Cantor’s Problem
Charles Sauerbier
November 2009

What is born on Sacred Cows is as much religion as is any theology.

Abstract: In 1891 a paper by Georg Cantor was published in which he addressed the relative cardinality of two sets, the set of integers and the set of real numbers, in effort to demonstrate that the two sets were of unequal cardinality. This paper offers a contrary conclusion to Cantor’s argument, together with implications to the theory of computation.

In 1891 a paper by Georg Cantor was published in which he addressed the relative cardinality of two sets, the set of integers and the set of real numbers, in effort to demonstrate that the two sets were of unequal cardinality. In that paper Cantor used what is referred to as a Diagonalization method. The essential aspect of Diagonalization and Cantor’s argument has been represented in numerous basic mathematical and computational texts with illustrations. This paper offers a contrary conclusion to Cantor’s argument, together with implications to the theory of computation.

It is important to acknowledge in any examination of Cantor’s argument that such is with benefit of hindsight provided by time and the expansion of human knowledge. It is equally important to keep in mind that Georg Cantor, in his time, was at the forefront of the emergence of set theory. The Diagonalization argument is one of Cantor’s works that has historically not been without its critics.

This paper makes use of basic principles and concepts that are presented in basic mathematical and computer science textbooks commonly accessible in university libraries.\footnote{See (Boas & Boas, 1997), (Cohen, 1966), (Hopcroft, Motwani, & Ullman, 1979), (Kozen, 1997), (Kozen, 2006), (Sipser, 2005), (Brioda & Williamson, 1989), (Ross & Wright, 1985).} The fundamental principles, arguments, and related materials are also widely available in approachable form via the internet.\footnote{\url{http://mathworld.wolfram.com/} is as of this writing a reliable source of reference where library access may be limited.}

Preliminaries:
We first dispose of misconceptions arising of terms. The word "infinite" is a fabrication of mental ether - a vacuous expression lacking mathematical precision - with many mystical connotations.

\footnote{See (Boas & Boas, 1997), (Cohen, 1966), (Hopcroft, Motwani, & Ullman, 1979), (Kozen, 1997), (Kozen, 2006), (Sipser, 2005), (Brioda & Williamson, 1989), (Ross & Wright, 1985).}
The property of sets with which we concern ourselves here is such that all sets belong to one of
two classes: (a) computationally-countable and (b) computationally-uncountable. For obvious
reasons we attach to these terms mental concepts that are as precise as our meager mastery of
language and words permits. We have prefaced the words “countable” and “uncountable” with
the word “computationally” as the prior have meanings commonly found in textbooks that are
inconsistent with our meaning here. We define a set to be computationally-countable if once
initiated, producing a monotonic sequence of integers, the process of counting - whatever that
might be - at some future time is assured to cease, and at which time all elements are account
for by an exclusive (i.e. 1-1) association with exactly one integer. We define a set to be
computationally-uncountable where once initiated, exclusive association being constant, the
process of counting does not cease.

We will use the term “denumerable” where we mean to infer the more conventional concept of
“countable” meaning an injection or bijection with the set of integers.

While our discussion may revolve around sets, in truth we discuss representations of sets in
mathematics. Regardless of whether the sets we discuss are of things tangible or intangible in
nature, any discussion (mathematical or otherwise) makes it necessary that all elements within
the set have some manner of representation expressed in symbolic form. Effectively, there
must be a 1-1 correspondence between the elements of a set and some representation of the
elements. The means conventionally used is association of set elements with representative
symbols or strings of symbols.

**Principal Set:**

The principal set around which our discussion will revolve we define as $S = \{ s \mid s \in \{0, 1\}^* \}$; the
set of all strings composed of elements from the set containing the symbols 0 (zero) and 1
(one). How many elements are in the set $S$? It is not unreasonable to conclude there exists at
least one element in $S$ that has the property of having length at least as great as all other
elements of $S$, to which we will assign the symbol sup$_S$. It is also not unreasonable to conclude
there exists at least one element in $S$ that has the property of having a non-zero length that is
not greater than that of any other element of $S$ having a non-zero length, to which we will
assign the symbol inf$_S$. If we construct a string $s'$ by appending (i.e. concatenating) the string
to which referred by inf$_S$ to either end of the string to which referred by sup$_S$ is the resulting
string $s'$ in $S$? By our definition of $S$ the resulting string must be in $S$. What is the length of $s'$? It
obviously must be as long as length-of-sup$_S$ + length-of-inf$_S$. Such makes $s'$ longer than
sup$_S$, in contradiction of our assumption sup$_S$ was at least as long as all other elements in $S$.
So sup$_S$ supplants sup$_S$. And we can then repeat the concatenation process ad infinitum.
Independent of whatever the cardinality of $S$ was presumed to be when we began, this process
will “count” our discovery of additional elements without ceasing. Such would seem to
reasonably assure the set S is computationally-uncountable; regardless any initial presumptions as to its cardinality.

**Relative Cardinality of Sets:**

Are there any well known sets in mathematics for which an injection into some set contained in S exists? The set Z of all integers we know to be - one and all - expressible as binary strings; that is, of strings composed of the symbols 0 (zero) and 1 (one). It also follows from our collective common definition of the equivalent representation of binary strings that S can be placed into 1-1 and onto correspondence with Z (i.e. a bijection).\[^{[3]}\] The bijection between Z and S is well covered by most relevant basic mathematical and computer science theory texts.\[^{[4]}\] In consequence of the bijection between Z and S it follows from the Schroeder-Bernstein Theorem that |Z|=|S|. It then follows that S is denumerable by definition.

A question likely more interesting is whether there exists any set for which no injection into S exists? Assume some set exists for which no injection into S exists. Let that set be represented by the symbol Q. We first partition all plausible sets into two disjoint subsets: Q1 and Q2. Let Q1 be those sets for which representation of elements is derived from a computationally-countable set of symbols. Let Q2 be those sets for which representation of elements is derived from a computationally-uncountable set of symbols.

Let Q′ be some set in Q1 defined above with representation of elements of Q′ derived of strings of symbols in some computationally-countable set P′. The symbols in P′ can be placed into a 1-1 correspondence with some set of integers in Z, thus in 1-1 correspondence with some set of elements of S. Then by simple substitution rewriting of their representation all elements of Q′ are mapped to elements of S. The consequence of said mapping is at least an injection from Q′ to a set contained in S. It follows then that for any set Q′ that |Q′| is not greater than |S|.

Let Q″ be some set in Q1 defined above with representation of elements of Q″ derived of strings of symbols in some computationally-uncountable set P″. By applying Cantor’s Diagonalization method, having established for any subset of P″ an injective assignment from S one can obtain yet another element in S to assign the next yet unassigned element in P″, using the assignments thus far made to assure the mapping is a proper injection (i.e. 1-1). The injection of the set of symbols into S, while not computable in total, arguably exists. It then follows of the same principles as held for Q′ above that S contains at least one set into which an injection from Q″ is formed through simple substitution rewriting of the symbols of each string.

---

[^{[3]}]: Note that the integers may also be represented by the finite set of symbols {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -} and {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -}.

[^{[4]}]: Conventionally shown within discussions of “bijection base-k enumerations” that are forms of k-adic numeric representations. See (Knuth, 1981) for discussion of general topic of positional numbers.
representation. This is not to say that the injection of $Q''$ into $S$ is necessarily readily or easily computable, only that means exists to establish its existence.

It can therefore be concluded that there exists no set $Q$ having cardinality greater than the cardinality of $S$, as any set $Q$ can be shown to be either in $Q_1$ or $Q_2$ as defined above, and thus “contained” by correspondence with a set contained in $S$ through an injection into $S$. The later is a result of considered importance to the conclusions that Cantor’s Problem does admit.

Now to the essential question of whether $R$ has an injection into any set contained in $S$. Since $R$ is represented by strings over the set $P' = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -, \}$, a computationally-countable (finite) set, it can be concluded that $S$ contains at least one set into which an injection from $R$ exists because $R$ is in $Q_1$ as later was defined above. Such a conclusion contradicts Cantor’s conclusion as to the relative cardinality of $R$ and $Z$; given that it has been shown that $Z$ has a bijection to $S$ by constructive proof and $R$ has at least an injection into $S$. If we assume that $R$ exists in a bijection relation to $S$, such grants at best $|R| = |Z|$. The latter gives definitive evidence that Cantor’s Continuum Hypothesis was correct in as much as no “infinite set” can exist having cardinality between $|Z|$ and $|R|$ where $|Z| = |R|$.

Cantor, in his paper and as delivered by other authors over time, supported by illustration which is unnecessary here and not represented, put forth the Diagonalization argument by constructing strings. The representations we have commonly observed used by other authors in presenting the argument has used strings formed of two letters $\{m, w\}$. It is evident from the presented arguments that the set of consideration is the set $C = \{c \mid c \in \{m, w\}^* \}$. The set $C$ is contained in $Q_1$, as defined above. By direct substitution of symbols $m \rightarrow 0$ and $w \rightarrow 1$, any string in $C$ can be rewritten to a string in $S$. Given that there does not exist in $C$ two strings $c_1, c_2$ such that $c_1 = c_2$, the rewriting of strings in $C$ produces by definition a bijection with $S$. This results in a contradiction for two reasons. The bijection with $S$ is such that $C$ is denumerable by the same means with which $S$ is placed into bijection relation with $Z$. While at the same time, $S$ is not denumerable if the Diagonalization argument is valid. The latter then stands in contradiction of the bijection between $Z$ and $S$.

It is not possible, absent forgoing our collective understanding of logic, that both premises can be true. One has to be false. The bijection of $Z$ and $S$ has an irrefutable proof available in many basic texts in mathematics and computer science, and is accepted common knowledge. Cantor’s Diagonalization has only an apparent paradox.

Laying aside for the moment Cantor’s conclusions arising from that apparent paradox, his argument presents the seed of which proof by induction can be carried forward independent of

---

[5] Note that a map of $R$ and hence $Z$ into $S$ can also be established by substitution on the representation by $x.0^*y0^*$ of elements of $R$, where $x$ and $y$ are elements of $Z$. In the case where $r$ is both element of $R$ and $Z$ the $x.y$ becomes $x.0^*$. 
Set Theory as we use it today. Cantor's failing, if it can be rightfully called such, was in not realizing at the time that his argument reveals his a priori assumption of complete enumeration to be false, and the consequent necessity of then adding the resulting string to the enumeration; then only to again discover another, ad infinitum. The same result as shown above of $S$, by means provided since the time of Cantor's paper. The conclusions Cantor offered are a non sequitur.

**BIJECTION OF Z AND R:**

Consider the set $K = \{ k \mid k \in \{ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \}^* \}$; the set of all strings of numeric symbols. It is important to note here that unlike numeric values two strings having differing numbers of leading zeros are not equivalent. From our results above the set $K$ is computationally-uncountable and has an injection into $S$. The subset of $S$ described by the injection from $K$ is by basic principles denumerable given that $S$ is denumerable. However, by the same means as is done for base-2 enumerations of $S$, the set $K$ can be shown to be in bijection relation with $Z$.

The bijection of $K$ with $S$ can also be established by observing that any $k$ in the set $K$ can be represented as $w = 0^*v$, where $v$ is an element of $K$ having no zeros on the left (leading zeros). It follows then that $v$ ranges over the set of all positive integers. The set of positive integers have an established translation to binary representation in $S$. That it follows as consequence that $w$ ranges over all of $S$ should then be evident to the reader, and thus defines a bijection between $K$ and $S$.

Now, given any $r$ element of $R$, the set of real numbers, it is possible to represent $r$ as $u.w$ where both $u, w$ are elements of $K$. However, given the bijection of $K$ with $S$ and thus $Z$ we can substitute the corresponding integers associate with the elements of $K$ for $u, w$. The last substitution provides us with an ordered pair of integers $(x, y)$. As a result of Cantor’s proof of the bijection relation between $Z$ and ordered pairs of integers the real numbers are shown to be in a bijection relation with $Z$. The Schroeder-Bernstein Theorem then establishes that $|R| = |Z|$. We again note the evident difficulties in computing the bijection.

**IMPLICATIONS IN THEORY OF COMPUTATION:**

Several subsequent conclusions in computational theory that apply Cantor’s Diagonalization method are as a consequence also without standing. The question of relative cardinality of the set $M$ of Turing Machines in relation to that of problems computable by Turing Machines is one such left without standing. It follows by Schroeder-Bernstein from the above results that $|M| = |S| = |Z| = |R|$; a result contrary to what is presently commonly accepted and taught. The generally held belief that there exist a greater number of problems than the number of Turing

---

[6] Commonly referred to as Cantor’s Pairing Function.
Machines is then also without standing, as both can now be shown to have a bijection with S. It then follows from the set of all problems being at least an injection into S that for any problem at least one Turing Machine could exist that computes it.

Cantor's method in argument, as too the limits of inference from it, remains important for reasons beyond the issue of relative cardinality of sets due to implications of both in the theory of computation that has arose since. His argument establishes the basis of an inductive proof that an injection between two computationally-uncountable sets is not denumerable by any computation that halts, as such produces a computationally-uncountable denumerable set. That result is also intuitively equitable.

An instance of a problem that is demonstrably definitive of basis for the Halting Problem in the order of both space and time provides a keystone to the partition of the set M of Turing Machines into M_H, Turing Machines that halt, and M_N, Turing Machines that do not halt. Enumeration of an injection between two uncountable sets is definitive for at least a set of problems, if not all problems, that do not halt. Assigning the uncountable set injection problem the name Cantor’s Problem (or C for short), and defining it such that for any partial injection of an uncountable set into another uncountable set an additional element mapping can be obtained by iterative application of Cantor’s Diagonalization method until all elements of one or both sets have been included within the injection. It can then be asserted, by principles provided in many basic texts, that: (1) C is an element of M by definition of both C and M; (2) given any m in M where m can be shown equivalent by reduction to C, m has computational complexity at least as difficult as C and thus cannot halt unless C also halts.

CONCLUSION:

Cantor’s efforts at proving relative cardinality of computationally-uncountable sets by Diagonalization was flawed in that it failed to recognize the construction simply proved the a priori assumption of a complete enumeration was incorrect by contradiction. Despite such flaw the methods remains important in that it provides constructive means to iteratively enumerate an injection between computationally-uncountable sets. The iterative enumeration being assured to not halt the problem provides a definitive basis for a class of problems assured not to halt, and may provide basis for the class of all problems that do not halt. The latter case could conceivably move the theoretical solution of the Halting Problem into the realm of plausibility.

REFERENCES:

Boas, R. P., & Boas, H. P. (1997). A Primer of Real Functions. Mathematical Association of America.

[7] If both sets are uncountable then the computation does not halt. Any surjection of a countable set onto an uncountable set is equivalent to a Cantor injection (between uncountable sets).
Brioda, J. G., & Williamson, G. (1989). *A Comprehensive Introduction to Linear Algebra*. Addison-Wesley.

Cohen, P. J. (1966). *Set Theory and the Continuum Hypothesis*. Addison-Wesley.

Hopcroft, J. E., Motwani, R., & Ullman, J. D. (1979). *Introduction to Automata Theory, Languages, and Computation*. Addison Wesley.

Knuth, D. (1981). *Art of Computer Programming, Volume 2: Seminumerical Algorithms*. Addison-Wesley.

Kozen, D. C. (1997). *Automata and Computability*. Springer.

Kozen, D. C. (2006). *Theory of Computation*. Springer.

Ross, K. A., & Wright, C. R. (1985). *Discrete Mathematics*. Printice Hall.

Sipser, M. (2005). *Introduction to the Theory of Computation*. Course Technology.