COMPUTABLE REALS AND CONSTRUCTIBLE SETS

JAMES TAYLOR
U4395303

1. Computable Reals

The purpose of this section is to describe computable reals, then prove some of the main properties of computable reals.
First we will define a computable real. Informally a computable real is a number that has some finite representation. In the physical sciences these numbers are of great interest because within the physical universe any number that can actually have some physical meaning will only contain a finite amount of information. This is equivalent to it having a finite representation. The computable reals are differentiated from the normal Cauchy reals because within the Zermelo-Fraenkel set theory with the Axiom of Countable Choice there exist numbers that contain an infinite amount of information.
We also define informally a computable function as a function f which could in principle be calculated using a mechanical calculation device given finite but unbounded time and storage space. In the language of computer science, this is to say that the function has an algorithm. An algorithm is taken to consist of a finite set of instructions that are described by a finite number of symbols which could in principle be computed by a person with pencil and paper and is strictly deterministic. Computable functions are vital to any formal description of computable real numbers.

Formal definitions
Before giving a formal definition of the computable reals we formally define a computable function. By Church's thesis, an function is effectively computable if it is partial recursive, which is to say that it can be defined from the initial functions

(i) The zero function $\mathbf{0}(n) = 0, \quad \forall n \in \mathbb{N}$
(ii) The successor function $n' = n + 1, \quad \forall n \in \mathbb{N}$
(iii) The projection function $U_i^k(\mathbf{m}) = m_i, \quad k \geq 1, \ i = 1, \ldots k$

using a finite number of applications of

(i) Substitution given by $f(\mathbf{m}) = g(h_0(\mathbf{m}), \ldots, h_l(\mathbf{m}))$
(ii) Primitive recursion given by $f(\mathbf{m}, 0) = g(\mathbf{m})$

(iii) $f(\mathbf{m}, n+1) = h(\mathbf{m}, n, f(\mathbf{m}, n))$

(iv) The μ-Operator given by $f(\mathbf{n}) = \mu m[g(\mathbf{m}, m) = 0]$ where $\mu m[g(\mathbf{m}, m) = 0] \Leftrightarrow g(\mathbf{n}, m_0) = 0$ and $(\forall m < m_0)[g(\mathbf{n}, m) \neq 0]$

Note that the μ-Operator is a search operation; it says to compute $g(\mathbf{n}, 0), g(\mathbf{n}, 1), \dots$ until we find $g(\mathbf{n}, m_0) = 0$. Then m_0 is the desired value.

Using this definition of computable functions it is easy to construct some simple example functions including the constant functions, addition and multiplication.

We are now able to properly define the computable real numbers. Formally a real number a is a computable real if it can be approximated by a computable function $f : \mathbb{N} \to \mathbb{Z}$ such that $f(n) = k$ where

$$\frac{k-1}{n} \leq a \leq \frac{k+1}{n} \tag{1}$$

There are three similar definitions which are also equivalent

- There exists a computable function $f : \mathbb{Q}^+ \to \mathbb{Q}$ which maps ϵ to r such that

$$|r - a| < \epsilon \tag{2}$$

- There is a computable sequence of rationals q_i which converge to a and such that

$$|q_i - q_{i+1}| < 2^{-i} \tag{3}$$

- There exists a computable Dedekind cut D converging to a, where a computable Dedekind cut is a computable function $D : \mathbb{Q} \to \{TRUE, FALSE\}$ such that

  $\exists r$ s.t. $D(r) = \text{true}$

  $\exists r$ s.t. $D(r) = \text{false}$

  $(D(r) = \text{true}) \wedge (D(s) = \text{false}) \Rightarrow r < s$

  $D(r) = \text{true} \Rightarrow \exists s > r$ s.t. $D(s) = \text{true}$

  The function D is unique for each $a \in$ computable reals.

Some of the basic properties therefore associated with the computable reals are that it is closed under addition, subtraction, multiplication and division. In fact, the computable reals are closed under any operation with a finite algorithm as if f is a computable function such that $f(\mathbf{0}) = c$ for c a computable real, and g is any finite operation, then $f(g(\mathbf{0}))$ will still be some finite algorithm, and so the output of $f(g(\mathbf{0}))$ will be computable.

Incompleteness of the Computable Reals

At first glance it would appear that the computable reals are complete as any Cauchy sequence $\{a_n\}$ of computable reals must converge to a computable real. The flaw in this way of thinking is that while each a_i will have a finite algorithm, there are a countably infinite number of such a_i's. This means that the sequence itself may not be computable, even if every element of the sequence is. If there is indeed no finite representation of the sequence, that is, no finite algorithm for generating the sequence, then the sequence would converge to a non-computable real. For example, if we let $c \in \mathbb{R}$ be non-computable such that the decimal expansion of c can be written $c = c_0.c_1c_2c_3\ldots$ then the sequence $\{a_n\} = \{a_i | a_i = c_0.c_1\ldots c_i\}$ converges to a non-computable real, but each element a_i is clearly computable as it is finite. Therefore the computable reals are incomplete. Note that we have assumed that there are reals that are not computable. This will be shown in later in the paper without reference to the incompleteness proof.

The order relation on the Computable Reals

The order relation on the computable reals is necessarily non-computable. If we let A be the algorithm which gives as output the computable real a, note that here we are using the second definition of computable real, then there is no computable function g which gives

$$g(A) = \begin{cases} 1 & a > 0 \\ 0 & a \leq 0 \end{cases}$$

Suppose the first N ϵ-approximations given by A are 0, and after this are given by some value greater than 0. Then if N is unknown but very large, it is not clear how long we must wait before the function outputs an ϵ-approximation which forces a to be positive. As the function g must produce an output in some finite time, after the first M approximations it will output $g(A) = 0$. But suppose then that $M < N$, then in fact $a > 0$. Therefore the order relation function is non-computable.

Countability of the Computable Reals

We now discuss the result attributed to Turing which is perhaps the most important feature of the computable reals. This result is that the cardinality of the computable reals is $\aleph_0$, that is, the computable reals are countable. At first glance, it might be supposed that the diagonal argument, used to show the reals are denumerable, could also be used to show that the computable reals are denumerable. The standard diagonal argument applied to computable reals essentially states, suppose by way of contradiction, that the computable sequences are countable. Then we can let I_n be the n'th computable sequence, and

let $I_n(m)$ be the m'th element of I_n. Let J be the computable sequence with $J(n) = 1 - I_n(n)$. Since J is computable, there exists some $k \in \mathbb{Z}$ such that $1 - I_n(n) = I_k(n)$ for all n. If we consider $n = k$, then $1 = 2Y_k(k)$. But then 1 is even, which is a contradiction.
The fallacy in this argument, here attributed to Turing, is the assumption that J is computable. For normal sequences this is not an issue as J is indeed a sequence, but it can be shown that it is incorrect to suppose that J is computable. This is because although there is a finite representation for each I_n, and thus a finite representation for $I_n(n)$, this does not imply there is a finite representation for J. There is in fact no possible finite representation as the problem of enumerating computable sequences is equivalent to Hilbert's halting problem . In the language of computer science, the halting problem is the question of determining whether a program will ever produce an output, or will simply continue forever without producing further output. If there was a computable method by which we could enumerate the computable sequences, then J would indeed be computable, but as we cannot enumerate the computable sequences by finite means, therefore the computable reals are not uncountable, that is, they are countable. As the Cauchy reals are uncountable, this gives us that most reals are non-computable. This result was used previously to show that the computable reals are incomplete.

Axiom of Choice and the construction of the Computable Reals

It is at this time that we wish to remove the Axiom of Countable Choice. This is the weakest of the choice axioms, in that it is implied by all stronger choice axioms, and is the statement that:

If S is a countable, disjoint set of non-empty sets, then there is a subset T of the union of S which has exactly one element in common with each member of S.

More simply put, this is equivalent to stating that given a countable set S we can arbitrarily choose elements of S to make any subset. In the case of the construction of the reals, this axiom is vital for the construction of the non-computable reals as, by definition, these cannot be computed by non-arbitrary means.
Mathematically and philosophically the axiom of choice is in some sense the least accepted of all the axioms in the Zermelo-Fraenkel system. Philosophically and physically the axiom is problematic as within the

finite physical universe it is not possible to have a process which necessarily requires an infinite amount of information. Mathematically the axiom is problematic in that it allows the construction of results that seem in direct opposition to what is expected or indeed realistic. The most famous example of this is the Banarch-Tarski Paradox which, simply put, states that it is possible to decompose the unit ball in $\mathbb{R}^3$ into five disjoint pieces, then recombine them using only translations and rotations into two balls of the same size. This is of course unacceptable if we wish to use mathematics as a model for the real world. As a result of this, mathematicians prefer to avoid the use of the axiom of choice is possible, and if this is not possible, then an attempt is made to use the weakest axiom of choice possible.

With this in mind, we decide to abandon the axiom of choice in the construction of the real numbers. So we construct our set $\mathbb{CR}$ as the convergents of Cauchy sequences of elements in $\mathbb{Q}$. So, for simplicity $\mathbb{Q}$ is taken to be the set of elements $\{\pm p/q | p, q \in \mathbb{N},\ \mathrm{GCD}(p, q) = 1\}$, where GCD is the greatest common divisor of two natural numbers. This formulation of $\mathbb{Q}$ has a natural bijection with the normal formulation of $\mathbb{Q}$, however for the purposes here it is more intuitive. If we let $\{a_n\}$ be a Cauchy sequence of elements in $\mathbb{Q}$ then we have that the sequence itself must be computable. To show this, suppose by way of contradiction that the sequence is not computable, that is, there exists no finite function which allows us to determine the elements of $\{a_n\}$. Then the elements of $\{a_n\}$ must have been chosen arbitrarily, but this requires the axiom of countable choice which we have decided will not apply. Therefore there must exist a finite function which allows us to determine the value of each $a_i \in \{a_n\}$. When such a function exists, we call the sequence computable.

We then have a Cauchy sequence with elements in $\mathbb{Q}$ such that each element of the sequence is determined by a computable function. But then we have that the convergent of this sequence, which will be in $\mathbb{R}$, is completely determined by this computable function. Therefore the convergent will be a computable real. As this holds for any sequence of elements in $\mathbb{Q}$, the set of convergents of all sequences will be at least a subset of the computable reals, and from Equation 2 we see that any computable real can be expressed in this fashion, albeit not uniquely. Therefore the set $\mathbb{CR}$ is in fact the set of computable reals.

Completeness of the Computable Reals

We now wish to show that, without the axiom of choice, the set of computable reals are, in fact, complete. Now, as usual a complete set is a set for which any Cauchy sequence $\{a_n\}$, that is, $\forall \epsilon > 0,\ \exists N_\epsilon \in \mathbb{Z}$ such that $m, n \geq N_\epsilon \Rightarrow |a_m - a_n| < \epsilon$, converges to a computable real. This proof works much the same as the previous construction of the computable reals. As we do not have the axiom of choice, any sequence will be computable, that is, there is a finite function which determines each element of $\{a_n\}$. But then the convergent will be determined by the computable function which generates $\{a_n\}$. This gives us that the convergent is computable, and so the Cauchy sequence converges to a computable real. Therefore the computable reals are complete if we discard the axiom of choice.

Final remarks

In this section we have attempted to give a reasonably rigorous definition of the computable reals and some of their main properties as these appear in the context of standard analysis. We have then gone on to remove the Axiom of Choice to allow a natural construction of the computable reals. With this construction we find that some of the fundamental properties of the computable reals change, primarily the completeness. However, as the proof of the countability still holds we do not have the issues associated with non-measurable sets which leads to the Banarch-Tarski paradox. However, we also recognise that this construction is not sufficient for pure mathematics as some fundamental results such as Tychonoff's Product Theorem do not hold. For applications in fields of science which deal only with finite applications, such as computer science, we believe that this construction may be adequate as a model for the reals. With this in mind, in the next section we build a simplified model of set theory for application in computer science.

2. Constructible Sets

In this section we discuss an alternative, constructivist set theory which may be useful for application in computer science and other physical fields. The motivation for this is basically that the Zermelo-Fraenkel system consists of a large number of separate axioms which seem to be somewhat arbitrary. This seems unusual given that the Zermelo-Fraenkel set theory is the basis for almost all mathematics, and through this, for much of the rigorous, formal science used today.

The Inclusion Function and the Natural Numbers
With this in mind we consider the original description of a set as a collection of elements. Then we define a set to be a collection of other sets, together with the inclusion operation whereby previously constructed sets can be 'included' into a set. This inclusion function can be more formally described as

Definition 1. *For any computable function f which assigns a boolean value to each previously constructed set, there exists a set $S = \{A|f(A) = 1\}$. Then f is the <u>inclusion function</u> which <u>generates</u> S.*

We also assume that it is possible to define an infinite class of computable functions, so long as each is computable. It is through this process that we are able to construct infinite sets. Throughout, we will use the nomenclature that A is a previously constructed set, and S is the set we are currently constructing.
Now, using nothing but this inclusion function we can go about constructing some simple sets. We first show that the empty set must exist. We start with no previously constructed sets. Then the inclusion function f has no sets A such that $f(A) = 1$; therefore $S_0 = \{\}$, that is, the empty set exists.
We will now go about constructing the natural numbers. Let $f(A) = 1,\ \forall A$. Then we have

$$S_1 = \{\{\}\}$$
$$S_2 = \{\{\}, \{\{\}\}\} = \{S_0, S_1\}$$
$$S_3 = \{S_0, S_1, S_2\} = \{\{\}, \{\{\}\}, \{\{\}, \{\{\}\}\}\}$$
$$S_n = \{S_0, S_1, \dots, S_{n-1}\}$$

This is precisely the definition of the natural numbers in normal set theory, so if we apply this particular inclusion function up to transfinite induction then S_ω, where ω is the first transfinite ordinal, is precisely the set of natural numbers. So we have successfully constructed the natural numbers within this constructivist framework.
We now wish to derive some simple operations which follow from the

inclusion function. First we will show that the inclusion function implies an 'exclusion' function which removes an element from a set. So, let S be some set. Then from the definition of a set, there exists a computable function such that $S = \{A | f(A) = 1\}$. Let g be a function such that

$$g(A) = \begin{cases} f(A) & A \neq A_0 \\ 0 & A = A_0 \end{cases}$$

Then g is the inclusion function which corresponds to the set S/A_0, and so it is possible to remove an element from a set, that is, 'exclusion' is possible.

Equivalence

With exclusion we can now define a 'size' of sets. We call the complexity of a set the number of single inclusions necessary to create the set, minus the number of single exclusions. That is,

Definition 2. *For $S = \{A_1, A_2, \ldots, A_n\}$, then the* <u>*complexity*</u> *of S is given by*

$$\|S\| = \sum_{i=1}^{n} \|A_i\| + n$$

For example, the sets corresponding to the natural numbers have $\|S_0\| = 0$, $\|S_1\| = 1$, $\|S_2\| = 3, \ldots,$

$$\begin{aligned} \|S_n\| &= n + \|S_{n-1}\| + \ldots + \|S_1\| \\ &= n + (n-1) + 2\|S_{n-2}\| + \ldots + 2\|S_1\| \\ &= n + \sum_{i=1}^{n-1} i - \sum_{i=1}^{n-1} i^2 \\ &= n + \frac{n^2(n-1)}{2} - \frac{n(n-1)(2n-1)}{6} \end{aligned}$$

using the formula for square pyramidal numbers.
For S containing a countably infinite number of elements, simply let $n = \omega$ and for all calculations use ordinal arithmetic. When S is infite, it is usually only necessary to be able to compare the complexity of sets rather than to find the actual value.

We can now define an equivalence relation such that the constructed sets S and T are equal if and only if they contain the same elements and if $\|S\| = \|T\|$. We note that this is indeed an equivalence relation as reflexivity, symmetry and transitivity hold as these hold for sets normally, so the first condition is satisfied, and they hold for ordinal arithmetic, so the second condition is satisfied. We have chosen to describe equality in this way so as to avoid some of the paradoxes of naive set theory.

Basic Results

We now will attempt to prove some simple results within the framework.

Theorem 1. *Sets cannot contain themselves.*

Proof. This result comes from the definition of equality of sets, and is in fact one of the main reasons equality was defined in such a manner. Suppose that $S = \{A_0, A_1, \ldots\}$. Then suppose there exists some i such that $A_i = S$ then $\|\{A_0, A_1, \ldots, A_{i-1}, S, A_{i+1}, \ldots\}\| > \|\{S\}\| = \|S\| + 1$. But this is a contradiction as $S = \{A_0, A_1, \ldots, A_{i-1}, S, A_{i+1}, \ldots\}$. Therefore there cannot exist a set which contains itself. Note that this proof relies heavily on the fact that complexity is discussed in terms of ordinal numbers and ordinal arithmetic, and not cardinal arithmetic. □

The main possible objection to this theorem is the set $S = \{S\}$ which appears as $S = \{\{\{\ldots\}\}\}$, as this set appears to contain itself. However we note that if $\|S\| = \omega$ then $\|\{S\}\| = \omega + 1$. But as $S = \{S\}$ therefore $\|S\| = \|\{S\}\|$, but $\omega \neq \omega + 1$. Therefore $S \neq \{S\}$. This is not really surprising as we have constructed S using S, which is not within the definition of a set. Therefore $S = \{S\}$ is not a constructible set. Note that there is indeed a constructible set $S = \{\{\{\ldots\}\}\}$ but that it is incorrect to write this as $S = \{S\}$.

Corollary 1. *There is no set of all sets.*

Proof. Suppose by way of contradiction that there exists a set of all sets, call it $V = \{A_0, A_1, \ldots\}$. That is, we have constructed every set which it is possible to construct, and taken the inclusion of all such sets. Then as V is a set, it must be in the set V, but this is impossible by the theorem. Therefore there exists no set of all sets. □

This is the answer to Cantor's paradox, and is in fact the same answer as that given by ZF. Cantor's paradox states that there is no largest cardinal, that is, there is no set of all sets. This was not seen by Cantor as a paradox, merely as a merely as a statement about the nature of sets. In ZF this position is agreed, there is indeed no set of all sets. Here we have shown that this result carries into constructible sets.

This corollary also gives a solution to the Burali-Forti paradox, which shows that there can be no largest ordinal. The Burali-Forti paradox states that the set of all ordinal numbers Ω must itself be considered an ordinal number, and so it has a successor given by $\Omega + 1$. But then $\Omega < \Omega + 1 \leq \Omega$ which is a paradox. This is solved for constructible sets as we cannot construct the set of all ordinals, merely the set of all ordinals which have been previously constructed. This circumvents the problem completely, however has the unfortunate property of not being able to construct the set of ordinals.

Some ZF Axioms

Now we will attempt to derive some of the axioms of Zermelo-Fraenkel set theory. We begin with the Axiom of Elementary Sets, which states

Theorem 2. *Axiom of Elementary Sets*
There is a set with no elements, called the empty set, and for any two sets a and b, there exist sets $\{a\}$ and $\{a, b\}$

Proof. We have already proved the existence of the empty set. So, suppose a and b are any previously constructed set. Then define the function

$$f_a(A) = \begin{cases} 1 & A = a \\ 0 & A = b \end{cases}$$

and

$$f_{a,b}(A) = \begin{cases} 1 & A = a \text{ or } A = b \\ 0 & A \neq a \text{ and } A \neq b \end{cases}$$

Then f_a generates the set $\{a\}$, and $f_{a,b}$ generates the set $\{a, b\}$. $\square$

Next we work with a modified version of the axiom of union. Normally this axiom is not restricted to finite sets, however it is deemed necessary for constructible set, so:

Theorem 3. *Axiom of Union*
If S is a finite constructible set, then the union of S is a constructible set.

Proof. If we let $S_A = \{A_1, \dots A_n\}$ be a set, where $A_i = \{a_{i1}, \dots, a_{im_i}\}$ then for A_i to be previously constructed set we must have that a_{ij} is also a previously constructed set. We also have that the function $f_{A,i}$ which generates A_i is computable. Then we have that the function $f_{S,a}$ for which

$$f_{S,a}(A) = \begin{cases} 1 & A = a_{ij} \text{ for some } i, j \in \mathbb{N} \\ 0 & A \neq a_{ij} \text{ for some } i, j \in \mathbb{N} \end{cases}$$

Then $f_{S,a}$ certaintly generates the union of S_A and must be computable as it is the finite union of computable functions and thus has finite representation. Unfortunately, we cannot prove this result for infinite S_A, as although each $f_{A,i}$ is computable, an infinite collection of them need not necessarily have finite representation. $\square$

Russel's Paradox

One of the primary motivations for the development of axiomatic set theory was the existence of paradox within naive set theory. Cantor's paradox has been discussed earlier. Another of the primary paradoxes of naive set theory is Russel's paradox which essentially states that the set S such that $S = \{A | A \notin S\}$. We show that this paradox does not hold in this framework by supposing by way of contradiction that there is a computable function f which generates S. Then f has assigned a boolean value to each previously constructed set A. But the set S is such that each set A is both in the set, and not in the set, therefore must have boolean value 1 and 0. But the function f assigns only one such value to each A. Therefore we have a contradiction, and so Russell's paradox does not hold.

The Computable Reals

We now return to constructing some important sets. First we will construct the set of rational numbers. We want a rational number to be an ordered pair of natural numbers, together with a sign value. That is, some rational number r can be written $r = (\pm 1, p, q) = \pm p/q$ for $p, q \in \mathbb{N}$, $q \neq 0$, $\mathrm{GCD}(p, q) = 1$. So we say for $r > 0$, $r = \{S_p, \{S_q, \{\}\}\}$ and for $r < 0$, $r = \{S_p, \{S_q, \{\}\}, \{\{\{\}\}\}\}$. We then have that these sets are indeed computable as they are a finite collection of finite sets. Then the set of all such r is indeed the set of rational numbers. Note that the formulation given has been chosen so that each value of r has a unique representation, and that this requires the condition $q \neq 0$.

We are now in a position to return to the original problem of computable reals. Recall that we originally defined computable reals to be computable sequences of rationals. As we now have a set of rationals we can do much the same construction tog ive the computable reals in constructible sets. First we must make sure that we can in fact construct an ordered infinite sequence. To do this we define the counting elements $\bar{1} = \{\{\{\}\}\}$, then $\bar{2} = \{\bar{1}\}, \ldots \bar{i} = \{i \bar{-} 1\}$. These elements are clearly constructible as they are finite, and are distinct from every rational. Then the set $\{\{A_1, \bar{1}\}, \{A_2, \bar{2}\}, \ldots\}$ can be viewed as an ordered infinite sequence. If we restrict A_i to the rationals then this becomes an infinite sequence of rationals, call this a_n. Now suppose a_n is a Cauchy sequence which converges to a. Then as the sequence is in fact a set, there must be a computable function which generates the set. Then the value a of the convergent is completely determined by the computable function which generates the set a_n. But this is precisely the definition of a computable real given in Section 1. Thus we have used our simplified version of set theory to construct the computable reals.

Final Remarks

In this section we have attempted to provide an alternative set theory, one based on constructivist principles. This theory is in many ways similar to Type theory, as the hierarchy of construction gives a natural way so that sets cannot be defined in terms of themselves, the essence of type theory. We have attempted to expound upon this theory, giving some preliminary results as well as constructing some important sets. Hopefully the links with computability theory are clear, as sets are here defined to be generated by computable functions. As such this model may be useful in computer science, where only the computable is of any real relevance.

References

[1] S. B. Cooper, *Computability Theory*, CRC Press, Florida, 2003.
[2] E. Borger, *Computability, Complexity, Logic*, Elsevier Science Publishers, New York, 1989.
[3] G. H. Moore, *Zermelo's Axiom of Choice: Its Origins, Development, and Influence*, Springer-Verlag, New York, 1982.
[4] T. J. Jech, *The Axiom of Choice*, North-Holland Publishing Company, Amsterdam, 1973.
[5] G. Cantor, *Contributions to the founding of the theory of Transfinite Numbers*, 1895, Translated and Re-printed by Dover Publications, New York.
[6] A. M. Turing, *On Computable Numbers with an Application to the Entscheidungsproblem* 1936.
[7] B. Andrews, *Introduction to Mathematical Analyis*, Course Notes: ANU, MATH2320, 2007.
[8] P. T. Johnstone, *Notes on Logic and Set Theory*, Course notes, 1986.