Encodings of Sets and Hypersets

Alberto Policriti

[1] University of Udine, Department of Mathematics and Informatics, Udine, Italy
[2] Istituto di Genomica Applicata, Udine, Italy

Abstract. We will present some results and open problems on an extension of the Ackermann encoding of Hereditarily Finite Sets into Natural Numbers. In particular, we will introduce and discuss a simple modification of the above mentioned Ackermann encoding, that should naturally generalize from Hereditarily Finite Sets to Hereditarily Finite Hypersets.

Keywords: Ackermann encoding, Hereditarily Finite Sets, Hereditarily Finite Hypersets.

Introduction

This work is related with an attempt to (sensibly) extend the following function (map) introduced by Ackermann in 1937 (see [Ack37]) :

Definition 1.

$$\mathbb{N}_A(x) = \Sigma_{y \in x} 2^{\mathbb{N}_A(y)} \tag{1}$$

The above map is a bijection between Hereditarily Finite Sets (denoted by HF: the collection of sets obtained starting from $\emptyset$ and closing with respect to finite set formation) and Natural Numbers (denoted by $\mathbb{N}$).

The usage of 2 as base for the exponentiations in Definition 1 allows us to see the binary expansion of the natural number $\mathbb{N}_A(x)$ as a *full description*—that is the list of its elements—of x in terms of $\mathbb{N}_A$: the presence of a 1 in position i of the binary expansion of $\mathbb{N}_A(x)$ is equivalent to saying that the element y such that $\mathbb{N}_A(y) = i$ belongs to x.

Example 1. $\mathbb{N}_A(\emptyset) = 0, \mathbb{N}_A(\{\emptyset\}) = 2^0 = 1, \mathbb{N}_A(\{\emptyset, \{\emptyset\}\}) = 2^0 + 2^1 = 3$, the binary code of $\mathbb{N}_A(\{\emptyset, \{\emptyset, \{\emptyset\}\}\})$ is 1001, that is 9.

$\mathbb{N}_A$ is a simple and natural encoding of sets, hence a powerful tool for representing and manipulating objects that are suitable to represent any kind of mathematical information.

Two sets are equal if and only if they have the same elements and the "translation" of this in terms of $\mathbb{N}_A$ corresponds to observing that two natural numbers are equal if and only if they have the same binary code. The previously mentioned set-theoretic principle for testing equality—a basic axiom in classical Set Theory called *extensionality*—can be applied only if the membership relation $\in$

does not admit cycles. A "set" u such that u belongs to itself, for example, cannot be tested for equality using extensionality against another set v: among the equalities to be checked we would need ... to take u into account! Nevertheless, non well-founded set theories (whose elements are called *hypersets*) are useful and very expressive tools. Especially in Informatics. They add the ability to represent circular phenomena by blending the basic set-theoretic machinery with the notion of bisimulation used in place of extensionality (see [Acz88]). Therefore, for example, it becomes important to find (fast) algorithms to compute hyperset-equality. In [PP04] a number of examples of usages and extensions of the Ackermann map are given, exploring the possibility of computing hyeperset equality by comparing Ackermann-like encodings.

Here we discuss a possible extension of $\mathbb{N}_A$ whose aim is to maintain formal elegance while using as codomain a number system larger than $\mathbb{N}$.

We mention the fact that, along the same line, we already proposed extensions $\mathbb{Q}_A$ of $\mathbb{N}_A$ mapping the collection $\overline{\mathsf{HF}}$ of rational hereditarily finite hypersets into *dyadic rational numbers* (see [DOPT10]). Dyadic rational numbers are rational numbers that can be denoted by finitely many (binary) digits and our proposal can be illustrated by a simple example: if $\mathbb{Q}_A(z) = 1010, 0111$, then z is the hyperset whose well-founded elements are the second and the fourth (that is those having Ackermann code equal to 2 and 4), while z's non well-founded elements are the second, the third, and the fourth. Clearly, to build $\mathbb{Q}_A$ an ordering of the hereditarily finite hypersets must enter into play. The set $\Omega = \{\Omega\}$ is—naturally—the first non well-founded set and, consequently, $\mathbb{Q}_A(\Omega) = 0, 1$.

The extension of Ackermann map discussed below is more *direct* than $\mathbb{Q}_A$, as it does not require any (somehow arbitrary) ordering of hereditarily finite sets. This feature opens the way to an usage of the newly proposed map for bisimulation computation based on code comparison, as well as to a large array of numerically-based (hyper)set manipulation techniques.

1   Extending Ackermann map

Consider the following definition, obtained from Definition (1) by simply adding a minus sign at exponent.

Definition 2.

$$\mathbb{R}_A(x) = \Sigma_{y \in x} 2^{-\mathbb{R}_A(y)} \tag{2}$$

As a first and very basic motivation to consider the above map, notice that the above definition allows a (unique) solution to the following equation:

$$x = 2^{-x} \tag{3}$$

To see this, it suffices to observe that the two curves $y = x$ and $y = 2^{-x}$ are increasing and decreasing, respectively, and intersect in the first quadrant of $\mathbb{R}$. Let $\boldsymbol{\Omega}$ be the solution of (3) over $\mathbb{R}$. It is not difficult to see that $\boldsymbol{\Omega} \notin \mathbb{Q}$.

Our first objective would be to show that (2) is injective on the collection of Hereditarily Finite Sets.

Conjecture 1. The function $\mathbb{R}_A$ is injective on HF.

A second, more challenging, point will consist in establishing the fact that $\mathbb{R}_A$ is injective on the full collection of rational hereditarily finite hypersets.

Conjecture 2. The function $\mathbb{R}_A$ is injective on $\overline{\mathsf{HF}}$.

We only have partial results related with the above conjectures that are mostly related with a study of the codes of the elements of the following sub-family of HF:

Definition 3. *The elements of the family $\mathcal{S}$ of super-singletons*

$$\mathcal{S} = \left\{ \{\emptyset\}^i \mid i \in \mathbb{N} \right\},$$

are defined recursively as follows: $\{\emptyset\}^0 = \emptyset$ and $\{\emptyset\}^{n+1} = \{\{\emptyset\}^n\}$.

Super-singletons were first introduced by Zermelo in [Zer08] as a set-theoretic representation for ordinals and they have been recently discussed by Kirby in [Kir13].

The following figure shows the disposition of the first few code values of super-sigletons (let s_i denote the code of the i-th super-singleton).
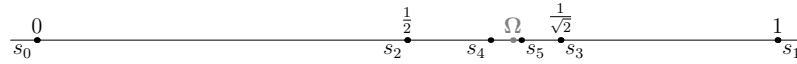

Fig. 1: The $\mathbb{R}_A$-code of the first 5 super-singletons

The $\mathbb{R}_A$-code of super-singletons is easily determined and seen to converge to the above defined value $\boldsymbol{\Omega}$.

Proposition 1. *The following hold:*

$$0 = s_0 < s_2 < \cdots s_{2i} < s_{2i+2} \cdots < \boldsymbol{\Omega} < \cdots s_{2i+3} < s_{2i+1} \cdots < s_3 < s_1 = 1,$$

and

$$\lim_{i \to \infty} s_{2i} = \lim_{i \to \infty} s_{2i+1} = \boldsymbol{\Omega}.$$

Proof. To see the above result it suffices to observe that $2^{-2^{-x}}$ is increasing and therefore, since $s_0 = 0 < s_2 = 1/2$ and since $s_1 = 1 > s_3 = 1/\sqrt{2}$, we have:

$$s_0 < s_2 < \cdots s_{2i} < 2^{-2^{-s_{2i}}} = s_{2i+2} \cdots$$

and

$$s_1 > s_3 > \cdots s_{2i+1} > 2^{-2^{-s_{2i+1}}} = s_{2i+3} \cdots .$$

Moreover, since $s_0 < \mathbf{\Omega} = 2^{-\mathbf{\Omega}}$ and since $s_{2i+2} = 2^{-2^{-s_{2i}}} < 2^{-2^{-\mathbf{\Omega}}} = \mathbf{\Omega}$, we have that all even-indexed super-singletons are smaller than $\mathbf{\Omega}$.

Similarly, all odd-indexed super-singletons are larger than $\mathbf{\Omega}$.

To conclude we must prove that both even and odd indexed sequences of super-singletons converge to $\mathbf{\Omega}$.

This is a consequence of the fact that $(2^{-x} - 2^{-y}) < (y - x)/2$, for all $x, y \in [1/2, 1]$. This, in turn, follows from Lagrange theorem stating that $(f(b) - f(a))/(b - a) = f'(z)$, for some $z \in (a, b)$. In fact, assuming $y > x$, the value $(2^{-x} - 2^{-y})/(y - x)$ is equal to $(2^{-z})' = -2^{-z} \ln(2)$, for some $z \in [x, y] \subseteq [1/2, 1]$, and this value is always smaller than $-1/2$. To conclude it is sufficient to consider the sequence for $i > 0$, with $x = s_{2i}$ and $y = s_{2i-1}$. $\square$

With some extra observations and using the above result we can prove the injectivity of $\mathbb{R}_A$ on the codes of arbitrary unions of super-singletons.

Definition 4. *Given j pairwise distinct indexes $i_1, \ldots, i_j$, let $s_{i_1, \ldots, i_j}$ be the code of $\{\emptyset\}^{i_1} \cup \cdots \cup \{\emptyset\}^{i_j}$, that is $s_{i_1} + \cdots + s_{i_j}$.*

Moreover, let

$$\mathcal{S}_{i_1, \ldots, \overline{i_j}} = \left\{ s_{i_1, \ldots, i_j, k} \mid k > i_j \right\}.$$

On the grounds of the above definition, we have that the codes of non-null super-singletons in $\mathcal{S}$ are in $\mathcal{S}_{\overline{0}}$. If we imagine (codes of) super-singletons in $\mathcal{S}_{\overline{0}}$ as obtained from the intersection of a *spiral* with the x-axis, as in the Figure 2,
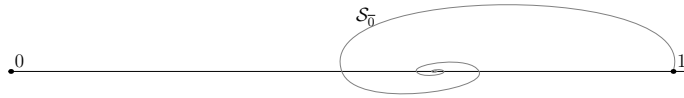

Fig. 2: The spiral of $\mathcal{S}_{\overline{0}}$

then the arrangement of the subsequent spirals can be seen to be a spiral of smaller and smaller spirals, as in Figure 3.

Notice that the points of convergence of all the above spirals—that is $\mathbf{\Omega} + s_i = \mathbb{R}_A(\Omega \cup \{\emptyset\}^i)$, for $i \geqslant 0$—are, in fact, codes of hypersets. This is not the case for the point of convergence of all the points of convergence, that turns out to be $2\mathbf{\Omega}$. Looking at the point of convergence of $2^{-(\mathbf{\Omega}+s_i)}$ for $i \geqslant 0$, one obtains the sequence of $s_{i+1}\mathbf{\Omega}$ that—no wonder—converges at $\mathbf{\Omega}^2$. Starting from the sequence of spirals $\mathcal{S}_{i,\overline{j}}$, whose points of convergence bring us at $3\mathbf{\Omega}$, by exponentiating we get to $\mathbf{\Omega}^3$, and so on.
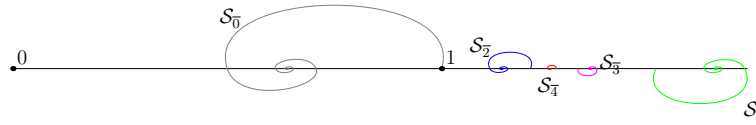
Fig. 3: The spirals of $\mathcal{S}_{\overline{0}}, \mathcal{S}_{\overline{1}}, \mathcal{S}_{\overline{2}}, \mathcal{S}_{\overline{3}}, \mathcal{S}_{\overline{4}}$

Proposition 2. *If* $\{i_1, \ldots, i_j\} \neq \{h_1, \ldots, h_k\}$, *then* $\mathcal{S}_{i_1, \ldots, \overline{i_j}} \cap \mathcal{S}_{h_1, \ldots, \overline{h_k}} = \emptyset$.

The above proposition is proved by first reducing the general case to the case in which $j = k$, since a padding of 0's on the left can always be performed. Then, proving that the leftmost difference among indexes in two elements belonging to $\mathcal{S}_{i_1, \ldots, \overline{i_j}}$ and $\mathcal{S}_{h_1, \ldots, \overline{h_j}}$, respectively, can never be compensated by the following differences.

By letting h_i be the set belonging to HF whose Ackermann code is i, that is such that $\mathbb{N}_A(h_i) = i$, an ordering among the elements in HF is naturally (!) induced by $\mathbb{N}_A$[3].

Looking at indexes that do not necessarily belong to the collection of super-singletons or to sums of such sets, the following result holds.

Proposition 3. *For all* $i \in \mathbb{N}$:

1. $\mathbb{R}_A(h_i) \neq \mathbb{R}_A(h_{i+1})$;
2. $\mathbb{R}_A(h_i) \neq \mathbb{R}_A(h_{i+2})$.

We can prove the above proposition by rather *ad-hoc* arguments based on the specific value that a difference between two subsequent codes can assume.

Conclusions

We proposed a new numerical encoding for hereditarily finite hypersets that is a natural extension of the celebrated Ackermann map establishing a bijection between natural numbers and hereditarily finite (well-founded) sets. Our proposed encoding differs from Ackermann's one only for a minus sign in the exponent. Such a small difference in the definition, however, radically changes the encoding that now maps the hypersets universe on real numbers. The map seems to have elegant analytical properties guaranteeing its injectivity on both well-founded and non well-founded hereditarily finite sets. Both injectivities, however, are

---

[3] Such ordering can be seen to correspond to the ordering holding on binary representation of natural numbers: given two strings of bits α and β, if the leftmost difference is such that a 1 appears in α, then $\alpha > \beta$.

just conjectured here and our opinion is that once proved $\mathbb{R}_A$ to be 1-1 on well-founded sets, Conjecture 2 could/should be attacked by proving that the code of a hyperset is the unique accumulation point of the codes of the well-founded sets obtained by its unfolding. Notice that this is the case for $\boldsymbol{\Omega}$, as well as for other cases briefly mentioned above.

References

[Ack37]   W. Ackermann, *Die Widerspruchfreiheit der allgemeinen Mengenlehre*, Mathematische Annalen **114** (1937), 305–315.

[Acz88]   P. Aczel, *Non-well-founded sets*, vol. 14 of CSLI Lecture Notes, Stanford, CA, 1988.

[DOPT10] G. D'Agostino, E. Omodeo, A. Policriti, and A. Tomescu, *Mapping hypersets into numbers (extended abstract)*, 12th Italian Conference on Theoretical Computer Science (ICTCS), 2010.

[Kir13]   L. Kirby, *Ordinal operations on graph representations of sets*, Math. Log. Q. **59** (2013), no. 1-2, 19–26.

[PP04]   C. Piazza and A. Policriti, *Ackermann Encoding, Bisimulations, and OBDDs*, Theory and Practice of Logic Programming **4** (2004), no. 5-6, 695–718.

[Zer08]   E. Zermelo, *Untersuchungen über die Grundlagen der Mengenlehre. I*, Mathematische Annalen **65** (1908), no. 2, 261–281 (German).