Goodstein's Theorem: Independence in Action

Pesara Amarasekera, Megan Holmes, Tyler Ng

April 10, 2024

1 Introduction

Reuben Goodstein (1912-1985), wrote his paper On the Restricted Ordinal Theorem in 1944, the main result being his famous theorem to be discussed in this paper. It was at the time not recognized as important to the level it is today. This is because his theorem is an example of an improvable L_A truth, showing that Godel's incompleteness theorem does not deal with truths that are only theoretical and irrelevant to modern mathematics. [1]

Introduction

A Goodstein Sequence starts with any natural number n. For simplicity, we will label each value in a Goodstein sequence n_k, for $k \geq 1$, and $n_1 = n$. To find the succeeding terms, we need to introduce hereditary base-k representation.

Hereditary Base-K Representation

We can start by recalling regular base-k representation. For some k, we write n as a sum of powers of k, with positive coefficients a_i for each term less than k:

$$n = a_m * k^m + a_{m-1} * k^{m-1} + a_{m-2} * k^{m-2} + ... + a_1 * b + a_0$$

Hereditary base-k representation expands upon this process and for each exponent, 1,...,m, we also apply base-k representation. For the exponents of these exponents, we repeat the process. And for their exponents, and so on. For example, 266 can be written in Base-2 form, as

$$266 = 2^8 + 2^3 + 2^1.$$

In Hereditary base-2 form, it would be

$$266 = 2^{2^{2^1+1}} + 2^{2^1+1} + 2^1$$

Creating a Sequence

For $k \geq 2$, the process of finding n_k starts with writing n_{k-1} in hereditary base k representation. Then every instance of k in this representation is replaced with $k+1$. The resulting value calculated after is then decreased by 1, and this yields n_k.

To formalize this process we can define the base bump function, $B_k(n)$. This function will replace all instances of k in the hereditary base-k representation of n with $k+1$. So taking our previous example,

$$B_2(266) = 3^{3^{3^1+1}} + 3^{3^1+1} + 3^1.$$

This means that a full Goodstein sequence for n can be written as: $n_1, B_2(n_1) - 1, B_3(n_2) - 1, ...$

With all of this introduced, we can now look at a simple example of a Goodstein sequence:
$3_1 = 3 = 2^1 + 1$
$3_2 = B_2(3_1) - 1 = 3^1 + 1 - 1 = 3$
$3_3 = B_3(3_2) - 1 = 4^1 - 1 = 3$
$3_4 = B_4(3_3) - 1 = 3 * 1 - 1 = 2$
$3_5 = B_5(3_4) - 1 = 2 * 1 - 1 = 1$

$3_6 = B_6(3_5) - 1 = 1 * 1 - 1 = 0$
This sequence has only 6 terms, terminating when $3_6 = 0$. This is the final key piece of information for Goodstein sequences, they terminate when $n_k = 0$. Observe this partial example:
$19_1 = 19 = 2^{2^{2^1}} + 2^1 + 1$
$19_2 = B_2(19_1) - 1 = 3^{3^{3^1}} + 3^1 + 1 - 1 = 3^{3^{3^1}} + 3^1 (\approx 7 * 10^{13})$
$19_3 = B_3(19_2) - 1 = 4^{4^{4^1}} + 4^1 - 1 = 4^{4^{4^1}} + 1 + 1 + 1 (\approx 1.3 * 10^{154})$
$19_4 = B_4(19_3) - 1 = 5^{5^{5^1}} + 1 + 1 + 1 - 1 = 5^{5^{5^1}} + 1 + 1 (\approx 2 * 10^{2184})$
$19_5 = B_5(19_4) - 1 = 6^{6^{6^1}} + 1 + 1 - 1 = 6^{6^{6^1}} + 1 (\approx 2.6 * 10^{36305})$
$19_6 = B_6(19_5) - 1 = 7^{7^{7^1}} + 1 - 1 = 7^{7^{7^1}}$

Here we see that the numbers in the sequence become extremely large, and can do so very quickly. By 19_6, there is no sign that the sequences is converging to 0. This is what is so remarkable about Goodstein sequences, they grow rapidly and deal with very large numbers, and can continue to grow for an unimaginably large number of terms. Despite this, Goodstein's Theorem states that all Goodstein sequences eventually terminate at zero. [2]

2 Proof Introduction

The general structure of most proofs of Goodstein's Theorem is as follows [3]:

1. For every Goodstein sequence, we can construct a sequence of ordinals that exists if and only if the particular Goodstein sequence exists. That is, for a particular Goodstein sequence, we may map each element of that sequence to an ordinal.

2. This corresponding sequence of ordinals is strictly decreasing. That is, given two successive elements of a particular Goodstein sequence g_n and g_{n+1}, their corresponding ordinals o_n and o_{n+1} in the corresponding ordinal sequence must satisfy $o_n > o_{n+1}$.

3. Any strictly decreasing sequence of ordinals must terminate. This is via the well ordering property of ordinals.

4. Hence, by 3, for any Goodstein sequence, the corresponding sequence of ordinals must terminate.

5. The ordinal sequence terminates if and only if the corresponding Goodstein sequence terminates. Therefore the Goodstein sequence must terminate. Therefore Goodstein's Theorem Part 1 is true.

2.1 Proof of 2: Intro/setup $\forall n(o_n > o_{n+1})$

Take an arbitrary Goodstein sequence, and select an arbitrary element of that Goodstein sequence g_n, for $n \in \mathbb{N}$.
To obtain the next element in the Goodstein sequence, write it in hereditary base $n+1$, replace each occurrence of $n+1$ with $n+2$, and subtract 1.
Suppose that g_n is written in as follows:

$$g_n = \sum_{i=0}^{k} (n+1)^{d_i} a_i = (n+1)^{d_0} a_0 + (n+1)^{d_1} a_1 + (n+1)^{d_2} a_2 ... + (n+1)^{d_k} a_k$$

where $\forall i(0 \leq a_i < n+1)$ and $d_0 > d_1 ... > d_k$. Each d_i is the exponent for each nonzero term of g_n written in base $n+1$, and $k+1$ is the total number of nonzero terms in the base $n+1$ representation of g_n. Each a_i is the coefficient in such a representation. So, for example, for

$$3^{3^3} + 2(3^3) + 1$$

we would have $n = 2$ $a_0 = 1$, $a_1 = 2$, $a_2 = 1$, $d_0 = 3^3$, $d_1 = 3$, $d_3 = 0$.
To obtain g_{n+2} then, one would apply two transformations to g_n:

(a) Apply the base bumping function: replace every occurrence of $n+1$ in the hereditary base $n+1$ form of g_n with $n+2$

(b) subtract 1 from the resulting number

That is,

$$g_{n+1} = (\sum_{i=0}^{k} a_i(n+2)^{f_i}) - 1$$
$$= a_0(n+2)^{f_0} + a_1(n+2)^{f_1} + a_2(n+2)^{f_2} ... + a_k(n+2)^{f_k} - 1$$

where each f_i is the number resulting from replacing each occurrence of $n+1$ in the hereditary base $n+1$ representation of d_i with $n+2$.

A key observation is that subtracting 1 from the resulting base $n+2$ representation does not affect any of the terms save for the very last nonzero term, $a_k(n+2)^{f_k}$. Note that $a_k(n+1)^{d_k}$ is the smallest nonzero term in the hereditary base $n+1$ representation for g_{n+1}, since it is defined as having the smallest exponent. Furthermore, since the base bumping function is strictly increasing, we have that $a_k(n+2)^{f_k}$ is also the smallest nonzero term in the base $n+2$ representation of g_{n+1}.
As the smallest nonzero term in the sum, $a_k(n+2)^{f_k} - 1 \geq 0$, and furthermore by definition of base notation, every other term in the base $n+2$ representation must remain the same after subtraction in (b).
The importance of this observation about notation becomes apparent when we note which ordinals g_n and g_{n+1} are mapped to.
Firstly, since every term in $a_0(n+1)^{d_0} + a_1(n+1)^{d_1} + a_2(n+1)^{d_2} ... + a_{k-1}(n+1)^{d_{k-1}}$ and $a_0(n+2)^{f_0} + a_1(n+2)^{f_1} + a_2(n+2)^{f_2} ... + a_{k-1}(n+2)^{f_{k-1}}$ is essentially the same save for the fact that every $n+1$ is now an $n+2$, these initial segments are mapped to the exact same ordinal. Therefore, the determining factor that makes o_n a different ordinal from o_{n+1} are the final terms, $a_k(n+1)^{f_k}$ and $a_k(n+2)^{f_k} - 1$. If $a_k(n+2)^{f_k} - 1$ always maps to a smaller ordinal than $a_k(n+1)^{d_k}$, then this is sufficient to show that $o_n > o_{n+1}$ We shall now prove this.

2.2 Main Proof

Define the ordinal mapping function $o^{m,\omega}(n)$ that takes as input the natural number n, writes it in hereditary base m notation, and replaces each occurrence of m with ω (note also the technical detail that each natural number coefficient is to the right of the ω in order to be of proper Cantor normal form). This is the function that will associate each natural number in our Goodstein sequence with an ordinal. Our claim is that for arbitrary elements in a Goodstein sequence g_n and g_{n+1}, $o^{n+1,\omega}(g_n) > o^{n+2,\omega}(g_{n+1})$.
It was shown in section 1.1 that given g_n, g_{n+1}, we have

$$o^{n+1,\omega}(g_n - a_k(n+1)^{d_k}) = o^{n+2,\omega}(g_{n+1} - (a_k(n+2)^{f_k} - 1))$$

or in plain language, if we ignore the last term of g_n and g_{n+1} when they are written in hereditary base $n+1$ and $n+2$ respectively, they are mapped to the exact same ordinal. Thus to show

$$o^{n+1,\omega}(g_n) > o^{n+2,\omega}(g_{n+1})$$

one need only show that

$$o^{n+1,\omega}(a_k(n+1)^{d_k}) > o^{n+2,\omega}(a_k(n+2)^{f_k} - 1)$$

since, as a general fact about ordinals, for ordinals α, β, γ, if $\alpha > \beta$ then $\gamma + \alpha > \gamma + \beta$. In plain language, we need only compare which ordinals the smallest term in each sum is mapped to.

2.2.1 A further simplification

Supposing $a_k > 1$, note that

$$o^{n+1,\omega}(a_k(n+1)^{d_k}) = \omega^{o^{n+1,\omega}(d_k)} a_k$$
$$= \omega^{o^{n+1,\omega}(d_k)}(a_k - 1) + \omega^{o^{n+1,\omega}(d_k)}$$

and that

$$o^{n+2,\omega}(a_k(n+2)^{f_k} - 1)$$

$$= o^{n+2,\omega}((a_k - 1)(n+2)^{f_k} + (n+2)^{f_k} - 1)$$

$$= \omega^{o^{n+2,\omega}(f_k)}(a_k - 1) + o^{n+2,\omega}((n+2)^{f_k} - 1)$$

Note the following equality:

$$\omega^{o^{n+1,\omega}(d_k)}(a_k - 1) = \omega^{o^{n+2,\omega}(f_k)}(a_k - 1)$$

which follows simply from the fact that d_k and f_k are written identically, save for the fact that each $n+1$ in d_k is replaced with an $n+2$ (and so the ordinal mapping function will in each case map to the same ordinal). Thus we conclude that the only determining factor for whether

$$o^{n+1,\omega}(a_k(n+1)^{d_k}) > o^{n+2,\omega}(a_k(n+2)^{f_k} - 1)$$

comes down to whether

$$\omega^{o^{n+1,\omega}(d_k)} > o^{n+2,\omega}((n+2)^{f_k} - 1) \tag{1}$$

or in other words, every case where $a_k > 1$ is fully dependent on the output of the ordinal mapping function for $a_k = 1$. Therefore going forth we shall only examine this case.

2.2.2 $a_k = 1$

In order to prove the strict inequality (1) we break into two cases:

1. $f_k \leq n+2$
2. $f_k > n+2$

Case 1: $f_k \leq n+2$

We derive via the pre-calculus formula for summing a finite geometric series the following equation:

$$(n+2)^{f_k} - 1 = (n+1)\sum_{i=0}^{f_k-1}(n+2)^i$$

For the derivation of this equation, see [†]. We may evaluate the ordinal mapping function as

$$o^{n+2,\omega}((n+1)\sum_{i=0}^{f_k-1}(n+2)^i)$$

$$= \omega^{f_k-1}(n+1) + \omega^{f_k-2}(n+1) + ...\omega^2(n+1) + \omega(n+1) + n + 1$$

Note that because $f_k \leq n+2$, each exponent in the sum of ordinals is a natural number (unchanged by the writing of it in hereditary base $n+2$ notation), and therefore unambiguously defined via ordinal exponentiation.
Let us bound this sum above, by replacing each $\omega^i(n+1)$ for $i < f_k - 1$ with an ordinal strictly greater than it, $\omega^{f_k-1}(n+1)$. Thus we obtain

$$\omega^{f_k-1}(n+1) + \omega^{f_k-2}(n+1) + ...\omega^2(n+1) + \omega(n+1) + n + 1$$

$$< \omega^{f_k-1}(n+1) + \omega^{f_k-1}(n+1) + ...\omega^{f_k-1}(n+1) + \omega^{f_k-1}(n+1) + \omega^{f_k-1}(n+1)$$

$$= \omega^{f_k-1}(n+1) * f_k$$

So in this case, the ordinal that $(n+2)^{f_k} - 1$ is mapped to is bounded above by $\omega^{f_k-1}(n+1) * f_k$. How does this compare to the ordinal that $(n+1)^{d_k}$ maps to?
Suppose that $f_k = n+2$. This immediately implies that $d_k = n+1$, implying

$$o^{n+1,\omega}((n+1)^{d_k}) = o^{n+1,\omega}((n+1)^{n+1})$$

$$= \omega^{\omega}$$

and $\omega^{f_k-1}(n+1)*f_k = \omega^{n+1}(n+1)(n+2)$ is clearly much, much smaller than this.
On the other hand, suppose $f_k < n+2$. Then we know that d_k must have been less than $n+1$ (if not, the base-bumping transformation would have produced an f_k that is at least equal to $n+2$). But for any $d_k < n+1$, the base-bumping transformation takes d_k to exactly d_k, without changing its value in any way, since there are no $n+1$s to replace with $n+2$. In other words, $f_k < n+2 \implies (d_k = f_k) \wedge (d_k < n+1)$. Therefore,

$$o^{n+1,\omega}((n+1)^{d_k}) = \omega^{d_k} = \omega^{f_k}$$

How does this compare to $\omega^{f_k-1}(n+1)*f_k$? Since ordinal multiplication is both associative and strictly increasing on the right argument, one could rewrite

$$\omega^{f_k} = \omega^{f_k-1}\omega$$

and then note that $\omega >> (n+1)*f_k$, and thus

$$\omega^{f_k} = \omega^{f_k-1}\omega >> \omega^{f_k-1}(n+1)*f_k$$

It is difficult to overstate just how much larger the left side of this inequality is than the right side. And with that, our proof for Case 1 is complete: the ordinal indeed decreases and inequality (1) is true in Case 1.

Case 2: $f_k > n+2$

We shall demonstrate that Case 2 reduces to Case 1.
As in Case 1, one makes use of basic pre-calculus results and rewrites

$$o^{n+2,\omega}((n+2)^{f_k} - 1)$$

$$= o^{n+2,\omega}((n+1)\sum_{i=0}^{f_k-1}(n+2)^i)$$

$$= \omega^{o^{n+2,\omega}(f_k-1)}(n+1) + \omega^{o^{n+2,\omega}(f_k-2)}(n+1) + ...\omega^{o^{n+2,\omega}(1)}(n+1) + (n+1)$$

Note that if $x > y$, we have $o^{n+2,\omega}(x) \geq o^{n+2,\omega}(y)$. Therefore, we may again replace each term in the sum with an ordinal greater than or equal to it to produce an upper bound:

$$o^{n+2,\omega}((n+2)^{f_k} - 1)$$

$$< \omega^{o^{n+2,\omega}(f_k-1)}(n+1)*f_k$$

If we can prove that

$$\omega^{o^{n+1,\omega}(d_k)} > \omega^{o^{n+2,\omega}(f_k-1)}(n+1)*f_k$$

then we will have proven (1) for case 2. Note that this inequality holds if

$$o^{n+1,\omega}(d_k) > o^{n+2,\omega}(f_k-1)$$

since ordinal exponentiation is strictly increasing. The extra coefficient of $(n+1)*f_k$ does not matter since it is merely a finite coefficient, whereas an increased exponent would indicate an extra factor of at least ω, which is vastly larger.
A crucial observation is that this is the exact same question we were asking originally, but for the exponent! f_k is once again simply the base-bumped version of d_k, from which we are subtracting 1. Note the similarity to asking how the ordinal corresponding to $(n+1)^{d_k}$ compares to ordinal that the base-bumped and subtracted by one $(n+2)^{f_k}-1$ maps to.
To answer this question, one once again splits into cases, with Case 1 concluding that yes, the ordinal decreases, and Case 2 deferring to yet another exponent higher. However, since the numbers in a Goodstein sequence are always finite, they have a finite number of exponents: eventually, at some level of the exponents, you must encounter Case 1.

Therefore, every Case 2 reduces to a Case 1.

Therefore, in every case, the smallest term in the hereditary base notation of g_{n+1} maps to a strictly lesser ordinal than g_n.

Therefore, the overall ordinal corresponding to g_{n+1} maps to a smaller ordinal than the one corresponding to g_n.

So to every Goodstein sequence one may associate a decreasing sequence of ordinals that exists if and only if the Goodstein sequence exists.

Since any decreasing sequence of ordinals must terminate, so too must the Goodstein sequence it corresponds to.

Goodstein sequences may only terminate at 0, therefore all Goodstein sequences eventually reach 0.
□

3 Independence

We now briefly state the independence of Goodstein's Theorem (GT) from Peano Arithmetic (PA).

We know from Gentzen that *transfinite induction* up to ϵ_0 implies the consistency of PA. This is due Gentzen's ordering of possible proof-trees in his sequent calculus for PA have the the order types of ordinals less than ϵ_0. We also note that GT seems to be connected to ordinal arithmetic, specifically with ordinals up to e_0.

The PA proofs which we order are finite objects; we can use Gödel numbering to encode them. So in ordering the proofs, we are thinking about an ordering of code numbers.

In a Gentzen-style argument, the assignment of a place in the ordering to a proof can be handled by primitive recursive functions (p.r. functions), and facts about the dependency relations between proofs at different points in the ordering can be handled by a p.r. function too. A theory in which we can have a version of Gentzen's proof will be one in which we can handle both p.r. functions and handle transfinite induction up to ϵ_0.

It turns out to be enough to have all p.r. functions as axioms together with transfinite induction just for simple quantifier-free formulas containing expressions for these p.r. functions. Such a theory is neither contained in PA (since it can prove the consistency of PA by Gentzen's method, which PA cannot), nor does it contain PA. Therefore, we can prove the consistency of PA by using a theory which is weaker in some respects and stronger in others.

We note the following independence statement due to Kirby and Paris [4].

Kirby-Paris *Goodstein's Theorem is not provable in PA (assuming PA is consistent)*

Goodstein's theorem, depends fundamentally on that there is no infinitely descending chain of ordinals which are sums of powers of ω (i.e. there is no infinite decreasing chain of ordinals less than ϵ_0). It is believable that GT+PA implies the fundamental result that transfinite induction up to ϵ_0 is sound.

We note that there are Natural Gödel numberings for sums of powers of ω, effectively allowing us to to reason about ordinals using their numerical codes. Hence this implies being able to prove GT in PA would conclude that PA can handle transfinite induction up to ϵ_0.

Therefore, if PA could prove Goodstein's theorem, it implies transfinite induction up to ϵ_0. Which in turn implies the consistency of PA, this is a contradiction of Gödel's Second Incompleteness Theorem. This means that GT must be independent of PA[2].
□

References

[1] H.E. Rose. "R. L. Goodstein". In: *Bulletin of the London Mathematical Society* (1988).

[2] Peter Smith. *An Introduction to Gödel's Theorems*. Logic Matters, 2020.

[3] R. L. Goodstein. "On the Restricted Ordinal Theorem". In: *The Journal of Symbolic Logic* (1944).

[4] Laurie Kirbie and Jeff Paris. "Accessible independence results for Peano arithmetic". In: *Bulletin of the London Mathematical Society* (198).

Appendix

[†] Proof that

$$a^n - 1 = (a-1)\sum_{i=0}^{n-1} a^i$$

Let

$$1 + a + a^2 + ... + a^{n-1} = S$$

$$\implies aS = a + a^2 + ... + a^n$$

$$\implies aS - S = a^n - 1$$

$$\implies (a-1)S = a^n - 1$$

$$(a-1)S = (a-1)\sum_{i=0}^{n-1} a^i$$

[††] Proof sketch: if $x > y$ then $o^{n+2,\omega}(x) > o^{n+2,\omega}(y)$
If $x > y$ then when written in base $n+2$ it must contain a coefficient that is greater than the corresponding coefficient in the base $n+2$ representation of y. In other words, for some $a_k(n+2)^{d_k}$ in x and $b_k(n+2)^{d_k}$ in y, we have $a_k > b_k$. Pick the coefficient corresponding to the largest exponent where x and y differ. The ordinal mapping would then produce at least one extra $(n+2)^{d_k}$ for the ordinal corresponding to x. This extra term in x, being the largest exponent that differs from y, is strictly greater than the sum of every other exponent in y that differs from x.