Very large finite numbers

Andrés E. Caicedo

Department of Mathematics
Boise State University

Graduate Student Seminar, October 15, 2008

Introduction

I would like to present two sets of results in finite combinatorics motivated by problems in mathematical logic. The common theme is that the results produce sequences of numbers that grow *very* fast.

I will be discussing regressive functions on dimension 2. This topic is part of the branch of combinatorics known as Ramsey theory.

I will also talk about Goodstein's function, an easy example of a computable function that grows faster than any function one can reasonably understand.

Introduction

I would like to present two sets of results in finite combinatorics motivated by problems in mathematical logic. The common theme is that the results produce sequences of numbers that grow *very* fast.

I will be discussing regressive functions on dimension 2. This topic is part of the branch of combinatorics known as Ramsey theory.

I will also talk about Goodstein's function, an easy example of a computable function that grows faster than any function one can reasonably understand.

Introduction

I would like to present two sets of results in finite combinatorics motivated by problems in mathematical logic. The common theme is that the results produce sequences of numbers that grow *very* fast.

I will be discussing regressive functions on dimension 2. This topic is part of the branch of combinatorics known as Ramsey theory.

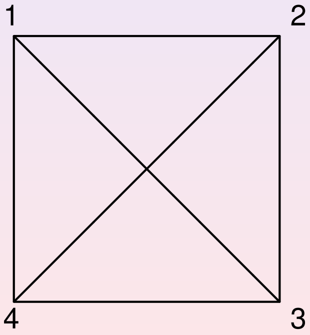
I will also talk about Goodstein's function, an easy example of a computable function that grows faster than any function one can reasonably understand.

Ramsey theory

Given a collection V of vertices, the complete graph on V consists of all the possible edges between these vertices.

Ramsey theory

Given a collection V of vertices, the complete graph on V consists of all the possible edges between these vertices.



Ramsey theory is concerned with substructures and partitions: You color each edge of the complete graph on V and look for a subset W of the set of vertices such that the complete subgraph on W is monochromatic.

Classical Ramsey theory fixes the number of colors in advance, and asks how large V needs to be given the size of the desired set W.

For example: Let's use only two colors. Then V needs to have size 6 to ensure that we can find a monochromatic *triangle*. In symbols, $r_3 = 6$.

Ramsey theory is concerned with substructures and partitions: You color each edge of the complete graph on V and look for a subset W of the set of vertices such that the complete subgraph on W is monochromatic.

Classical Ramsey theory fixes the number of colors in advance, and asks how large V needs to be given the size of the desired set W.

For example: Let's use only two colors. Then V needs to have size 6 to ensure that we can find a monochromatic *triangle*. In symbols, $r_3 = 6$.

Ramsey theory is concerned with substructures and partitions: You color each edge of the complete graph on V and look for a subset W of the set of vertices such that the complete subgraph on W is monochromatic.

Classical Ramsey theory fixes the number of colors in advance, and asks how large V needs to be given the size of the desired set W.

For example: Let's use only two colors. Then V needs to have size 6 to ensure that we can find a monochromatic *triangle*. In symbols, $r_3 = 6$.

This is usually presented in the following way:

> *In any party of six people, there are always three who know each other, or three who do not know one another.*

Five does not suffice: Color the lines of a pentagon "dark" or "transparent" as below, and notice no monochromatic triangles appear.

This is usually presented in the following way:

> *In any party of six people, there are always three who know each other, or three who do not know one another.*

Five does not suffice: Color the lines of a pentagon "dark" or "transparent" as below, and notice no monochromatic triangles appear.

This is usually presented in the following way:

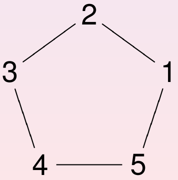
> *In any party of six people, there are always three who know each other, or three who do not know one another.*

Five does not suffice: Color the lines of a pentagon "dark" or "transparent" as below, and notice no monochromatic triangles appear.

This is usually presented in the following way:

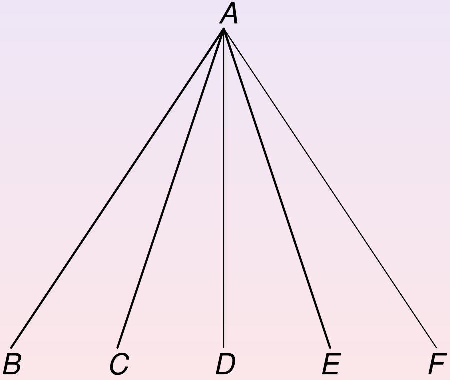
> *In any party of six people, there are always three who know each other, or three who do not know one another.*

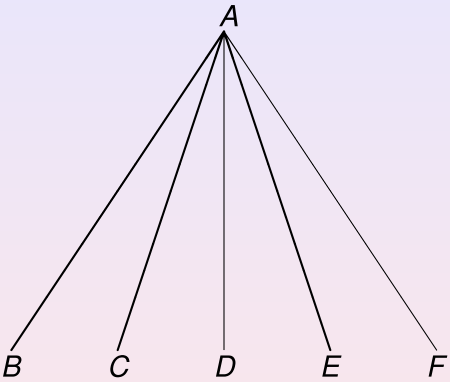
Five does not suffice: Color the lines of a pentagon "dark" or "transparent" as below, and notice no monochromatic triangles appear.



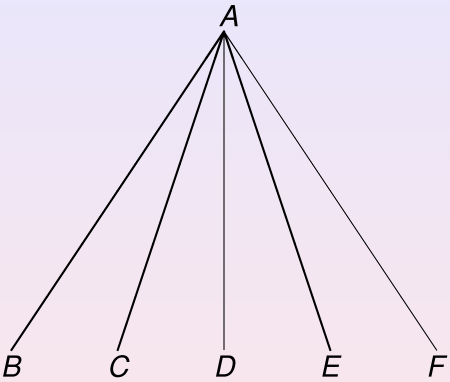
To see that six works, fix one of the vertices, call it A. There are five more vertices, B–F. Necessarily, three of the lines connecting A to these vertices must have the same color.

To see that six works, fix one of the vertices, call it A. There are five more vertices, B–F. Necessarily, three of the lines connecting A to these vertices must have the same color.





Now notice that no line between B, C, E can be "dark", or we obtain a monochromatic dark triangle. But then, if these three lines are light, we obtain a monochromatic light triangle.



Now notice that no line between B, C, E can be "dark", or we obtain a monochromatic dark triangle. But then, if these three lines are light, we obtain a monochromatic light triangle.

Similarly, using only two colors, V needs to have size 18 to ensure that we can find $W \subseteq V$ of size 4, with all the edges among vertices in W of the same color. In symbols, $r_4 = 18$.
The exact value of r_5 is not known. We know that $43 \leq r_5 \leq 49$.

On the other hand, although exact values are not known, we have a good understanding of the rate of growth of r_n, the n^{th} *Ramsey number*, which turns out to be *exponential in n*:

$$2^{n/2} \leq r_n \leq 2^{2n}.$$

Similarly, using only two colors, V needs to have size 18 to ensure that we can find $W \subseteq V$ of size 4, with all the edges among vertices in W of the same color. In symbols, $r_4 = 18$. The exact value of r_5 is not known. We know that $43 \leq r_5 \leq 49$.

On the other hand, although exact values are not known, we have a good understanding of the rate of growth of r_n, the n^{th} *Ramsey number*, which turns out to be *exponential in n*:

$$2^{n/2} \leq r_n \leq 2^{2n}.$$

Similarly, using only two colors, V needs to have size 18 to ensure that we can find $W \subseteq V$ of size 4, with all the edges among vertices in W of the same color. In symbols, $r_4 = 18$. The exact value of r_5 is not known. We know that $43 \leq r_5 \leq 49$.

On the other hand, although exact values are not known, we have a good understanding of the rate of growth of r_n, the n^{th} *Ramsey number*, which turns out to be *exponential in n*:

$$2^{n/2} \leq r_n \leq 2^{2n}.$$

Regressive Ramsey numbers

The work I want to concentrate on studies a similar function R_n, but now we allow the number of colors to increase. To describe the numbers R_n, it is convenient to imagine the vertices in V numbered $1, 2, \ldots, n$. I write ij for the edge between vertices i and j.

Now we use whole numbers as colors. The color of an edge $1i$ is always 0. The color of an edge $2i$ (with $2 < i$) can be 0 or 1. The color of an edge $3i$ (with $3 < i$) can be 0, 1, or 2. And so on. We call these colorings regressive.

Regressive Ramsey numbers

The work I want to concentrate on studies a similar function R_n, but now we allow the number of colors to increase. To describe the numbers R_n, it is convenient to imagine the vertices in V numbered $1, 2, \ldots, n$. I write ij for the edge between vertices i and j.

Now we use whole numbers as colors. The color of an edge $1i$ is always 0. The color of an edge $2i$ (with $2 < i$) can be 0 or 1. The color of an edge $3i$ (with $3 < i$) can be 0, 1, or 2. And so on. We call these colorings regressive.

Now, instead of a mono-chromatic "polygon," we look at min-homogeneous sets. (Homogeneous sets of more than two elements might not exist.)

Let $W \subset V$, say $W = \{i_1, i_2, \dots, i_k\}$ where $i_1 < \cdots < i_k$. We say that W is min-homogeneous for a given regressive coloring if all edges $i_1 i_2, i_1 i_3, \dots, i_1 i_k$ receive the same color $c < i_1$; all the edges $i_2 i_3, i_2 i_4, \dots, i_2 i_k$ receive the same color $d < i_2$ (that may be different from c); and so on.

Let R_n be the smallest number of vertices that V needs to have to guarantee that for any regressive coloring there is $W \subset V$ of size n and min-homogeneous.

Now, instead of a mono-chromatic "polygon," we look at min-homogeneous sets. (Homogeneous sets of more than two elements might not exist.)

Let $W \subset V$, say $W = \{i_1, i_2, \dots, i_k\}$ where $i_1 < \cdots < i_k$. We say that W is min-homogeneous for a given regressive coloring if all edges $i_1 i_2, i_1 i_3, \dots, i_1 i_k$ receive the same color $c < i_1$; all the edges $i_2 i_3, i_2 i_4, \dots, i_2 i_k$ receive the same color $d < i_2$ (that may be different from c); and so on.

Let R_n be the smallest number of vertices that V needs to have to guarantee that for any regressive coloring there is $W \subset V$ of size n and min-homogeneous.

Now, instead of a mono-chromatic "polygon," we look at min-homogeneous sets. (Homogeneous sets of more than two elements might not exist.)

Let $W \subset V$, say $W = \{i_1, i_2, \dots, i_k\}$ where $i_1 < \cdots < i_k$. We say that W is min-homogeneous for a given regressive coloring if all edges $i_1 i_2, i_1 i_3, \dots, i_1 i_k$ receive the same color $c < i_1$; all the edges $i_2 i_3, i_2 i_4, \dots, i_2 i_k$ receive the same color $d < i_2$ (that may be different from c); and so on.

Let R_n be the smallest number of vertices that V needs to have to guarantee that for any regressive coloring there is $W \subset V$ of size n and min-homogeneous.

For example, $R_3 = 3$, because 12 and 13 always receive color 0, so 23 can receive any color, and $\{1, 2, 3\}$ is min-homogeneous.

Similarly, $R_4 = 5$ and we can always find a min-homogeneous set of the form $\{1, 2, a, b\}$ with $3 \leq a < b \leq 5$.

For example, $R_3 = 3$, because 12 and 13 always receive color 0, so 23 can receive any color, and $\{1, 2, 3\}$ is min-homogeneous.

Similarly, $R_4 = 5$ and we can always find a min-homogeneous set of the form $\{1, 2, a, b\}$ with $3 \leq a < b \leq 5$.

The numbers R_n were originally studied by Kanamori and McAloon. Their result has two parts.

Theorem (Kanamori-McAloon)

For any $n \in \mathbb{N}$, R_n exists.

To state the second part of their theorem, we need the notion of primitive recursive function.

The numbers R_n were originally studied by Kanamori and McAloon. Their result has two parts.

Theorem (Kanamori-McAloon)

For any $n \in \mathbb{N}$, R_n exists.

To state the second part of their theorem, we need the notion of primitive recursive function.

The numbers R_n were originally studied by Kanamori and McAloon. Their result has two parts.

Theorem (Kanamori-McAloon)

For any $n \in \mathbb{N}$, R_n exists.

To state the second part of their theorem, we need the notion of primitive recursive function.

We define the primitive recursive functions by starting with the basic functions $f : \mathbb{N}^k \to \mathbb{N}$ consisting of constant functions, the successor function $S(n) = n + 1$, and projections $f(n_1, \ldots, n_k) = n_i$.

We *close* these functions under composition, and recursion: If $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^{k+2} \to \mathbb{N}$ are primitive recursive, so is $f : \mathbb{N}^{k+1} \to \mathbb{N}$ given by

$$f(\vec{x}, y) = \begin{cases} g(\vec{x}) & \text{if } y = 0, \\ h(\vec{x}, n, f(\vec{x}, n)) & \text{if } y = n + 1. \end{cases}$$

To understand this definition, it is perhaps useful to see a few examples.

We define the primitive recursive functions by starting with the basic functions $f : \mathbb{N}^k \to \mathbb{N}$ consisting of constant functions, the successor function $S(n) = n + 1$, and projections $f(n_1, \dots, n_k) = n_i$.

We *close* these functions under composition, and recursion: If $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^{k+2} \to \mathbb{N}$ are primitive recursive, so is $f : \mathbb{N}^{k+1} \to \mathbb{N}$ given by

$$f(\vec{x}, y) = \begin{cases} g(\vec{x}) & \text{if } y = 0, \\ h(\vec{x}, n, f(\vec{x}, n)) & \text{if } y = n + 1. \end{cases}$$

To understand this definition, it is perhaps useful to see a few examples.

We define the primitive recursive functions by starting with the basic functions $f : \mathbb{N}^k \to \mathbb{N}$ consisting of constant functions, the successor function $S(n) = n + 1$, and projections $f(n_1, \dots, n_k) = n_i$.
We *close* these functions under composition, and recursion: If $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^{k+2} \to \mathbb{N}$ are primitive recursive, so is $f : \mathbb{N}^{k+1} \to \mathbb{N}$ given by

$$f(\vec{x}, y) = \begin{cases} g(\vec{x}) & \text{if } y = 0, \\ h(\vec{x}, n, f(\vec{x}, n)) & \text{if } y = n + 1. \end{cases}$$

To understand this definition, it is perhaps useful to see a few examples.

The following functions are primitive recursive:

1. $f(x, y) = x + y$. We can describe it in the format above noting that

$$f(x, y) = \begin{cases} x & \text{if } y = 0, \\ (x + n) + 1 & \text{if } y = n + 1. \end{cases}$$

2. $f(x, y) = xy$.
3. $f(x, y) = x^y$.
4. $f(x, y) = x^{x^{x^{\cdot^{\cdot^{\cdot^{x}}}}}}$, where the tower has height y.
5. Etc.

The following functions are primitive recursive:

1. $f(x, y) = x + y$. We can describe it in the format above noting that

$$f(x,y) = \begin{cases} x & \text{if } y = 0, \\ (x+n)+1 & \text{if } y = n+1. \end{cases}$$

2. $f(x, y) = xy$.
3. $f(x, y) = x^y$.
4. $f(x, y) = x^{x^{x^{\cdot^{\cdot^{\cdot^{x}}}}}}$, where the tower has height y.
5. Etc.

The following functions are primitive recursive:

1. $f(x, y) = x + y$. We can describe it in the format above noting that

$$f(x,y) = \begin{cases} x & \text{if } y = 0, \\ (x+n)+1 & \text{if } y = n+1. \end{cases}$$

2. $f(x, y) = xy$.
3. $f(x, y) = x^y$.
4. $f(x, y) = x^{x^{x^{\cdot^{\cdot^{\cdot^{x}}}}}}$, where the tower has height y.
5. Etc.

The following functions are primitive recursive:

1. $f(x, y) = x + y$. We can describe it in the format above noting that

$$f(x, y) = \begin{cases} x & \text{if } y = 0, \\ (x + n) + 1 & \text{if } y = n + 1. \end{cases}$$

2. $f(x, y) = xy$.
3. $f(x, y) = x^y$.
4. $f(x, y) = x^{x^{x^{\cdot^{\cdot^{\cdot^{x}}}}}}$, where the tower has height y.
5. Etc.

The following functions are primitive recursive:

1. $f(x,y) = x + y$. We can describe it in the format above noting that

$$f(x,y) = \begin{cases} x & \text{if } y = 0, \\ (x+n)+1 & \text{if } y = n+1. \end{cases}$$

2. $f(x,y) = xy$.
3. $f(x,y) = x^y$.
4. $f(x,y) = x^{x^{x^{\cdot^{\cdot^{\cdot^{x}}}}}}$, where the tower has height y.
5. Etc.

The following functions are primitive recursive:

1. $f(x,y) = x + y$. We can describe it in the format above noting that

$$f(x,y) = \begin{cases} x & \text{if } y = 0, \\ (x+n)+1 & \text{if } y = n+1. \end{cases}$$

2. $f(x,y) = xy$.
3. $f(x,y) = x^y$.
4. $f(x,y) = x^{x^{x^{\cdot^{\cdot^{\cdot^{x}}}}}}$, where the tower has height y.
5. Etc.

We can now state the second part of the Kanamori-McAloon result. Say that $f : \mathbb{N} \to \mathbb{N}$ grows faster than $g : \mathbb{N} \to \mathbb{N}$ or that f eventually dominates g, iff $f(n) > g(n)$ for all but finitely many values of n.

Theorem (Kanamori-McAloon)

The function $f(n) = R_n$ grows faster than any primitive recursive function.

Their proof used tools of mathematical logic and did not produce any explicit bounds.

Several mathematicians tried to provide explicit proofs of the Kanamori-McAloon result; among others, Kojman, Shelah, Blanchard, and Weiermann.

We can now state the second part of the Kanamori-McAloon result. Say that $f : \mathbb{N} \to \mathbb{N}$ grows faster than $g : \mathbb{N} \to \mathbb{N}$ or that f eventually dominates g, iff $f(n) > g(n)$ for all but finitely many values of n.

Theorem (Kanamori-McAloon)

The function $f(n) = R_n$ grows faster than any primitive recursive function.

Their proof used tools of mathematical logic and did not produce any explicit bounds.

Several mathematicians tried to provide explicit proofs of the Kanamori-McAloon result; among others, Kojman, Shelah, Blanchard, and Weiermann.

We can now state the second part of the Kanamori-McAloon result. Say that $f : \mathbb{N} \to \mathbb{N}$ grows faster than $g : \mathbb{N} \to \mathbb{N}$ or that f eventually dominates g, iff $f(n) > g(n)$ for all but finitely many values of n.

Theorem (Kanamori-McAloon)

The function $f(n) = R_n$ grows faster than any primitive recursive function.

Their proof used tools of mathematical logic and did not produce any explicit bounds.

Several mathematicians tried to provide explicit proofs of the Kanamori-McAloon result; among others, Kojman, Shelah, Blanchard, and Weiermann.

We can now state the second part of the Kanamori-McAloon result. Say that $f : \mathbb{N} \to \mathbb{N}$ grows faster than $g : \mathbb{N} \to \mathbb{N}$ or that f eventually dominates g, iff $f(n) > g(n)$ for all but finitely many values of n.

Theorem (Kanamori-McAloon)

The function $f(n) = R_n$ grows faster than any primitive recursive function.

Their proof used tools of mathematical logic and did not produce any explicit bounds.

Several mathematicians tried to provide explicit proofs of the Kanamori-McAloon result; among others, Kojman, Shelah, Blanchard, and Weiermann.

Definition

Ackermann's function $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is defined "by double recursion" as follows:

- $A(0, m) = m + 1.$
- $A(n, 0) = A(n - 1, 1)$ for $n > 0.$
- $A(n, m) = A(n - 1, A(n, m - 1))$ for $n, m > 0.$

Let $A_n = A(n, \cdot)$. For example, $A_0(m) = m + 1$, $A_1(m) = m + 2$, $A_2(m) = 2m + 3$, A_3 grows like 2^m, A_4 grows like a tower of two's of length m, etc.

Definition

Ackermann's function $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is defined "by double recursion" as follows:

- $A(0, m) = m + 1$.
- $A(n, 0) = A(n-1, 1)$ for $n > 0$.
- $A(n, m) = A(n-1, A(n, m-1))$ for $n, m > 0$.

Let $A_n = A(n, \cdot)$. For example, $A_0(m) = m + 1$, $A_1(m) = m + 2$, $A_2(m) = 2m + 3$, A_3 grows like 2^m, A_4 grows like a tower of two's of length m, etc.

Definition

Ackermann's function $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is defined "by double recursion" as follows:

- $A(0, m) = m + 1$.
- $A(n, 0) = A(n - 1, 1)$ for $n > 0$.
- $A(n, m) = A(n - 1, A(n, m - 1))$ for $n, m > 0$.

Let $A_n = A(n, \cdot)$. For example, $A_0(m) = m + 1$, $A_1(m) = m + 2$, $A_2(m) = 2m + 3$, A_3 grows like 2^m, A_4 grows like a tower of two's of length m, etc.

Definition

Ackermann's function $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is defined "by double recursion" as follows:

- $A(0, m) = m + 1$.
- $A(n, 0) = A(n - 1, 1)$ for $n > 0$.
- $A(n, m) = A(n - 1, A(n, m - 1))$ for $n, m > 0$.

Let $A_n = A(n, \cdot)$. For example, $A_0(m) = m + 1$, $A_1(m) = m + 2$, $A_2(m) = 2m + 3$, A_3 grows like 2^m, A_4 grows like a tower of two's of length m, etc.

Definition

Ackermann's function $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is defined "by double recursion" as follows:

- $A(0, m) = m + 1$.
- $A(n, 0) = A(n-1, 1)$ for $n > 0$.
- $A(n, m) = A(n-1, A(n, m-1))$ for $n, m > 0$.

Let $A_n = A(n, \cdot)$. For example, $A_0(m) = m + 1$, $A_1(m) = m + 2$, $A_2(m) = 2m + 3$, A_3 grows like 2^m, A_4 grows like a tower of two's of length m, etc.

Theorem

Ackermann's function is not primitive recursive. In fact, $A_n(n)$ grows faster than any primitive recursive function.

Ackermann's original definition used three variables. The version given above was introduced by Rafael Robinson and Rózsa Péter, and has become the standard example of a computable function that is not primitive recursive.

Theorem

Ackermann's function is not primitive recursive. In fact, $A_n(n)$ grows faster than any primitive recursive function.

Ackermann's original definition used three variables. The version given above was introduced by Rafael Robinson and Rózsa Péter, and has become the standard example of a computable function that is not primitive recursive.

Let $g(n, m)$ be the least l such that if $V = \{m, m+1, \ldots, l\}$ then any regressive coloring of V admits a min-homogeneous W of size n. For example, $g(2, m) = m + 1$ and $g(3, m) = 2m + 1$.

The n^{th} regressive Ramsey number R_n is $g(n, 1) = g(n - 1, 2)$.

Let $g(n, m)$ be the least l such that if $V = \{m, m+1, \ldots, l\}$ then any regressive coloring of V admits a min-homogeneous W of size n. For example, $g(2, m) = m + 1$ and $g(3, m) = 2m + 1$. The n^{th} regressive Ramsey number R_n is $g(n, 1) = g(n - 1, 2)$.

Set $g^0(n,m) = m$ and $g^{k+1}(n,m) = g(n, g^k(n,m))$. We then have:

Theorem

$g(n+1, m+1) \geq g(n, g(n+1, m) + 1)$. In particular, for $n \geq 2$ and $m \geq 1$, $g(n,m) \geq A_{n-1}(m-1)$.

Set $g^0(n, m) = m$ and $g^{k+1}(n, m) = g(n, g^k(n, m))$. We then have:

Theorem

$g(n+1, m+1) \geq g(n, g(n+1, m) + 1)$. In particular, for $n \geq 2$ and $m \geq 1$, $g(n, m) \geq A_{n-1}(m-1)$.

Theorem

1. *For $m \geq 2$, $g(4, m) \leq 2^m(m+2) - 2^{m-1} + 1$.*
2. *For all n there is a constant c_n such that $g(n, m) < A_{n-1}(c_n m)$ for all m.*

Thus $g(n, \cdot)$ grows like A_{n-1} and $g(\cdot, m)$ grows like Ackermann's function.

Theorem

1. *For* $m \geq 2$, $g(4, m) \leq 2^m(m+2) - 2^{m-1} + 1$.
2. *For all* n *there is a constant* c_n *such that* $g(n, m) < A_{n-1}(c_n m)$ *for all* m.

Thus $g(n, \cdot)$ grows like A_{n-1} and $g(\cdot, m)$ grows like Ackermann's function.

Theorem

1. *For $m \geq 2$, $g(4, m) \leq 2^m(m+2) - 2^{m-1} + 1$.*
2. *For all n there is a constant c_n such that $g(n, m) < A_{n-1}(c_n m)$ for all m.*

Thus $g(n, \cdot)$ grows like A_{n-1} and $g(\cdot, m)$ grows like Ackermann's function.

The proofs I have obtained are explicit, considerably improve the previously known bounds, and identify the right rate of growth of the numbers R_n. They provide us with a combinatorial proof of a result formerly obtained by non-constructive means. They also identify the first example of a naturally occurring function whose rate of growth is *Ackermannian*.

Similar problems can be studied in dimension 3 or larger, and many interesting questions related to the computation of explicit bounds remain.

The proofs I have obtained are explicit, considerably improve the previously known bounds, and identify the right rate of growth of the numbers R_n. They provide us with a combinatorial proof of a result formerly obtained by non-constructive means. They also identify the first example of a naturally occurring function whose rate of growth is *Ackermannian*.

Similar problems can be studied in dimension 3 or larger, and many interesting questions related to the computation of explicit bounds remain.

Goodstein's function

Definition

The depth-1 base b representation of $n \in \mathbb{N}$ is just the usual base b representation of n:

$$n = b^{m_1} n_1 + \cdots + b^{m_k} n_k$$

where $m_1 > \cdots > m_k \geq 0$ and $1 \leq n_i < b$ for each i.
The depth-$(m+1)$ representation is obtained by replacing each m_i with their depth-m base b representation.

Goodstein's function

Definition

The depth-1 base b representation of $n \in \mathbb{N}$ is just the usual base b representation of n:

$$n = b^{m_1} n_1 + \cdots + b^{m_k} n_k$$

where $m_1 > \cdots > m_k \geq 0$ and $1 \leq n_i < b$ for each i.
The depth-$(m+1)$ representation is obtained by replacing each m_i with their depth-m base b representation.

For example, the depth-1 base 2 representation of 266 is $2^8 + 2^3 + 2^1$, its depth-2 base 2 representation is $2^{2^3} + 2^{2^1+1} + 2^1$ and so its depth-3 (or higher) base 2 representation is

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

For any n and b, as m increases, the depth-$(m+1)$ base b representations of n eventually stabilize.

Definition

We call this stable representation the complete base b representation of $n \in \mathbb{N}$.

For example, the depth-1 base 2 representation of 266 is $2^8 + 2^3 + 2^1$, its depth-2 base 2 representation is $2^{2^3} + 2^{2^1+1} + 2^1$ and so its depth-3 (or higher) base 2 representation is

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

For any n and b, as m increases, the depth-$(m+1)$ base b representations of n eventually stabilize.

Definition

We call this stable representation the complete base b representation of $n \in \mathbb{N}$.

For example, the depth-1 base 2 representation of 266 is $2^8 + 2^3 + 2^1$, its depth-2 base 2 representation is $2^{2^3} + 2^{2^1+1} + 2^1$ and so its depth-3 (or higher) base 2 representation is

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

For any n and b, as m increases, the depth-$(m+1)$ base b representations of n eventually stabilize.

Definition

We call this stable representation the complete base b representation of $n \in \mathbb{N}$.

Definition

The change of base function $R_b : \mathbb{N} \to \mathbb{N}$ takes a natural number n, and then replaces every b with $b+1$ in the complete base b representation of n.

For example,

$$R_2(266) = 3^{3^{3+1}} + 3^{3+1} + 3 = 443426488243037769948249630619149892887.$$

Definition

The change of base function $R_b : \mathbb{N} \to \mathbb{N}$ takes a natural number n, and then replaces every b with $b+1$ in the complete base b representation of n.

For example,

$$\begin{aligned} R_2(266) &= 3^{3^{3+1}} + 3^{3+1} + 3 \\ &= 443426488243037769948249630619149892887. \end{aligned}$$

Definition

The Goodstein Sequence beginning with n, $(n)_k$, is defined by $(n)_1 = n$ and for $k \geq 1$,

$$(n)_{k+1} = \begin{cases} R_{k+1}((n)_k) - 1 & \text{if} \quad (n)_k > 0 \\ 0 & \text{if} \quad (n)_k = 0. \end{cases}$$

For example, the sequence for $n = 3$ is $3, 3, 3, 2, 1, 0, 0, \ldots$

Definition

The Goodstein Function $\mathcal{G} : \mathbb{N} \to \mathbb{N}$ is defined to be the smallest number k for which $(n)_k = 0$.

Clearly, $\mathcal{G}$ is computable.

Definition

The Goodstein Sequence beginning with n, $(n)_k$, is defined by $(n)_1 = n$ and for $k \geq 1$,

$$(n)_{k+1} = \begin{cases} R_{k+1}((n)_k) - 1 & \text{if} \quad (n)_k > 0 \\ 0 & \text{if} \quad (n)_k = 0. \end{cases}$$

For example, the sequence for $n = 3$ is $3, 3, 3, 2, 1, 0, 0, \ldots$

Definition

The Goodstein Function $\mathcal{G} : \mathbb{N} \to \mathbb{N}$ is defined to be the smallest number k for which $(n)_k = 0$.

Clearly, $\mathcal{G}$ is computable.

Definition

The Goodstein Sequence beginning with n, $(n)_k$, is defined by $(n)_1 = n$ and for $k \geq 1$,

$$(n)_{k+1} = \begin{cases} R_{k+1}((n)_k) - 1 & \text{if} \quad (n)_k > 0 \\ 0 & \text{if} \quad (n)_k = 0. \end{cases}$$

For example, the sequence for $n = 3$ is $3, 3, 3, 2, 1, 0, 0, \ldots$

Definition

The Goodstein Function $\mathcal{G} : \mathbb{N} \to \mathbb{N}$ is defined to be the smallest number k for which $(n)_k = 0$.

Clearly, $\mathcal{G}$ is computable.

Definition

The Goodstein Sequence beginning with n, $(n)_k$, is defined by $(n)_1 = n$ and for $k \geq 1$,

$$(n)_{k+1} = \begin{cases} R_{k+1}((n)_k) - 1 & \text{if} \quad (n)_k > 0 \\ 0 & \text{if} \quad (n)_k = 0. \end{cases}$$

For example, the sequence for $n = 3$ is $3, 3, 3, 2, 1, 0, 0, \ldots$

Definition

The Goodstein Function $\mathcal{G} : \mathbb{N} \to \mathbb{N}$ is defined to be the smallest number k for which $(n)_k = 0$.

Clearly, $\mathcal{G}$ is computable.

Theorem (Goodstein)

The function $\mathcal{G}$ is well defined, i.e., $\mathcal{G}(n)$ exists for all n.

Here are the first few values of $\mathcal{G}$:

$$\mathcal{G}(0) = 1$$

$$\mathcal{G}(1) = 2$$

$$\mathcal{G}(2) = 4$$

$$\mathcal{G}(3) = 6$$

Theorem (Goodstein)

The function $\mathcal{G}$ is well defined, i.e., $\mathcal{G}(n)$ exists for all n.

Here are the first few values of $\mathcal{G}$:

$$\mathcal{G}(0) = 1$$

$$\mathcal{G}(1) = 2$$

$$\mathcal{G}(2) = 4$$

$$\mathcal{G}(3) = 6$$

Theorem (Goodstein)

The function $\mathcal{G}$ is well defined, i.e., $\mathcal{G}(n)$ exists for all n.

Here are the first few values of $\mathcal{G}$:

$$\mathcal{G}(0) = 1$$

$$\mathcal{G}(1) = 2$$

$$\mathcal{G}(2) = 4$$

$$\mathcal{G}(3) = 6$$

Theorem (Goodstein)

The function $\mathcal{G}$ is well defined, i.e., $\mathcal{G}(n)$ exists for all n.

Here are the first few values of $\mathcal{G}$:

$$\mathcal{G}(0) = 1$$

$$\mathcal{G}(1) = 2$$

$$\mathcal{G}(2) = 4$$

$$\mathcal{G}(3) = 6$$

Theorem (Goodstein)

The function $\mathcal{G}$ is well defined, i.e., $\mathcal{G}(n)$ exists for all n.

Here are the first few values of $\mathcal{G}$:

$$\mathcal{G}(0) = 1$$

$$\mathcal{G}(1) = 2$$

$$\mathcal{G}(2) = 4$$

$$\mathcal{G}(3) = 6$$

$$\mathcal{G}(4) = 3 \cdot 2^{402653211} - 2 \approx 6.895 \times 10^{121210694}.$$

The number of digits of $\mathcal{G}(5)$ is much larger than $\mathcal{G}(4)$. The number of elementary particles in the universe is estimated to be (significantly) below 10^{90}.

To analyze this function one is required to step out of the natural numbers and consider the more general notion of ordinals.

$$\mathcal{G}(4) = 3 \cdot 2^{402653211} - 2 \approx 6.895 \times 10^{121210694}.$$

The number of digits of $\mathcal{G}(5)$ is much larger than $\mathcal{G}(4)$. The number of elementary particles in the universe is estimated to be (significantly) below 10^{90}.

To analyze this function one is required to step out of the natural numbers and consider the more general notion of ordinals.

$$\mathcal{G}(4) = 3 \cdot 2^{402653211} - 2 \approx 6.895 \times 10^{121210694}.$$

The number of digits of $\mathcal{G}(5)$ is much larger than $\mathcal{G}(4)$. The number of elementary particles in the universe is estimated to be (significantly) below 10^{90}.
To analyze this function one is required to step out of the natural numbers and consider the more general notion of ordinals.

The ordinals are obtained by continuing the natural numbers into the *transfinite*. We have two operations:

- Given any ordinal α, one can add 1 to it to obtain its successor ordinal, $\alpha + 1$.
- Or we can look at a collection of ordinals without a maximum, and add its least upper bound to it. These are called limit ordinals.

It is customary to call ω the first infinite ordinal. The first few ordinals are $0, 1, 2, \dots,$ and then:

$$\omega, \omega + 1, \dots, \omega \cdot 2, \omega \cdot 2 + 1, \dots, \omega \cdot 3, \dots, \omega^2, \dots, \omega^3, \dots, \omega^\omega, \dots$$

The ordinals are obtained by continuing the natural numbers into the *transfinite*. We have two operations:

- Given any ordinal α, one can add 1 to it to obtain its successor ordinal, $\alpha + 1$.
- Or we can look at a collection of ordinals without a maximum, and add its least upper bound to it. These are called limit ordinals.

It is customary to call ω the first infinite ordinal. The first few ordinals are $0, 1, 2, \dots,$ and then:

$$\omega, \omega + 1, \dots, \omega \cdot 2, \omega \cdot 2 + 1, \dots, \omega \cdot 3, \dots, \omega^2, \dots, \omega^3, \dots, \omega^\omega, \dots$$

The ordinals are obtained by continuing the natural numbers into the *transfinite*. We have two operations:

- Given any ordinal α, one can add 1 to it to obtain its successor ordinal, $\alpha + 1$.
- Or we can look at a collection of ordinals without a maximum, and add its least upper bound to it. These are called limit ordinals.

It is customary to call ω the first infinite ordinal. The first few ordinals are $0, 1, 2, \ldots$, and then:

$$\omega, \omega + 1, \ldots, \omega \cdot 2, \omega \cdot 2 + 1, \ldots, \omega \cdot 3, \ldots, \omega^2, \ldots, \omega^3, \ldots, \omega^\omega, \ldots$$

The ordinals are obtained by continuing the natural numbers into the *transfinite*. We have two operations:

- Given any ordinal α, one can add 1 to it to obtain its successor ordinal, $\alpha + 1$.
- Or we can look at a collection of ordinals without a maximum, and add its least upper bound to it. These are called limit ordinals.

It is customary to call ω the first infinite ordinal. The first few ordinals are $0, 1, 2, \dots$, and then:

$$\omega, \omega + 1, \dots, \omega \cdot 2, \omega \cdot 2 + 1, \dots, \omega \cdot 3, \dots, \omega^2, \dots, \omega^3, \dots, \omega^\omega, \dots$$

The ordinals are obtained by continuing the natural numbers into the *transfinite*. We have two operations:

- Given any ordinal α, one can add 1 to it to obtain its successor ordinal, $\alpha + 1$.
- Or we can look at a collection of ordinals without a maximum, and add its least upper bound to it. These are called limit ordinals.

It is customary to call ω the first infinite ordinal. The first few ordinals are $0, 1, 2, \dots,$ and then:

$$\omega, \omega + 1, \dots, \omega \cdot 2, \omega \cdot 2 + 1, \dots, \omega \cdot 3, \dots, \omega^2, \dots, \omega^3, \dots, \omega^\omega, \dots$$

We denote by ϵ_0 the first ordinal α such that $\alpha = \omega^\alpha$.
Any ordinal $\alpha < \epsilon_0$ can be written in a unique way as $\alpha = \omega^\beta(\gamma + 1)$ where $\beta < \alpha$. Define for limit $\alpha < \epsilon_0$ an increasing sequence $d(\alpha, n)$ that converges to α by setting

$$d(\alpha, n) = \omega^\beta \gamma + \begin{cases} \omega^\delta n & \text{if } \beta = \delta + 1, \\ \omega^{d(\beta, n)} & \text{if } \beta \text{ is limit.} \end{cases}$$

For example, $d(\omega, n) = n$, $d(\omega^3, n) = \omega^2 \cdot n$, $d(\omega^{\omega^\omega}, n) = \omega^{\omega^n}$.

We denote by ϵ_0 the first ordinal α such that $\alpha = \omega^\alpha$. Any ordinal $\alpha < \epsilon_0$ can be written in a unique way as $\alpha = \omega^\beta(\gamma + 1)$ where $\beta < \alpha$. Define for limit $\alpha < \epsilon_0$ an increasing sequence $d(\alpha, n)$ that converges to α by setting

$$d(\alpha, n) = \omega^\beta\gamma + \begin{cases} \omega^\delta n & \text{if } \beta = \delta + 1, \\ \omega^{d(\beta,n)} & \text{if } \beta \text{ is limit.} \end{cases}$$

For example, $d(\omega, n) = n$, $d(\omega^3, n) = \omega^2 \cdot n$, $d(\omega^{\omega^\omega}, n) = \omega^{\omega^n}$.

We denote by ϵ_0 the first ordinal α such that $\alpha = \omega^\alpha$. Any ordinal $\alpha < \epsilon_0$ can be written in a unique way as $\alpha = \omega^\beta(\gamma + 1)$ where $\beta < \alpha$. Define for limit $\alpha < \epsilon_0$ an increasing sequence $d(\alpha, n)$ that converges to α by setting

$$d(\alpha, n) = \omega^\beta \gamma + \begin{cases} \omega^\delta n & \text{if } \beta = \delta + 1, \\ \omega^{d(\beta, n)} & \text{if } \beta \text{ is limit.} \end{cases}$$

For example, $d(\omega, n) = n$, $d(\omega^3, n) = \omega^2 \cdot n$, $d(\omega^{\omega^\omega}, n) = \omega^{\omega^n}$.

We denote by ϵ_0 the first ordinal α such that $\alpha = \omega^\alpha$. Any ordinal $\alpha < \epsilon_0$ can be written in a unique way as $\alpha = \omega^\beta(\gamma + 1)$ where $\beta < \alpha$. Define for limit $\alpha < \epsilon_0$ an increasing sequence $d(\alpha, n)$ that converges to α by setting

$$d(\alpha, n) = \omega^\beta \gamma + \begin{cases} \omega^\delta n & \text{if } \beta = \delta + 1, \\ \omega^{d(\beta, n)} & \text{if } \beta \text{ is limit.} \end{cases}$$

For example, $d(\omega, n) = n$, $d(\omega^3, n) = \omega^2 \cdot n$, $d(\omega^{\omega^\omega}, n) = \omega^{\omega^n}$.

The Löb-Wainer fast growing hierarchy $(f_\alpha)_{\alpha<\epsilon_0}$ of functions $f : \mathbb{N} \to \mathbb{N}$ is:

Definition

1. $f_0(n) = n + 1$.
2. For $\alpha < \epsilon_0$, $f_{\alpha+1}(n) = f_\alpha^n(n)$, where the superindex indicates that f_α is iterated n times.
3. For limit $\alpha < \epsilon_0$, $f_\alpha(n) = f_{d(\alpha,n)}(n)$.

For example: $f_1(n) = 2n$, $f_2(n) = n2^n$ and $f_3(n)$ is (significantly) larger than a stack of powers of two of length n; $f_\omega(n) = f_n(n)$ grows like Ackermann's function $A_n(n)$.

The Löb-Wainer fast growing hierarchy $(f_\alpha)_{\alpha<\epsilon_0}$ of functions $f : \mathbb{N} \to \mathbb{N}$ is:

Definition

1. $f_0(n) = n + 1$.
2. For $\alpha < \epsilon_0$, $f_{\alpha+1}(n) = f_\alpha^n(n)$, where the superindex indicates that f_α is iterated n times.
3. For limit $\alpha < \epsilon_0$, $f_\alpha(n) = f_{d(\alpha,n)}(n)$.

For example: $f_1(n) = 2n$, $f_2(n) = n2^n$ and $f_3(n)$ is (significantly) larger than a stack of powers of two of length n; $f_\omega(n) = f_n(n)$ grows like Ackermann's function $A_n(n)$.

The Löb-Wainer fast growing hierarchy $(f_\alpha)_{\alpha<\epsilon_0}$ of functions $f : \mathbb{N} \to \mathbb{N}$ is:

Definition

1. $f_0(n) = n + 1$.
2. For $\alpha < \epsilon_0$, $f_{\alpha+1}(n) = f_\alpha^n(n)$, where the superindex indicates that f_α is iterated n times.
3. For limit $\alpha < \epsilon_0$, $f_\alpha(n) = f_{d(\alpha,n)}(n)$.

For example: $f_1(n) = 2n$, $f_2(n) = n2^n$ and $f_3(n)$ is (significantly) larger than a stack of powers of two of length n; $f_\omega(n) = f_n(n)$ grows like Ackermann's function $A_n(n)$.

The Löb-Wainer fast growing hierarchy $(f_\alpha)_{\alpha<\epsilon_0}$ of functions $f : \mathbb{N} \to \mathbb{N}$ is:

Definition

1. $f_0(n) = n + 1$.
2. For $\alpha < \epsilon_0$, $f_{\alpha+1}(n) = f_\alpha^n(n)$, where the superindex indicates that f_α is iterated n times.
3. For limit $\alpha < \epsilon_0$, $f_\alpha(n) = f_{d(\alpha,n)}(n)$.

For example: $f_1(n) = 2n$, $f_2(n) = n2^n$ and $f_3(n)$ is (significantly) larger than a stack of powers of two of length n; $f_\omega(n) = f_n(n)$ grows like Ackermann's function $A_n(n)$.

The Löb-Wainer fast growing hierarchy $(f_\alpha)_{\alpha<\epsilon_0}$ of functions $f : \mathbb{N} \to \mathbb{N}$ is:

Definition

1. $f_0(n) = n + 1$.
2. For $\alpha < \epsilon_0$, $f_{\alpha+1}(n) = f_\alpha^n(n)$, where the superindex indicates that f_α is iterated n times.
3. For limit $\alpha < \epsilon_0$, $f_\alpha(n) = f_{d(\alpha,n)}(n)$.

For example: $f_1(n) = 2n$, $f_2(n) = n2^n$ and $f_3(n)$ is (significantly) larger than a stack of powers of two of length n; $f_\omega(n) = f_n(n)$ grows like Ackermann's function $A_n(n)$.

Peano Arithmetic PA is the formal theory that captures the intuitive notion of "finite mathematics." If a result is not provable in PA, it must use infinite objects in an essential way.

Theorem (Wainer)

1. *Each f_α is strictly increasing.*
2. *If $\alpha < \beta < \epsilon_0$ then f_α is eventually dominated by f_β.*
3. *Each f_α is computable, and provably total in PA.*
4. *Any computable f provably total in PA is eventually dominated by some f_α, $\alpha < \epsilon_0$.*

Peano Arithmetic PA is the formal theory that captures the intuitive notion of "finite mathematics." If a result is not provable in PA, it must use infinite objects in an essential way.

Theorem (Wainer)

1. *Each f_α is strictly increasing.*
2. *If $\alpha < \beta < \epsilon_0$ then f_α is eventually dominated by f_β.*
3. *Each f_α is computable, and provably total in* PA.
4. *Any computable f provably total in* PA *is eventually dominated by some f_α, $\alpha < \epsilon_0$.*

Peano Arithmetic PA is the formal theory that captures the intuitive notion of "finite mathematics." If a result is not provable in PA, it must use infinite objects in an essential way.

Theorem (Wainer)

1. *Each f_α is strictly increasing.*
2. *If $\alpha < \beta < \epsilon_0$ then f_α is eventually dominated by f_β.*
3. *Each f_α is computable, and provably total in PA.*
4. *Any computable f provably total in PA is eventually dominated by some f_α, $\alpha < \epsilon_0$.*

Peano Arithmetic PA is the formal theory that captures the intuitive notion of "finite mathematics." If a result is not provable in PA, it must use infinite objects in an essential way.

Theorem (Wainer)

1. *Each f_α is strictly increasing.*
2. *If $\alpha < \beta < \epsilon_0$ then f_α is eventually dominated by f_β.*
3. *Each f_α is computable, and provably total in* PA.
4. *Any computable f provably total in* PA *is eventually dominated by some f_α, $\alpha < \epsilon_0$.*

Peano Arithmetic PA is the formal theory that captures the intuitive notion of "finite mathematics." If a result is not provable in PA, it must use infinite objects in an essential way.

Theorem (Wainer)

1. *Each* f_α *is strictly increasing.*
2. *If* $\alpha < \beta < \epsilon_0$ *then* f_α *is eventually dominated by* f_β.
3. *Each* f_α *is computable, and provably total in* PA.
4. *Any computable* f *provably total in* PA *is eventually dominated by some* f_α, $\alpha < \epsilon_0$.

Definition

Let $R_n^\omega(m)$ be the change of base function replacing each n by ω in the complete base n representation of m.

For example, since $266 = 3^{3+2} + 3^2 \cdot 2 + 3 + 2$, then

$$R_3^\omega(266) = \omega^{\omega+2} + \omega^2 2 + \omega + 2.$$

Definition

Let $R_n^\omega(m)$ be the change of base function replacing each n by ω in the complete base n representation of m.

For example, since $266 = 3^{3+2} + 3^2 \cdot 2 + 3 + 2$, then

$$R_3^\omega(266) = \omega^{\omega+2} + \omega^2 2 + \omega + 2.$$

Theorem

Let

$$n = 2^{m_1} + 2^{m_2} + \cdots + 2^{m_k}$$

where $m_1 > m_2 > \cdots > m_k$. *Let* $\alpha_i = R_2^{\omega}(m_i)$. *Then*

$$\mathcal{G}(n) = f_{\alpha_1}(f_{\alpha_2}(\ldots(f_{\alpha_k}(3))\ldots)) - 2.$$

For example, $\mathcal{G}(266) = f_{\omega^{\omega+1}}(f_{\omega+1}(6)) - 2$, because $266 = 2^{2^{2+1}} + 2^{2+1} + 2^1$ and $f_1(3) = 6$.
Similarly,

$$\mathcal{G}(4) = f_{\omega}(3) - 2 = f_3(3) - 2 = 3 \cdot 2^3 \cdot 2^{3 \cdot 2^3} \cdot 2^{3 \cdot 2^3 \cdot 2^{3 \cdot 2^3}} - 2.$$

Theorem

Let

$$n = 2^{m_1} + 2^{m_2} + \cdots + 2^{m_k}$$

where $m_1 > m_2 > \cdots > m_k$. *Let* $\alpha_i = R_2^{\omega}(m_i)$. *Then*

$$\mathcal{G}(n) = f_{\alpha_1}(f_{\alpha_2}(\ldots(f_{\alpha_k}(3))\ldots)) - 2.$$

For example, $\mathcal{G}(266) = f_{\omega^{\omega+1}}(f_{\omega+1}(6)) - 2$, because $266 = 2^{2^{2+1}} + 2^{2+1} + 2^1$ and $f_1(3) = 6$.

Similarly,

$$\mathcal{G}(4) = f_\omega(3) - 2 = f_3(3) - 2 = 3 \cdot 2^3 \cdot 2^{3 \cdot 2^3} \cdot 2^{3 \cdot 2^3 \cdot 2^{3 \cdot 2^3}} - 2.$$

Theorem

Let

$$n = 2^{m_1} + 2^{m_2} + \cdots + 2^{m_k}$$

where $m_1 > m_2 > \cdots > m_k$. *Let* $\alpha_i = R_2^\omega(m_i)$. *Then*

$$\mathcal{G}(n) = f_{\alpha_1}(f_{\alpha_2}(\dots(f_{\alpha_k}(3))\dots)) - 2.$$

For example, $\mathcal{G}(266) = f_{\omega^{\omega+1}}(f_{\omega+1}(6)) - 2$, because $266 = 2^{2^{2+1}} + 2^{2+1} + 2^1$ and $f_1(3) = 6$.
Similarly,

$$\mathcal{G}(4) = f_\omega(3) - 2 = f_3(3) - 2 = 3 \cdot 2^3 \cdot 2^{3 \cdot 2^3} \cdot 2^{3 \cdot 2^3 \cdot 2^{3 \cdot 2^3}} - 2.$$

We obtain at once:

Corollary

1. *(Goodstein) $\mathcal{G}$ is total.*
2. *(Kirby-Paris) Item (1) is not a theorem of* PA.

THE END

We obtain at once:

Corollary

1. *(Goodstein)* $\mathcal{G}$ *is total.*
2. *(Kirby-Paris) Item (1) is not a theorem of* PA*.*

THE END

We obtain at once:

Corollary

1. *(Goodstein)* $\mathcal{G}$ *is total.*
2. *(Kirby-Paris) Item (1) is not a theorem of* PA.

THE END

We obtain at once:

Corollary

1. *(Goodstein)* $\mathcal{G}$ *is total.*
2. *(Kirby-Paris) Item (1) is not a theorem of* PA.

THE END