. If $n | m$, $0 \le q < n$, and $p \equiv q \bmod n$, then

$$CA \mid\!- MOD\{m,p\} \to MOD\{n,q\}.$$

Proof: Suppose $\underline{A} \models CA; MOD\{m,p\}$, $k1 = m/n$, and $p = n*k2 + q$. So $B(\underline{A})$ can be partitioned into m sets of the same size

$$b(1,1), \ldots, b(1,n), b(2,1), \ldots, b(k1,n)$$

and p atoms

$$a(1,1), \ldots, a(k2,n), c.1, \ldots, c.q$$

For $1 \le i \le n$, let

$$B.i = \bigvee \{b(j,n) : 1 \le j \le k1\} \vee \bigvee \{a(j,n) : 1 \le j \le k2\}$$

Since $\underline{A} \models DISJU$, $\underline{A} \models B.i \approx B.j$, for all i and j between 1 and n. Furthermore,

$$B(\underline{A}) = \bigvee \{B.i : 1 \le i \le n\} \vee c.1 \vee \ldots \vee c.q$$

So $\underline{A} \models MOD\{n,q\}$.

Lemma 3.3.11. If $0 \le p < q < m$, then

$$CA \mid\!- MOD\{m,p\} \to \neg MOD\{m,q\}.$$

Proof: Suppose $\underline{A} \models (MOD\{m,p\} \;\&\; MOD\{m,q\})$:

Then $\underline{A} \models x1 \vee \ldots \vee x.m \vee a1 \vee \ldots \vee a.p = I$

and $\underline{A} \models y1 \vee \ldots \vee y.m \vee b1 \vee \ldots \vee b.q = I$

where the a's and b's are atoms and the x's (y's) are disjoint sets of the same size (in $\underline{A}$).

Let $X = x1 \vee \ldots \vee x.m$

$Y = y1 \vee \ldots \vee y.m$

$A = a1 \vee \ldots \vee a.p$

$$B = b_1 \vee \ldots \vee b_p$$

$$B' = b_{p+1} \vee \ldots \vee b_q$$

We claim that $y_1 < x_1$. For if $y_1 \approx x_1$, then $(X \vee A) \approx (Y \vee B)$ and if $x_1 < y_1$, then $(X \vee A) < (Y \vee B)$, by DISJU; neither is possible since $(Y \vee B) \subset I = (X \vee A)$. So $y_i < x_i$ for $1 \leq i \leq m$. Since $\underline{A} \models$ REP<, there is a proper subset y'_i of x_i which is the same size as y_i. Let $z_i = x_i - y'_i$

$$Y' = y'_1 \vee \ldots \vee y'_m$$

$$Z = z_1 \vee \ldots \vee z_m$$

So, $Y' \vee Z = X = I - A$

and $Y' \vee B' = I - B$.

But $A \approx B$, since each is the disjoint union of p atoms. Thus $(I - A) \approx (I - B)$, by RC≈, so $(Y' \vee Z) \approx (Y \vee B')$. But $Y' \approx Y$, since for each component of Y there is a component of Y' of the same size. So $Z \approx B'$, by RC≈.

But Z must be larger than B', for B' is the union of fewer than m atoms while Z is the union of m non-empty sets. So the original supposition entails a contradiction.

Theorem 3.3.12

(a) If f is finite, then CA(f) is consistent iff f is congruous.

(b) CA(f) is consistent iff f is congruous.

(c) CAI(f) is consistent iff f is congruous.

Proof:

(a) (<-) Follows from 3.3.9a since CA(f) $\subseteq$ CS(f).

(->) Supposing that f is incongruous, there exist i, j, and k such that k = gcd(i,j), k$\nmid$(f(i)-f(j)). We will show that

(*) CA |- ¬(MOD{i,f(i)} & MOD{j,f(j)})

from which it follows that CA(f) is inconsistent. Let p and q be such that

$0 \leq p,q < k$,

$f(i) \equiv p \bmod k$, and

$f(j) \equiv q \bmod k$.

By lemma 3.3.10, we have

(1) CA |- (MOD{i,f(i)} -> MOD{k,p}), and

(2) CA |- (MOD{j,f(j)} -> MOD{k,q})

since k|i and k|j. Since k$\nmid$(f(i)-f(j)), p $\neq$ q. So lemma 3.3.11 yields

(3) CA |- (MOD{k,p} -> ¬MOD{k,q}).

Finally, from (1), (2) and (3), we may conclude (*).

(b) and (c) follow from (a) as in 3.3.9.

Corollary 3.3.13. CAI(f) is consistent iff CSI(f) is consistent.

Proof: Immediate from 3.3.9c and 3.3.12c.

It might help to review our strategy before presenting the difficult parts of the proof that CA $\equiv$ CS. The main objective is (1), which follows from (2) by 3.3.1b.

(1) CA $\vdash$ CS

(2) Every completion of CA is consistent with CS.

We already know that the finite completions of CA are consistent with CS (see 3.3.5c) and that if CAI(f) is consistent, the CSI(f) is also consistent (see 3.3.13). So (2) is a consequence of (3).

(3) If T is a completion of CAI, then T $\equiv$ CAI(f) for some total, congruous f.

To establish (3), it is sufficient to prove (4) because every completion of CAI entails CAI(f) for some total f.

(4) If f is total and congruous, CAI(f) is complete.

To prove (4), we invoke the prime model test: if T is model complete and T has a prime model, then T is complete (see A.3.3). So (4) follows from (5) and (6).

(5) If f is total and congruous, then CAI(f) has a prime model.

(6) For any f, CAI(f) is model complete.

Finally, since any extension of a model complete theory is also model complete (see A.3.7a), we can infer (6) from (7).

(7) CAI is model complete.

So (1) follows from (5) and (7).

The proof outlined here will be carried out in chapter 5. But first we consider a simpler theory, PSIZE, which deals only with sizes of sets and ignores boolean relations. Chapter 4 formulates PSIZE and establishes that it is model complete, a result we need for showing that CAI is model complete.

4 THE PURE THEORY OF CLASS SIZES

CS is about sets; it makes claims about sets in terms of their boolean relations and their size relations. In this chapter, we identify a theory PCS (the pure theory of class sizes), which is not about sets, but only about sizes of sets. "Sizes", here, are equivalence classes of sets having the same size. PCS is formulated in the language of size relations, L(<).

PCS is worth examining in its own right, for if "number theory" is taken to mean the theory of cardinal numbers, then PCS is our version of number theory. But our main reason for introducing PCS is to aid the proof that CA is model complete. For this reason, we will give only a sketchy treatment of PCS itself.

Section 1 defines PCS and develops a set of axioms, PCA, for PCS, as follows: for each model, $\underline{A}$, of BASIC, we define a "size-model", $\underline{S}(\underline{A})$, whose domain consists of equivalence classes drawn from A under the same-size relation; PCS is the set of statements true in $\underline{S}(\underline{A})$ for any standard finite model, $\underline{A}$. PCA consists of a theory, PSIZE, which holds in $\underline{S}(\underline{A})$ whenever $\underline{A} \models$ BASIC and a set of divisibility principles.

Using some results about model theory in section 2, Section

3 establishes that PCA is model complete, the main result of this chapter and the only result needed for subsequent proofs. This is done by reducing PCA to the theory Zgm, whose models are Z-groups taken modulo some specific element.

Finally, section 4 indicates how PCA could be shown to axiomatize PCS. The method is the same as that outlined in chapter 3 to show that CA ≡ CS.

4.1 SIZE-MODELS AND THE PURE THEORY OF CLASS-SIZES

Definition 4.1.1. Suppose $\underline{A} \models$ BASIC. Then,

(a) If x is a member of A, then $\delta(x,\underline{A})$ is the size of x in $\underline{A}$:

$$\delta(x,\underline{A}) = \{y: \underline{A} \models x \approx y\}$$

(b) $\underline{S}(\underline{A})$, the size model for $\underline{A}$ is the interpretation of L(<) whose domain is $\{\delta(x,\underline{A}): x \in A\}$ where

1) $\underline{S}(\underline{A}) \models \delta(x,\underline{A}) \approx \delta(y,\underline{A})$ iff $\underline{A} \models x \approx y$,

2) $\underline{S}(\underline{A}) \models \delta(x,\underline{A}) < \delta(y,\underline{A})$ iff $\underline{A} \models x < y$,

3) $\underline{S}(\underline{A}) \models \mathrm{SUM}(\delta(x,\underline{A}), \delta(y,\underline{A}), \delta(z,\underline{A}))$ iff $\underline{A} \models \mathrm{SUM}(x, y, z)$, and

4) $\underline{S}(\underline{A}) \models \mathrm{UNIT}(\delta(x,\underline{A}))$ iff $\underline{A} \models \mathrm{ATOM}(x)$

(c) We shall use "!" as a one-place operator to be read as "the complementary size of":

$$\underline{S}(\underline{A}) \models (y = x!) \quad \text{iff} \quad \underline{S}(\underline{A}) \models \mathrm{SUM}(x,y,\bar{1})$$

Notice that these interpretations are well defined because the predicates are satisfied by elements of A in virtue of their sizes. For example, if $\underline{A} \models \mathrm{SUM}(x,y,z)$ and $\underline{A} \models (z \approx w)$, then $\underline{A} \models \mathrm{SUM}(x,y,w)$.

Definition 4.1.2. PCS, <u>the pure theory of class sizes</u>, consists of all sentences of L(<) which are true in the size model of every finite standard interpretation of L(C<).

Definition 4.1.3. PSIZE consists of the following axioms:

Order axioms:

IRREF $\neg (x < x)$

TRANS $(x < y) \mathbin{\&} (y < z) \rightarrow (x < z)$

UNIQ= $x \approx y \leftrightarrow x = y$

MIN $\emptyset \leq x$

MAX $x \leq I$

TRICH $x < y \mid x \approx y \mid y < x$

Unit axioms: $UNIT(x) \leftrightarrow (y < x \leftrightarrow y = 0)$

$(Ex)\ UNIT(x)$

Sum axioms:

IDENT $SUM(x, \emptyset, x)$

COMM $SUM(x, y, z) \leftrightarrow SUM(y, x, z)$

MONOT $SUM(x1, y, z1) \mathbin{\&} SUM(x2, y, z2) \rightarrow (x1 < x2 \leftrightarrow z1 < z2)$

ASSOC $SUM(x, y, w1) \mathbin{\&} SUM(w1, z, w) \mathbin{\&} SUM(y, z, w2) \rightarrow SUM(x, w2, w)$

EXIST+ $(Ez) SUM(x, y1, z) \mathbin{\&} y2 \leq y1 \rightarrow (Ez) SUM(x, y2, z)$

EXIST- $x \leq z \rightarrow (Ey) SUM(x, y, z)$

COMP $SUM(x, x!, I)$

Fact 4.1.4. If A |= BASIC, then S(A) |= PSIZE

PSIZE fails to axiomatize PCS for the same reason that BASIC fails to axiomatize CS: the lack of divisibility principles. We offer PCA as an axiomatic version of PCS:

Definition 4.1.6. PCA = PSIZE \/ {ADIV{n}: n > 0}.

Fact 4.1.7. If A |= CA, then S(A) |= PCA.

If A is a finite standard interpretation of BASIC with n atoms, then the elements of S(A) can be regarded as the sequence

0, 1, ..., n

with the usual ordering, where

S(A) |= UNIT(x) iff x = 1

and

S(A) |= SUM(i,j,k) iff (i+j) = k $\leq$ n

Once n is fixed, this is the only interpretation allowed by the axioms PSIZE. Notice, in particular, that MONOT+ is needed to rule out the interpretation in which SUM(i,j,k) is satisfied just in case (i+j) = k mod (n+1).

4.2 SOME MODEL THEORY.

We shall use theorem 4.2.1a to show that PCAI is model complete and we shall use theorem 4.2.1b to show, in chapter 5, that CAI is model complete.

Theorem 4.2.1.

(a) (Monk) If T satisfies (*), then T is model complete.

(*) If $\underline{A} \models T$, $\underline{B} \models T$, $\underline{A} \subseteq \underline{B}$, and $\underline{C}$ is a finitely generated submodel of $\underline{B}$, then there is an isomorphism, f, of $\underline{C}$ into $\underline{A}$ such that if $x \in C \wedge A$, then $f(x) = x$.

(b) If T is model complete and L(T) has no function symbols, then T satisfies (*).

Proof:

(a) See Monk, p.359.

(b) If L(T) has no function symbols, then any finitely generated structure over L(T) is finite. So suppose C contains a1, ..., a.n (from A) and b1, ... b.m (from (B - A)). Let ø1 be the diagram of $\underline{C}$ and obtain ø2 from ø1 by substituting the variable 'x.i' for each constant a.i and the variable 'y.i' for each constant b.i. Finally, obtain ø3 prepending to ø2 an existential quantifier for each y.i. So ø3 is a primitive formula.

$\underline{B} \models \phi3(a1,\ldots,a.n)$, so $\underline{A}$ does also, by fact A.3.5d. So, to obtain the desired isomorphism, map the a.i's into themselves and map the b.i's into a sequence of elements of A which can stand in for the existentially quantified variables of $\phi3$.

We shall refer to (*) of 4.2.1 as "Monk's condition" and to the mapping, f, as a "Monk mapping."

We use Monk's theorem in chapter 5 to infer the model completeness of the theory CA from that of PCA. But the model completeness of PCA is established by a method given below to infer the model-completeness of one theory from the model-completeness of another, along with fact 4.2.2.

Fact 4.2.2.

(a) If T1 is model complete and T2 $\vdash$ T1, then T2 is also model complete.

(b) If T is model complete in L, and L' is an expansion of L by adjoining new individual constants, then T is model complete in L' (Monk, p. 355).

Definition 4.2.3. Suppose L1 and L2 are first order languages and L12 = L1 - L2. A (<u>simple</u>) <u>translation</u> of L1 into L2 is a function, τ, which:

(1) assigns to the universal quantifier a (quantifier-free) formula, τ:A, of L2 with exactly

one free variable,

(2) assigns to each n-place predicate, P, in L12 a (quantifier free) formula, τ:P, of L2 with exactly n-free variables, and

(3) assigns to each n-place function symbol, O, in L12 a (quantifier free) formula, τ:O, of L2 with exactly (n+1) free variables.

Definition 4.2.4. If τ is a translation of L1 into L2, then τ extends to all formulae of L1 as follows:

(a) Predicates and function symbols of L2 are translated into themselves.

(b) τ:(ø1 ¦ ø2) = τ:ø1 ¦ τ:ø2

(b) τ:(ø1 ¦ ø2) = τ:ø1 ¦ τ:ø2

τ:(¬ø) = ¬(τ:ø)

τ:((x)ø) = (x)(τ:A(x) -> ø)

τ:(Exø) = Ex(τ:A(x) & ø)

Definition 4.2.5. If τ is a translation of L1 into L2,

(a) The "functional assumptions of τ" are the sentences:

(x1)...(x.n)(τ:A(x1) & ... & τ:A(x.n) -->

(Ey1)(τ:A(y1) & (y2)(τ:A(y2) -->

(τ:O(x1, ..., xn,y2) <-->y1 = y2))

where O is a function symbol of L1 but not L2.

(b) The "existential assumption of" τ is:

Ex τ:A(x)

The functional assumptions of a translation say that the formulas which translate function symbols yield unique values within the relevant part of the domain when given values in the relevant part of the domain. The relevant part of the domain, here, is the set of elements which satisfy the interpretation of the universal quantifier. The existential assumption of a translation says that that subdomain is non-empty. Notice that the existential and functional assumptions of a translation are sentences of L2.

A translation is a mapping from symbols to formulae, but it induces, in an obvious way, a mapping from interpretations of L2 into interpretations of L1:

Definition 4.2.6. If $\mathfrak{T}$ is a translation from L1 into L2 and $\underline{B}$ is an interpretation of L2 which satisfies the existential and functional assumptions of $\mathfrak{T}$, then $\mathfrak{T}(\underline{B})$ is the interpretation, $\underline{A}$, of L1 such that:

(a) The domain of $\underline{A}$ is the set of elements of B which satisfy $\mathfrak{T}$:A,

(b) $\underline{A}$ interprets all predicates and function-symbols common to L1 and L2 in the same way that $\underline{B}$ does, and

(c) $\underline{A}$ interprets all predicates and function symbols in L12 in accordance with the translations assigned by τ, that is:

$$\underline{A} \models P(x) \text{ iff } \underline{B} \models \tau{:}P(x)$$

$$\underline{A} \models (y = O(x)) \text{ iff } \underline{B} \models \tau{:}O(x,y)$$

Finally, we can formulate a condition on theories that allows us to infer the model completeness of one from the model completeness of the other:

Definition 4.2.7.

(a) If τ is a translation from L(T1) to L(T2), then T1 is <u>τ-reducible</u> to T2 iff for every model $\underline{A}$ of T1 there is a model $\underline{B}$ of T2 such that $\underline{A} = \tau(\underline{B})$.

(b) T1 is (<u>simply</u>) <u>reducible</u> to T2 iff there is a (simple) translation, τ, for which T1 is τ-reducible to T2. (c) T1 is <u>uniformly</u> <u>τ-reducible</u> to T2 iff for any models $\underline{A}1$ and $\underline{B}1$, where $\underline{A}1 \models T1$, $\underline{B}1 \models T1$, and $\underline{A}1 \subseteq \underline{B}1$, there exist models $\underline{A}2$ and $\underline{B}2$, such that $\underline{A}2 \models T2$, $\underline{B}2 \models T2$, $\underline{A}2 \subseteq \underline{B}2$, $\underline{A}1 = \tau(\underline{A}2)$ and $\underline{B}1 = \tau(\underline{B}2)$.

Lemma 4.2.8. Suppose that T1 is τ-reducible to T2 and that $\underline{A}1 = \tau(\underline{A}2)$. Then, for any primitive formula, ϕ, of L1 and any sequence, $\underline{x}$, $\in$ A1,

$$\underline{A}1 \models \phi(\underline{x}) \text{ iff } \underline{A}2 \models \tau{:}\phi(\underline{x})$$

Proof:

Suppose

$$\phi(\underline{x}) = Ey1...y.n\ \phi'(\underline{x})$$

where ϕ' is a conjunction of atomic formulae and negations of atomic formulae.

Then $\underline{A}1 \models \phi(\underline{x})$

iff $\underline{A}1 \models Ey1...y.n\ \phi'(\underline{x},y1,....,y.n)$

iff $\underline{A}1 \models \phi'(\underline{x},b1,....,b.n)$ $(b.i \in A1)$

iff $\underline{A}2 \models \tau{:}\phi'(\underline{x},b1,....,b.n)$

iff $\underline{A}2 \models Ey1...y.n(\tau{:}A(y1)\ \&\ ...\ \&\ \tau{:}A(y.n)$

$\&\ \tau{:}\phi'(\underline{x},y1,....,y.n)\ )$

iff $\underline{A}2 \models \tau{:}\phi(\underline{x})$

Theorem 4.2.9. If T2 is model complete, τ is a simple translation from L(T2) to L(T1), and T1 is uniformly τ-reducible to T2, then T1 is also model complete.

Proof: By A.3.5d it is enough to show that given models $\underline{A}1$ and $\underline{B}1$ of T1, where $\underline{A}1 \subseteq \underline{B}1$ and a primitive formula, ϕ:

if $\underline{B}1 \models \phi(\underline{x})$ for $\underline{x} \in A$,

then $\underline{A}1 \models \phi(\underline{x})$.

Since T1 is uniformly τ-reducible to T2, there are models of T2, $\underline{A}2 \subseteq \underline{B}2$ where $\underline{A}1 = \tau(\underline{A}2)$ and $\underline{B}1 = \tau(\underline{B}2)$.

since $\underline{B}1 \models \phi(\underline{x})$, by assumption

$\underline{B}2 \models \tau{:}\phi(\underline{x})$, by lemma 4.2.8.

So $\underline{A}2 \models \tau{:}\phi(\underline{x})$, since T2 is model complete,

and $\underline{A}1 \models \phi(x)$, by lemma 4.2.8.

We shall now define several theories, all more or less familiar, which will serve as stepping stones in showing that our theory of size is model complete.

Definition 4.2.10.

(a) The theory of abelian groups with identity has the following axioms:

(1) $x + (y + z) = (x + y) + z$

(2) $x + y = y + x$

(3) $x + 0 = x$

(4) $(Ey)(x + y = 0)$

(b) The theory of cancellable abelian semigroups with identity consists of (1), (2), and (3) above and:

(4') $x + y = x + z \rightarrow y = z$

(c) The axioms of simple order are:

$x \leq y \ \& \ y \leq z \rightarrow x \leq z$

$x \leq y \ \& \ y \leq x \rightarrow x = y$

$x \leq x$

$x \leq y \mid y \leq x$

(d) The theory of Z-groups, Zg, has the following axioms:

(1) The axioms for abelian groups with identity, (2) The axioms for simple order,

(3) $y \leq z \rightarrow x + y \leq x + z$,

(4) 1 is the least element greater than 0, and

(5) $(x)(Ey)(ny = x \mid \ldots \mid ny = x + (n - 1))$,

for each positive n, where "ny" stands for "y+...+y" (n times).

(e) The theory of N-semigroups has the following axioms:

(1) The axioms for cancellable abelian semi-groups with identity,

(2) - (5): as for Zg, and

(6) $0 \leq x$.

(f) The theory of Z-groups modulo I consists of the following axioms:

(1),(2),(4), and (5): as for Zg,

(3) $(y \leq z \ \& \ x \leq x + z) \rightarrow x + y \leq x + z$,

(6) $0 \leq x$

(7) $x \leq I$

(The theory of Z-groups is taken from Chang and Keisler, p.291).

Fact 4.2.11.

(a) Zg is the complete theory of <Z,+,0,1,<= >

(Chang and Keisler, p.291).

(b) Zg is model complete (Robinson and Zakon).

Theorem 4.2.12.

(a) The theory of N-semigroups is model complete.

(b) The theory of Z-groups modulo I is model complete.

Proof:

(a) Every abelian semigroup with cancellation can be isomorphically embedded in an abelian group (see Kurosh, pp.44-48). It is clear from the construction in Kurosh that if the semigroup is ordered, the abelian group in which it is embedded may also be ordered and that the elements of the semigroup will be the positive elements of the group. Moreover, the (rough) divisibility of the elements in the semigroup will also be carried over to the group.

Consequently, the theory of N-semigroups is uniformly reducible to the the theory of Z-groups by the translation:

$\tau:A = \text{'}0 \leq x\text{'}$

Since the latter is model complete, so is the former, by 4.2.9.

(b) First, consider the theory of N-semigroups in the language which contains, besides the constant symbols in the original theory, an individual constant, I. The theory of N-semigroups is model-complete in this language, by 4.2.2b.

We claim that the theory of Z-groups modulo I is uniformly reducible to this new theory by the following translation:

$\mathfrak{T}$:A = '$x \leq I$'

$\mathfrak{T}$:+ = '$(x+y = z) \mid (x+y = I+z)$'

(The construction: given a model of Zgm, stack up omega-many copies of the model, assigning interpretations in the obvious way. Need to show that the result is an N-semigroup and that the original model is isomorphic to the first copy of itself.)

In the next section, we use theorem 4.2.9, to show that the theory PSIZE is model complete, by reducing it to the theory of Z-groups mod I; in chapter 5, we use Monk's theorem to show that the theory CA is model complete.

4.3 PCA IS MODEL COMPLETE.

To show that PCA is model complete, we shall reduce it to Zgm, the theory of Z-groups with addition taken modulo some constant (see 4.2.11). The model completeness of PCA then follows from the model completeness of Zgm (fact 4.2.12) and theorem 4.2.9. Specifically, we shall show that every model of PCA is the τ-image of a model of Zgm, where τ is the following translation:

Definition 4.3.1. Let τ be the translation from L(PSIZE) to L(Zgm) where:

(a) τ:A = 'x = x'

(b) τ:SUM(x,y,z) = 'x + y = z & x $\leq$ z'

(c) τ:UNIT(x) = 'x = 1'

(d) τ:x! = 'x + y = I'

(e) τ:(x < y) = 'x $\leq$ y & x $\neq$ y'

(f) τ:$\emptyset$ = 'x = 0'

Given a model, A, of PCA, we can construct a model, B, of Zgm directly:

Definition 4.3.2. If A |= PSIZE, then Zgm(A) is the interpretation, B, of L(Zgm) in which:

(a) B = A

(b) B |= x ≤ y iff A |= (x < y ; x = y)

(c) B(I) = A(I)

(d) B(0) = A(∅)

(e) B |= (x + y = z) iff

A |= SUM(x,y,z)

or A |= EaEw(UNIT(a) & SUM(x!,y!,w) & SUM(w,a,z!))

(f) B |= (x = 1) iff A |= UNIT(x)

(It is obvious that A = 𝔗(Zgm(A)).)

Fact 4.3.3 establishes that (e) gives a functional interpretation of '+'. This is what 4.3.3j says. Theorem 4.3.8 establishes that B |= Zgm, on the basis of the intervening facts: 4.3.3 deals with the model A of PSIZE; 4.3.4 deals with the corresponding model, Zgm(A); 4.3.6 verifies some connections between A and Zgm(A).

Fact 4.3.3. The following are theorems of PSIZE.

('Sm(x,y)' abbreviates '(Ez)SUM(x,y,z)'.)

(a) UNIQUE+:

SUM(x,y,z1) & SUM(x,y,z2) -> z1 = z2

(b) UNIQUE-:

SUM(x,y1,z) & SUM(x,y2,z) -> y1 = y2

(c) ASSOC2:

SUM(x,y,w1) & SUM(w1,z,w) ->

(Ew2)(SUM(y,z,w2) & SUM(x,w2,w))

(d) x!! = x

(e) x < y <-> y! < x!

(f) x < y & y < y! -> x < x!

(g) Sm(x,y) ¦ Sm(x!,y!)

(h1) SUM(x,y,z) -> SUM(z!,y,x!)

(h2) SUM(x,y,z) -> SUM(x,z!,y!)

(i) SUM(x,y,z1) & SUM(x!,y!,z2!) -> (z1 = z2! = I)

(j) (Ez1)SUM(x,y,z1) <->

¬(EaEwEz2)(UNIT(a) & SUM(x!,y!,w) & SUM(w,a,z2!))

Proof: COMM is used freely in these proofs, without citation.

a. Suppose SUM(x,y,z1)
and SUM(x,y,z2)
So z1 < z2 iff z2 < z1 iff x < x, by MONOT+
so z1 ≈ z2, by IRREF and TRICH
so z1 = z2, by UNIQUE≈.

b. Suppose SUM(x,y1,z)
and SUM(x,y2,z)
So y1 < y2 iff y2 < y1 iff z < z, by MONOT+
so y1 = y2, by IRREF, TRICH, and UNIQUE≈.

c. Suppose SUM(x,y,w1)
and SUM(w1,z,w)
But SUM(Ø,y,y), by IDENT
and Ø ≤ x, by MIN
so y ≤ w1, by MONOT+ and (b)
so (Ew2)SUM(y,z,w2), by EXIST+
so SUM(x,w2,w), by ASSOC.

d. We know SUM(x,x!,I)
and SUM(x!!,x!,I), by COMP
So x = x!!, by UNIQUE-.

e->. Suppose x < y
Assume x! = y!
But SUM(y,y!,I), by COMP
so SUM(y,x!,I)
but SUM(x,x!,I), by COMP

```
            so x = y, by UNIQUE-
           but this contradicts IRREF
            so ¬(x! = y!)
        Assume x! < y!
           But SUM(y,y!,I), by COMP
            so (Ew1)SUM(y,x!,w1), by EXIST+
           and w1 < I, by MONOT+
          Also (Ew2)SUM(x,x!,w2), by EXIST+, since x < y
           and w2 < w1, by MONOT+
            so w2 < I, by TRANS
           and w2 = I, by UNIQUE+
           but this contradicts IRREF
            so ¬(x! < y!).
         Hence y! < x!, by TRICH.
 <-. Immediate from (->), given (d).
f.   Suppose x < y
           and y < y!
            So y! < x!, by (e)
            so x < x!, by TRANS
g.   We know SUM(x,x!,I), by COMP
         so if y ≤ x!
          then (Ez)SUM(x,y,z), by EXIST+
            so Sm(x,y)
        and if x! < y
          then y! < x!!, by (e)
            so y! < x, by (d)
            so Sm(y!,x!), by EXIST+
```

h1. Suppose SUM(x,y,z)

We know SUM(z,z!,I), by COMP

so SUM(y,z!,w)

and SUM(x,w,I) for some w, by (c)

But w = x!, by COMP and UNIQUE-

so SUM(y,z!,x!)

so SUM(z!,y,x!), by COMM.

h2. Suppose SUM(x,y,z)

so SUM(y,x,z), by COMM

so SUM(z!,x,y!), by (h1)

so SUM(x,z!,y!), by COMM.

i. Suppose SUM(x,y,z1)

and SUM(x!,y!,z2!)

So SUM(z1!,y,x!), by (h1)

but SUM(y,y!,I), by COMP

so SUM(z1!,I,z2!), by ASSOC

but SUM(Ø,I,I), by IDENT

so Ø < z1! iff I < z2!, by MONOT+

so z1! = Ø, by MAX and MIN

and z2! = I, by UNIQUE+

so z1 = z2! = I.

j->. Suppose SUM(x,y,z1)

So if SUM(x!,y!,w)

then w = z1 = I, by (i)

But SUM(I,Ø,I), by IDENT

so if SUM(w,a,z2!)

then Ø < a iff I < z2', by MONOT

so ∅ = a, by MAX and MIN

so ¬UNIT(a)

<-. Suppose ¬Sm(x,y)

then SUM(x!,y!,w), for some w, by (g)

But if w = I,

then x! = y!!

and y! = x!!, by COMP and UNIQUE-

so SUM(x,y,w), by (d)

but this contradicts our original assumption

so w < I, by MAX.

Hence SUM(w,w!,I)

and w! ≠ ∅

So a ≤ w!, for any atom, a

and Sm(w,a), by EXIST+

so SUM(w,a,z2) for some z2

In addition to the theorems of PSIZE listed in 4.3.3, we require a battery of tedious facts about the model B. We shall state these in terms of a model, B*, an expansion of both A and B, which interprets two additional operators, as follows:

Definition 4.3.4. Given a model, <u>A</u>, of PSIZE and <u>B</u> = Zgm(<u>A</u>), <u>B</u>* is the expansion of <u>B</u> induced by the definitions in 4.3.2 together with (a) and (b):

(a) <u>B</u>* |= y = -x iff <u>B</u> |= x + y = 0

(b) <u>B</u>* |= z = x - y iff <u>B</u>* |= z = (x + (-y))

Claim: For each x in B there is a unique y such that

<u>B</u> |= x + y = 0:

Proof:

If x = 0, then <u>B</u> |= x + y = 0 iff y = 0.

->. If <u>B</u> |= 0 + y = 0

then <u>A</u> |= SUM(Ø,y,Ø), by 4.3.2e,

since ¬(Ea)(Ew)(UNIT(a) & SUM(Ø!,y!,w)

& SUM(w,a,z!))

(since Ø!, i.e. I, would have to be < z!)

but <u>A</u> |= SUM(Ø,Ø,Ø), by IDENT

so y = 0, by UNIQUE-.

<-. Obvious.

If x > 0, there is a unique y which satisfies:

(*) SUM(x!,a,y), where UNIT(a)

since there is a unique unit and <u>A</u> |= UNIQUE+

But <u>B</u> |= x + y = 0 iff <u>A</u> |= (*)

->. If <u>B</u> |= x + y = 0

then <u>A</u> |= SUM(x!,y!,w) & SUM(w,a,0!)

since A, surely, doesn'<u>t</u> |= SUM(x,y,0)

but 0! = I

so w = a!, by COMP and UNIQUE-so A |= SUM(x!,y!,a!)

so A |= SUM(x!,a,y), by 4.3.3d and h2.

<-. If A |= (*)

then A |= SUM(x!,y!,a!), by 4.3.3h2

and A |= SUM(a!,a,0!), by COMP

so B |= x + y = 0, by 4.3.2e

Given that we have functional interpretations of '+' and unary '-', (b) also yields a functional interpretation.)

Fact 4.3.5. The following statements hold in B*:

(a) -(x+y) = -x + -y

(b) -(x-y) = y-x

(c) (x+y)-y = x

(d) -(-x) = x

(e) x ≠ 0 & x < y -> -y < -x

Proof: Omitted.

Fact 4.3.6. B* satisfies the following:

(a) Sm(a,b) <-> SUM(a,b,a+b)

(b) c ≤ b <-> SUM(c,b-c,b)

(c) SUM(a,b,c) & 0 < b -> SUM(a,-c,-b)

SUM(a,b,c) & 0 < a -> SUM(-c,b,-a)

(d) ¬Sm(a,b) -> Sm(-a,-b) ¦ b = -a

Proof:

a->. Suppose Sm(a,b)

so SUM(a,b,w) for some w

so (a+b) = w, by the definition of '+'

so SUM(a,b,a+b)

<-. Obvious.

b->. Suppose c ≤ b

then SUM(c,w,b), for some w, by EXIST-

so b = (c+w), by (a) and UNIQUE+

so (b-c) = (c+w)-c

so (b-c) = w, by 4.3.5c

so SUM(c,b-c,b)

<-. Suppose SUM(c,b-c,b)

but SUM(c,0,c) by IDENT

so 0 ≤ (b-c) <-> b ≤ c, by MONOT and UNIQUE+

so b ≤ c, by MIN

c->. Suppose SUM(a,b,c)

and 0 < b

then c = (a+b) by (a) and UNIQUE+

so (c-b) = (a+b)-b

so (c-b) = a, by 4.3.5c

so -b = a-c, by 4.3.5c

so SUM(a,-c,-b) if Sm(a,-c), by (a)

But if a = 0

then Sm(a,-c) by IDENT

and if a ≠ 0,

then a < c, since 0 < b, by MONOT and UNIQUE+

so -c < -a, by 4.3.5e

so -c ≤ a! (see 4.3.4)

but Sm(a,a!), by IDENT

so Sm(a,-c), by EXIST+.

(The second form follows immediately by COMM)

d->. Suppose ¬Sm(a,b)

but b = (a+b)-a, by 4.3.5c

so ¬SUM(a,(a+b)-a,a+b), by (a)

so ¬(a ≤ a+b), by (b)

so (a+b) < a, by TRICH

so if (a+b) ≠ 0

then -a < -(a+b), by 4.3.5e

so -a < -a + -b, by 4.3.5a

so SUM(-a,(-a + -b) - (-a), -a + -b), by (b)

but (-a + -b) - (-a) = -b, by 4.3.5c

so SUM(-a,-b,-a + -b)

so Sm(-a,-b)

and if (a+b) = 0

then b = -a

We can now show that $\underline{B} \models Zgm$. The only real difficulty arises in verifying that addition is associative. We need the following lemma.

Lemma 4.3.7. $\underline{B} \models (*)$.

$$(*)\ (a+c) + (b-c) = (a+b)$$

Proof: we shall establish (*) for successively more general cases.

Case 1. Sm(a,b) and $c \leq b$

We know SUM(b−c,c,b), by 4.3.6b

and SUM(b,a,a+b), by 4.3.6a

so Ew(SUM(c,a,w) & SUM(b−c,w,a+b)) by ASSOC2

so w = c+a, by 4.3.6a and UNIQUE+

and a+b = (b−c) + (c+a), for the same reasons

so a+b = (a+c) + (b−c).

Case 2. Sm(a,b) and Sm(a,c)

If $c \leq b$, case 1 applies directly.

So assume $b < c$.

then a+c = (a+b) + (c−b), by case 1

so (a+c) − (c−b) = (a+b), by 4.3.5c

so (a+c) + (−(c−b)) = a+b, by 4.3.4b

so (a+c) + (b−c) = (a+b), by 4.3.5b

Case 3. Sm(a,b)

If Sm(a,c), case 2 applies.

So assume ¬Sm(a,c)

and thus b < c, by EXIST+

So SUM(b,c-b,c), by 4.3.6b

and 0 < c-b, by IDENT and MONOT+

So SUM(b,-c,-(c-b)), by 4.3.6c

so SUM(b,-c,b-c), by 4.3.5b

so SUM(-c,b,b-c) (3a)

Either Sm(-a,-c)

or c = -a, by 4.3.6d, since ¬Sm(a,c).

Assume Sm(-a,-c)

then SUM(-a,-c,(-a) + (-c)), by 4.3.6a

so SUM(-a,-c,-(a+c)), by 4.3.5a

and 0 < -a

So SUM(-(-(a+c)),-c,-(-a)), by 4.3.6c

so SUM(a+c,-c,a), by 4.3.5d (3b)

Assume c = -a

then a+c = 0

and a = -c

so, again, (3b) obtains.

Hence SUM(a+c,-c,a) (3b)

and SUM(-c,b,b-c) (3a)

and SUM(a,b,a+b) since Sm(a,b)

So SUM(a+c,b-c,a+b), by ASSOC

so (a+c) + (b-c) = (a+b), by 4.3.6a

Case 4. Whenever.

Suppose $\neg Sm(a,b)$, for otherwise case 3 applies.

Either $Sm(-a,-b)$

or $b = -a$, by 4.3.6d

Assume $b = -a$.

$$
\begin{aligned}
\text{Then } a+b &= 0 \\
\text{and } b-c &= (-a)-c \\
&= (-a)+(-c) \\
&= -(a+c)
\end{aligned}
$$

$$
\begin{aligned}
\text{So } (a+b) + (b-c) &= (a+c) + -(a+c) \\
&= 0 \\
&= (a+b).
\end{aligned}
$$

Assume $Sm(-a,-b)$.

$$
\begin{aligned}
\text{Then } -a + -b &= (-a + -c) + (-b - -c), \text{ by case 3} \\
\text{so } -(a+b) &= -(a+c) + -(b-c), \text{ by 4.3.5a} \\
&= -((a+c) + (b-c)), \text{ by 4.3.5a} \\
\text{so } a+b &= (a+c) + (b-c), \text{ by 4.3.5d.}
\end{aligned}
$$

Theorem 4.3.8. $\underline{B} \models Zgm$.

Proof. The only axiom for abelian groups that needs further verification is associativity. We prove this using lemma 4.3.7:

$$
\begin{aligned}
x + (y+z) &= (x + y) + ((y+z) - y), \text{ by 4.3.7} \\
&= (x + y) + z, \text{ by 4.3.5c}
\end{aligned}
$$

The ordering axioms of Zgm are satisfied in $\underline{B}$ because $\underline{B}$ uses the same ordering as $\underline{A}$, $\underline{A}$ satisfies PSIZE, and PSIZE includes the same ordering axioms. Similarly, $\underline{B}$

satisfies axioms (iv), (vi), and (vii) of Zgm because A satisfies MIN, MAX, and the UNIT axioms of PSIZE.

Axiom (iii) of Zgm is:

$$y \leq z \;\&\; x \leq x+z \rightarrow x+y \leq x+z$$

If B |= $x \leq x+z$, then A |= SUM(x,z,x+z)

So if B |= $y < z$, then A |= SUM(x,y,x+y), by EXIST+

and A |= $x+y \leq x+z$, by MONOT

so B |= $x+y \leq x+z$.

Finally, the divisibility of elements in B required by axiom (v) of Zgm is guaranteed by the fact that A satisfies the divisibility principles of PCA.

So, we have shown that PCA is ₮-reducible to Zgm. The reduction is obviously uniform, since each model A of PCA has the same domain as its Zgm-model. So we may conclude that PCA is model complete.

Theorem 4.3.9.

(a) PCA is model complete.

(b) PCA satisfies Monk's condition.

Proof:

(a) Apply theorem 4.2.9.

(b) Immediate from (a) and 4.2.1b.

4.4 REMARKS ON PROVING THAT PCA = PCS.

To prove that PCA $\equiv$ PCS, we could follow the method outlined at the end of chapter 3 for showing that CA $\equiv$ CS. that is to say: we already know that PCS |- PCA, so we need only prove (1), which follows from (2) by 3.3.1b.

(1) PCA |- PCS

(2) Every completion of PCA is consistent with PCS.

But (2) is equivalent to the conjunction of (2a) and (2b).

(2a) Every finite completion of PCA is consistent with PCS.

(2b) Every infinite completion of PCA is consistent with PCS.

The finite completions of PCA are just the (categorical) theories PCA;EXACTLY{n}. But PCA;EXACTLY{n} is true in S(A), where A is the finite standard interpretation of L(C<) containing n atoms. So, the finite completions of PCA are consistent with PCS. (Formally, we would have to redefine "EXACTLY" in terms of units rather than atoms.)

Letting PCAI = PCA + INF, we see that (2b) is a consequence of (3a) and (3b).

(3a) If PCAI(f) is consistent, then PCSI(f) is consistent.

(3b) If T is a completion of PCAI, then T $\equiv$ PCAI(f), for some f.

To prove (3a), we would have to prove analogues of 3.3.9 through 3.3.13 for PCA and PCS. This seems straightforward, but tedious. The trick is to show that enough axioms about size relations have been incorporated in PCA to establish the entailments among MOD statements.

To prove (3b), it is sufficient to demonstrate (4), because every completion of PCAI entails PCAI(f) for some total f:

(4) If f is total, then PCAI(f) is complete.

But PCA is model complete, so only (5) remains to be shown.

(5) If f is total and congruous, then PCAI(f) has a prime model.

We won't construct prime models for the extensions of PCAI. It's apparent that the size models of the prime models for CAI(f) would do nicely. Alternatively, the construction could be duplicated in this simpler case.

Page 126 was not included in the original document.

5.1 CA IS MODEL COMPLETE.

We shall show that CA is model complete by showing that it satisfies Monk's criterion (see 4.3.1). So, given assumption 5.1.1, we want to prove 5.1.2:

Assumption 5.1.1. $\underline{A}$ |= CA, $\underline{B}$ |= CA, $\underline{A} \subseteq \underline{B}$, and $\underline{C}$ is a finitely generated substructure of $\underline{B}$.

Theorem 5.1.2. There is an isomorphic embedding, f: $\underline{C}$ ==> $\underline{A}$, where f(x) = x if x is in C /\ A.

We will use the Monk mappings that exist for PCA as a guide in constructing Monk mappings for CA. The existence of Monk mappings for PCA tells us that we can find elements with the right sizes. DISJU and REP< then allow us to find elements with those same sizes that fit together in the right way.

Strictly speaking, $\underline{S}(\underline{A})$ is not a submodel of $\underline{S}(\underline{B})$, so we cannot apply Monk's theorem directly. But, let

$\underline{S}(X,\underline{B})$ = the submodel of $\underline{S}(\underline{B})$ whose domain is

$\{\delta(x,\underline{B}): x \in X\}$

Then, clearly,

$\underline{S}(\underline{A}) \cong \underline{S}(A,\underline{B}) \subseteq \underline{S}(\underline{B})$, and

$\underline{S}(\underline{C}) \cong \underline{S}(C,\underline{B})$, a finitely generated submodel of $\underline{S}(\underline{B})$

Monk's theorem applies directly to $\underline{S}(A,\underline{B})$, $\underline{S}(\underline{B})$ and $\underline{S}(C,\underline{B})$, so we may conclude 5.1.3:

Fact 5.1.3. There is an isomorphic embedding, $g: \underline{S}(\underline{C}) \Longrightarrow \underline{S}(\underline{A})$, where $g(\delta(x,\underline{B})) = \delta(x,\underline{A})$ for all x in $C \wedge A$.

That is to say, the sizes of elements in $\underline{C}$ can be embedded in the sizes of elements of $\underline{A}$. It remains to be shown that the elements of C themselves can be mapped into A by a function, f, which preserves boolean relations as well as size relations.

Notice that $\underline{C}$ is a finite, and hence atomic, boolean algebra, though its "atoms" need not be atoms of $\underline{B}$; indeed, if $\underline{B}$ is infinite, there must be some atoms of $\underline{C}$ which are not atoms of $\underline{B}$ since the union of all atoms of $\underline{C}$ is the basis of $\underline{B}$.

Definition 5.1.4. d is a molecule iff $d \in C \wedge A$ and no proper subset of $d \in C \wedge A$.

Notice that $C \wedge A$ is a boolean algebra whose basis is the same as the basis of C. So every atom of C is included in some molecule. The embedding, f, has to map each molecule into itself. Moreover, f has to be determined by its values on the atoms of $\underline{C}$, since f must preserve unions. In fact, the atoms of $\underline{C}$ can be partitioned among the molecules. So, if

$$d = b.1 \vee \dots \vee b.n$$

where d is a molecule and b.1,...., b.n are the atoms of $\underline{C}$ contained in d, then d must also be the (disjoint) union of f(b1),....,f(b.n). If this condition is satisfied, then f will preserve boolean relations. Of course, f must also select images with appropriate sizes.

Proof of 5.1.2. Given a molecule, d, let b1, ..., b.n be all atoms of C contained in d. For each b.i, let c.i be some member of g(δ(b.i,$\underline{B}$)). These elements will be elements of A, since g yields sizes in $\underline{A}$, whose members are in A. These elements have the right sizes, as we will show below, but they are in the wrong places. We have no guarantee that they are contained in the molecule, d. So we still have to show that there are disjoint elements of A, a1, ..., a.n, whose union is d and whose sizes are the same as b1, ..., b.n, respectively. Well,

$$d = b1 \vee \dots \vee b.n$$

so $\quad \underline{C} \models \mathrm{SUM}(b1,\dots,b.n,d)$ because $\underline{C} \models \mathrm{DISJU}$

so $\underline{S}(\underline{C}) \models \mathrm{SUM}(\delta(b1,\underline{B}),\dots,\delta(b.n,\underline{B}),\delta(d,\underline{B}))$

so $\underline{S}(\underline{A}) \models \mathrm{SUM}(g(\delta(b1,\underline{B})),\dots,g(\delta(b.n,\underline{B})),g(\delta(d,\underline{B})))$

so $\underline{A} \models SUM(c1, \ldots, c.n, d)$

since each c.i is in $g(\delta(b.i, \underline{B}))$

and d is in $g(\delta(d, \underline{B})) = \delta(d, \underline{A})$

so the existence of a1, ..., a.n, as above, are guaranteed because $\underline{A} \models$ DEF+.

Now, let f(b.i) = a.i for each b.i in the molecule d. Repeating this procedure for each molecule yields a value of f for each atom of $\underline{C}$. Finally, if $x \in C$ is non-atomic, then

$x = b1 \vee \ldots \vee b.k$, where each b.i is atomic

So let

$f(x) = f(b1) \vee \ldots \vee f(b.k)$

We claim that f satisfies all the requirements of theorem 5.1.2: Boolean relations are preserved by f because the function is determined by its values on the atoms of $\underline{C}$; f maps elements of $C \wedge A$ into themselves because the set of atoms contained in each of these molecules is mapped into a disjoint collection of elements of A whose union is the same molecule; so, we need only show that f preserves size relations. To do so, we invoke lemma 5.1.5, below.

(a) $\underline{C} \models x < y$ iff $\underline{A} \models f(x) < f(y)$

(b) $\underline{C} \models x \approx y$ iff $\underline{A} \models f(x) \approx f(y)$

(c) $\underline{C} \models SUM(x, y, z)$ iff $\underline{A} \models SUM(f(x), f(y), f(z))$

(d) $\underline{C} \models UNIT(x)$ iff $A \models UNIT(f(x))$

Proof of (a):(The others are similar.)

$\underline{C} \models x < y$ iff $\underline{S}(\underline{B}) \models \delta(x,\underline{B}) < \delta(y,\underline{B})$

iff $\underline{S}(\underline{A}) \models g(\delta(x,\underline{B})) < g(\delta(y,\underline{B}))$

iff $\underline{S}(\underline{A}) \models \delta(f(x,\underline{A})) < \delta(f(y,\underline{A}))$, by 5.1.5

iff $A \models f(x) < f(y)$

So, we may conclude

Theorem 5.1.6. CA is model complete.

Lemma 5.1.5 For all x in C,

$$\delta(f(x),\underline{A}) = g(\delta(x,\underline{B}))$$

Proof: Suppose

$$x = b1 \lor \cdots \lor b.n,$$

where each b.i is an atom of C.

So, $\underline{B} \models SUM(b1, \ldots, b.n, x)$, since $\underline{B} \models DISJU$

so $\underline{S}(\underline{B}) \models SUM(\delta(b1,\underline{B}), \ldots, \delta(b.n,\underline{B}), \delta(x,\underline{B}))$

so $\underline{S}(\underline{A}) \models SUM(g(\delta(b1,\underline{B})), \ldots, g(\delta(b.n,\underline{B})), g(\delta(x,\underline{B})))$

but for atomic b, f(b) was chosen so that

$$g(\delta(b,\underline{B})) = \delta(f(b),\underline{A})$$

so $\underline{S}(\underline{A}) \models SUM(\delta(f(b1),\underline{A}), \ldots, \delta(f(b.n),\underline{A}), g(\delta(x,\underline{B})))$

but $\underline{A} \models SUM(f(b1), \ldots, f(b.n), f(x))$,

since $\underline{A} \models DISJU$

so $\underline{S}(\underline{A}) \models SUM(\delta(f(b1),\underline{A}), \ldots, \delta(f(b.n),\underline{A}), \delta(f(x),\underline{A}))$

but, sums are unique in $\underline{S}(\underline{A})$,

so, $g(\delta(x,\underline{B})) = \delta(f(x),\underline{A})$

5.2 PRIME MODELS FOR CAI(F).

For each total, congruous remainder function, f, we want to find a prime model, Q(f), for CAI(f). All of these prime models can be defined over the class, Q, of sets near quasi-congruence classes (see section 3.2). Of course, for different remainder functions, we have to assign different size relations over Q. Section 5.2.1 defines the structures Q(f) and verifies that each satisfies the respective theory CAI(f); section 5.2.2 defines, for each model of CAI(f), a submodel, or "shell"; section 5.2.3 shows that Q(f) is isomorphic to the shell of any model of CAI(f).

5.2.1 The models Q(f).

The construction here is just a more elaborate version of the construction of Q in chapter 3 (see 3.2.19). The old model, Q, turns out to be Q(f), where f(n) = 0 for all n. As in the case of Q, the models Q(f) and their copies in arbitrary models of CAI are the unions of chains of smaller models.

The sizes assigned to elements of Q to induce Q(f) for a total, congruous remainder function f are more elaborate than those used in the definition of Q (see 3.2.22), but they are employed in substantially the same way:

Definition 5.2.1.

(a) A <u>size</u> is an ordered pair, $\langle \bar{\beta}, \bar{\delta} \rangle$, of rational numbers.

(b) If $\theta 1 = \langle \bar{\beta}1, \bar{\delta}1 \rangle$ and $\theta 2 = \langle \bar{\beta}1, \bar{\delta}1 \rangle$, then

i) $\theta 1 < \theta 2$ iff $\bar{\beta}1 < \bar{\beta}2$ or

$$(\bar{\beta}1 = \bar{\beta}2 \text{ and } \bar{\delta}1 < \bar{\delta}2)$$

ii) $\theta 1 + \theta 2 = \langle \bar{\beta}1 + \bar{\beta}2, \bar{\delta}1 + \bar{\delta}2 \rangle$

(cf. 3.2.22)

To assign sizes for $\underline{Q}(f)$ we rely on the representation of sets in Q which is defined in 3.2.24. $\underline{Q}(f)$, unlike $\underline{Q}$, assigns different sizes to the n-congruence classes for a given n:

Definition 5.2.2. If f is total and congruous, then

(a) if $x = \{n*\underline{k} + i\}$, then

$$\theta(f,x) = \langle 1/n, 1/n \rangle \qquad \text{if } i < f(n)$$

and $\theta(f,x) = \langle 1/n, (1/n) - 1 \rangle$ if $f(n) \leq i$

(b) if $x \in QC(n)$, so x is the disjoint union

$$x.1 \vee \ldots \vee x.k$$

of n-congruence classes, then

$$\theta(f,x) = \theta(f,x.1) + \ldots + \theta(f,x.k).$$

(c) if $x \in Q$, so x can be represented as

$$(C(x) \vee D1(x)) - D2(x)$$

as in 3.2.24, then

$$\theta(f,x) = \theta(f,C(x)) + \langle 0, cd(D1(x)) - cd(D2(x)) \rangle$$

Intuitively, all n-congruence classes are assigned sizes $\langle 1/n, \bar{a} \rangle$, but $\bar{a}$ is no longer 0 in all cases, as in $\underline{Q}$. Instead, the first f(n) n-congruence classes are each one atom larger than the remaining (n - f(n)) n-congruence classes.

The desired models of CAI may now be defined:

Definition 5.2.3. If f is total and congruous, then $\underline{Q}(f)$ is the model, A, with domain Q in which

$\underline{A} \models (x = y)$ iff $\theta(f,x) = \theta(f,y)$

$\underline{A} \models (x < y)$ iff $\theta(f,x) < \theta(f,y)$

$\underline{A} \models SUM(x,y,z)$ iff $\theta(f,z) = \theta(f,x) + \theta(f,y)$

(cf. 3.2.27)

To verify that the structure $\underline{Q}(f)$ is indeed a model of CAI(f), for total and congruous f, we exhibit each such model as the union of a chain of models.

Definition 5.2.4. If f is total and congruous, then $\underline{Q}(f,n)$ is the submodel of $\underline{Q}(f)$ whose domain is Q(n)

Fact 5.2.5. If f is total and congruous and $n > 0$, then $\underline{Q}(f,n) \models$ BASIC.

Proof: The proof can be obtained from the proof of theorem 3.2.29 by substituting:

'$\underline{Q}(f,n)$' for '$\underline{Q}(n)$',

'$\theta(f,x)$' for '$\theta(x)$',

'$\bar{\beta}(f,x)$' for '$\bar{\beta}(x)$',

and '$\bar{\delta}(f,x)$' for '$\bar{\delta}(x)$'.

The following notion is helpful in understanding our constructions.

Definition 5.2.6. Suppose $\underline{A}$ |= BASIC, x ∈ A, and $0 \leq m < n$. Then an <u>(n,m)-partition of</u> x <u>in</u> $\underline{A}$ is a sequence x1,...,x.n, where

(1) a. If $0 < i < j \leq m$, then x.i ≈ x.j,

b. If $m < i < j \leq n$, then x.i ≈ x.j,

c. If $0 < i \leq m < j \leq n$, then SUM(x.j,a,x.i) for any atom, a.

(2) If $0 < i < j \leq n$, then x.i /\ x.j = ∅, and

(3) x = x1 \/ ... \/ x.n

In other words, x is partitioned among n infinite, pairwise disjoints sets which are roughly the same size: each of the first m is one atom larger than each of the remaining (n-m). If $0 < i \leq m$, x.i is called a "charmed n-factor of x"; for $i > m$, x.i is a "common n-factor of x".

The sequence {$\underline{Q}(f,n)$} does not constitute a chain of models. For example, $\underline{Q}(f,3)$ is not an extension of $\underline{Q}(f,2)$. But this sequence does harbor a chain of models:

Fact 5.2.7. If n < m, then Q(f,n!) c Q(f,m!).

Proof: It will be clearer, as well as easier, to establish this by example than by formal proof. Letting n = 2 and m = 3, we want to show that the 2-congruence classes have the same size relations in Q(f,6) as they do in Q(f,2), for any f. The other elements of Q(f,2) will then fall into place, since size relations are determined by the representation of a set, x, as C(x), D1(x), and D2(x).

Suppose that f(2) = 0, so that

Q(f,2) |= {2k} = {2k + 1}

Since f is congruous, f(6) = 0, 2, or 4. If f(6) = 0, then all of the 6-congruence classess are common. If f(6) = 2, then |_6k| and |_6k + 1_| are the only charmed 3-congruence clases. If f(6) = 4, then all of the 3-congruence classes are charmed except {6n + 4} and {6n + 5}.

In any case, {2n} will include the same number of charmed 3-congruence classes as {2n + 1}, so

Q(f,6) |= {2k} = {2k + 1}

Suppose, however, that f(2) =1, so that

Q(f,2) |= {2n} is one atom larger than {2n + 1}

Here, f(6) = 1, 3, or 5, since f is congruous. In any case, { 2n_} contains exactly one more charmed 3-congruence class than {2n + 1}, so

$\underline{Q}(f,6) \models \{2n\}$ is one atom larger than $\{2n + 1\}$

So it goes in general.

Fact 5.2.7 allows us to regard $\underline{Q}(f)$ as the union of a chain of models:

Fact 5.2.8. If f is total and congruous, then

$$\underline{Q}(f) = \bigvee \{ \underline{Q}(f,n!) : n > 0\}$$

Fact 5.2.9. If f is total and congruous, then

(a) $\underline{Q}(f) \models$ BASIC

(b) $\underline{Q}(f) \models$ ADIV{k}, for $k > 0$

(c) $\underline{Q}(f) \models$ MOD{n,f(n)}, for $n > 0$

(d) $\underline{Q}(f) \models$ CAI(f)

Proof:

(a) BASIC is a universal-existential theory; so it is preserved under unions of chains (see A.2.5).

(b) Each n-congruence class can be partitioned into k (n*k)-congruence classes.

(c) Clearly, $\underline{Q}(f,n) \models$ MOD{n,f(n)}.
But MOD{n,f(n)} is an existential sentence, so it is preserved under extensions.

(d) Immediate from (a), (b), and (c).

5.2.2 Shells of models.

To embed the model Q(f) into an arbitrary model, A, of CAI(f), we must find a "smallest" submodel, B, of A which satisfies CAI. Clearly, the basis of A, call it x0, must be included in B, since the symbol 'I' must refer to the same set in the submodel as it does in the model. But if the basis of A is in B and B |= CAI, then B must contain two disjoint sets of roughly the same size whose union is the basis of A. Pick such a pair, x1 and x2, to include in B. (Whether these are exactly the same size or differ by an atom is determined by whether A |= MOD{2,0} or A |= MOD{2,1}.)

B must also satisfy ADIV{3}. We can aim for this by placing in B three disjoint sets, x11, x12, and x13, whose union is x1 and another three disjoint sets, x21, x22, and x23, whose union is x2. The existence of such sets is assured because A |= ADIV{3}. Again, the exact size relations among these sets will be determined by which MOD-principles are satisfied in A. Notice that by insuring that x1 and x2 are divisible by three, we also guarantee that x0 is divisible by three: the three unions

    x11 \/ x21, x12 \/ x22, x31 \/ x32

will be roughly the same size and exhaust x0.

We can continue this process indefinitely, dividing each set introduced at stage n into (n+1) roughly equal subsets at stage (n+1). This will produce an infinite tree, bearing sets. The deeper a node is in this tree, the smaller the set it bears and the greater the number of successors among which this set will be partitioned.

This great tree of sets will not form a boolean algebra, for it will not be closed under finite unions. A boolean algebra could be obtained by including both the node-sets and their finite unions, but this would still not be an atomic boolean algebra, which is what we are looking for. We can't correct for this problem by including in B all atoms of $\underline{A}$: $\underline{A}$ may have uncountably many atoms while $\underline{B}$, if it is to be a prime model, must be countable. We leave the solution of this problem to the formal construction below.

The formal proof will proceed as follows: First, we define a tree, i.e. a set of nodes, on which we shall hang both the components of the successive partitions described above and the atoms which will find their way into the submodel being constructed. Second, we present the construction which, given a model $\underline{A}$ of CAI(f), assigns a "node-set" A(P) and a "node-atom", a(P), to each node, P. Third, we define the shell of $\underline{A}$ as the submodel of A generated by the collection of node sets and node atoms. In the next subsection, we show that the shell of $\underline{A}$ is isomorphic to Q(f).

First, the tree:

Definition 5.2.10.

(a) A *node* is a finite sequence $\langle n_1, \ldots, n_k \rangle$, where $k > 0$ and for all $i \leq k$, $n_i < i$. (The letters P and R will be used as variables ranging over nodes.)

(b) If $P = \langle n_1, \ldots, n_k \rangle$, then

1. The *length* (or *depth*) of P, $|P|$, is k,
2. If $1 \leq i \leq k$, then $P(i) = n_i$, and
3. $P'm = \langle n_1, \ldots, n_k, m \rangle$

(c) P *extends* R iff $|P| > |R|$ and
if $1 \leq i \leq |R|$, then $P(i) = R(i)$

(d) P *0-extends* R iff P extends R and
if $|R| < i \leq |P|$, then $P(i) = 0$

Nodes can be regarded as the vertices of an infinite tree in which $\langle 0 \rangle$ is the root and P dominates R iff R extends P. Notice that the number of immediate descendants of a node grows as the depth of the node increases.

We shall now assign a set to each node by repeatedly partitioning the basis of *A*. At the same time we shall assign an atom to each node.

Construction 5.2.11. Suppose $\underline{A} \models CAI(f)$. For each node, P, we define A(P), an infinite element of $\underline{A}$, and a(P), an atomic element of $\underline{A}$, as follows; note that this construction makes arbitrary choices at a number of points:

(a) Let A(<0>) be the basis of $\underline{A}$, and

let a(<0>) be any atom of $\underline{A}$.

(b) Suppose A(P) and a(P) have been chosen.

Let $m = |P|$ and let

$$k = (f((m+1)!) - f(m!)) \,/\, (m!)$$

(Recall that f must be congruous if CAI(f) has a model, so k is an integer. Now, let

$$A(P^\bullet 0), \ldots, A(P^\bullet m)$$

be an (m+1,k)-partition of A(P) if A(P) is common or an (m+1,k+1)-partition of A(P) if A(P) is charmed. Fact 5.2.12 guarantees that such partitions exist. In either case, choose A(P•0) so that it contains a(P). This is always possible because A(P) contains a(P).

(c) Let a(P•0) = a(P)

If $0 < i \leq m$, let a(P•i) be any atomic subset of A(P•i).

The sets A(P) will be referred to as "node-sets" and the atoms a(P) as "node-atoms".

Fact 5.2.12. Suppose $\underline{A} \models CAI(f)$, $n > 0$, $m > 0$, and

$$k = f(n*m) - f(n))/n.$$ Then

(a) x has an (m,k)-partition iff $\underline{A} \models Mod\{m,k\}(x)$,

(b) every common n-factor of A has an (m,k)-partition,

(c) every charmed n-factor of A has an (m,k+1)-partition.

Proof:

(a) Obvious.

(b), (c). If (b) holds, then so does (c), since each charmed n-factor is one atom larger than each common n-factor.

Furthermore, (b) and (c) must hold for some k, since all common (charmed) n-factors are the same size and satisfy the same Mod{m,k} predicate.
So, suppose that each common n-factor has k charmed m-factors and (m-k) common m factors.

By (a), $\underline{A}$ has f(n) charmed n-factors and (n-f(n)) common n-factors. Partitioning each of the n-factors into m subsets of roughly the same size yields an (n*m)-partition of $\underline{A}$; the charmed m-factors of the n-factors of $\underline{A}$ are the charmed (n*m)-factors of $\underline{A}$ and the common m-factors of the n-factors of $\underline{A}$ are the common (n*m)-factors of $\underline{A}$.

Each of the common n-factors has k charmed m-factors and each of the charmed n-factors has (k+1) charmed m-factors. In all, then, there are

$$(n - f(n))*k + f(n)*(k+1),$$

i.e. $n*k + f(n)$,

charmed m-factors among the n-factors of A.

But, by (a) again, there are f(n*m) charmed (n*m)-factors of A. So

$$f(n*m) = n*k + f(n)$$

and $k = (f(n*m) - f(n))/n$.

The following list of facts are provided mainly to help the reader understand this construction. All of them can be established by induction on the depth of nodes.

Fact 5.2.13.

(a) $A(P) \subset A(R)$ iff R extends P or $P = R$.

(b) If $A(P) = A(R)$, then $P = R$

(c) There are $n!$ nodes (and hence node-sets) of depth n.

(d) If $i \neq j$, then $A(P'i) \wedge A(P'j) = 0$

(e) Any two node sets of the same depth are disjoint.

(f) Each node-set is the disjoint union of its immediate descendants.

(g) Each node-set is the (disjoint) union of all of its descendants of a given depth.

(h) $a(P) \subset A(R)$ iff P extends R.

(i) $a(P) = a(R)$ iff $P = R$ or one of P and R 0-extends the other.

(j) Every node-set contains infinitely many node-atoms.

(k) For any n, the node-sets of depth n form an $(n!, f(n!))$-partition of the basis of A.

Proof of (k):

For each n, let (*) be the claim that

(*) the node-sets of depth n form an
(n!,f(n!))-partition of the basis of A.

If n = 1, then n! = 1, f(n!) = 0; A(<0>) = the basis of A. So (*) holds because any set is a <1,0>-partition of itself.

Assume that (*) holds for n. Then (*) also holds for (n+1): each node set of depth n has (n+1) immediate descendants, so there are (n!)*(n+1) = (n+1)! node sets of depth (n+1).

Furthermore, f(n!) of the n-factors are charmed and (n! - f(n!)) are common. By fact 5.2.12, each charmed n-factor has an (n+1,k+1)-partition and each common n-factor has an (n+1,k)-partition, where

k = (f(n!*(n+1)) - f(n!))/(n!)

(substituting 'n!' for 'n' and '(n+1)' for 'm').

So, there are

(k+1)*f(n!) charmed (n+1)-factors from the
charmed n-factors
and k*(n! - f(n!)) charmed (n+1)-factors from

the common n-factors.

In all, then, the number of charmed n-factors is:

$$(k+1)*f(n!) + k*(n! - f(n!))$$
$$= k*f(n!) + f(n!) + k*(n!) - k*f(n!)$$
$$= f(n!) + k*(n!)$$
$$= f(n!) + f((n!)*(n+1)) - f(n!)$$
$$= f((n!)*(n+1))$$
$$= f((n+1)!).$$

So, the (n+1) factors of the n-factors of the basis form an ((n+1)!, f((n+1)!))-partition of the basis. That is to say, (*) holds for (n+1).

Given a collection of node-sets, A(P), and node-atoms, a(P), from a model, A, of CAI(f), we can now construct a submodel, B, of A which is isomorphic to Q(f).

Definition 5.2.14. Suppose f is total and congruous, A |= CAI(f) and {A(P),a(P) : P is a node} are a collection of node sets and node atoms of A produced by construction 5.2.11. Then the submodel of A generated by {A(P),a(P)} is a shell of A.

For the remainder of this chapter, we will regard as fixed:

f - a total, congruous remainder function,

A - a model of CAI(f),

{A(P)}, {a(P)} - a collection of node sets and node atoms produced by 5.2.11

and B - the shell of A generated from {A(P), a(P)}

To show that B is isomorphic to Q(f), we need a sharper characterization of the elements of B. Recall that each member, x, of Q can be represented as:

$$(C(x) \cup D1(x)) - D2(x)$$

where C(x) is a quasi-congruence class and D1, D2 are finite sets. We may require that:

$$D1(x) \cap C(x) = D1(x) \cap D2(x) = \emptyset$$

and that $D2(x) \subseteq C(x)$

and, if we do so, the representation is unique (see 3.2.24).

We can obtain a similar representation for elements of B: the node sets play the role of (some of) the congruence classes; finite unions of node sets correspond to the quasi-congruence classes; finite sets of node- atoms correspond to the finite subsets in Q.

Definition 5.2.15.

(a) x is a quasi-nodal set of A iff it is the union of finitely many node sets (iff it is the union of finitely many node sets at a given level).

(b) x is a finite A-set iff it is a finite set of node atoms.

(c) If x ∈ A, y ∈ A, then x is A-near y iff both (x-y) and (y-x) are finite A-sets.

(cf. 3.2.14 - 3.2.18)

Still following in the footsteps of chapter 3, we can characterize B as the collection of sets A-near quasi-nodal sets. Analogues of 3.2.15 through 3.2.18 obtain for A-nearness.

Fact 5.2.16.

(a) x ∈ B iff x is A-near some quasi-nodal set of A.

(b) B is an atomic boolean algebra whose atoms are the node atoms, a(P).

(c) If x ∈ B, then x has a unique representation as

$$(C(x) \vee D1(x)) - D2(x),$$

where C(x) is a quasi-nodal set disjoint from D1(x) and including D2(x), both of which are finite A-sets.

Proof:

(a) (->) B is generated from node sets and node atoms via the boolean operations \/ , /\ , and -, each of which preserves A-nearness to quasi-nodal sets.

(<-) B must contain finite unions of node-sets as well as A-finite sets; so it must also contain sets obtained by adding or removing A-finite sets from quasi-nodal sets.

(b) The proof parallels that of theorem 3.2.21 exactly.

(c) Let C(x) be *the* quasi-nodal set which is A-near x. (cf. 3.2.23); let D1(x) = (x - C(x)); and let D2(x) = (C(x) - x) (cf. 3.2.24).

5.2.3 <u>The embeddings</u>.

To embed <u>Q</u>(f) into <u>A</u>, we first describe Q in terms of node sets and node atoms. In effect, we are performing the construction 5.2.12 on <u>Q</u>(f), but with two differences: first, we are stipulating which (n,m)-partitions to use at each level; second, we are selecting node atoms so that every singleton in Q is the node atom for some node. By satisfying this latter condition, we can be assured that the shell of <u>Q</u>(f) will be <u>Q</u>(f) itself.

Definition 5.2.17. Suppose P is a node. Then

(a) If |P| = k, then

$$Q(P) = \{ (k!)*\underline{n} + \text{sum}\{P(i)*(i-1)! : 1 \le i \le k\} \}$$

(b) The <u>depth of</u> Q(P) = |P|.

Examples:

(a) Q(<0>) = {n}.

(b) Q(<0,1>) = {2n+1}.

(c) Q(<0,0>) = {2n}.

(d) Q(<0,1,0>) = {6n+1}.

(e) Q(<0,1,2>) = {6n+5}.

(f) Q(<0,0,2>) = {6n+4}.

Definition 5.2.18.

(a) I(P) = the least n ∈ Q(P).

(b) q(P) = {I(P)}.

Fact 5.2.19.

(a) If $P = \langle n_1, \ldots, n_k \rangle$, then

$I(P) = \text{sum} \{ ((i-1)!) * n_i : 1 \leq i \leq k \}$

(b) $I(P) = I(R)$ iff $P = R$, P 0-extends R or R 0-extends P.

(c) For every n, there's a P such that $n = I(P)$.

(d) For every n, there are infinitely many P such that $n = I(P)$.

(e) For every n, there's a P such that $n = I(R)$ iff $R = P$ or R 0-extends P.

Fact 5.2.20.

(a) At each depth, n, the $I(P)$ take on all and only values less than $n!$.

(b) If $I(P) = k$, then all nodes along the left-most branch descending from P also have value k. These are the only nodes below P with value k.

(c) Indeed, every natural number will be the value of all and only those nodes along the left-most branch descending from some node.

So, though for a given n there will be infinitely many nodes P such that $I(P) = n$, we can associate with each natural number a shortest (i.e. shallowest) node for which $I(P) = n$.

Definition 5.2.21. $\bar{X}(n)$ = the shortest P such that

$$I(P) = n.$$

Fact 5.2.22.

(a) $I(\bar{X}(n)) = n$

(b) P extends $\bar{X}(I(P))$

(c) $\bar{X}(\bar{\imath}(\bar{X}(n))) = \bar{X}(n)$

(d) $\bar{X}(I(P)) = P$ iff $P = \langle 0 \rangle$ or $P(|P|) \neq 0$.

(i.e. a node, P, will be the highest node with a certain value just in case P is not the leftmost immediate descendant of its parent.)

Note: Each of the points listed in Fact 5.2.13 hold for the sets Q(P) and q(P). That is to say, the Q(P) can be regarded as node sets and the q(P) as node atoms for any model $\underline{Q}(f)$. Notice, especially, that 5.2.13k holds.

We may finally define the embedding of $\underline{Q}(f)$ into $\underline{A}$:

Fact 5.2.23. If $x \in Q$, there is a unique $y \in B$ such that:

$$(P)(q(P) \subseteq x \leftrightarrow a(P) \subseteq y)$$

Proof:

If there is any such y, there is at most one, by fact 5.2.16. To show that there is such a y, suppose first that $x \in QC$. So there is some n such that:

$$x = x1 \vee \cdots \vee x.k$$

where each x.i is an (n!)-congruence class and hence a node-set in Q(f). So

$$x = Q(P1) \vee \cdots \vee Q(P.k).$$

Now, let

$$y = A(P1) \vee \cdots \vee A(P.k).$$

So $y \in B$ and:

q(P) c x

iff q(P) c Q(P.i) for some i

iff P extends P.i, for some i, by 5.2.13h

iff a(P) c A(P.i) for some i

iff A(P) c y.

If x is not a quasi-congruence class, then

$$x = (x' \vee D1(x)) - D2(x),$$

where x' is a quasi-congruence class. So, let y' be the element of B corresponding to x', as described above, and let

$$y = (y' \vee \{a(P) : q(P) \subseteq D1(x)\})$$
$$- \{a(P) : q(P) \subseteq D2(x)\}).$$

Definition 5.2.24. If $x \in Q$, let g(x) be the $y \in B$ such that:

$$(P)(a(P) \subseteq y \leftrightarrow q(P) \subseteq x)$$

We shall call g the <u>nodal embedding</u> of Q into A.

Fact 5.2.25.

(a) If $x \in Q$, then

$$g(x) = (g(C(x)) \vee g(D1(x))) - g(D2(x))$$

(b) g is one-one.

(c) g maps Q onto B.

Proof:

(a) Immediate from the proof of 5.2.23.

(b) Suppose $g(x) = y = g(x')$. Then

$$(P)(q(P) \subseteq x \leftrightarrow a(P) \subseteq y \leftrightarrow q(P) \subseteq x')$$

But every integer is I(P), for some P, so $x = x'$.

(c) Obvious.

Theorem 5.2.26. If g is the nodal embedding of Q into $\underline{A}$, then g is an isomorphism of $\underline{Q}(f)$ onto B.

Proof:

(1) $\underline{Q}(f) \models (x \subseteq y)$

iff $(P)(q(P) \subseteq x \rightarrow q(P) \subseteq y)$

iff $(P)(a(P) \subseteq g(x) \rightarrow a(P) \subseteq g(y))$, by 5.2.24

iff $\underline{B} \models (g(x) \subseteq g(y))$, by 5.2.16.

(2) g can be shown to preserve $\emptyset$, I, unions, intersections, relative complements, and proper subsets as in (1).

(3) $\underline{Q}(f) \models (x \approx y)$

iff $\theta(f,x) = \theta(f,y)$

iff $nc(x,n) - (cd(D1(x)) - cd(D2(x)))$

$= nc(y,n) - (cd(D1(y)) - cd(D2(y)))$

where $nc(z,n)$ = the number of charmed node sets of level n contained in $C(z)$

and n = the least k such that x and y are both unions of node sets of depth k.

iff $nc(g(x),n) - (cd(g(D1(x)) - cd(g(D2(x)))$

$= nc(g(x),n) - (cd(g(D1(x)) - cd(g(D2(x)))$

since $Q(P)$ is charmed iff $A(P)$ is charmed and g preserves boolean relations

iff $\underline{A} \models (g(x) \approx g(y))$

iff $\underline{B} \models (g(x) \approx g(y))$, since $\underline{B} \subseteq \underline{A}$

(4) $\underline{Q}(f) \models (x < y)$

iff $\underline{Q}(f) \models (x \approx x' \ \& \ x' \subset y)$ for some $x' \in Q$

since $\underline{Q}(f) \models$ REP<

iff $A \models (g(x) \approx g(x') \;\&\; g(x') \subset g(y))$, by 2,3

iff $\underline{A} \models (g(x) < g(y))$, since $\underline{A} \models$ SUBSET, INDISC$\approx$

iff $\underline{B} \models (g(x) < g(y))$, since $\underline{B} \subseteq \underline{A}$

(5) $\underline{Q}(f) \models SUM(x,y,z)$

iff $\underline{Q}(f) \models (x' \vee y' = z \;\&\; x \approx x' \;\&\; y \approx y'$

$\&\; x' \wedge y' = \emptyset)$, for some $x', y' \in Q$

since $\underline{Q}(f) \models$ DEF+

iff $\underline{B} \models (g(x') \vee g(y') = g(x) \;\&\; g(x) \approx g(x')$

$\&\; g(y) \approx g(y') \;\&\; g(x') \wedge g(y') = \emptyset)$, by 2,3

iff $\underline{A} \models (g(x') \vee g(y') = g(x) \;\&\; g(x) \approx g(x')$

$\&\; g(y) \approx g(y') \;\&\; g(x') \wedge g(y') = \emptyset)$

iff $\underline{A} \models SUM(g(x),g(y),g(z))$, since $\underline{A} \models$ DEF+

iff $\underline{B} \models SUM(g(x),g(y),g(z))$, since $\underline{B} \subseteq \underline{A}$

So, each model of CAI(f) has a submodel isomorphic to $\underline{Q}(f)$, and we may conclude:

Corrolary 5.2.27. If f is total and congruous, then CAI(f) has a prime model.

5.3 SUMMARY

We can now draw our final conclusions about CA, CS, and their completions.

Theorem 5.3.1. If f is a total, congruous remainder function, then CAI(f) is complete and consistent.

Proof:

CAI(f) is consistent because f is congruous, by 3.3.12c. Since CAI(f) is model complete, by 5.1.6, and has a prime model, by 5.2.27, the prime model test, A.3.3, applies. So CAI(f) is complete.

Corollary 5.3.2. If T is a completion of CAI, then T $\equiv$ CAI(f) for some congruous f.

Proof:

For each n > 0, T |- MOD{n,i} for exactly one i, 0 $\leq$ i < n: Since T is complete, T |- MOD{n,i} or T |- ¬MOD{n,i} for each such i. But if T |- (¬MOD{n,0} & ... & ¬MOD{n,n-1}), then T is inconsistent, since T |- CAI and CAI |- DIV{n}. Hence T |- MOD{n,i} for at least one i, 0 $\leq$ i < n.

But suppose T |- MOD{n,i} and T |- MOD{n,j}, where 0 $\leq$ i $\neq$ j < n. Again, T would be inconsistent, for CA |- MOD{n,i} -> ¬ MOD{n,j} (see lemma 3.3.11).

So, let f(n) = m iff T |- MOD{n,m}. Then T |- CAI(f) and, since CAI(f) is complete, CAI(f) |- T.

Corollary 5.3.3.

(a) Every completion of CA is consistent with CS.

(b) CA $\equiv$ CS.

Proof:

(a) Follows from 5.3.2 and 3.3.13.

(b) Follows from (a) and 3.3.1b, given that CS |- CA.

Theorem 5.3.4.

(a) For n > 0, CS;EXACTLY{n} is decidable.

(b) CS is decidable.

Proof:

(a) CS;EXACTLY{n} |- ∅ iff <u>F</u>.n |= ∅.

But <u>F</u>.n is a finite model.

(b) To determine whether CS |- ∅, alternate between generating theorems of CA and testing whether <u>F</u>.n |= ¬∅.

Theorem 5.3.5.

(a) CSI has 2**$\bar{w}$ completions.

(b) For total f, CSI(f) is decidable iff f is decidable.

Proof:

(a) There are 2**ω remainder functions whose domain is the set of prime numbers. Each such function is congruous, so each corresponds to a consistent extension of CSI. By Lindenbaum's lemma each of these extensions has a consistent and complete extension.

(b->) If f is decidable, then CAI(f) is recursively enumerable. But CAI(f) is complete, so it is decidable.

(b<-) To calculate f(n), see which sentence MOD{n,m} is in CSI(f).

Theorem 5.3.6. There is no sentence ø such that T = CA;ø is consistent and T only has infinite models.

Proof: If ø is true only in infinite models of CA, then ¬ø is true in all finite models of CA, so ¬ø ∈ CS. But CA ≡ CS, so CA;ø is inconsistent.

Theorem 5.3.7. CS is not finitely axiomatizable.

Proof: Suppose CS |- ø. So CA |- ø and, by compactness, (BASIC \/ T) |- ø for some finite set of

ADIV principles, T = {ADIV(n) : n ∈ J}. Let K = {n : every prime factor of n is a member of J}.

Let $\underline{A}$ be a model whose domain, A, is \/ {Q(k) : k ∈ K}, in which size relations are determined in accordance with the size function θ defined in 3.2.26. We claim the following without proof:

(1) $\underline{A} \models$ BASIC

(2) $\underline{A} \models$ ADIV(j) for all j ∈ J.

(3) $\underline{A} \not\models$ ADIV(k) for any k ∈ K.

By (1) and (2) $\underline{A} \models$ (BASIC \/ T), so $\underline{A} \models \emptyset$. But by (3), $\underline{A} \not\models$ CS. Hence $\emptyset \not\models$ CS.

6 SETS OF NATURAL NUMBERS

CS has finite standard models since it consists of sentences true in all such models. CS has infinite models $\underline{Q}(n)$ and $\underline{Q}(f)$. In this chapter we will show that CS has infinite standard models over P(N).

An ordering of P(N) that satisfies CS will not necessarily appear reasonable. For example, some such orderings say that there are fewer even numbers than there are prime numbers (see below, 6.2.13). To rule out such anomalies, we introduce a principle, OUTPACING, in section 1. OUTPACING mentions the natural ordering of N and applies only to subsets of N. Section 2 establishes that OUTPACING can be satisfied jointly with any consistent extension of CS is a model whose domain is P(N). So CS; OUTPACING does not fix the size relations over P(N).

6.1 THE OUTPACING PRINCIPLE.

Throughout this chapter, 'x' and 'y' will range over P(N).

Definition 6.1.1. x <u>outpaces</u> y iff

$$(En)(m)(m > n \rightarrow x\{m\} > y\{m\})$$

That is to say, x outpaces y just in case: given any sufficiently large initial segment of N, the restriction of x to that initial segment is larger than the corresponding restriction of y. Notice that the size comparison between the two restricted sets will always agree with the comparison of their normal cardinalities since all initial segments of N are finite.

We employ this notion to state a <u>sufficient</u> condition for one set of natural numbers to be larger than another:

OUTPACING: If x outpaces y, then x > y.

The general motivation behind this principle should be familiar. We extrapolate from well-understood finite cases to puzzling infinite cases. But we should also emphasize, again, that this extrapolation cannot be done in any straightforward, mechanical way without risking contradiction. We cannot, for example, strengthen the conditional to a biconditional, thus:

(1) x > y iff x outpaces y

This revised principle conflicts with CS, for outpacing is not a quasi-linear ordering. For example, neither {2n} nor {2n+1} outpaces the other since each initial segment {0,...,2n+1} of N contains n evens and n odds. But the two are discernible under outpacing, since {2n} outpaces {2n+2} while {2n+1} does not. There is another point that underlines the need for care in extrapolating from finite cases to infinite cases: we cannot just use (2):

(2) If, given any finite subset z of N, x restricted to z is larger than y restricted to z, then x > y.

Though (2) is true, its antecedent is only satified when y is a proper subset of x.

So, there are many statements that assert of infinite cases what is true of finite cases. Some of these conflict with one another. Others are too weak to be helpful. It is doubtful whether there is any mechanical way to decide which of these statements are true. The best we can do is propose plausible theories, determine whether they are consistent, and see how far they go.

Definition 6.1.0. A is an outpacing model iff

A = P(N),

A |= BASIC, and

A |= OUTPACING.

(There is a slight difficulty in saying that an interpretation of L(C<) "satisfies OUTPACING". Since OUTPACING involves the smaller-than relation over N, it cannot be expressed in L(C<). We shall finesse this problem by regarding OUTPACING as the (very large) set of sentences

{ b < a : a outpaces b}.

Fact 6.1.1. Every outpacing model satisfies the following:

(a) {2n} > {3n}

(b) {3n} > {4n} > {5n} > ...

(c) If k > 0, then {k*n} > {n**2}

Proof:

(a) {2n} has at least (k-1)/2 members less than or equal to k, for any given k. {3n} has at most (k/3) + 1 such members. If k > 4, then (k-1)/2 > (k/3) + 1.

(b) Similar to (a).

(c) Note that if m = k**2, both {k*n} and {n**2} have exactly k members < m. For m > k*(k+1), N{m} will have more members in {k*n} than in {n**2}.

Theorem 6.1.2. Every outpacing models satisfies

(a) $\{2n\} \geq \{2n+1\}$

(b) $\{2n+1\} \geq \{2n+2\}$

(c) If $\{2n\} > \{2n+1\}$, then $\{2n+1\} \approx \{2n+2\}$.

Proof:

(a) By TRICH, it is sufficient to show that $\{2n\}$ is not smaller than $\{2n+1\}$. If it were, then, by REP<, there is a y such that $y \approx \{2n\}$ and $y \subset \{2n+1\}$. But any proper subset of $\{2n+1\}$ is outpaced by, and hence smaller than, $\{2n\}$: Let k be the least odd number not in y. Then $\{2n\}$ leads y at (k+1) and y never catches up. So there is no y such that $\{2n\} \approx y \subset \{2n+1\}$.

(b) Similar to proof of (a).

(c) Note that $\{2n\} = \{2n+2\} \vee \{0\}$ and that BASIC $\vdash$ (*).

$$(*)\ (y \subset x \ \&\ z \leq y \ \&\ ATOM(z') \ \&\ z \subset x \ \&\ x = z \vee z') \rightarrow y \approx z.$$

Letting $x = \{2n\}$, $y = \{2n+1\}$, $z = \{2n+2\}$ and applying (*), $\{2n+1\} \approx \{2n+2\}$.

The two alternatives left open in 6.1.2 correspond to the possibilities that N may be odd or even: if $\{2n\} \approx \{2n+1\}$, then N is even; if $\{2n+1\} \approx \{2n+2\}$, then N is odd. In section 6.2 we show that both of these possibilities can be realized in standard models over P(N). Here, we generalize 6.1.2 to similar cases, including other congruence classes.

Definition 6.1.3. $\langle x,y\rangle$ is an <u>alternating pair</u> iff x and y are infinite and for all $i > 0$, $x(i) < y(i) < x(i+1)$.

Fact 6.1.4. If $\langle x,y\rangle$ is an alternating pair, then in any outpacing model:

$$x \approx y \text{ or } x > y \approx (x - x(1)).$$

Proof: The argument for 6.1.2 applies here since the only facts about $\{2n\}$ and $\{2n+1\}$ used hold by virtue of these sets forming an alternating pair.

Theorem 6.1.5.

(a) If $0 \leq i < j < k$, then $\langle\{k*\underline{n} + i\},\{k*\underline{n} + j\}\rangle$ is an alternating pair.

(b) For a given $k > 0$, let $A.i = \{k*\underline{n} + i\}$ for each $i < k$. Then there is a p, $0 < p \leq k$ such that

(i) If $i < j < p$, then $A.i \approx A.j$,

(ii) If $p \neq k$, $A.0 \approx A.p \cup \{0\}$, and

(iii) If $p \leq i < k$, then $A.i \approx A.p$.

(See example below).

Proof:

(a) $A.i(n) = k*n + i$, $A.j(n) = k*n + j$, $A.i(n+1) = k*n + i + k$, and $i < j < k + i$.

(b) If $A.0 > A.i$ for some i, let p be the least such i; otherwise let $p = k$.

(i) For $0 \leq i < p$, $\langle A0, A.i \rangle$ is an alternating pair so $A0 > A.i$ or $A0 \simeq A.i$ by 6.1.4. But $A0 \leq A.i$ by the selection of p, so $A0 \simeq A.i$. (i) follows by TRANS-.

(ii) Immediate from 6.1.4 since $\langle A.0, A.p \rangle$ is an alternating pair and $A.0 > A.p$.

(iii) $A.0 > A.p \geq A.i$ if $i \geq p$. So $A.0 > A.i$. Hence $A.i \simeq A.0 - \{0\} \simeq A.p$. So $A.i \simeq A.p$.

Example: Let $k = 4$, so $A.i = \{4n+i\}$ for $i = 0,1,2,3$. Then one of the following situations obtains:

(1) $A.0 \simeq A.1 \simeq A.2 \simeq A.3 > \{4n+4\}$

(2) $A.0 > A.1 \simeq A.2 \simeq A.3 \simeq \{4n+4\}$

(3) $A.0 \simeq A.1 > A.2 \simeq A.3 \simeq \{4n+4\}$

(4) $A.0 \simeq A.1 \simeq A.2 > A.3 \simeq \{4n+4\}$

6.2 MODELS OF CS AND OUTPACING.

In this section, we construct models of CS over P(N) that satisfy OUTPACING.

Outpacing models will be constructed out of finite models of CS by a technique which is very much like the "ultraproduct construction" common in model theory, though the application here demands some important differences.

Definition 6.2.1.

(a) L(N) is the first order language which results from adding to L(C<), as individual constants, a name for each subset of N.

(b) $\underline{A}$.n is the finite standard interpretation of L(N) over P(N{n}) in which

$\underline{A}.n(\underline{a}) = a \wedge N\{n\} = a\{n\}$ for each $a \subseteq N$.

Definition 6.2.2. (cf. Monk, Def. 18.15, p.318) If X is a set and $F \subseteq P(X)$, then

(a) F has the <u>finite intersection property</u> iff the intersection of any finite subset of F is non-empty.

(b) F is a <u>filter</u> over X iff

(i) $F \neq \emptyset$

(ii) If $a \in F$ and $a \subseteq b$, then $b \in F$, and

(iii) If $a \in F$ and $b \in F$, then $a \wedge b \in F$.

(c) F is an <u>ultrafilter</u> over X iff

(i) F is a filter over X,

(ii) $X \in F$, and

(iii) if $Y \subseteq X$, then $Y \in F$ or $(X - Y) \in F$.

(d) An ultrafilter, F, over X is <u>principal</u> iff there is some $x \in F$ such that $F = \{a \subseteq X : x \in a\}$; otherwise F is <u>non-principal</u>.

Fact 6.2.3.

(a) A non-principal ultrafilter contains no finite sets. (See Bell and Slomson, Ch.6, lemma 1.3, p.108).

(b) A non-principal ultrafilter over X contains all cofinite subsets of X.

(c) If $F \subseteq P(X)$ and F has the finite intersection property, then there is an ultrafilter over X which includes F. (See Monk, Prop. 18.18, p.319).

(d) If $Y \subseteq X$ and Y is infinite, then there is a non-principal ultrafilter over X which contains Y.

Definition 6.2.4. If F is an ultrafilter over N, then $\underline{A}(F)$ is the interpretation of $L(C<)$ in which

(i) $A(F) = P(N)$

(ii) $\underline{A}(F) \models \underline{a} < \underline{b}$ iff $\{k : a[k] < b[k]\} \in F$, and similarly for other size predicates.

(iii) Boolean symbols receive the usual interpretation.

Our main result is that if F is non-principal, then $\underline{A}(F)$ is an outpacing model.

Theorem 6.2.5. If F is a non-principal ultrafilter over N and $\underline{A} = \underline{A}(F)$, then

(a) If t is a term of L(N), then

A.k(t) = A(t) /\ A.k = A(t){n}

(b) If ø is a quantifier free formula of L(N), then

$\underline{A}$ |= ø iff {k: $\underline{A}$.k |= ø} ∈ F

(c) If ø is a universal formula of L(N) and

$\underline{A}$.k |= ø, for every k, then $\underline{A}$ |= ø.

(d) $\underline{A}$ |= REP<.

(e) $\underline{A}$ |= BASIC.

(f) $\underline{A}$ |= ADIV{n}, for every n.

Proof:

(a) By induction on the structure of t:

(1) If t is a constant, t = a for a c N.

So A.k(t) = a{k} by 6.2.1b.

(2) If t = 't1 \/ t2',

A.k(t) = A.k(t1) \/ A.k(t2)

= (A(t1) /\ A.k) \/ (A(t2) /\ A.k)

= (A(t1) \/ A(t2)) /\ A.k

The proofs for intersections and relative complements are similar.

(b) By induction on the structure of ø:

(1) If ø = 'a c b', then $\underline{A}$ |= ø iff a c b.

If a c b, then there is a k ∈ b but not ∈ a.

So if $n > k$, $a\{n\} \subset b\{n\}$. Hence,

$\{n: A.n \models \phi\}$ is cofinite and, by 4.2.3b, in F.

Conversely, if $\{n: a\{n\} \subset b\{n\}\} \in F$, then it is infinite. So there cannot be a k in a but not in b; otherwise a{n} would not be included in b{n} for any n greater than k. So $a \subset b$. But, clearly $a \neq b$, so a is a proper subset of b.

(2) If ϕ = 'a = b', then

$\underline{A} \models \phi$ iff $a = b$

iff $a\{n\} = b\{n\}$ for all n

iff $\{n: A.n \models \phi\} = N$

iff $\{n: A.n \models a = b\} \in F$

since if $A.n \not\models$ 'a = b', and $k > n$,

then $A.k \not\models$ 'a = b'.

(3) $\underline{A} \models (a < b)$ iff $\{k: \underline{A}.k \models (a < b)\} \in F$.

(immediate from 6.2.4b.)

(4) If ϕ is non-atomic,

$\underline{A} \models (\phi 1 \,\&\, \phi 2)$ iff $\underline{A} \models \phi 1$ and $\underline{A} \models \phi 2$

iff $\{k: \underline{A}.k \models \phi 1\} \in F$

and $\{k: \underline{A}.k \models \phi 2\} \in F$

iff $\{k: \underline{A}.k \models (\phi 1 \,\&\, \phi 2)\} \in F$

$\underline{A} \models \neg\phi 1$ iff not$(\underline{A} \models \phi 1)$

iff $\{k: \underline{A}.k \models \phi 1\}$ is not in F

iff $\{k: \underline{A}.k \models \neg\phi 1\} \in F$

since F is an ultrafilter.

(c) Suppose $\underline{A}$.k |= (x)ø(x), for all k

then $\underline{A}$.k |= ø(a), for all a, for all k

so $\underline{A}$.k |= ø(a), for all k, for all a

so $\underline{A}$ |= ø(a), for all a, by (b)

so $\underline{A}$ |= (x)ø(x)

(d) Suppose $\underline{A}$ |= (a < b).

We want to construct a', a subset of b, for which $\underline{A}$ |= (a = a').

Let K = {k: a[k] < b[k]}; so K ∈ F.

Let K = {k1,....,k.i,...}, where the k.i's are in strictly increasing order.

Let a0 = ø

a.(i+1) = a.i \/ {the n greatest members of b[k.(i+1)] which are not in a.i},

where n = cd(a[k.(i+1)]) - cd(a.i)

let a' = \/ {a.i}

Then a' $\subseteq$ b, since each a.i draws its new members from b.

Claim: If k ∈ K, then a'[k] = a[k].

Hence: $\underline{A}$ |= a' = a, since they are the same size over some set which contains K, and is, thus, in F.

(e) Immediate from (c) and (d): BASIC is equivalent to a set of universal sentences, together with ATOM and REP<. ($\underline{A}$ |= ATOM because it contains all singletons of natural numbers.)

(f) Given n > 0, x, an infinite subset of N, and i < n, let x.i = {x(k*n + i - 1): k ∈ N}.

So, the n sets, x.i, partition x. Furthermore, these form an "alternating n-tuple", in the manner of the congruence classes modulo n (see theorem 6.1.5). So, as in 6.1.5, these sets are approximately equal in size and $\underline{A}$ |= Div{n}(x).

Corollary 6.2.6. If F is a non-principal ultrafilter over N, then

(a) $\underline{A}$(F) |= CAI, and

(b) $\underline{A}$(F) |= CSI

Proof:

(a) Immediate from 6.2.5e and 6.2.5f.

(b) Immediate from (a) and 5.3.3b.

Perhaps a remark is in order: The proof of 6.2.6 is modeled on the usual "ultraproduct construction", but isn't quite the same. In the usual construction (see Bell and Slomson, pp. 87-92), a model is built by first taking the product of all the factors (in this case, the A.k), which results in a domain whose elements are functions from the index-set (N, here) to elements of the factors. These functions are then gathered together into equivalence classes (by virtue of agreeing "almost everywhere", i.e. on some member of the filter) and the reduced ultraproduct is defined by interpreting the language over these equivalence classes. The model so constructed, which we'll call "Pr(A.k)/F", has the handy property that it satisfies any formula which is satisfied by "almost all" factors, and certainly any formula which is satisfied in all of the factors. This is handy because, given that each of the $\underline{A}$.k |= CS, we can immediately conclude that Pr(A.k)/F also satisfies CS.

Unfortunately, Pr(A.k)/F isn't the model we wanted: for its elements are not subsets of N, but equivalence classes of functions from N to finite subsets of N. There is, indeed, a "natural mapping" from subsets of N to such elements, and this mapping would allow us to identify a model over P(N) as a submodel of Pr(A.k)/F; but only a submodel. So, had we constructed the reduced ultraproduct, we would have then been able to infer that the part of that model which held our interest satisfied all universal formulas of CS; we still would have had to resort to special means to show that the non-universal formulas were likewise satisfied.

Fortunately, these special means were available; the only non-universal axioms of CAI could be verified in the constructed model more or less directly, and the completeness proof of the last chapter, allowed us to infer that all formulas true in all of the factors are true in the model A(F), after all.

Notice that, given any F, there will be many cases where $\underline{A}$(F) |= $\underline{a} <_r \underline{b}$ even though b doesn't outpace a. This will happen whenever $\{k : \underline{A}.k \models a\{k\} < b\{k\}\} \in F$ but isn't cofinite. Consider a familiar example: Let $a = \{2n+1\}$ and let $b = \{2n\}$. Then $a\{k\} < b\{k\}$ iff $k \in \{2n\}$, for if we count the even numbers and the odd numbers up to some even number, there will always be one more even and if we count up to some odd number, there will be the same number of evens and odds. $\{2n\}$ is neither finite nor cofinite, so it may or may not be in F. If $\{2n\} \in F$, then $\underline{A}(F) \models (\{2n+1\} < \{2n\})$. Otherwise $\{2n+1\} \in F$, so $\underline{A}(F) \models (\{2n\} = \{2n+1\})$.

The construction of a model $\underline{A}(F)$ from any non-principal ultrafilter, F, suggests that there are many outpacing models unless different ultrafilters yield the same model. We will first show that this qualification is not needed.

> Theorem 6.2.7. If F1 and F2 are distinct non-principal ultrafilters over N, then $\underline{A}(F1) \not\cong \underline{A}(F2)$. (Proof below.)

To show this, we will show that the presence of a set in an ultrafilter makes a direct, "personalized" contribution to the model $\underline{A}(F)$. Putting this in another way, there is a set of decisions that must be made in constructing an outpacing model; each decision may go either way, though the decisions are not independent of one another. Furthermore, each decision is made for a model $\underline{A}(F)$ by the presence or absence of a particular set in F.

Definition 6.2.8. If $x \subseteq N$, then $x+ = \{i+1 : i \in x\}$.

A pair of sets $\langle x, x+\rangle$ can sometimes be an alternating pair, but this is not always the case.

Fact 6.2.9. $\langle x, x+\rangle$ is an alternating pair iff x is infinite and there is no $n \in x$ such that $(n+1) \in x$, i.e. no two consecutive numbers are in x.

Nevertheless, pairs $\langle x, x+\rangle$ are like alternating pairs in the following way:

Lemma 6.2.10. If x is infinite, $x1 = (x - x+)$, $x2 = (x+ - x)$, F is a non-principal ultrafilter, and $\underline{A} = A(F)$, then

(a) $\langle x1, x2\rangle$ is an alternating pair.

(b) $\underline{A} \models (x \approx x+)$ or $\underline{A} \models (x > x+ \approx (x - x(1)))$.

(c) $\underline{A} \models (x > x+)$ iff $x \in F$.

Proof:

(a) Let a <u>run</u> of x be a maximal consecutive subset of x. (So $\{2n\}$ has only one-membered runs, while $(N - \{10n\})$ has only nine-membered runs.) So $x1(n)$ is the first element in the $\underline{n}$th run of x and $x2(n)$ is the first element after the $\underline{n}$th run of x.

(b) We know from (a) that $x2 \approx x1$ or $x2 \approx (x1 - x1(1))$.

But the disjoint union of (x /\ x+) with x1 or x2, respectively, yields x or x+. So (b) follows from DISJU. (c) We need only prove (*), which we shall do

(*) cd(x{n}) > cd(x+{n}) iff n ∈ x

informally: x{n} first becomes greater than x+{n} where n = x(1) since x(1) is not ∈ x+ because (x(1) - 1) not ∈ x. Throughout the first run of x, x retains its lead, losing this lead at the least n, n not ∈ x (for (n-1) ∈ x, so n ∈ x+). This pattern repeats during successive runs of x.

We can now prove our theorem.

Proof of 6.2.7. Without loss of generality, we can suppose there is a set, x, such that x ∈ F1 and x not ∈ F2. By 6.2.10c, $\underline{A}$(F1) |= x > x+ and $\underline{A}$(F2) |= x ≈ x+.

Theorem 6.2.7 allows us to improve upon some previous results. For example, we can show that either of the alternatives in 6.2.10b can be obtained for any alternating pair.

Theorem 6.2.11.

(a) If neither x nor y outpaces the other, then there is an ultrafilter, F, such that $\underline{A}$(F) |= x ≈ y.

(b) If <x,y> is an alternating pair, then there is an ultrafilter, F, such that $\underline{A}$(F) |= x > y.

Proof:

(a) Let $J = \{k : cd(x\{k\}) = cd(y\{k\})\}$. Since neither x nor y outpaces the other, J is infinite. By 6.2.3d, let F be a non-principal ultrafilter which contains J. Then $\underline{A}(F) \models x \approx y$.

(b) If $\langle x,y \rangle$ is an alternating pair, so is $\langle y,(x - x(1)) \rangle$. So, by (a) there is an F such that $\underline{A}(F) \models (y \approx (x - x(1)))$. But then $\underline{A}(F) \models (x > y)$.

Theorem 6.2.12. Every infinite completion of CS has an outpacing model.

Proof: Recall that every infinite completion of CS is equivalent to CSI(f) for some total and congruous remainder function, f. (See 3.6.2)

Given such an f, let $G.k = \{k*\underline{n} + f(k) + 1)\}$ for each $k > 0$. Note that (*) holds for each k:

(*) $G.k = \{n : \underline{A}.n \models MOD\{k,f(k)\}\}$

Let $G = \{G.k : k > 0\}$.

The intersection of any finite subset, H, of G is infinite. If H is a finite subset of G, then $H = \{G.k : k \in J\}$, where J is some finite subset of $N+$. So $H = \{n+1 :$ if $k \in J$, then $n \equiv f(k) \bmod k\}$. But f is congruous, so the restriction of f to the finite domain J has infinitely many solutions (see 3.3.8a).

Since the intersection of any finite subset of G is infinite, there is a non-principal ultrafilter F such that G $\subseteq$ F. By (*), $\underline{A}$(F) |= MOD{k,f(k)}, so $\underline{A}$(F) |= CSI(f).

On the basis of 6.2.11 we noted that there are even and odd outpacing models; we can now extend that observation to moduli other than 2. More specifically, note that all of the possibilities listed in 6.1.5 for the relative sizes of the k-congruence classes are, in fact, obtainable in outpacing models.

This section has explored the existence and variety of outpacing models. Three comments are in order before we turn to the common structure of outpacing models.

First, even if T is an infinite completion of CS, there is no unique outpacing model which satisfies T. This would be true only if fixing the relative sizes of congruence classes determined whether x ≃ x+ or x > x+ for every x $\subseteq$ N (see 6.2.10c). That all such choices are not determined by a remainder theory can be seen intuitively, perhaps, by considering x = {n**2}: Any finite set of congruences has infinitely many solutions that are squares and infinitely many that aren't; so whether $x \in F$ is an independent choice. Note also that there are only (2**ω) remainder functions while there are 2**(2**ω) non-principal ultrafilters over N, each yielding a different outpacing model (see Bell and Slomson, Ch.6, Theorem 1.5).

Second, it is not clear whether every outpacing model can be obtained by the construction of 6.2.4. Lemma 6.2.10c may suggest that any outpacing model, $\underline{A}$, is $\underline{A}$(F) for F = F.A = {x : $\underline{A}$ |= x > x+}, but it shouldn't. To establish that $\underline{A}$ = $\underline{A}$(F.A), both (1) and (2) are necessary.

(1) If $\underline{A}$ is an outpacing model, then F.$\underline{A}$ is a non-principal ultrafilter.

(2) If F.$\underline{A}$ = F.$\underline{B}$, then $\underline{A}$ = $\underline{B}$.

I have not been able to prove (1) or (2). If (1) is false, then clearly $\underline{A} \neq \underline{A}$(F.$\underline{A}$). But even if (1) is true, two outpacing models may agree about all pairs <x,x+> but disagree elsewhere. At most one of them is obtainable by our construction. So (1) and (2) are open problems.

Finally, though it's probably extraneous to show that OUTPACING is independent, we will do so.

Theorem 6.2.13. There are standard models of CS over P(N) which do not satisfy OUTPACING.

Proof: Suppose that $\prec$ is a linear ordering of N, under which N forms an ω sequence. Then we could define x $\prec$-outpaces y, thus:

$$En(m)(n \prec m \rightarrow cd(\{k : k \in x \ \& \ k \prec m\}) > cd(\{k : k \in y \ \& \ k \prec m\}))$$

and OUTPACING$\prec$, thus:

If x $\prec$-outpaces y, then x > y.

Modifying 6.2.4, we could produce standard models of CS over P(N) which satisfy OUTPACING$\prec$ and these won't, in general, satisfy OUTPACING.

Suppose, for example, that $\prec$ is the ordering:

p(1), q(1), ..., p(k), q(k),....

where p is the set of primes and q is its complement. In any OUTPACING$\prec$ model, p and q will be nearly the same size, and the evens and odds are in outpacing models. But the evens are much smaller than q, so p > {2n} and OUTPACING is false in OUTPACING$\succ$ models.

6.3 DENSITY AND SIZE

When number theorists talk about the sizes of sets of natural numbers, they don't content themselves with speaking of the (Cantorian) cardinalities of these sets. Since they often want to compare infinite subsets of N, they need a more discriminating notion.

One notion they use is "asymptotic density". The aysmptotic density of a set, x, of natural numbers is the limit, if there is one, of cd(x{n})/n as n grows. For example, the asymptotic density of {2n} is 1/2. From now on, we shall use the term "density" instead of "asymptotic density".

In this section, we compare the ordering of P(N) given by density to the orderings given by CS and OUTPACING.

Definition 6.3.1.

(a) fr(x,i) = cd(x{i})/i, the fraction of numbers less than or equal to i that are members of x.

(b) If $x \subseteq y \neq \emptyset$, then p(x,y), <u>the density of</u> x <u>in</u> y, is the limit, if it exists, of

fr(x,i)/fr(y,i)

as i goes from y(1) to infinity. That is,

p(x,y) = r iff (d)(d > 0 --> (En)(i)(i > n --> -d < fr(x,i)/fr(y,i) < d))

(c) <u>The density of</u> x, p(x), is the density of x in N, if x has a density in N.

(d) If $x \subset y \neq \emptyset$, then x <u>converges in</u> y, cvg(x,y), iff x has a density in y; otherwise x <u>diverges in</u> y, dvg(x,y). (e) x <u>converges</u>, cvg(x), iff x converges in N; x <u>diverges</u> iff x diverges in N.

Fact 6.3.2.

(a) If x is finite, p(x) = 0

(b) If x is cofinite, p(x) = 1

(c) p({2n}) = 1/2

p({4n},{2n}) = 1/2

p({4n}) = 1/4

(d) If cvg(x,y) and cvg(y,z), then cvg(x,z)

and p(x,z) = p(x,y)*p(y,z)

Fact 6.3.3.

(a) There are divergent sets.

(b) If $0 \leq r \leq 1$, there is a set with density r.

(c) If $0 \leq r \leq 1$ and y is infinite, then there is a set with density r in y.

Proof:

(a) Let x = {i: (En)(10**2n $\leq$ i < 10**(2n+1))},

so x contains all numbers between 0 and 9, between 100 and 999, between 10000 and 99999, and so forth. If n > 1, fr(x,10**2n) $\leq$ .1 and fr(x,10**(2n+1)) $\geq$ .9. So fr(x,k) cannot have a limit.

(b) Suppose r is given. Construct the set x as follows:

$$x0 = \emptyset$$

$$x.(i+1) = \begin{cases} x.i & \text{if } cd(x.i)/(i+1) \geq r \\ (x.i);(i+1) & \text{if } cd(x.i)/(i+1) < r \end{cases}$$

$$x = \bigvee \{x.i\}$$

(c) Modify the construction for (b) in the obvious ways.

Theorem 6.3.4. Suppose that both x and y converge in z. Then, if p(x,z) < p(y,z), y outpaces x.

Proof:

Let b = (p(y,z) - p(x,z))/3

n1 = the least n such that for all i > n,

$$-b < (fr(x,i)/fr(z,i)) - p(x,z) < b$$

n2 = the least n such that for all i > n,

$$-b < (fr(y,i)/fr(z,i)) - p(y,z) < b$$

Then, for any i > (n1 + n2),

$$fr(x,i)/fr(z,i) < p(x,z) + b$$

and $$fr(y,i)/fr(z,i) > p(y,z) - b$$

But p(x,z) + b < p(y,z) - b,

so f(x,i)/f(z,i) < f(y,i)/f(z,i)

so f(x,i) < f(y,i)

so cd(x{i}) < cd(y{i}).

We can use the relation between density and outpacing to draw conclusions about densities that have nothing to do with outpacing, as in theorem 6.3.5.

Theorem 6.3.5. If $p(x,z1) < p(y,z1)$ and both x and y converge in z2, then $p(x,z2) \leq p(y,z2)$.

Proof: Since $p(x,z1) < p(y,z1)$, y outpaces x. But if $p(x,z2) > p(y,z2)$, then x outpaces y, so $p(x,z2) \leq p(y,z2)$.

Notice that we cannot strengthen the consequent of 6.3.5 to say that $p(x,z2) < p(y,z2)$: let z1 be the set of primes, let x contain every third member of z1, and let y be (z1 - x). Then $p(x,z1) = 1/3$ and $p(y,z1) = 2/3$, but $p(x,N) = p(y,N) = 0$.

Theorem 6.3.4 implies that in any outpacing model, sets with distinct densities will have distinct sizes. Even if two sets have the same density, they will differ in size if they have different densities in some common set. So, from facts 6.3.3(b) and (c), we can begin to appreciate how precise an ordering outpacing models provide:

Fact 6.3.6. If A is an outpacing model, then
(a) there are uncountably many sizes of sets in A, and
(b) if $0 \leq r \leq 1$, then even among sets with density r, there are uncountably many sizes in A.

Proof:

(a) follows immediately from fact 6.3.3b and theorem 6.3.4. (b) Let x be an infinite set with density r, let y be an infinite subset of x, where $p(y,x) = 0$, and let $z = (x - y)$. There are uncountably many subsets of y with distinct densities in y, though $p(y1,x) = 0$, for any y1 c y.

Suppose, now, that y1 c y,

y2 c y, and

$p(y1,y) < p(y2,y)$.

Then $\underline{A}$ |= y1 < y2, by 6.3.4

so $\underline{A}$ |= (z \/ y1) < (z \/ y2), by DISJU

But p(z \/ y1) = p(z \/ y2) = p(x), since z was obtained by removing from x a set with density 0 relative to x.

6.3.1 The Convexity Problem.

It's tempting to infer from these results that the extremely fine ordering of sets by size (or, rather, any such ordering which is realized in an outpacing model) is both a refinement and a completion of the ordering suggested by density: a refinement because it preserves all differences in size which are captured by the notion of density, a completion because all sets are located in a single, linear ordering of sizes.

But the situation is really not so clear. It is evident that the size-ordering over P(N) in any outpacing model is a refinement of the ordering by cardinality: if x has a smaller cardinal number than y, then x is smaller than y, though two sets with the same cardinal number may have different sizes. We can regard the cardinality of a set in P(N) as determined by its size, though different sizes may yield the same cardinal number. We shall express this fact by saying that, at least when we focus on P(N), cardinality is a function of size. (Note: it is not at all clear that this is true in any power set.)

Now, we want to know whether the density of a set is a function of its size. It turns out that a negative answer is compatible with our theory (CS and OUTPACING), while an affirmative answer may or may not be consistent. First, we will give a more precise formulation of this problem; second, we will show that a negative answer is consistent; finally, we'll discuss the consistency of an affirmative answer. In passing, we'll explain why this is called "the convexity problem."

Consider (1) and (1a):

(1) If $x \simeq y$ and $p(x) = r$, then $p(y) = r$.

(1a) If $x \simeq y$ and $p(x) = r$ and cvg(y), then $p(y) = r$.

(1a) is an immediate consequence of theorem 6.3.4. For if y converges, there is some $r2 = p(y)$; if $r < r2$, then $x < y$ and if $r2 < r$, then $y < x$; but $y \simeq x$, so $r2 = r$.

So, the questionable part of (1) can be expressed as (2):

(2) If $x \simeq y$ and x converges, then y converges.

Theorem 6.3.4 insures that (1) if and only if (2). Recalling that sets may have the same density even if they differ in size, we may consider two additional formulations:

(3) If $p(x) = p(y)$ and $x < z < y$, then $p(z) = p(x)$.

(4) If $p(x) = p(y)$ and $x < z < y$, then z converges.

(3) and (4) are equivalent for the same reasons that (1) and (2) are equivalent.

Fact 6.3.7. An outpacing model satisfies (1) just in case it satisfies (3).

Proof:

(->) Suppose $x \doteq y$ and $p(x) = r$.

If $x = \emptyset$ or $x = N$, then $y = x$, so $p(y) = p(x)$

To apply (3), we need to find two sets, x1 and x2, such that

$$p(x1) = p(x2) = r = p(x)$$

and

$$x1 < y < x2.$$

Assuming that $x \neq \emptyset$ and $x \neq N$, let

$$x1 = x - \{a\}, \text{ for some } a \in x$$

and

$$x2 = x;b, \text{ for some } b \text{ not} \in x.$$

Then

$$x1 < x < x2, \text{ by SUBSET}$$

and

$$x1 < y < x2, \text{ by INDISC}^{\doteq}.$$

Since adding or removing a single element has no effect on the density of a set,

$$p(x1) = p(x) = p(x2).$$

So, by (3), $p(y) = r$.

(<-) Suppose $x < y < z$ and $p(x) = p(z)$.

By REP<, there are two sets, x' and y', where

$$x' \subset y' \subset z, \quad x \cong x', \quad y \cong y'$$

Since $x \cong x'$ and $p(x) = p(z)$, (1) guarantees that $p(x') = p(z)$. But then $p(z - x') = 0$, so $p(z - y') = 0$. But $y' = z - (z - y')$, so $p(y') = p(z)$. So, by (1), $p(y) = p(z)$.

The question at hand is called the "convexity problem" because of the formulation in (3). In geometry, a figure is convex if, given any two points in the figure, any point between them (i.e. on the line segment from one to the other) is also in the figure. Applying this notion in the obvious way to the size ordering, (3) says that the class of sets having a given density is convex. By theorems 6.3.4 and 6.3.7, (1), (2), and (4) say the same thing.

We regret that all we know about (3) is that it may be false:

Theorem 6.3.8. The negation of (1) is satisfied in some outpacing model.

Proof: Let x be a set with density r and let y be a divergent set which neither outpaces, nor is outpaced by, x. Let K be

$$\{ n : x[n] = y[n] \}$$

K is infinite, so K is a member of some non-principal ultrafilter, F.

$\underline{A}(F) \models (x \approx y)$, so (1) is false in $\underline{A}(F)$.

Open problem: Is (1) consistent with CS and OUTPACING?

APPENDIX

A.1 NOTATION

A.1.1 Predicate logic.

The formal theories discussed in this thesis are theories with standard formalization, in the sense of (Tarski, p.5). That is, they are formalized in first-order predicate logic with identity and function symbols. The following notation is used for the predicate calculus:

&	-	and
¦	-	or
->	-	if...then
<->	-	if and only if
¬	-	not
=	-	identical
≠	-	not identical
(Ex)	-	there is an x
(x)	-	for all x

Conjunctions and disjunctions of sets of sentences are represented by

& {ø : ...ø - - -}

and ¦ {ø : ...ø - - -}

Here, as elsewhere, we don't bother to use corner- quotes.

Subscripts on variables appear on the line, rather than below it, thus

x1, x2, ...

Where confusion might arise, in particular where the subscripts are not numerals, a dot is used, thus

x.1, x.2, x.i, x.n

Double subscripts are displayed in parentheses":

x(1,1), x(n,k)

A first order language is determined by its non-logical constant symbols, in the usual way. These may be predicates, individual constants, or function symbols. We shall specify languages, in the style of Chang and Keisler, as a set of constant symbols. In most cases, the ranks of the symbols will accord with their familiar uses and we will not bother to stipulate them.

We use parentheses in the conventional manner to indicate the argument places of predicates. So,

SUM(x,y,z)

is a three-place predicate.

A "schematic function" is a function whose range is a set of formulae. DIV, for example (see 3.2.8), is such a function. DIV maps natural numbers into sentences. The arguments of such functions are indicated within square brackets. So, for any n,

DIV[n]

refers to a particular DIV sentence: the one with n disjuncts.

Similarly, MOD is a 2-argument schematic function and for every pair of natural numbers, n and m, there is a MOD sentence, MOD{n,m}.

Schematic functions may also have predicates as values. In this case, the notation for schematic functions and the notation for predicate arguments are combined. For example, Div is such a function and, for any n,

Div{n}(x)

is a 1-place Div predicate.

A.1.2 Set-theory.

For first-order languages with boolean operations and predicates, we use the following notation:

I - the universe

∅ - the null element

x \/ y - the union of x and y

x /\ y - the intersection of x and y

x ⊆ y - x is a subset of y

x ⊂ y - x is a proper subset of y

x - y - the relative complement of y, in x

These symbols are also used for set-theoretic relations, outside of first-order languages. In addition, we use the following:

x ∈ y - x is a member of y

P(x) - the power set of x

{x: F(x)} - the set of x which are F

\/ x - the union of all members of x

/\ x - the intersection of all members of x

x;a - the union of x and {a} (from Enderton)

<x1,...,x.n> - the n-tuple whose i.th member is x.i

N - the set of natural numbers: {0,1,...}

N+ - N - {0}

Z - the set of integers

ω - the smallest infinite cardinal

2**ω - 2 to the ω

A.1.3 Arithmetic

For arithmetic, we use the following: ("i" through "n" range over the natural numbers and "I", "J", and "K" range over sets of natural numbers, unless otherwise specified.)

i+j - i plus j

i-j - i minus j

i*j - i times j

i**j - i to the j-th power

i! - i factorial

i | j - i divides j

gcd(i,j) - the greatest common divisor of i and j

n $\equiv$ m mod j - n is congruent to m modulo j

For sets of natural numbers, we use the following notations:

{...n...} = {k: (En)(k = ...n...)}

For example,

{2n} = the set of even numbers

{n**3} = the set of cubes

When more than one variable appears within the brackets, one is designated (by underlining) as that which corresponds to 'n' above. Thus,

{i*$\underline{n}$ + j} = the set of k congruent to j mod i, but

{i*n + $\underline{j}$} = the set of numbers $\geq$ i*n

Finally,

x{k} = {i: i $\in$ x and 0 $\leq$ i $\leq$ k}

and x(i) = the i-th element of x.

A.1.4 Greek letters

Upper case greek letters are not used. Lower case greek letters are approximated in the following ways:

ø - phi

θ - theta

õ - sigma

ã - alpha

b̃ - beta

d̃ - delta

t̃ - tau

ṽ - nu

x̃ - chi

w̃ - omega

p̃ - rho

A.2 MODEL THEORY

This section lists the model-theoretic notions and results assumed in the text and presents our notation for these notions.

An interpretation, $\underline{A}$, of a first order language, L, consists of a domain, A, and a function which assigns to each individual constant of L a member of A; to each n-place predicate of L, a set of n-tuples over A; and to each n-argument function symbol of L, an n-ary function defined over all n-tuples of A and yielding values in A. We use underlined, upper case letters to denote models and the same letters, without underlining, to denote the domains of those models.

We assume the notion of "satisfaction" as defined in (Chang and Keisler, section 1.3) and use

$\underline{A} \models \phi$ for "$\underline{A}$ satisfies ϕ" or "ϕ is true in $\underline{A}$"

We assume familiarity with the following notions:

A.2.1. Familiar notions (models)

(a) $\underline{A}$ is a submodel of $\underline{B}$; $\underline{A} \subseteq \underline{B}$

(b) $\underline{B}$ is an extension of $\underline{A}$; $\underline{A} \subseteq \underline{B}$

(c) the submodel of $\underline{B}$ generated by X, where X is a subset of B

(d) $\underline{A}$ is isomorphic to $\underline{B}$; $\underline{A} \cong \underline{B}$

(e) f is an isomorphic embedding of $\underline{A}$ into $\underline{B}$

(f) B is an elementary extension of A

(g) {A.i} is a chain of models;

(h) A = \/ {A.i}; the union of a chain of models

A.2.2. Familiar notions (theories)

(a) A theory is a set of first-order sentences

(b) The language of theory T; L(T)

(c) T proves ø; T |- ø

T1 proves T2; T1 |- T2

(d) T is complete

(e) T is consistent

(f) T is categorical

(g) T1 is equivalent to T2; T1 ≡ T2

A.2.3. Well known facts

(a) If T |- ø, then some finite subset of T proves o. (Compactness)

(b) If T is complete and T;ø is consistent, then T |- ø.

(c) If T is categorical, then T is complete.

A.2.4. Fairly familiar notions

(a) T is existential iff it is equivalent to a set of prenex sentences, none of which have universal quantifiers. (b) T is universal iff it is equivalent to a set of prenex sentences, none of which have existential quantifiers.

(c) T is universal-existential iff it is equivalent to a set of sentences in prenex form, each of which has all of its universal quantifiers preceding all of its existential quantifiers.

(d) ϕ is a primitive formula iff ϕ is an existential formula in prenex form whose matrix is a conjunction of atomic formulas and negations of atomic formulas.

A.2.5. Fairly familiar facts

(a) If T is existential, $\underline{A} \models T$, and $\underline{A} \subseteq \underline{B}$, then $\underline{B} \models T$.

(b) If T is universal and $\underline{A} \models T$, then any submodel of $\underline{A}$ also satisfies T.

(c) If T is universal existential and $\underline{A}.i \models T$, for all $i \geq 0$, then $\bigvee \{\underline{A}.i\} \models T$.

A.3 MODEL COMPLETENESS

This section presents the definition of and basic results concerning model completeness so that we don't have to pause over them in chapters 4 and 5.

There are several ways of showing that a theory is complete; we use only one. The method we use is based on A. Robinson's notion of "model completeness".

Definition A.3.1. T is model complete iff T is consistent and for any two models A and B of T, $\underline{A} \subseteq \underline{B}$ iff A is an elementary submodel of B (Monk, p.355).

A model complete theory isn't necessarily complete, unless the theory has a prime model:

Definition A.3.2. $\underline{A}$ is a prime model of T if $\underline{A}$ is a model of T and $\underline{A}$ can be embedded in any model of T (Monk, p.359).

Fact A.3.3. If T is model complete and T has a prime model, then T is complete.

To show that a theory is model complete we rely, directly or indirectly, on a theorem of Monk's (see 4.2.1a), which in turn is based on some often-cited equivalences for model completeness:

Definition A.3.4.

(a) The <u>A-expansion of</u> L is the result of adding to L a constant for each element of A.

(b) The <u>diagram</u> of $\underline{A}$ is the set of atomic sentences and negations of atomic sentences of the $\underline{A}$-expansion of the language of A which are true in $\underline{A}$.

Fact A.3.5. The following are equivalent:

(a) T is model complete.

(b) For every model $\underline{A}$ of T and every A-expansion L' of L, T + (L' - diagram of $\underline{A}$) is complete.

(c) If $\underline{A}$ and $\underline{B}$ are models of T, $\underline{A} \subseteq \underline{B}$, $\emptyset$ is a universal formula, $\underline{x}$ is in A, and $\underline{A} \models \emptyset(\underline{x})$, then $\underline{B} \models \emptyset(\underline{x})$.

(d) If $\underline{A}$ and $\underline{B}$ are models of T, $\underline{A} \subseteq \underline{B}$, $\emptyset$ is a primitive formula, $\underline{x}$ is in A, and $\underline{B} \models \emptyset(\underline{x})$, then $\underline{A} \models \emptyset(\underline{x})$.

(Monk, p.356)

Index of Models, Theories, and Statements

Bibliography

Bell and Slomson. Models and Ultraproducts: An Introduction. Amsterdam: North Holland, 1971.

Cantor, G. Contributions to the Founding of the Theory of Transfinite Numbers. New York: Dover, 1955.

Chang and Keisler. Model Theory. Amsterdam: North Holland, 1973.

Dauben, J.W. Georg Cantor: His Mathematics and Philosophy of the Infinite.

Enderton, H. A Mathematical Introduction to Logic.

Godel, K. "What is Cantor's Continuum Problem"

Griffin, H. Elementary Theory of Numbers. New York: McGraw-Hill, 1965

Kurosh, A.G. Lectures in General Algebra. New York: Pergamon Press, 1965.

Monk, D. Mathematical Logic. Springer-Verlag, 1975.

Robinson, A. and Zakon, E. Transactions of the American

Mathematical Society 96(1960), pp. 222-236.

Russell, B. *Introduction to Mathematical Philosophy*

----------- *The Principles of Mathematics*, 2nd ed.,
New York: W.W. Norton, 1937.

Tarski, A. *Undecidable Theories*. Amsterdam:
North Holland, 1968.