The design of greedy number representations

In the following we restrict ourselves for convenience to the design of greedy binary number representations of nonnegative numbers, however all definitions can easily be generalized to greedy number representations with other sets of digits. For naming we will follow the traditions in number representation so far as most as possible. Number representations like the factorial number representations are excluded because they need an infinite number of digits to represent all nonnegative numbers. The greedy number representations must satisfy the condition that each finite nonnegative number can be represented by a finite string of digits, so any infinite representation, like p-adic number representation is excluded.

All we need is some basic algebraic structures like equivalence relations and partial ordering relations, as well as regular languages in formal language theory. Apart from the syntax, the (denotational) semantics of a number representation is important as well.

Although we started this study as a challenge in combining basic mathematics and language theory, it could contribute to coding theory as well, as from the fact that from coding theory, a (denotational) meaning becomes important as well. As from language theory in general, we have to deal with a syntax level, a semantic level as well as a pragmatic level. All three levels will occur below in a special sence.

The so called A-numbers in this paper refer to Neil Sloane's On-line Encyclopedia of Integer Sequences [1].

History

The tally representation of numbers can be considered to be the oldest greedy representation known, although it was not intended that way in the past. In fact, it becomes only a greedy representation if we use the symbol 0 for the empty string as we will see by an example below. The first modern greedy representation was published in 1951 by the Dutch mathematician Gerrit Lekkerkerken [2], however, in his paper he referred already to Dr E. Zeckendorf (most likely in some form of private communication), and as turned out later, Zeckendorfs first ideas go back to 1939. In 1972 Edouard Zeckendorf published his contribution to the greedy number representations as well [3]. Lekkerkerker showed in his paper that for the greedy Zeckendorf representation of numbers, the arithmetic mean of the fraction of 1's in the representation tends to ½ (1 – 1/√5), i.e., the constant A244847.

Definitions

Basic statement:

Any binary greedy representation G of nonnegative numbers is a subset of the set of binary fixed radix (base = 2) representations, i.e., $G \subset L(0|1(0|1)^*)$.

Here, $0|1(0|1)^*$ is the regular expression of the set of binary fixed radix representations. Note that the representation 0 for zero is an exception; the 0 notation replaces the empty string in traditional number representation as it will follow from the meaning function M below.

Beside binary representations, which use the alphabet {0, 1}, we can easily extend the definitions to alphabets with higher cardinality, like the ternary alphabet {0, 1, 2} with regular expression 0|(1|2)(0|1|2|)* for ternary fixed radix representations.

Definition of meaning function and weights:

In order to define a meaning of a (greedy) representation, we need an ordered set of weights W_i, such that the meaning of the number representation $b_n\ b_{n-1} \dots b_1\ b_0 \in G$, $M_\mathbf{W}(\mathbf{b})$, is given by $M(b_n\ b_{n-1} \dots b_1\ b_0) = b_n\ W_n + b_{n-1}\ W_{n-1} + \dots + b_1\ W_1 + b_0\ W_0$. The ordered set of weihts, $(W_0, W_1, W_2, \dots)$ will be denoted by **W**.

For binary fixed radix representation we have $W_0 = 1$ and $W_{i+1} = 2\ W_i$ for $i \geq 0$. We will denote this special set of weights by $\mathbf{W}_2$, it's meaning is represented by A007088 in the sense that n is the meaning of A007088(n). For the well-known Zeckendorf representations, A014417 and A104326, we have $W_0 = 1$, $W_1 = 2$ and $W_{i+1} = W_i + W_{i-1}$ for $i \geq 1$.

The main challenge in designing a (greedy) representation will be designing a suitable set of weights **W** to define a (greedy) representation. For this reason we mention Brown's completeness criterion [4]:

The sequence **W** of weights is considered complete if and only if $W_0 = 1$, and $W_{n+1} \leq 1 + (b-1)\ \Sigma_{0\leq k\leq n}\ W_k$ for $n \geq 0$.

Here b refers to "base", i.e., for b=2 we are dealing with binary representations with alphabet {0, 1} and for b=3 we are dealing with ternary representations with alphabet {0, 1, 2}, and so on.

In fact, Brown's completes criterion is to strong; any finite permutation of a weight sequence **W** satisfying this criterion will satisfy also as we will see below.

By definition, an ordered set of weights **W** satisfies if and only if for each nonnegative integer n, there exists a $\mathbf{b} \in L(0|1(0|1)^*)$ such that $M_\mathbf{W}(\mathbf{b}) = n$. However, for any set of weights that is suitable for defining a greedy representation, there exists a pair $\mathbf{b}_1$ and $\mathbf{b}_2$ ($\mathbf{b}_1, \mathbf{b}_2 \in L(0|1(0|1)^*)$) such that $\mathbf{b}_1 \neq \mathbf{b}_2$ and $M_\mathbf{W}(\mathbf{b}_1) = M_\mathbf{W}(\mathbf{b}_2)$. This later will be part of the definition for a representation to be greedy. I.e., in general a greedy representation will in general need substantial more (binary) digits than any non-greedy representation.

Note that at syntax level, an alphabet is just a non-empty finite set of symbols. When the semantic meaning function becomes involved, the alphabet becomes an ordered set of symbols; i.e. for binary representation the ordered set (0, 1), with 1 > 0.

Definition of interpretation function:

Next we need an interpretation function $I_\mathbf{W}$ (with respect to some meaning function $M_\mathbf{W}$), such that $I_\mathbf{W}(n)$ gives the set of all **b** such that $M_\mathbf{W}(\mathbf{b}) = n$, i.e., $I_\mathbf{W}(n) = \{\ \mathbf{b} \in L(0|1(0|1)^* \mid M_\mathbf{W}(\mathbf{b}) = n\ \}$.

Definition of condensed representations:

As for number representation in general, for all n, $n \geq 0$, $I(n) \neq \varnothing$, i.e., each n must have at least one representation in L(0|1(0|1)*). A representation is called condensed if for all $n \geq 0$, $|I_W(n)| = 1$. This definition is only partial, as we will see from a pragmatic reason below given by a counterexample.

An example of a condensed representation is given by **W** = (2, 1, 4, 8, …). Of course any finite permutation of the set of powers of 2 will lead to a condensed representation. Note that the condition finite in finite permutation is essential.

Example 1.

Suppose we have an infinite permutation given by **W** is the concatenation (⊕) of all even powers of 2 and all odd powers of 2, i.e. **W** = $(2^{2n} \mid n = 0, 1, 2, \ldots) \oplus (2^{2n+1} \mid n = 0, 1, 2, \ldots)$, representing the finite number 2 would lead to an infinite long strings for some finite numbers, which is in contradiction with a restriction we gave before.

The above example represents a pragmatic restriction; finite permutations are allowed, but infinite permutations are excluded!

Any condensed representation as defined above including the pragmatic restriction, can be used as a reference for defining a greedy representation, including its denotational meaning, as we will see below.

Definition of the weight product P_W:

The number of different representations with respect to some meaning function M_W and its weight function **W** is given by the product $P_W = \Pi_{n \geq 0} \mid I_W(n) \mid$.

As mentioned already, in case $P_W = 0$, there exists a number n with no representation, so **W** is not suitable, and in case $P_W = 1$ we are dealing with a condensed representation if some pragmatic conditions are satisfied. If $P_W = \infty$, **W** is suitable for defining a greedy representation. This leads us to considerate the possibility that P_W satisfies $1 < P_W < \infty$. Although we consider these possibilities as less interesting, we propose to name representations based on a weight **W** satisfying this condition as stretched representations.

Definition of stretched representations:

For any stretched representation, the weight function **W** results in a weight product P_W, satisfying $1 < P_W < \infty$.

In fact, we are dealing with an equivalence relation ~, by $\mathbf{b}_1 \sim \mathbf{b}_2 ::= M_W(\mathbf{b}_1) = M_W(\mathbf{b}_2)$. Each class can be represented either by its meaning n, or by a representing vector **b** such that $M_W(\mathbf{b}) = n$. Once we have defined a representing vector **b**, this vector will be denoted by **r**. More specific, **r**(n) will be the representing vector of the class $I_W(n)$, such that **r**(n) is a unique representation.

Definition of partial ordering relation $\geq_2$:

Next we need is a partial ordering relation, $\geq_2$. Let R_1 and R_2 be two (binary) representations with respect to some weight **W**, $R_1 \geq_2 R_2$ is defined as, for all $\mathbf{r}_1 \in R_1$ and $\mathbf{r}_2 \in R_2$, $M_2(\mathbf{r}_1) \geq M_2(\mathbf{r}_2)$. The result is that there exists a lattice with respect to some weight **W**. In many cases we have a bounded lattice with a maximal and a minimal (top and bottom, or upper and lower bound) element, represented by R_T and $R_\perp$ respectively.

Here, we need some more comments. First of all, the set of all representations (greedy, stretched as well as condensed), will lead to some (unbounded) lattice. However, for each weight sequence **W**, we have a sub lattice, which has in general a top and a bottom element. A necessary condition here is that for all n, $| I_w(n) |$ is finite, otherwise, the top element will not exist. A counterexample which is unbounded will be given below.

We decided to call the representations that are greedy and are either top or bottom representations greedy top and greedy bottom representations. This could be confusing with respect to the traditional naming of minimal and maximal Zeckendorf representations [5] in the sense that the minimal Zeckendorf representation is an example of a greedy top representation, and the maximal Zeckendorf representation is an example of a greedy bottom representation. The naming of minimum and maximum Zeckendorf representations where based however on the number of 1's in the representation, and not on the ordering relation $\geq_2$ proposed here.

The choice for the meaning function M_2 in the definition for the partial ordering relation seems to be an obvious one, but, on the other hand, any meaning function of a condensed representation could have been used as well.

If r(n) is the representation of n in a greedy top representation R_T, $M_2(r(n))$ is the maximum value of $\{M_2(r) \mid r \in I(n)\}$. If r(n) is the representation of n in a greedy bottom representation $R_\perp$, $M_2(r(n))$ is the minimum value of $\{M_2(r) \mid r \in I(n)\}$.

If R_T and $R_\perp$ both exist, any representation R has a complimentary representation R^c, in particular $R_T^c = R_\perp$. Moreover, if R is a condensed representation, $R = R^c = R_T = R_\perp$.

Conjecture 1.

With respect to some weights **W**, based on some linear recurrence relation, the greedy top and greedy bottom representation, if they exist, will always be a regular language.

For binary greedy representations, the weights **W** based on some linear recurrence relation must of course satisfy the condition $\text{Lim}_{i\to\infty} W_{i+1}/W_i \leq 2$. Further, there must exist a k such that $W_k = 1$. These conditions are necessary to assure that $I_W(n) \neq \varnothing$ for all $n \geq 0$. However, **W** = (1, 2, 3, 5, …) as applied Fibonacci numbers for the Zeckendorf representations, gives a minimal set from the Fibonacci sequence, but **W** = (1, 1, 2, 3, 5, …) or **W** = (0, 1, 1, 2, 3, 5, …) will satisfy as well for the design of binary greedy representations with the definitions given here.

In the above conjecture, if we replace M_2 in the definition of the partial ordering relation by the meaning function of some other condensed representation, the conjecture is believed to be hold as well.

Example 2.

Let **W** be given by the all 1's sequence A000012. The greedy representations formed by this **W** leads to a lattice with only a lower bound. This lower bound is given by A002275, i.e., the tally representation where 0 represents the empty string. The regular language is given by 0|11*.

Conjecture 2.

If with respect to some weights **W**, the greedy top and/or greedy bottom representations exists, the set of palindromic representations is a context free language.

The table below shows the most important sequences in the OEIS [1] related to condensed and greedy number representations.

W [a]	$\mid l_W(n) \mid$	M_W(A007088)	$R_\perp$	$R_\top$	comment
A000012[0:]	1,∞,∞, …		A002275	dne [b]	Tally
A000079[0:]	A000012		A007088	A007088	binary condensed
A000045[2:]	A000119	A022290	A104326	A014417	Zeckendorf
A000930[2:]		A350311	A350312	A350215	Based on Narayana’s cows sequence
A001045[1:]	A026465				Based on Jacobsthal numbers
A000244[0:]	A000012	- [c]	A007089	A007089	ternary condensed
A007598[1:]	A147561	- [c]			squared Fibonacci [5]
A000302[0:]	A000012	- [c]	A007090	A007090	quaternary condensed
A000331[0:]	A000012	- [c]	A007091	A007091	quinary condensed
A056570[1:]	A147561	- [c]			cubed Fibonacci [5]

Table 1.

Notes:

[a] The weights sequence is given by the OEIS A number, where the suffix [n:] means that the sequence starts at index n.

[b] dne stands for “does not exist”.

[c] This cell is not relevant here, because the sequence A007088 refers only to binary representations.

Acknowledgement

Thanks to Hesper Weirauch and Kevin Ride for helpful reflections, especially on naming.

[1] N.J.A. Sloane, “*On-line Encyclopedia of Integer Sequences*”, https://oeis.org.

[2] C.G. Lekkerkerker, “*Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci*” (in Dutch), Mathematisch Centrum (Amsterdam), ZW 1951 – 30 (1951); Simon Stevin 29 (1952), pp. 190-195.

[3] E. Zeckendorf, “*Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres Lucas*” (in French), Bull. Soc. R. Sci. Liège 41 (1972), pp. 179-182.

[4] L.J. Brown Jr., “*Note on Complete Sequences of Integers*”, American Mathematical Monthly 68 (1969), pp. 557-560.

[5] Ron Knott, "*Using the Fibonacci numbers to represent whole numbers*", http://www.maths.surrey.ac.uk/hosted-sites/R.Knott/Fibonacci/fibrep.html

A.H.M. Smeets, ton.smeets2@gmail.com

Maastricht (Netherlands), Jan 22 2022.