Formal Methods Pete Manolios
Ordinals handout Fall 2008

Throughout these notes you will find "observations." Observations are exercises which you should think about if they are not obvious to you, but you need not turn in the solutions.

# 1 Initial Notation and Definitions

$\mathbb{N}$ and $\omega$ both denote the natural numbers, *i.e.*, $\{0, 1, \ldots\}$. The ordered pair whose first component is $i$ and whose second component is $j$ is denoted $\langle i, j\rangle$. $[i..j]$ denotes the closed interval $\{k \in \mathbb{N} \ :\ i \le k \le j\}$; parentheses are used to denote open and half-open intervals, *e.g.*, $[i..j)$ denotes the set $\{k \in \mathbb{N} \ :\ i \le k < j\}$.

$R$ is a *binary relation* on set $S$ if $R \ \subseteq\ S \times S = \{\langle x, y\rangle : x, y \in S\}$. We abbreviate $\langle s, w\rangle \in R$ by $sRw$. A function is a relation such that $xRy$ and $xRw$ implies $y = w$.

Function application is sometimes denoted by an infix dot "." and is left associative. That is, $f.x$ is the unique $y$ such that $xfy$. This allows us to use the curried version of a function when it suits us, *e.g.*, we may write $f.x.y$ instead of $f(x, y)$.

From highest to lowest binding power, we have: parentheses, function application, binary relations (*e.g.*, $sBw$), equality ($=$) and membership ($\in$), conjunction ($\wedge$) and disjunction ($\vee$), implication ($\Rightarrow$), and finally, binary equivalence ($\equiv$). Spacing is used to reinforce binding: more space indicates lower binding.

$\langle Qx \ :\ r \ :\ b\rangle$ denotes a quantified expression, where $Q$ is the quantifier, $x$ the bound variable, $r$ the range of $x$ (**true** if omitted), and $b$ the body. We sometimes write $\langle Qx \in X : r : b\rangle$ as an abbreviation for $\langle Qx : x \in X \quad \wedge \quad r : b\rangle$, where $r$ is **true** if omitted, as before.

Cardinality of a set $S$ is denoted by $|S|$. $\mathcal{P}(S)$ denotes the powerset of $S$.

A function from $[0..n)$, where $n$ is a natural number, is called a *finite sequence* or an *n-sequence*.

What are numbers as mathematical objects? von Neumann proposed the following: $0 = \emptyset, 1 = \{0\}, 2 = \{0, 1\}, \ldots$, so $n = [0..n)$. Thus an $n$-sequence is a function from $n$.

An *$\omega$-sequence* is a function from $\omega$. We may sometimes refer to $\omega$-sequences as infinite sequences, but as we will see there are infinite sequences that are "longer" than $\omega$-sequences.

When we write $x \in \sigma$, for a sequence $\sigma$, we mean that $x$ is in the range of $\sigma$.

# 2 Binary Relations

Let $B, C$ be binary relations on set $S$. $B|_A$ denotes $B$ *left-restricted* to the set $A$, *i.e.*, $B|_A = \{\langle x, y\rangle : xBy \quad \wedge \quad x \in A\}$.

Some important definitions follow.

- $B$ is *reflexive* if $\langle \forall x \in S :: xBx \rangle$.
- $B$ is *irreflexive* if $\langle \forall x \in S :: \neg(xBx) \rangle$.
- $B$ is *transitive* if $\langle \forall x, y, z \in S :: xBy \quad \wedge \quad yBz \quad \Rightarrow \quad xBz \rangle$.
- $B$ is a *preorder* (also called a *quasi-order*) if it is reflexive and transitive.
- The identity relation, $B^0$, is $\{\langle x, x \rangle : x \in S\}$.
- The *composition* of $B$ and $C$ is denoted $B; C$ and is the set $\{\langle b, c \rangle : \langle \exists x :: bBx \quad \wedge \quad xCc \rangle\}$.
- For all natural numbers $i$, $B^{i+1}$ is $B^i; B$.

**Observation 1** *Prove the following.*

1. *$B$ is reflexive iff $B^0 \subseteq B$.*
2. *$B^1 = B$.*
3. *$B$ is transitive iff $B^2 \subseteq B$.*

We now continue with the definitions.

- $B$ is *symmetric* if $\langle \forall x, y \in S :: xBy \quad \Rightarrow \quad yBx \rangle$.
- A preorder that is also symmetric is an *equivalence relation*.
- $B$ is *asymmetric* if $\langle \forall x, y \in S :: xBy \quad \Rightarrow \quad \neg(yBx) \rangle$.
- $B$ is *antisymmetric* if $\langle \forall x, y \in S :: xBy \quad \wedge \quad yBx \quad \Rightarrow \quad x = y \rangle$.
- A preorder that is antisymmetric is a *partial order*.
- If $B$ is a partial order, $\langle S, B \rangle$ is a *poset*.
- The *inverse* of $B$ is denoted $B^{-1}$ and is $\{\langle x, y \rangle : yBx\}$.

**Observation 2** *Prove the following.*

1. *$B$ is symmetric iff $B^{-1} \subseteq B$.*
2. *$B$ is antisymmetric iff $B \cap B^{-1} \subseteq B^0$.*

If $B$ is an equivalence relation, for each $x \in S$, it induces an *equivalence class* $[x]_B = \{y : xBy\}$. The *quotient* $S/B$ is $\{[x]_B : x \in S\}$.

**Observation 3** *Prove the following.*

1. *If $B$ is an equivalence relation, then $[x]_B$ and $[y]_B$ are either identical or disjoint.*

*2. If $C$ is a preorder, then*

   *(a) $B = \{\langle x, y\rangle : xCy \quad \wedge \quad yCx\}$ is an equivalence relation.*

   *(b) $\langle S/B, \preccurlyeq\rangle$ is a poset, where $\preccurlyeq$ is defined as follows:*
   *$[x]_B \preccurlyeq [y]_B \qquad \equiv \qquad xCy$.*

We now continue with the definitions.

- $B$ is *total* (also called *linear* or *connected*) if $\langle \forall x, y \in S :: xBy \quad \vee \quad yBx\rangle$.
- A *total order* is a partial order that is total.
- If $B$ is a total order, $\langle S, B\rangle$ is a *toset*.
- An $\alpha$-sequence $\langle a_0, a_1, a_2, \ldots\rangle$, where $\alpha \in \omega \quad \vee \quad \alpha = \omega$, is *decreasing* in $B$ if $\langle \forall i : i + 1 \ \in \ \alpha : a_{i+1} B a_i\rangle$.
- $B$ is *terminating* (also called *well-founded*) if there is no decreasing $\omega$-sequence in $B$.
- If $B$ is terminating, then $\langle S, B\rangle$ is a *well-founded structure*.
- The *strict part* of a relation $B$ is $\{\langle x, y\rangle : xBy \quad \wedge \quad x \neq y\}$.
- $B$ is a *strict partial order* if it is the strict part of some partial order. Strict total orders are defined in an analogous way.
- A *well order* is a strict total order that is well-founded.
- If $B$ is a well order, $\langle S, B\rangle$ is a *woset*.
- For $T \subseteq S$:
  - If $(m \in T \quad \wedge \quad \langle \forall x \in T :: xBm \quad \Rightarrow \quad x = m\rangle)$, then $m$ is a *minimal* element of $T$ (under $B$).
  - If $(m \in T \quad \wedge \quad \langle \forall x \in T :: mBx\rangle)$, then $m$ is the *least* element of $T$ (under $B$).
  - If $(m \in S \quad \wedge \quad \langle \forall x \in T :: mBx\rangle)$, then $m$ is a *lower bound* of $T$ (under $B$).
  - The notions of *maximal*, *greatest*, and *upper bound* are defined dually, *e.g.*, $m$ is a maximal element of $T$ under $B$ iff $m$ is a minimal element of $T$ under $B^{-1}$.

**Observation 4** *Prove the following.*

1. *$B$ is total iff $B \cup B^{-1} = S \times S$.*
2. *$B$ is a strict partial order iff it is irreflexive and transitive.*

3. *If $\prec$ is a strict partial order and* $x \preccurlyeq y \quad \equiv \quad x \prec y \quad \vee \quad x = y$ *then* $\preccurlyeq$ *is a partial order.*

4. *If $\preccurlyeq$ is a preorder and* $x \prec y \quad \equiv \quad x \preccurlyeq y \quad \wedge \quad \neg(y \prec x)$ *then* $\prec$ *is a strict partial order.*

5. *$B$ is a strict total order iff*

   (a) *$B$ is irreflexive.*

   (b) *$B$ is transitive.*

   (c) $\langle \forall x, y \in S :: xBy \quad \vee \quad yBx \quad \vee \quad x = y \rangle$.

6. *$B$ is a well order iff it is well-founded and* $\langle \forall x, y \in S :: xBy \quad \vee \quad yBx \quad \vee \quad x = y \rangle$.

**Observation 5** *Prove the following.*

1. *Prove that $\langle S, \prec \rangle$ is a well-founded structure iff all non-empty subsets of $S$ have a minimal element under $\prec$.*

2. *Prove that $\langle S, \prec \rangle$ is a woset iff all non-empty subsets of $S$ have a least element.*

Given a set $U$ (the "universe"), $X \subseteq U$, and a property $P$ which is satisfied by some subsets of $U$, the $P$-sets, we say that $C$ is the $P$-closure of $X$ if $C$ is the least $P$-set which includes $X$. If the $P$-sets include $U$ and are closed under arbitrary intersections, we say that the $P$-sets of $U$ form a *closure system.* If the $P$-sets of $U$ form a closure system, then the $P$-closure of $X$ always exists. It is $\cap\{Y \subseteq U : X \subseteq Y \quad \wedge \quad Y \text{ is a } P\text{-set}\}$.

**Observation 6** *Prove the following, where $U = S \times S$.*

1. *The reflexive relations form a closure system.*

2. *The irreflexive relations do not form a closure system.*

3. *The symmetric relations form a closure system.*

4. *The asymmetric relations do not form a closure system.*

5. *The antisymmetric relations do not form a closure system.*

6. *The transitive relations form a closure system.*

We can therefore speak of the reflexive closure, or the symmetric closure, or the transitive closure, or the reflexive, transitive closure, etc. $B^+$ denotes the transitive closure of $B$ and $B^*$ denotes the reflexive, transitive closure of $B$. This same notation is used in regular languages.

## 3 Induction and Recursion

Mathematical induction works because the natural numbers (with the usual ordering) are a well-founded: if some property fails to hold for all naturals, it fails for some minimal $n$, but holds for all smaller numbers, which is exactly what we prove doesn't happen. We can extend this idea to more general sets. The *principle of well-founded induction* states: If $\langle W, \prec\rangle$ is a well-founded structure,

$$\text{(WFI)} \quad \langle \forall w \in W :: P.w\rangle \equiv \langle \forall w \in W :: \langle \forall v : v \prec w : P.v\rangle \Rightarrow P.w\rangle$$

**Observation 7** *Show that (weak) mathematical induction is a special case of well-founded induction.*

**Observation 8** *Show that strong mathematical induction (course of values induction) is a special case of well-founded induction.*

**Observation 9** *Let $\prec$ be a binary relation on $W$. Show that* WFI *holds iff $\prec$ is terminating.*

**Observation 10** *Prove that if a relation is well-founded iff its transitive closure is well-founded.*

**Observation 11** *Prove that if a relation $\prec$ on $S$ is well-founded, then so is $\prec_n$ on $n$-tuples of elements from $S$, where $n$ is a positive natural number and $\prec_n$, the* lexicographic *version of $\prec$, is defined as follows: $\prec_1 = \prec$ and for $n > 1$, $\langle x_n, x_{n-1}, \ldots, x_1\rangle \prec_n \langle y_n, y_{n-1}, \ldots, y_1\rangle$ iff $x_n \prec y_n$ or ($x_n = y_n$ and $\langle x_{n-1}, \ldots, x_1\rangle \prec_{n-1} \langle y_{n-1}, \ldots, y_1\rangle$).*

**Observation 12** *Is the dictionary order well-founded?*

Induction on wosets is called *well-ordered induction* or *transfinite induction.*

It turns out, that as a consequence of the *axiom of choice*, which states: the cartesian product of a non-empty family of non-empty sets is non-empty, we have that for any set $S$, there is a relation $\prec$ s.t. $\langle S, \prec\rangle$ is a woset. Note the remarkable consequence: we can well-order any set and can thus apply induction to any set.

Induction can be used to justify recursive definitions. A general principle of recursive definitions follows. If

1. $\langle W, \prec\rangle$ is a well-founded structure; and
2. $g$ is a binary function that maps any $w \in W$ and any function from $\{v : v \prec w\}$ to $W$ into $W$.

Then, the following is satisfied by exactly one function on $W$.

$$\text{(WFD)} \qquad f.x = g(x, \{\langle y, f.y\rangle : y \prec x\})$$

Note that $f$ is defined in terms of itself, but for any $x$, $f.x$ depends only on $f.y$ for $y \prec x$. The idea is that since $\prec$ is terminating, the dependencies can be unrolled until minimal elements are reached, thus the above equation defines a unique function.

**Observation 13** *How would you go about proving the above?*

Let us examine how to use the above principle to show that recursive equations are meaningful. First, here is an example of why such a principle is needed. Consider the following "definition":

```
(defun foo (x)
  (1+ (foo x)))
```

We get that `(foo x)` = `(1+ (foo x))`, which leads to `0` = `1`. Thus, one can introduce inconsistencies if not careful. If we think of `foo` as a function, it does not terminate. Showing that definitions are meaningful amounts to showing that they terminate.

Here is a sequence of ACL2 events culminating in a proof of `nil` from the above equation.

```
(defstub foo (*) => *)

(defaxiom foo-def (equal (foo x) (1+ (foo x))))

(defthm contradiction
  (implies (equal x (1+ x)) nil)
  :rule-classes nil)

(thm nil
     :hints (("goal"
              :use ((:instance contradiction (x (foo x)))
                    (:instance foo-def)))))
```

Consider the following definition.

```
(defun foo (x)
  (if (zp x)
      0
    (1+ (foo (1- x)))))
```

Recall the definition of `zp`:

```
(defun zp (x)
  (if (integerp x) (<= x 0) t))
```

Why is `foo` a proper definition? Because it terminates. How do we apply (WFD)? Well, $W$ is the set of ACL2 objects (this is going to remain the same no matter what function we define). We can define $\prec = \{\langle x, x+1\rangle : x \in \mathbb{N}\}$.

$g.x$ is 0 if `(zp x)` and othewise is 1+`(foo` $x-1$`)`. Note that $x-1$ is the only ACL2 object $\prec x$, as required by the definition of (WFD).

There are many choices for $\prec$, *e.g.*, $\prec = \{\langle x, y\rangle : x, y \in \mathbb{N} \quad \wedge \quad x < y\}$ is another. In this case, $g$ is defined in the same way; $g$ only depends on $x-1$, as before, even though it has available to it the value of `foo` on $[0..x)$.

Another way of describing the process follows. To show that a recursive definition is meaningful, define a *measure function*, $m$, a function that maps $W$ into a well-founded structure $\langle T, \prec\rangle$, and to show that in every recursive call, $m$ decreases. For the above example, let $m$ map ACL2 objects to the naturals under the usual ordering, $<$. Every non-natural is mapped to 0. In the recursive call, `x` is greater than 0 and `foo` is called on `(1- x)`, thus $m(x-1) < m.x$ and `foo` terminates.

Let's look at another definition, which gives highlights the use of measure functions in ACL2.

```
(defun upto (i j)
  (declare (xargs :measure (nfix (- j i))))
  (if (and (integerp i)
           (integerp j)
           (< i j))
      (1+ (upto (1+ i) j))
    0))
```

Here, we are counting up, but, even so, `upto` is a terminating function, as the measure function above shows.

# 4 Ordinals

We will now start to set up the machinery for ordinals. We start with numbers and work our way up.

von Neumann proposed defining the natural numbers in set theory as follows: $0 = \emptyset, 1 = \{0\}, 2 = \{0, 1\}, \ldots$. Equivalently, we can characterize $\omega$ as the least set $S$ s.t. $\emptyset \in S$ and $n \in S \Rightarrow n^+ \in S$, where the *successor* of $n$, $n^+$ is defined to be $n \cup \{n\}$ for any set $n$ (note: $n < m = n \subset m = n \in m$).

**Definition 1** *For a relation $B$, let pred$(x, B)$ be defined as: $\{y : yBx\}$. When $B$ is clear from context, we sometimes write $s.x$ for pred$(x, B)$, which is called the* initial segment *of $x$ under $B$.*

We can show that each $n \in \omega$ is well-ordered by $\in$ as is $\omega$ and that for each $n \in \omega, n = s.n$. Go through the first few numbers to convince yourself.

We are now ready to deal with ordinals.

**Definition 2** *$\alpha$ is an* ordinal *if $\langle \alpha, \prec\rangle$ is well-ordered for some $\prec$ and for all $\beta \in \alpha, \beta = s.\beta$.*

Recall that $B$ is a well order iff it is well-founded and trichotomy holds, *i.e.*, $\langle \forall x, y \in S :: xBy \quad \vee \quad yBx \quad \vee \quad x = y \rangle$.

Let $Ord.\alpha$ denote that $\alpha$ is an ordinal. We know of a few ordinals already: the natural numbers and $\omega$.

**Lemma 1** *If $Ord.\alpha$ and $\beta, \gamma \in \alpha$, then $\beta \in \gamma \quad \equiv \quad \beta \prec \gamma \quad \equiv \quad \beta \subset \gamma$.*

We have shown that $Ord.\alpha \quad \Rightarrow \quad \prec = \in = \subset$.

**Lemma 2** *$W$ is well ordered by $\in$ iff $\langle \forall x, y \in W :: x = y \quad \vee \quad x \in y \quad \vee \quad y \in x \rangle$.*

**Corollary 1** *$Ord.\alpha \quad \equiv \quad \langle \forall \beta, \gamma \in \alpha :: \beta = \gamma \quad \vee \quad \beta \in \gamma \quad \vee \quad \gamma \in \beta \rangle \quad \wedge \quad \langle \forall \beta \in \alpha :: s.\beta = \beta \rangle$.*

**Lemma 3** *$Ord.\alpha \Rightarrow Ord.\alpha^+$.*

**Lemma 4** *$Ord.\alpha \quad \wedge \quad \beta \in \alpha \quad \Rightarrow \quad \beta \subset \alpha$.*

**Definition 3** *A set $x$ is* transitive *if $\langle \forall y, z :: z \in y \quad \wedge \quad y \in x \quad \Rightarrow \quad z \in x \rangle$, which is equivalent to $\langle \forall y \in x :: y \subset x \rangle$.*

**Lemma 5** *$Ord.\alpha$ iff $\alpha$ is transitive and well-ordered by $\in$.*

**Lemma 6** *$Ord.\alpha \Rightarrow \langle \forall \beta : \beta \in \alpha : Ord.\beta \rangle$*

**Lemma 7** *$Ord.\alpha \wedge Ord.\beta \Rightarrow (\alpha \in \beta) \vee (\beta \in \alpha) \vee (\alpha = \beta)$*

**Lemma 8** *If $A$ is a set of ordinals, then $\cup A$ is an ordinal.*

**Lemma 9** *$\neg \langle \exists x :: \langle \forall y :: Ord.y \quad \Rightarrow \quad y \in x \rangle \rangle$.*

This is the Burali-Forti paradox, which is similar to Russell's paradox and the cause is the same: we have to construct sets with care. There is no set that contains all the ordinals, but if one is careful, it is beneficial to think about the collection of all ordinals. Such collections which are too large to be sets are called *classes*. Another example of a useful class is $\mathbf{V} = \{x : \mathbf{true}\}$, the class of all sets. Expressions using classes can be turned into expressions that do not use classes, *e.g.*, the expression $\mathbf{V} = \mathbf{On}$ is best thought of as an abbreviation for $\langle \forall x :: \mathbf{true} \quad \equiv \quad Ord.x \rangle$, which is $\langle \forall x :: Ord.x \rangle$, a false statement.

## 4.1 V

$\mathbf{V}$, the universe of sets is obtained by iterating the power set operation over all the ordinals.

$$\mathbf{V}_0 = \emptyset$$

$$\mathbf{V}_\alpha = \bigcup \{\mathcal{P}(\mathbf{V}_\beta) : \beta \in \alpha\}$$

$$\mathbf{V} = \bigcup_{\alpha \in \mathbf{On}} \mathbf{V}_\alpha$$

$\emptyset$

Here is a graphical representation of $\mathbf{V}$. We start with the empty set and expand the universe at each stage.

$\mathbf{V}_1 = \{\emptyset\}$, $\mathbf{V}_2 = \{\emptyset, \{\emptyset\}\}$, $\mathbf{V}_3 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$, and so on.

## 4.2 Wosets

If $S, T$ are sets and $A, B$ are relations, then $\langle S, A\rangle$, $\langle T, B\rangle$ are *isomorphic*, denoted $\langle S, A\rangle \cong \langle T, B\rangle$, if there is a bijection $f : S \to T$ such that for any $x, y \in S$, $xAy$ iff $(f.x)B(f.y)$.

If $A$ is a relation on $T$ and $S \subseteq T$ a set, then we denote $A \cap (S \times S)$ by $\mathit{res}(A, S)$. If $x \in S$, then we denote $\mathit{res}(A, \mathit{pred}(x, A))$ by $\mathit{res}(A, x)$. Note that if $\langle S, A\rangle$ is a woset, then so is $\langle \mathit{pred}(x, A), \mathit{res}(A, x)\rangle$ for every $x \in S$.

**Lemma 10** *If $\langle S, A\rangle$ is a woset, then $\langle \forall x \in S :: \langle S, A\rangle \not\cong \langle \mathit{pred}(x, A), \mathit{res}(A, x)\rangle\rangle$.*

**Lemma 11** *If $\langle S, A\rangle$ and $\langle T, B\rangle$ are wosets and isomorphic, then the isomorphism between them is unique.*

**Lemma 12** *If $\langle S, A\rangle \cong \langle T, B\rangle$ and $\langle T, B\rangle \cong \langle U, C\rangle$, then $\langle S, A\rangle \cong \langle U, C\rangle$.*

**Lemma 13** *If $\langle S, A\rangle$ and $\langle T, B\rangle$ are wosets, then exactly one of the following holds.*

1. $\langle S, A\rangle \cong \langle T, B\rangle$
2. $\langle \exists x \in S :: \langle \mathit{pred}(x, A), \mathit{res}(A, x)\rangle \cong \langle T, B\rangle\rangle$
3. $\langle \exists y \in T :: \langle S, A\rangle \cong \langle \mathit{pred}(y, B), \mathit{res}(B, y)\rangle\rangle$

**Lemma 14** *If $\mathit{Ord}.\alpha$, $\mathit{Ord}.\beta$, and $\alpha \cong \beta$, then $\alpha = \beta$.*

**Lemma 15** *If $\langle S, A\rangle$ is a woset, then there is a unique ordinal $\alpha$ such that $\langle S, A\rangle \cong \alpha$.*

This shows that using ordinals we can order the elements of any set.

## 4.3 Ordinal Arithmetic

We now have the machinery to develop ordinal arithmetic.

We start with addition; the informal idea is that we can combine two wosets by listing the elements of the first woset in order, followed by the elements of the second woset in order. More rigorously for wosets $N, M$, construct wosets $N', M'$ where $N'$ is the set of pairs $\langle 0, n\rangle$ for $n \in N$ and $M'$ is the set of pairs $\langle 1, m\rangle$ for $m \in M$. Clearly $N', M'$ are disjoint, $N \cong N'$ and $M \cong M'$. This shows that we may assume that we have disjoint sets. Now we can define the wosum of two wosets $\langle N, \prec_N\rangle, \langle M, \prec_M\rangle$ to be the woset $\langle N \cup M, \prec_N \cup \prec_M \cup (N \times M)\rangle$. We now define ordinal addition for ordinals $\alpha, \beta$ as follows: let $A$ and $B$ be disjoint wosets such that $A \cong \alpha$ and $B \cong \beta$, and let $C$ be the wosum of $A$ and $B$. The ordinal sum $\alpha + \beta$ is the ordinal $\gamma$ such that $\gamma \cong C$. For ordinal addition we have: 0 is the identity, associativity holds, but commutativity does not, *e.g.*, $\omega + 1 \neq 1 + \omega = \omega$.

The woproduct of two wosets $A$ and $B$ is the result of adding $A$ to itself $B$ times; hence we define the woproduct of wosets $A$ and $B$ as $A \times B$ with the reverse lexicographic order. The ordinal product $\alpha\beta$ is the ordinal $\gamma$ such that $\gamma \cong (A \times B)$ where $A \cong \alpha$ and $B \cong \beta$. For ordinal multiplication we have: 1 is the identity, $0\alpha = 0 = \alpha 0$, associativity holds, left distribution holds, commutativity fails ($2\omega = w \neq \omega 2$), and right distribution fails ($(1 + 1)\omega$).

Ordinals containing a maximal element are called *successor* ordinals; the other ordinals, except for $\emptyset$, are called *limit* ordinals. Note that if $\alpha$ is a limit ordinal, $\alpha = \cup\alpha$.

We can also define ordinal exponentation as follows: $\alpha^0 = 1$, $\alpha^{\beta+1} = \alpha^\beta\alpha$, and if $\beta$ is a limit ordinal, $\alpha^\beta$ is $\bigcup_{\gamma\in\beta} \alpha^\gamma$. For ordinal exponentation we have: $0^\alpha = 0$ $(0 \prec \alpha)$, $1^\gamma = 1, \alpha^{\beta+\gamma} = \alpha^\beta\alpha^\gamma, \alpha^{\beta\gamma} = (\alpha^\beta)^\gamma$, but $(\alpha\beta)^\gamma = \alpha^\gamma\beta^\gamma$ does not hold, *e.g.*, $(2 \cdot 2)^\omega = 4^\omega = \omega$, but $2^\omega \cdot 2^\omega = \omega \cdot \omega = \omega^2$.

The first few ordinals are: $0,\ 1, \cdots, \omega,\ \omega + 1, \cdots, \omega 2,\ \omega 2 + 1, \cdots$. In this way we get $\omega,\ \omega 2,\ \omega 3, \cdots, \omega^2,\ \omega^2 + 1, \cdots, \omega^2 + \omega,\ \omega^2 + \omega + 1, \cdots, \omega^2 + \omega 2$, $\omega^2 + \omega 2 + 1, \cdots, \omega^2 + \omega 3,\ \cdots, \omega^2 + \omega 4,\ \cdots, \omega^2 2,\ \cdots, \omega^2 3,\ \cdots, \omega^3, \omega^4,\ \cdots, \omega^\omega$, $\cdots, \omega^{(\omega^\omega)},\ \cdots, \omega^{(\omega^{(\omega^\omega)})},\ \cdots, \epsilon_0,\ \epsilon_0 + 1,\ \cdots, \epsilon_0 + \omega,\ \cdots,\ \epsilon_0 + \omega 2,\ \cdots,\ \epsilon_0 + \omega^2$, $\cdots,\ \epsilon_0 + \omega^\omega,\ \cdots,\ \epsilon_0 2, \cdots,\ \epsilon_0\omega,\ \cdots,\ \epsilon_0\omega^\omega,\ \cdots, \epsilon_0^2, \cdots\ \cdots\ \cdots, \Omega, \cdots\ \cdots\ \cdots$ .

Notice that simplifying expression such $7 + \omega + \omega^2$ requires some thought. The above is equal to $\omega^2$. Imagine coming up with an algorithm for simplifying ordinals such as:

$(\omega^{(\omega+1)^2}) \cdot (\omega^{(\omega+\omega\cdot 12)\cdot(\omega^{\omega+1}+\omega^2\cdot 9)})$

How do we know that we can simplify such things?

We can prove that $\epsilon_0$ is a countable ordinal, and that ordinals can be put into Cantor normal form.

## 4.4 Cardinals

Cardinal numbers are used to measure the size of sets. They differ from ordinals in that order is not important, just size. One will be able to show that all ordinals of the same size form a set, thus a natural representative for sets of that size is the minimal ordinal. This is what is done, *i.e.*, an ordinal is a cardinal iff it is an element of every other ordinal of the same size. All of the natural numbers are cardinals, but all the ordinals from $\omega, \ldots, \epsilon_0$ and beyond, all the way up to (but not including $\Omega$), are of the same size, thus the only cardinal in this collection is $\omega$.

For sets $X, Y$, by $X \precsim Y$ we denote that there is an injection (a 1-1 function) from $X$ into $Y$. $X \approx Y$ denotes that there is a bijection (a 1-1, onto function) from $X$ onto $Y$.

By the axiom of choice (every set can be well-ordered), for any set $X$, there is a least ordinal $\alpha$ such that $X \approx \alpha$. $\alpha$ is the cardinality of $X$, also denoted $|X|$. We say that $\alpha$ is a *cardinal* iff $|\alpha = \alpha|$.

Cardinal arithmetic is significantly different from ordinal arithmetic. Cardinal addition is defined as follows: $\kappa + \lambda = |(\kappa \times \{0\} \cup \lambda \times \{1\})|$, thus $\kappa + \lambda = \lambda + \kappa$. Cardinal multiplication is defined as follows $\kappa \cdot \lambda = |(\kappa \times \lambda)|$, thus $\kappa \cdot \lambda = \lambda \cdot \kappa$. There are some surprises, *e.g.*, for any infinite cardinal $\kappa, \kappa \cdot \kappa = \kappa$. We do not have time to delve into this further (unfortunately).

# 5 ACL2 Ordinals

The ACL2 ordinals are essentially the ordinals less than $\epsilon_0$. ACL2 includes function for recognizing ordinals, comparing ordinals, and performing various arithmetic operations on ordinals. Look at the online documentation and try it out.