Implementing relations and functions without pairs

M. Randall Holmes

June 12, 2009

1 Introduction

This paper begins with a brief survey of notions of ordered pair. The prehistory of the notion is covered, standard pair definitions are introduced, and a couple of refinements due to the author are given as well.

The primary purpose of the paper is to present a partial implementation of the notion of function (in the simple theory of types with three types) which does not make use of the notion of ordered pair. This extends unpublished work of Henrard (see [?] in which he gave a complete implementation of the notion of cardinality of a set in the simple theory of types with three types. An intended application of this was to Grishin’s consistent fragment $NF_3$ (see [?]) of Quine’s “New Foundations” ([?]), in which Henrard’s and the author’s developments are applicable, though there will be little or no occasion to discuss this in this paper.

Limiting results are also presented. It is obvious that the limitation to three types makes sense: in four types, the usual Kuratowski pair is definable and a theory of functions and relations is readily developed along standard lines. We show that in the presence of a modest amount of choice, an ordered pair can be defined in three types which can be used to support a standard development of relations and functions. We also show, using Frankel-Mostowski permutation techniques, that there can be no complete theory of functions which works in general in three types: the partial theory of functions we present will be seen to be essentially the best possible in the absence of choice principles.

2 A survey of pairs

The ordered pair in set theory is not a notion of independent interest so much as a tool for the implementation of relations and functions as sets. The notion has a prehistory: for example, Bertrand Russell knew perfectly well that a relation should be presented as a set of couples (ordered pairs), but in *Principia Mathematica* ([?], hereinafter abbreviated *PM*) he and Whitehead found it necessary
to present a theory with types of \( n \)-ary relations for each \( n \), and in PM the ordered pair is defined as a species of binary relation!

The first definition of an ordered pair as a set was given by Norbert Wiener in 1914 (\cite{Wiener}). He defined the pair \((x, y)\) as \(\{\{x\}, \emptyset, \{\{y\}\}\}\). Wiener pointed out that the availability of this pair meant that relation types could be dispensed with in favor of set types only, and so one could have a linear hierarchy of set types instead of the elaborate type hierarchy of PM (at least if one also followed Ramsey in abandoning the predicativity restrictions of PM) (\cite{Ramsey}).

We briefly define the simple theory of types \( TST \). This is a multi-sorted first-order theory with types indexed by the natural numbers. Each variable is understood to have a type and we may sometimes (but will not always) adorn a variable of type \( i \) with the superscript \( i \). The primitive predicates of the theory are equality and membership. Well-formed atomic formulas are equations \( x = y \) in which \( x \) and \( y \) are variables of the same type and membership statements \( x \in y \) in which the type of \( y \) is the successor of the type of \( x \).

The axioms schemes of \( TST \) are

**Extensionality:** All sentences of the form \((\forall x y. x = y \leftrightarrow (\forall z. z \in x \leftrightarrow z \in y))\) are axioms.

**Comprehension:** All universal closures of formulas of the form \((\exists A. (\forall x. x \in A \leftrightarrow \phi))\) are axioms, where \( \phi \) is any formula in which \( A \) does not appear free.

It is usual to also include axioms of Infinity and Choice (it takes some work to determine the best form for these axioms), but we do not regard these as an essential part of the theory here. We will have occasion to consider subtheories \( TST_n \) of \( TST \): the language of \( TST_n \) is restricted to formulas in which no variable of type \( \geq n \) appears. \( TST_n \) is called \( \text{"type theory with } n \text{ types"} \).

The usefulness of the simple theory of types was made apparent by Wiener’s pair definition of 1914, though it seems that such theories were not clearly presented until about 1930 (see \cite{Kuratowski} for a discussion of the history of this theory). We will frame our discussion of pairs in terms of \( TST \), though the pairs we mention are also usable in Zermelo set theory and its extensions.

The Wiener pair has the defining property of an ordered pair

**Defining property of the ordered pair:** \((x, y) = (z, w) \rightarrow x = z \land y = w.\)

In \( TST \), if \( x \) and \( y \) are of type \( n \), \((x, y)\) is of type \( n + 3 \). Any logical relation (defined by a formula) on type \( n \) is implementable in type \( n + 4 \) as a set of Wiener pairs. Of course this includes functional relations, so the notation \( y = f(x) \) can be defined in \( TST \), with \( x \) and \( y \) understood to be of the same type and \( f \) four types higher (if the Wiener pair is understood to be used).

The Wiener pair is not usually used. It has been superseded by the Kuratowski pair \((x, y) = \{\{x\}, \{x, y\}\}\) (first presented in \cite{Kuratowski}). The Kuratowski pair is just two types higher than its projections, and if it is used as the official pair,
then relations and functions are three types higher than the elements of their domains and ranges.

The Wiener pair does have the merit that the proof that it is an ordered pair is easier. But this is not a major advantage, as the proof of the defining property of the ordered pair needs only be done once for each implementation, after which the details can be disregarded.

This last thought can conveniently be expressed by saying in the language of theoretical computer science that the ordered pair is an abstract data type. In this connection we briefly present a spurious but nonetheless interesting argument of Adrian Mathias([?])

Mathias claims (perhaps tongue in cheek) that our belief that the ordered pair is an abstract data type implies the Axiom of Replacement, in the context of Zermelo set theory.

Suppose that \( \forall x. (\exists! y. \phi(x,y)) \) (the formula \( \phi(x,y) \) describes a logical function). We define \( F_\phi(x) \) as the unique \( y \) such that \( \phi(x,y) \). We define a pair \( [x,y] \) as \( (x,F_\phi(y)) \), where \( (x,y) \) is the usual Kuratowski pair. It is clear that \( [x,y] \) satisfies the defining property of an ordered pair. Since this is a pair, we expect to have a cartesian product operation \( A \otimes B = \{ [a,b] \mid a \in A \land b \in B \} \) (reserving \( A \times B \) for the usual cartesian product using the Kuratowski pair). Now \( A \otimes A \) is a set of Kuratowski pairs, and in fact a set function (in the sense defined in terms of the Kuratowski pair), and we can prove in Zermelo set theory that its range is a set: but its range is \( F_\phi " A \), so we have established Replacement.

The argument is spurious, and the bad move should be fairly easy to identify from the above brief presentation. There is no reason to believe that \( A \otimes A \) is a set. The point is that the defining property of the ordered pair is not the only component of the “abstract data type interface” for the ordered pair. It is also necessary for a notion of ordered pair to be useful for the cartesian product \( A \times B = \{ (a,b) \mid a \in A \land b \in B \} \) to be a set, and, for any set of ordered pairs \( R \), for the class \( \{ x \mid (\exists y. (x,y) \in R \lor (y,x) \in R) \} \), the field of \( R \), to be a set. We can then see that the “Mathias pair” of the previous paragraph is not a legitimate pair (or at least that there is no reason to expect it to be). These are not illegitimate adjuncts to the notion of ordered pair, because of our intended use of the ordered pair: we use the pair not for its own sake (as the basis of a theory of lists, perhaps) but for the sake of implementing relations and functions, so it is reasonable to include conditions on set formation from ordered pairs.

We present another rather baroque definition of the ordered pair, due to Quine (in [?]), which has the merit in the context of type theory that \( (x,y) \) is of the same type as \( x \) and \( y \). The Quine pair has a corresponding merit in ZF: the rank of a Quine pair \( (x,y) \) in the cumulative hierarchy is the same as the maximum of the ranks of \( x \) and \( y \), whereas the rank of the corresponding Kuratowski pair is two higher. A further oddity of the Quine pair (in any context) is that every set is an ordered pair if this pair definition is used. If the Quine pair is used we have the interesting equation \( V_\alpha \times V_\alpha = V_\alpha \) for each infinite \( \alpha \).

Quine’s definition of the pair depends on having an implementation of the
natural numbers (and so on the Axiom of Infinity holding). For all sets \( x \), define \( \sigma(n) = n + 1 \) for each natural number \( n \) and \( \sigma(x) = x \) for all other sets. Define \( \sigma_1(x) = \sigma x \) and \( \sigma_2(x) = \sigma x \cup \{0\} \). Notice that \( \sigma_1 \) and \( \sigma_2 \) are bijections with disjoint ranges. We can define a map \( \sigma_3(x) = \{n - 1 \mid n \in x \cap \mathbb{N}\} \cup x - \mathbb{N} \) for which it is easily seen that \( \sigma_3(\sigma_1(x)) = \sigma_3(\sigma_2(x)) = x \). Now we define \((x, y)\) as \( \sigma_1 x \cup \sigma_2 y \). If we define \( \pi_1(x) \) as \( \sigma_3 \{y \in x \mid 0 \notin y\} \) and \( \pi_2(x) \) as \( \sigma_3 \{y \in x \mid 0 \in y\} \), it is readily seen that \( x = (\pi_1(x), \pi_2(x)) \) from which it is seen that the Quine pair satisfies the defining property of the ordered pair and that moreover every set is a Quine pair. In the context of the theory of types, it is also easily seen that \((x, y)\) is of the same type as \( x \) and \( y \). [The Quine pair is defined in \( \text{TST} \) only in types two higher than a type in which the natural numbers are defined: with the usual definition of natural number, there are natural numbers of each type \( n \geq 2 \), and so Quine pairs in each type \( n \geq 4 \).]

The principal application of the Quine pair is that it is the best pair to use in the context of Quine’s set theory “New Foundations”, where it has strong advantages in respect of technical convenience, but we have noted that it could be used in \( \text{ZF} \) and has a nice technical property in relation to the cumulative hierarchy. As we observed above, the baroque nature of the definition of the pair makes no difference once the basic properties are established.

The author has defined a pair which is one type higher than its projections and requires a weaker assumption than Infinity to work: define \((x, y)\) as \( \{\{x', a, b\}, \{x', c, d\}, \{y', e, f\}, \{y', g, h\} \mid x' \in x \land y' \in y\} \), where \( a, b, c, d, e, f, g, h \) are eight distinct objects.

The Kuratowski pair makes sense in \( \text{TST}_3 \), but pairs of type 0 objects are of the highest type 2, so it is useless there for theories of functions and relations: in \( \text{TST}_4 \) the Kuratowski pair would support a full theory of relations and functions on type 0 objects. For the Wiener pair we can say the same things with respect to \( \text{TST}_4 \) and \( \text{TST}_5 \), as one more type is needed. The author’s pair with type displacement one is definable in \( \text{TST}_3 \) but uselessly, as a pair of type 1 objects is of type 2, the highest type. In \( \text{TST}_4 \), this pair can be used to give a full theory of relations and functions on type 1. The picture for the Quine pair is more complicated as the hidden details of an implementation of the natural numbers would also need to be examined: we will not give details here. We will see below that no notion of pair is possible that will work in general to implement a theory of functions and relations in \( \text{TST}_3 \).

### 3 Implementing relations and functions without ordered pairs

In this section we consider ways of implementing relations and functions without using a notion of ordered pair. We work in \( \text{TT}_3 \), the theory of types with three types labelled 0,1,2. The definitions we give could also be used in other theories of course.

**Definition:** A logical relation is a formula \( \phi(x, y) \). A logical function is a logical
relation $\phi(x, y)$ with the property that $\phi(x, y) \land \phi(x, z) \rightarrow y = z$.

The notions of logical relation and function are useful for expressing what it means to implement a relation or function.

We first note that the unordered pair $\{x, y\}$ has some usefulness. A relation can be presented as a set of unordered pairs if it is symmetric. If $(\forall xy, \phi(x, y) \rightarrow \phi(y, x))$ (the logical relation $\phi(x, y)$ is symmetric), then $\phi(x, y) \leftrightarrow \{x, y\} \in \{\{x, y\} \mid \phi(x, y)\}$, so $R_\phi = \{x, y\} \in \{\{x, y\} \mid \phi(x, y)\}$ gives an implementation of symmetric relations $R$ on type 0 objects as set of unordered pairs of type 0 objects (and so as type 2 sets).

More important is the representation of a reflexive transitive relation (quasi-orders) by the set of its segments. This includes both equivalence relations and partial orders as special cases. If $(\forall orders) by the set of its segments. This includes both equivalence relations and type 0 objects (and so as type 2 sets).

The informal idea is that there is no problem with finding $f_\phi(x)$ as long as neither $x$ nor $f_\phi(x)$ belongs to a finite cycle in $\phi(x, y)$. For in this case we can define $f_\phi(x)$ as the unique $y$ such that $c_1(x, y) = c_1(y)$. This fails in two and only two ways: if $c_1(x) - \{x\}$ is not $c_1(y)$ for any $y$, which is equivalent to $x$ belonging to a finite cycle, or if $c_1(x) - \{x\} = c_1(y) = c_1(z)$ for some $y \neq z$, which is equivalent to $y$ belonging to a finite cycle with more than one element. We can determine $f_\phi(x)$ in two additional cases (if a certain convention
is adopted): if $c_\phi(x) = x$, there are two possibilities: either $f_\phi(x) = x$ (x is in a finite cycle of size 1) or x is in the range but not the domain of $f_\phi$. We adopt the convention that $f_\phi(x)$ is defined as x for all x not in the intended domain. Further, if $c_\phi(x) = c_\phi(y) = \{x, y\}$, it is clear that $f_\phi(x) = y$ and $f_\phi(y) = x$.

This motivates the following provisional definition of “function application”.

**Definition:** Let $f$ be a set. Define $c_f(x)$ as the intersection of all elements of $f$ which contain x. We stipulate that $c_f(x) \in f$ for all $x \in \bigcup^2 f$, or $f(x)$ is in all cases undefined.

**Definition:** We say that an element $C$ of $f$ is a cycle in $f$ iff $\emptyset \neq C = c_f(x)$ for each $x \in C$. We stipulate that for each element $A$ of $f$ it is that case either that $A$ a cycle in $f$ or that $A = c_f(x_A)$ for a uniquely determined $x_A$, or that $f(x)$ is undefined in all cases.

**Definition:** As long as for each element $A$ of $f$ it is that case either that $A$ a cycle in $f$ or that $A = c_f(x_A)$ for a uniquely determined $x_A$, we define $f(x) = x$ if $\{x\}$ is a cycle in $f$, $f(x) = y$ if $\{x, y\}$ is a cycle in $f$ with two elements, and otherwise $f(x) = y$, where $y$ is the unique element (if there is one) such that $c_f(x) - \{x\} = c_f(y)$.

**Definition:** We call a set $f$ a simple function iff $f(x)$ is defined for each $x \in \bigcup^2 (f)$.

It is straightforward to establish that any logical function with no cycle of size greater than 2 nor element mapped into a cycle of size 2 from outside is represented by a simple function in this sense, and that all simple functions have these characteristics.

[the paper needs a development of the theory of finitude in $TST_3$ and proofs that cycles in logical functions are finite, etcetera]

This is not our best representation of functions in $TST_3$, but it is already enough to present a complete theory of cardinality of type 1 sets. Suppose that $\phi(x, y)$ defines a logical bijection from $A$ to $B$ (i.e., both $\phi(x, y)$ and $\phi(y, x)$ are logical functions, $\{x \mid (\exists y. \phi(x, y)) = A\}$ and $\{y \mid (\exists x. \phi(x, y)) = B\}$). Observe further that any finite cycle in $f_\phi$ will be in $A \cap B$. Define $\phi^*(x, y)$ as “$\phi(x, y)$ and $x$ is not in a cycle in $f_\phi$ or $x = y$ and $x$ is in a cycle in $f_\phi$”. It should be clear from what goes before that $\phi^*(x, y)$ is definable, that $\phi^*(x, y)$ defines a logical bijection from $A$ to $B$, and that $f_\phi$ has no cycles of size greater than 1. Thus $\phi^*(x, y)$ is represented by a simple function (modified to send elements of $B - A$ to themselves). So we have shown that there is a logical bijection from $A$ to $B$ iff there is a simple function from $A$ to $B$ representing such a bijection (though the latter may not represent the same bijection we started with). Thus we can define $A \sim B$ as “there is a simple function $f$ such that $\bigcup^2 (f) = A \cup B$, for each $a \in A$, $f(a) \in B$, and for each $b \in B$, there is exactly one $a \in A$ such that $f(a) \in B$ [noting that if $b \notin A$ we will also have $f(b) = b$].

Further, we can prove in much the usual way that equinumerousness is an equivalence relation on sets. The usual proof needs to be modified because
the composition of two simple functions may have finite cycles; however, if the
composition is not representable by a simple function it is nonetheless a bijection
and can be modified as above to obtain a simple function which is bijective with
the same intended domain and range. The Schröder-Bernstein theorem can also
be proved in this context. So we obtain an adequate theory of cardinality in
TST,

The development of a theory of cardinality in TST has already been carried
out by the Belgian mathematician Henrard in unpublished work since described
in [?]. His approach was somewhat different though similar in spirit. He repres-
ented bijections using sets of unordered pairs: the intention is that f will be
represented by the set of unordered pairs \{x, f(x)\}.

Let f be a set of unordered pairs with the following properties: if \{x\} \in f,
then no other element of f contains x as an element. No element of \bigcup^2(f)
belongs to more than two elements of f. The closure of any element of f under
the relation of nonempty intersection is called a chain. A chain f has x as an
endpoint if there is one element of f containing x as an element. Prove that a
chain has no more than two endpoints. f represents a bijection from A to B iff
\bigcup^2(f) = A \cup B, A \Delta B is exactly the set of endpoints of chains with more than
one element, and any chain with two endpoints has one in A \setminus B and one in
B \setminus A. This, Henrard’s development, also defines equinumerousness, but is a
bit vaguer about the definition of actual bijections.

We now extend our definition of function somewhat farther, and indeed
about as far as possible, as the last section will suggest.

The basic idea is that the only problem in defining any function lies in the
finite cycles. Let \phi(x, y) be a logical function. Suppose that for each cycle C in
\phi(x, y) with more than one element we can choose two elements C_1, C_2, with
C_1 \in C, C_2 \in C [it is sufficient to choose one element C_1 from each cycle C
and set C_2 = f_\phi(C_1)]. We define \phi^*(x, y) as holding iff x is not equal to any
C_i and \phi(x, y), or if x is equal to some C_i and y = x: in short, we modify the
function by making each C_1 and C_2 a fixed point. It is then evident that \phi^*
is represented by a simple function \phi^*_. F_\phi is the union of \phi^* _ and the set of
cycles in f.

[fix notation in the next paragraph]

We explain how to determine f_\phi(x) from F_\phi and x. First, we point out that
the cycles in \phi can be recovered from F_\phi: a set C in F_\phi is a cycle in \phi iff it
is the union of two disjoint elements of F_\phi (these will be the orbits from f(C_1)
forward to C_2 and f(C_2) forward to C_1). So we can recover f_\phi_. We then define
f_\phi(x) as f_\phi^*(x) if x does not belong to a cycle in \phi or if x belongs to a cycle and
f_\phi^*(x) \neq x. If x belongs to a cycle and f_\phi^*(x) = x, define f_\phi(x) as the unique
y which belongs to the other element of the 2-partition of the cycle to which x
belongs and is not an image under f_\phi^*.

The limitation on this definition is the possibility of defining the choice of
representatives C_1, C_2. If there is a definable linear order of the universe then
we will always be able to do this (and under weaker conditions: the ability to
choose two unordered elements from any disjoint collection of finite sets is what
is needed). We will see in the next section that there cannot be a way of doing this which will work in all cases.

It is further useful to note that if there is a definable linear order on the universe there is a way to define an ordered pair. Note that the existence of a definable linear order is expressible in $TST_3$ in terms of the existence of a certain kind of set, since any definable quasi-order is representable by a set. So, suppose that a linear order $\leq$ has been defined. Choose ten distinct elements $a, b, c, d, e, a', b', c', d', e'$ and define $(x, y)$ as \{x, y\}$\Delta$\{a, b, c, d, e\} if $x \leq y$ and \{x, y\}$\Delta$\{a', b', c', d', e'\} otherwise. [Choosing ten distinct objects is unproblematic because if the universe is finite it is well-ordered and so the treatment of functions above works]. It is straightforward to verify that this satisfies the defining property of an ordered pair. Further, since $(x, y)$ is type 1 for $x$ and $y$ of type 0, we can define relations and functions in the standard way as type 2 sets of ordered pairs. So there is no reason to use the pair-free treatment if even a modest amount of choice is available.

In the choice-free context, there are some situations where we can see that the function definition given above will be useful. For example we get an adequate treatment of isomorphism between well-orderings (order type) in the choice free context, because the fields of isomorphisms between well-orderings have well-orderings on them (obviously!) which can be used to make choices from finite cycles.

4 There is no complete function definition in three types

In this section we argue that we essentially cannot do better (in the absence of choice) than we did in the previous section, using a Frankel-Mostowski technique.

We start with the natural model of $TST_4$ with type 0 implemented using the standard natural numbers. We consider type 0 as partitioned into the sets $A_n = \{3n, 3n + 1, 3n + 2\}$. We define $f_n(x)$ as sending $3n$ to $3n + 1$, $3n + 1$ to $3n + 2$, $3n + 2$ to $3n$, and fixing all natural numbers. We define $G$ as the smallest group containing all maps $f_n$. We restrict our attention to the collection of all sets with finite support in the group $G$: this will be a model of $TST_4$ in which choice fails rather badly.

We define $x \not\approx y$ as holding iff $y \neq x \land (\exists n. f_n(x) = y)$. Notice that $f$ is a function and invariant under $G$. The function in type 3 implementing $f$ as usual using Kuratowski pairs exists in the new model. The new function definition in type 2 fails completely because there cannot be a choice set from the 3-cycles in the function $f$. Why do I conclude that there can’t be any function definition scheme in $TST_3$ using this construction?

It’s in the slides and I do not fully understand it...

The basic claim is this: there can be no formula $\phi(x, y, f^*)$ which means $y = f(x)$, because in all but finitely many cases $(x, y)$ will have the same relation
as \((y, x)\) to the parameter \(f^*\) in case \(y = f(x)\). What is the basis for this claim? Notice that \((x, y)\) and \((y, x)\) do exist as sets.

For any set \(X\) of type 1, and all but finitely many cycles \(\{a, b, c\}\), all of the elements of the orbit or none of them belong to \(X\).

For any set \(Y\) of type 2, and all but finitely any cycles \(\{a, b, c\}\), and each set \(y\) not meeting the cycle, either all of \(y \cup \{a\}\), \(y \cup \{b\}\) and \(y \cup \{c\}\) belong to \(Y\) or none of them, and either all of \(y \cup \{a, b\}\), \(y \cup \{b, c\}\) and \(y \cup \{a, c\}\) belong to \(Y\) or none of them.

Ah, the real point: any set which has finite support in \(G\) in \(TST_3\) also has finite support in the group which permutes each cycle freely. This is not true in \(TST_4\).

Suppose \(\phi(f^*, \{\{a\}\}, \{a, b\}\})\) holds. Then in all but finitely many cases \(\phi(f^*, \{\{b\}\}, \{b, a\}\})\) holds as well (because free permutation (not just oriented permutation) of the cycles leaves everything invariant). Thus \(\phi(f^*, \{\{a\}\}, \{a, b\}\})\) cannot implement \(a = f(b)\).

However, in the model of \(TST_3\) obtained by dropping the top type of the given Frankel-Mostowski model of \(TST_4\), we can use the binary predicate \(y = f(x)\) to define sets: it is a logical relation in this context. Clearly this is a logical function. If there were any general way to implement logical functions in \(TST_3\) it could be used to implement this particular logical function, and we have just shown that this one cannot be implemented.