B Background Set Theory

Introduction

The aims of this appendix are to make clear to the reader how much knowledge of set theory is needed to understand this book, to catalogue the notation used that may not be standard and to present a proof of an important result due to Rieger that is not easily found elsewhere.

The reader will need to have seen something of the development of axiomatic set theory presented in textbooks such as (Enderton 1977, Halmos 1960). A summary of this material may be found in Chapter I of the excellent book (Kunen 1980). Chapters III and IV of that book form a convenient reference for additional material that it would be good for the reader to have seen. Certainly any reader who has read those chapters will find little difficulty with the contents of this book. Another worthwhile reference is (Shoenfield 1977).

I make free use of classes in this book, although I claim to be working informally in the axiomatic set theory, ZFC^-. The reader unfamiliar with this strategy should consult one of the above references. In part three familiarity with some of the language of category theory is needed. Very little standard category theory is really required, but the reader has to be prepared to consider functors on the superlarge category of classes. I found the book (Adamek 1983) helpful because it contains an investigation of certain types of functor on the category of sets.

Notation

The examples in chapter 1 of this book make use of the standard set theoretical representation of the natural numbers and ordered pairs. So the sets 0, $\{\emptyset\}, \{\emptyset, \{\emptyset, \{\emptyset\}\}, \ldots$ are used to represent the natural numbers 0, 1, 2,. . . and in general the natural number n is represented by the set $\{m \mid m < n\}$ of natural numbers less than n. The ordered pair (a, b) is represented, as usual, by the set $\{\{a\}, \{a, b\}\}$, and the ordered n-tuple $(a_1, a_2, \ldots, a_{n-1}, a_n)$ can be represented, in terms of ordered pairs as $(a_1, (a_2, \ldots, (a_{n-1}, a_n)))$.

Many of the standard operations on sets carry over in a natural way to classes. So, for classes $A_1, \ldots A$, we have the classes

$$A_1 \cup \cdots \cup A_n,$$
$$A_1 \cap \cdots \cap A_n,$$
$$A_1 \times \cdots \times A_n$$

defined in the expected way. It will also be useful to have their disjoint union,

$$A_1 + \cdots + A_n = (\{1\} \times A_1) \cup \cdots \cup (\{n\} \times A_n).$$

For classes A, B their set difference will be written $A - B = \{x \in A \mid x \notin B\}$. The universal class of all sets is V. The power-class of a class A is the class $powA = \{x \in V \mid x \subseteq A\}$ of all subsets of A.

A relation is a class of ordered pairs; i.e. R is a relation if $R \subseteq V \times V$. If R is a relation then xRy is written for $(x, y) \in R$ and the inverse of R is the relation $R^{-1} = \{(y, x) \mid xRy\}$. A relation R has domain $domR = \{x \mid xRy$ for some $y\}$ and range $ranR = \{y \mid xRy$ for some $x\}$. The relational composition of relations R and S is the relation

$$R \mid S = \{(x, z) \mid xRy \;\&\; yRz \text{ for some } y\}.$$

The membership relation $\in$ is the class $\{(x, y) \mid x \in y\}$, and for each class A put $\in_A = \in \cap (A \times A)$.

For classes A, B a function $f : A \to B$ is a relation $f \subseteq A \times B$ such that for each $a \in A$ there is a unique $b \in B$ such that afb. This unique b is written fa or also $f(a)$. If $X \subseteq A$ then the restriction of f to X is the map $f \restriction X : X \to B$, given by $f \restriction X = f \cap (X \times B)$. If $Y \subseteq B$ then it's inverse image under f is $f^{-1}Y = \{x \in A \mid fx \in Y\}$. If $A \subseteq B$ and $f : A \to B$ such that $fx = x$ for all $x \in A$ then f is an inclusion map and is written $f : A \hookrightarrow B$. If $f : A \to B$ and $g : B \to C$ then their function composition $g \circ f : A \to C$ is given by

$$(g \circ f)x = g(fx) \quad \text{for all } x \in A.$$

If A is a class and I is a set then A^I is the class of all the functions $f : I \to A$.

If A is a class of sets then

$$\bigcup A = \{x \mid x \in a \text{ for some } a \in A\},$$
$$\bigcap A = \{x \mid x \in a \text{ for all } a \in A\}.$$

For each class I a family of classes, A_i for $i \in I$, indexed by the class I can be represented as a relation $A \subseteq I \times V$, with $A_i = \{x \mid iAx\}$ for each $i \in I$. Given such a family of classes

$$\bigcup_{i \in I} A_i = \{x \mid x \in A_i \text{ for some } i \in I\},$$
$$\bigcap_{i \in I} A_i = \{x \mid x \in A_i \text{ for all } i \in I\},$$
$$\sum_{i \in I} A_i = \bigcup_{i \in I} \{i\} \times A_i,$$

and if I is a set,

$$\prod_{i \in I} A_i = \{f \in (\bigcup_{i \in I} A_i)^I \mid fi \in A_i \text{ for all } i \in I\}.$$

Occasionally it is convenient to consider mathematical structures having a proper class A as universe. It is usual to keep to the usual tupling notation $(A, \ldots)$ for such a structure, even though the standard definition of tuples only applies to sets. This can be understood as the class $A + R + \cdots$, using the disjoint union operation. A class that is actually a set is also called a small class, and a structure whose universe is small is called a small structure. All the functions and relations that make up a small structure will also be small.

In part III set continuous operators Φ are used. These assign a class ΦX to each class X. Because of the set continuity property the operator can be represented as the class

$$\hat{\Phi} = \{(a, x) \mid a \in \Phi x\},$$

as then for each class X

$$\Phi X = \{a \mid a \hat{\Phi} x \text{ for some } x \in powX\}.$$

Well-Foundedness

A relation R is well-founded if there is no infinite sequence a_0, a_1,. . such that $a_{n+1} R a_n$ for n = 0, I,. ... A set a is well-founded if there is no infinite sequence $a_0, a_1, \ldots$ such that $a_0 \in a$ and $a_{n+1} \in$ a, for $n = \emptyset$, **1, ...** V_{wf} is the class of all the well-founded sets.

A class A is transitive if A $\subseteq pow A$; i.e. every element of A is a subset of ***A.*** For transitive classes ***A we*** have the following principles, provided that the elements of A are all well-founded sets.

Set Induction on *A:*
For any class ***B if***

$$a \subseteq B \implies a \in B \text{ for all } a \in A$$

then $A \subseteq B$.

Set Recursion on *A:*
To uniquely define $F : A \to V$ ***it suffices*** to define Fa ***in*** terms of $F \upharpoonright a$ ***for*** each a $\in A$.

The following important result plays a special role in chapter 1.

Mostowsky s Collapsing Lemma:
If ***R is*** a well-founded relation on the ***set A*** then ***there is*** a unique ***function*** $f : A \to V$ ***such that for*** all a $\in A$

$$fa = \{fx \mid xRa\}.$$

The assumption that the class ***A*** is a set can be dropped provided it is assumed instead that $\{x \mid xRa\}$ is a set for each a $\in$ ***A;*** i.e. that ***(A, R)*** is a system in the sense of chapter 1.

We will use the standard von Neumann treatment of the ordinals, so that an ordinal α is identified with the set $\{\beta \mid \beta < \alpha\}$ of it's predecessors. So the class ***On*** of ordinals is defined to be the class of well-founded transitive sets, all of whose elements are also transitive.

The Axiomatisation of Set Theory

We take a standard first order language for set theory that just has the binary predicate symbols '=' and '$\in$'. We assume a standard axiomatisation of first order logic with equality. Also we use the standard abbreviations for the restricted quantifiers

$$\forall x \in a \cdots \stackrel{\text{def}}{=} \forall x(x \in a \to \cdots),$$
$$\exists x \in a \cdots \stackrel{\text{def}}{=} \exists x(x \in a \;\&\; \cdots).$$

In the following list of non-logical axioms for ZFC^- ***we*** have avoided the use of any other abbreviations.

Extensionality:
$$\forall z(z \in a \leftrightarrow z \in b) \to a = b$$

Pairing:
$$\exists z[\, a \in z \;\&\; b \in z \,]$$

Union:
$$\exists z(\forall x \in a)(\forall y \in x)(y \in z)$$

Powerset:
$$\exists z \forall x[\, (\forall u \in x)(u \in a) \to x \in z \,]$$

Infinity:
$$\exists z[\ (3s\ \in z)\forall y \neg(y \in x) \;\&\; (\forall x \in z)(\exists y \in z)(x \in y)\ ]$$

Separation:
$$\exists z \forall x[\, x \in z \text{ f-1 } x \in a \;\&\; \varphi \,]$$

Collection:
$$(\forall x \in a)\exists y\ \varphi \to \exists z(\forall x \in a)(\exists y \in z)\ \varphi$$

Choice:
$(\forall x \in a)\exists y(y \in x)$
$\& \ (\forall x_1 \in a)(\forall x_2 \in a)[\ \exists y(y \in x_1 \;\&\; y \in 52) \to x_1 = x_2]$
$\to \exists z(\forall x \in a)(\exists y \in x)(\forall u \in x)[\ u \in z \leftrightarrow u = y\ ]$

The choice **axiom** is abbreviated *AC*. Separation and Collection are schemes in which φ can be any formula in which the variable z does not occur free. *ZFC* is ZFC^- together with the following axiom.

Foundation:

$$\exists x(x \in a) \rightarrow (\exists x \in a)(\forall y \in x)\neg(y \in a).$$

This axiom is abbreviated *FA*.

ZFC has usually been formulated using the axiom scheme of replacement rather than the collection scheme used here. This makes no difference to the theorems of *ZFC,* but it probably does to the theorems of ZFC^-, as while each instance of replacement can easily be proved from collection, the usual proof of each instance of collection in *ZFC* makes essential use of *FA*. I prefer to take the apparently stronger collection scheme.

Global Choice and Quotients

When working with classes it is sometimes convenient to be able to use a global form of *AC*. The form that is used in this book is

$$V \cong On.$$

This expresses that there is a bijection between the universe and the class On of ordinals. This axiom cannot be formulated in the language of set theory alone but an additional predicate symbol is needed for the bijection and the axiom schemes of ZFC^- need to be extended to the larger language. A fairly cavalier approach to the use of *AC* is taken in this book. The stronger global form is used whenever it appears needed. One use of global choice is in the formation of the quotient of a class by an equivalence relation. In many situations this use can be avoided. If R is an equivalence relation on the class A we will call $f : A \rightarrow B$ a quotient of A with respect to R if f is surjective and for all $a_1, a_2 \in A$

$$a_1 R a_2 \iff f a_1 = f a_2.$$

Using global *AC* a quotient can be obtained as follows. The bijection between V and On determines a well-ordering of V. For each $a \in A$ let fa be the least set b in the well-ordering such that $b \in A$ and aRb. When A is a set, or more generally when each equivalence class $\{x \mid xRa\}$ is a set, we can follow the familiar procedure of defing fa to be the equivalence class of a. This method works in ZF^-; i.e. *ZFC* without *FA* or *AC*. For equivalence relations on a class A in general there is a trick to get a quotient, due to Dana Scott, that makes essential use of *FA*. The trick is to define fa to be the subset of the equivalence class $\{x \mid xRa\}$ consisting of those elements of the equivalence class having the least possible rank in the cumulative hierarchy of well-founded sets. In ZFC^- this trick is no longer available, but often a slight variation of the trick will work. For example if A is the class of linearly ordered sets and R is the isomorphism relation between linearly ordered sets then if $a \in A$ we can let fa be the set of linear orderings of the ordinal α that are isomorphic to the linearly ordered set a, where α is the least possible ordinal for which there is such a linear ordering of α. This works because by *AC* every set is in one-one correspondence with an ordinal.

Rieger s Theorem

Here we will prove the result that gives a general method for giving interpretations of ZFC^-. In order to interprete the language of set theory all that is needed is a class M for the variables to range over and a binary relation $\in_M \subseteq M \times M$ to interprete the predicate symbol '∈'. Now any system M, in the sense of chapter 1, determines the binary relation $\in_M$ given by

$$a \in_M b \iff a \in b_M.$$

We will show that this gives an interpretation of all the axioms of ZFC^- provided that the system is full. Recall from chapter 3 that a system M is full if for each set $x \subseteq M$ there is a unique $a \in M$ such that $x = a_M$. In the following we will let x^M be this unique $a \in M$.

Rieger s Theorem:
Every full system is a model of ZFC^-.

Proof: Let M be a full system. We will consider each axiom of ZFC^- in turn.

- Extensionality: Let $a, b \in M$ such that

  $$M \models \forall x(x \in a \leftrightarrow x \in b).$$

  Then $a_M = b_M$, so that $a = (a_M)^M = (b_M)^M = b$ and hence $M \models a = b$.

- Pairing: If $a, b \in M$ then $c = \{a, b\} \in M$ is such that $M \models (a \in c \;\&\; b \in c)$.
- Union: Let $a \in M$. Then $\bigcup\{y_M \mid y \in a_M\}$ is a subset x of M so that if $c = x^M \in M$ then $M \models \forall y \in a \forall z \in y(z \in c)$.
- Powerset: If $a \in M$ then $c = \{x^M \mid x \subseteq a_M\}^M \in M$ is such that

$$M \models \forall x[\,\forall z \in x(z \in a) \rightarrow x \in c\,].$$

- Infinity: Let

$$\begin{cases} \Delta_0 = \emptyset^M \\ \Delta_{n+1} = ((\Delta_n)_M \cup \{\Delta_n\})^M \text{ for } n = 0, 1, \ldots \end{cases}$$

Then $\Delta_n \in M$ for each natural number n, so that

$$\Delta_\omega = \{\Delta_n \mid n = 0, 1, \ldots\}^M \in M$$

is such that

$$M \models [\,\Delta_0 \in \Delta_\omega \;\&\; \forall y(y \notin \Delta_0)\,]$$

and

$$M \models \forall x \in \Delta_\omega \exists y \in \Delta_\omega (x \in y).$$

- Separation: Let $a \in M$ and let φ be a formula containing at most x free and perhaps constants for elements of M. Then

$$c = \{b \in a_M \mid M \models \varphi[b/x]\}^M \in M$$

is such that

$$M \models \forall x(x \in c \leftrightarrow x \in a \;\&\; \varphi).$$

- Collection: Let $a \in M$ and let φ be a formula containig at most x and y free and perhaps constants for elements of M. Suppose that

$$M \models \forall x \in a \exists y\, \varphi.$$

Then

$$\forall x \in a_M \exists y[\, y \in M \;\&\; M \models \varphi\,].$$

By the collection axiom scheme there is a set b such that

$$\forall x \in a_M \exists y \in b[\, y \in M \;\&\; M \models \varphi\,].$$

As $b \cap M$ is a subset of M we may form $c = (b \cap M)^M \in M$ such that

$$M \models \forall x \in a \exists y \in c\, \varphi.$$

- Choice: Let $a \in M$ such that

$$M \models \forall x \in a \exists y (y \in x)$$

and

$$M \models (\forall x_1, x_2 \in a)[\, \exists y(y \in x_1 \;\&\; y \in x_2) \rightarrow x_1 = x_2\,].$$

Then

$$\forall x \in a_M\; x_M \neq \emptyset$$

and for all $x_1, x_2 \in a_M$

$$(x_1)_M \cap (x_2)_M \neq 0 \Longrightarrow x_1 = x_2.$$

Thus $\{x_M \mid x \in a_M\}$ is a set of non-empty pairwise disjoint sets. Hence by the axiom of choice there is a set b such that for each $x \in a_M$ the set $b \cap x_M$ has a unique element $c_x \in M$. Hence $c = \{c_x \mid x \in a_M\} \in M$ such that

$$M \models \forall x \in a \exists y \in x \forall u \in x[\, u \in c \leftrightarrow u = y\,].$$

□