# On guessing whether a sequence has a certain property

Samuel Alexander
The Ohio State University
Department of Mathematics
231 West 18th Avenue,
Columbus OH 43210
USA
alexander@math.ohio-state.edu

**Abstract**

A concept of "guessability" is defined for sets of sequences of naturals. Eventually, these sets are thoroughly characterized. To do this, a nonstandard logic is developed, a logic containing symbols for the ellipsis as well as for functions without fixed arity.

## 1 Motivation

Suppose Alice and Bob are playing a game. Alice is reading a fixed sequence, one entry at a time. Bob is trying to guess whether 0 is in the sequence. He can revise his guess with each new revealed entry, and he wins if his guesses converge to the correct answer. He has an obvious strategy: always guess no, until 0 appears (if ever), then guess yes forever. The set of sequences containing 0 is guessable.

Suppose, instead, Bob is trying to guess whether Alice's sequence contains *infinitely many* zeroes. We will see there is no strategy, not even if Bob has unlimited computation power. The set of sequences with infinitely many zeroes is unguessable.

A sequence $f : \mathbb{N} \to \mathbb{N}$ is *onto* if $\forall m \, \exists n \, f(n) = m$. This definition uses nested quantifiers: quantifiers appear in the scope of other quantifiers. Is it possible to give an alternate definition without nested quantifiers? The answer is "no", but how to prove it? We will give a proof of a very strong negative answer, strong in the sense that nested quantifiers cannot be eliminated even in an extremely powerful language. Of course, the technique generalizes to a wide class of sets of sequences, not just the onto sequences.

# 2 Basics

Let $\mathbb{N}^{\mathbb{N}}$ be the set of sequences $f : \mathbb{N} \to \mathbb{N}$, and let $\mathbb{N}^{<\mathbb{N}}$ be the set of finite sequences.

**Definition 1.** A function $G : \mathbb{N}^{<\mathbb{N}} \to \{0, 1\}$ *guesses* (and is a *guesser* for) a set $S \subseteq \mathbb{N}^{\mathbb{N}}$ if for every $f : \mathbb{N} \to \mathbb{N}$, there exists some $m > 0$ such that for all $n > m$,

$$G(f(0), \ldots, f(n)) = \begin{cases} 1, & \text{if } f \in S; \\ 0, & \text{if } f \notin S. \end{cases}$$

A set $S \subseteq \mathbb{N}^{\mathbb{N}}$ is *guessable* if it has a guesser.

The next test is very useful for showing nonguessability, though its converse is not true.

**Theorem 1.** Let $S \subseteq \mathbb{N}^{\mathbb{N}}$. Suppose that, for every finite sequence $g \in \mathbb{N}^{<\mathbb{N}}$, there are sequences $g_1, g_2 \in \mathbb{N}^{\mathbb{N}}$ extending $g$ with $g_1 \in S$ and $g_2 \notin S$. Then $S$ is nonguessable.

*Proof.* Suppose $S$ has a guesser $G$. I will define a sequence $f : \mathbb{N} \to \mathbb{N}$ such that $G(f(0), \ldots, f(n))$ fails to converge, which violates the definition of guesser.

Clearly $S \neq \emptyset$, so let $s_1 : \mathbb{N} \to \mathbb{N}$ be some sequence in $S$. By definition of guesser, we can find some $x_1$ such that $G(s_1(0), \ldots, s_1(x_1)) = 1$. Let $f(0) = s_1(0), \ldots, f(x_1) = s_1(x_1)$.

Inductively, suppose $x_1 < \cdots < x_k$ and $f(0), \ldots, f(x_k)$ are defined such that $G(f(0), \ldots, f(x_i)) \equiv i \pmod 2$ for $i = 1, \ldots, k$. By the theorem's hypothesis, we can find some $s_{k+1} : \mathbb{N} \to \mathbb{N}$, extending the finite sequence $(f(0), \ldots, f(x_k))$, such that $s_{k+1}$ is in $S$ iff $k+1 \equiv 1 \pmod 2$. By definition of guesser, find $x_{k+1} > x_k$ such that $G(s_{k+1}(0), \ldots, s_{k+1}(x_{k+1})) \equiv k+1 \pmod 2$. Let $f(x_k + 1) = s_{k+1}(x_k + 1), \ldots, f(x_{k+1}) = s_{k+1}(x_{k+1})$.

This defines sequences $f : \mathbb{N} \to \mathbb{N}$ and $x_1 < x_2 < \cdots$ with the property that $G(f(0), \ldots, f(x_i)) \equiv i \pmod 2$ for every $i > 0$. This contradicts that $G(f(0), \ldots, f(n))$ is supposed to converge. □

Using the above test, we can immediately confirm, for example, the set of sequences containing infinitely many zeroes is nonguessable, as is the set of onto sequences.

**Remark 1.** Theorem 1 is constructive up to certain choices. Starting with a set $S$ satisfying the hypotheses of Theorem 1 and naively trying to guess it, and being systematic in the choices from the proof, can lead to the creation of a concrete sequence which thwarts the naive guessing attempt. In an informal sense, it should be especially difficult for someone not in the know to guess whether the resulting sequence lies in $S$. And the more sophisticated the futile guessing attempt, the more difficult the resulting sequence becomes. For some explicit examples, see sequences A082691, A182659, and A182660 in Sloane's OEIS [5].

Tsaban and Zdomskyy also briefly mention a somewhat similar notion of guessable sets in their paper [6].

# 3 A Logic for Ellipses

Because guessers are functions which do not have "arity" in the usual sense, instead being defined on the whole space $\mathbb{N}^{<\mathbb{N}}$ of finite sequences, and since we care so much about expressions like $G(f(0), \ldots, f(n))$, we will extend logic to mesh better with these sorts of expressions. I assume familiarity with basic first-order logic, which Enderton [2] has written about extensively, as has Bilaniuk [1].

**Definition 2.** A *language with ellipses* is a standard language of first-order logic, with a constant symbol $\mathbf{0}$, together with a set of *function symbols of arity* $\mathbb{N}^{<\mathbb{N}}$ and a special logical symbol $\cdots_x$ for every variable $x$.

To avoid confusion, we will use $\cdots_x$ for the syntactical symbol and $\ldots$ for meta-ellipses. For example, $G(s(\mathbf{0}), \ldots, s(\mathbf{2}))$ is a meta-abbreviation for

$$G(s(\mathbf{0}), s(\mathbf{1}), s(\mathbf{2})),$$

different than $G(s(\mathbf{0}), \cdots_x, s(\mathbf{2}))$ which has no counterpart in classical logic.

**Definition 3.** If $\mathscr{L}$ is a language with ellipses, then the *terms* of $\mathscr{L}$ (and their free variables) are defined inductively:

1. For any variable $x$, $x$ is a term and $FV(x) = \{x\}$.
2. For any constant symbol $c$, $c$ is a term and $FV(c) = \emptyset$.
3. If $f$ is a function symbol of arity $n$ or arity $\mathbb{N}^{<\mathbb{N}}$, and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term with free vars $FV(t_1) \cup \cdots \cup FV(t_n)$.
4. If $G$ is an $\mathbb{N}^{<\mathbb{N}}$-ary function symbol, and $u, v$ are terms, and $x$ is a variable, then $G(u(\mathbf{0}), \cdots_x, u(v))$ is a term with free variables

$$(FV(u)\backslash\{x\}) \cup FV(v).$$

The *well-formed formulas* of $\mathscr{L}$ are defined as usual from these terms. Term substitution is defined by the usual induction with two new cases:

- If $y \neq x$ then

$$G(u(\mathbf{0}), \cdots_x, u(v))(y|t) = G(u(y|t)(\mathbf{0}), \cdots_x, u(y|t)(v(y|t))).$$

- $G(u(\mathbf{0}), \cdots_x, u(v))(x|t) = G(u(\mathbf{0}), \cdots_x, u(v(x|t)))$.

A *model* for a language with ellipses $\mathscr{L}$ is a model $\mathscr{M}$ for the classical part of $\mathscr{L}$, together with a function $G^{\mathscr{M}} : \mathscr{M}^{<\mathbb{N}} \to \mathscr{M}$ for each $\mathbb{N}^{<\mathbb{N}}$-ary function symbol $G$ in $\mathscr{L}$. However, defining how an arbitrary model evaluates terms is difficult. We will only be interested in one very specific family of models, where there is no trouble evaluating terms.

**Definition 4.** The following models lie at the heart of all later results.

- $\mathscr{L}_{\max}$ is the language with ellipses which contains a constant symbol $\mathbf{n}$ for every $n \in \mathbb{N}$, an $n$-ary function symbol $\tilde{w}$ for every function $w : \mathbb{N}^n \to \mathbb{N}$ $(n > 0)$, an $n$-ary predicate symbol $\tilde{p}$ for every subset $p \subseteq \mathbb{N}^n$ $(n > 0)$, an $\mathbb{N}^{<\mathbb{N}}$-ary function symbol $\tilde{G}$ for every function $G : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N}$, and one additional unary function symbol $\mathbf{f}$.

- For every function $f : \mathbb{N} \to \mathbb{N}$, $\mathscr{M}_f$ is the model for the language $\mathscr{L}_{\text{max}}$ with universe $\mathbb{N}$, which interprets $\mathbf{n}$ as $n$ for every $n$, $\tilde{w}$ as $w$ for every $w : \mathbb{N}^n \to \mathbb{N}$, $\tilde{p}$ as $p$ for every $p \subseteq \mathbb{N}^n$, and $\tilde{G}$ as $G$ for every $G : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N}$, and which interprets $\mathbf{f}$ as $f$.

If $n \in \mathbb{N}$ then $\bar{n}$ denotes the numeral $\mathbf{n}$ of $n$.

**Definition 5.** Let $f : \mathbb{N} \to \mathbb{N}$. The semantics of $\mathscr{M}_f$ are defined as follows. Let $s$ be any assignment from the variables to $\mathbb{N}$.

- $(\mathscr{M}_f, s)$ interprets terms $t$ into naturals $t^{\mathscr{M}_f,s}$, or $t^s$ if there is no ambiguity, according to the usual inductive definition, with one new case:
  - If $u, v$ are terms and $x$ is a variable and $G$ is an $\mathbb{N}^{<\mathbb{N}}$-ary function symbol, then

$$G(u(\mathbf{0}), \cdots_x, u(v))^s = G^{\mathscr{M}_f}\left(u(x|\mathbf{0})^s, \ldots, u\left(x\left|\overline{v^s}\right.\right)^s\right).$$

- For example, the interpretation of $\tilde{G}(\mathbf{f}(x)(\mathbf{0}), \cdots_x, \mathbf{f}(x)(\mathbf{99}))$ is

$$G(f(0), \ldots, f(99)),$$

  while the interpretation of $\tilde{G}(\mathbf{f}(x)(\mathbf{0}), \cdots_x, \mathbf{f}(x)(\mathbf{f}(y)))$ is

$$G(f(0), \ldots, f(f(y^s))).$$

- From here, the remaining semantics of $\mathscr{M}_f$ are defined as usual.

In classical logic, every term with no free variables has the property that its interpretation in any model depends only on finitely many values of the interpretations of the function symbols in that model. For example, the interpretation of $5 + (2 \cdot 3)$ depends only on one value of $\cdot$ and one value of $+$. Similar properties are true of our $\mathscr{M}_f$ models.

**Lemma 2.** Suppose $u$ is a term with no free variables, and $c$ is a constant symbol. For any $f : \mathbb{N} \to \mathbb{N}$, $\mathscr{M}_f \models u = c$ iff there is some $k$ such that whenever $g : \mathbb{N} \to \mathbb{N}$ extends $(f(0), \ldots, f(k))$, $\mathscr{M}_g \models u = c$ and to check whether $\mathscr{M}_g \models u = c$ using the inductive definition of semantics for $\mathscr{M}_g$, it is not necessary to query $g(i)$ for any $i > k$.

*Proof.* $(\Rightarrow)$ Induction on complexity of $u$.

- Since $u$ has no free variables, $u$ cannot be a variable. If $u$ is a constant symbol, the lemma is trivial.

- Suppose that $u$ is $h(u_1, \ldots, u_n)$ for some $n$-ary (or $\mathbb{N}^{<\mathbb{N}}$-ary) function symbol $h$ other than $\mathbf{f}$, and some terms $u_1, \ldots, u_n$ with no free variables. If $\mathscr{M}_f \models h(u_1, \ldots, u_n) = c$, then there are $a_1, \ldots, a_n \in \mathbb{N}$ such that $h^{\mathscr{M}_f}(a_1, \ldots, a_n) = c^{\mathscr{M}_f}$ and $\mathscr{M}_f \models u_i = \bar{a}_i$ for $i = 1, \ldots, n$. Since $\bar{a}_i$ is a constant symbol, by induction find $k_1, \ldots, k_n$ such that for any $i = 1, \ldots, n$ and any $g : \mathbb{N} \to \mathbb{N}$ with $g(0) = f(0), \ldots, g(k_i) = f(k_i)$, $\mathscr{M}_g \models u_i = \bar{a}_i$, and checking this by definition of semantics does not require querying $g(j)$ for any $j > a_i$. Then $k = \max\{k_1, \ldots, k_n\}$ works (using the fact $h^{\mathscr{M}_g}$ does not depend on $g$ since $h$ is not $\mathbf{f}$).

- Next, suppose $u$ is $\mathbf{f}(v)$ where $v$ is a term with no free variables. If $\mathscr{M}_f \models \mathbf{f}(v) = c$ then there iss $a \in \mathbb{N}$ such that $f(a) = c^{\mathscr{M}_f}$ and $\mathscr{M}_f \models v = \bar{a}$. Since $\bar{a}$ is a constant symbol, by induction find $k_0$ such that whenever $g(0) = f(0), \ldots, g(k_0) = f(k_0)$, then $\mathscr{M}_g \models v = \bar{a}$, and checking $\mathscr{M}_g \models v = \bar{a}$ does not require querying $g(i)$ for any $i > k_0$. Let $k = \max\{k_0, a\}$. Suppose $g(0) = f(0), \ldots, g(k) = f(k)$. Then $\mathbf{f}(v)^{\mathscr{M}_g} = \mathbf{f}^{\mathscr{M}_g}(v^{\mathscr{M}_g}) = g(a) = f(a) = c^{\mathscr{M}_f}$. So $\mathscr{M}_g \models \mathbf{f}(v) = c$, and to check so, we only had to query $g(a)$ in addition to any queries we had to make to check $\mathscr{M}_g \models v = \bar{a}$, so we did not have to query $g(i)$ for any $i > k$.

- Finally, suppose $u$ is $G(v(\mathbf{0}), \cdots_x, v(w))$ where $v, w$ are terms, $x$ is a variable, $FV(w) = \emptyset$, $FV(v) \subseteq \{x\}$, and $G$ is an $\mathbb{N}^{<\mathbb{N}}$-ary function symbol. If $\mathscr{M}_f \models G(v(\mathbf{0}), \cdots_x, v(w)) = c$ then

$$G^{\mathscr{M}_f}\left(v(x|\mathbf{0})^{\mathscr{M}_f}, \ldots, v\left(x\left|\overline{w^{\mathscr{M}_f}}\right.\right)^{\mathscr{M}_f}\right) = c^{\mathscr{M}_f}.$$

  Since $\mathscr{M}_f \models w = \overline{w^{\mathscr{M}_f}}$, find some number $k_{-1}$ such that whenever $g$ extends $(f(0), ..., f(k_{-1}))$, $\mathscr{M}_g \models w = \overline{w^{\mathscr{M}_f}}$ and checking so does not require queries beyond $g(k_{-1})$. Since $\mathscr{M}_f \models v(x|\mathbf{i}) = \overline{v(x|\mathbf{i})}$ for $i = 0, \ldots, w^{\mathscr{M}_f}$, find $k_0, \ldots, k_{w^{\mathscr{M}_f}}$ such that for each $i = 0, \ldots, w^{\mathscr{M}_f}$, if $g(0) = f(0), \ldots, g(k_i) = f(k_i)$ then $\mathscr{M}_g \models v(x|\bar{i}) = \overline{v(x|\bar{i})}$ can be confirmed without querying $g$ beyond $g(k_i)$.

  Let $k = \max\{k_{-1}, k_0, \ldots, k_{w^{\mathscr{M}_f}}\}$. Suppose $g(0) = f(0), \ldots, g(k) = f(k)$. Then $\mathscr{M}_g \models w = \overline{w^{\mathscr{M}_f}}$, so $w^{\mathscr{M}_g} = w^{\mathscr{M}_f}$. Similarly $v(x|\mathbf{i})^{\mathscr{M}_g} =$

$v(x|\mathbf{i})^{\mathscr{M}_f}$ for $i = 0, \ldots, w^{\mathscr{M}_f}$. And $G^{\mathscr{M}_g} = G^{\mathscr{M}_f}$. It follows that

$$\mathscr{M}_g \models G(v(\mathbf{0}), \cdots_x, v(w)) = c,$$

and checking so does not require any queries to $g(j)$ for any $j > k$.

($\Leftarrow$) Suppose there is some $k$ so that whenever $g$ extends $(f(0), \ldots, f(k))$ then $\mathscr{M}_g \models u = c$. In particular, $f$ itself extends $(f(0), \ldots, f(k))$, so $\mathscr{M}_f \models u = c$. □

**Corollary 3.** Let $\phi$ be a quantifier-free sentence. For any $f : \mathbb{N} \to \mathbb{N}$, $\mathscr{M}_f \models \phi$ iff there is some $k$ such that for every $g : \mathbb{N} \to \mathbb{N}$ extending $(f(0), \ldots, f(k))$, $\mathscr{M}_g \models \phi$, and in checking $\mathscr{M}_g \models \phi$ by the inductive definition of semantics, we never need to query $g(i)$ for any $i > k$.

*Proof.* By induction on the complexity of $\phi$.

- Suppose $\phi$ is $u = v$ for terms $u, v$ with no free variables. Assume $\mathscr{M}_f \models u = v$. Then $\mathscr{M}_f \models u = \overline{u^{\mathscr{M}_f}}$ and $\mathscr{M}_f \models v = \overline{u^{\mathscr{M}_f}}$. By Lemma 2, find $k$ big enough that whenever $g(0) = f(0), \ldots, g(k) = f(k)$, then $\mathscr{M}_g \models u = \overline{u^{\mathscr{M}_f}}$ and $\mathscr{M}_g \models v = \overline{u^{\mathscr{M}_f}}$ and both facts can be confirmed without querying $g$ beyond $g(k)$. For any such $g$, $\mathscr{M}_g \models u = v$, verifiable with no additional $g$-queries. The converse is trivial.

- Next, suppose $\phi$ is $\tilde{p}(u_1, \ldots, u_n)$ for an $n$-ary predicate symbol $\tilde{p}$ and terms $u_1, \ldots, u_n$ with no free variables. Then $\phi$ is equivalent (in every $\mathscr{M}$.) to $\tilde{g}(u_1, \ldots, u_n) = \mathbf{1}$ where $g$ is the characteristic function of $p$, so we are done by the previous case.

- Suppose $\phi$ is $\phi_1 \wedge \phi_2$. Assume $\mathscr{M}_f \models \phi$. Inductively, find $k_1$ and $k_2$ such that if $g$ extends $(f(0), \ldots, f(k_i))$ then $\mathscr{M}_g \models \phi_i$ is verifiable with no $g$-queries beyond $g(k_i)$. Then any $g$ extending $(f(0), \ldots, f(\max\{k_1, k_2\}))$ has $\mathscr{M}_g \models \phi$, verifiable without querying beyond $g(\max\{k_1, k_2\})$. The converse is trivial.

- The cases of other propositional connectives are similar.

□

If $s$ is an assignment from the variables of a language onto the universe of the language, and if $x$ is a variable, and $n$ is a number, then $s(x|n)$ denotes the assignment which is identical to $s$ except that it maps $x$ to $n$. Similarly if a model is understood by context and $c$ is a constant symbol then $s(x|c)$ denotes the assignment identical to $s$ except that it maps $x$ to the interpretation of $c$ in the model.

**Lemma 4.** *(The Weak Substitution Lemma)* For a formula $\phi$, an assignment $s$, and a constant symbol $c$, and for any $f : \mathbb{N} \to \mathbb{N}$, $\mathscr{M}_f \models \phi[s(x|\overline{c^s})]$ iff $\mathscr{M}_f \models \phi(x|c)[s]$.

*Proof.* By the inductive argument used to prove the full Substitution Lemma in classical logic, most of which we omit. But there are tricky new cases for our new terms.

**Claim:** For any terms $u, v$, constant symbol $c$, variables $x \neq y$, and assignment $s$,

$$G(u(\mathbf{0}), \cdots_x, u(v))(y|c)^s = G(u(\mathbf{0}), \cdots_x, u(v))^{s(y|c)}.$$

The details are (using the induction hypothesis repeatedly) as follows:

$$\begin{aligned} G(u(\mathbf{0}), \cdots_x, u(v))(y|c)^s &= G\left(u(y|c)(\mathbf{0}), \cdots_x, u(y|c)(v(y|c))\right)^s \\ &= G^{\mathscr{M}_f}\left(u(y|c)(x|\mathbf{0})^s, \ldots, u(y|c)\left(x \,\middle|\, \overline{v(y|c)^s}\right)^s\right) \\ &= G^{\mathscr{M}_f}\left(u(x|\mathbf{0})^{s(y|c)}, \ldots, u\left(x \,\middle|\, \overline{v^{s(y|c)}}\right)^{s(y|c)}\right) \\ &= G(u(\mathbf{0}), \cdots_x, u(v))^{s(y|c)}. \end{aligned}$$

**Claim:** For any terms $u, v$, constant symbol $c$, and variable $x$ and assignment $s$,

$$G(u(\mathbf{0}), \cdots_x, u(v))(x|c)^s = G(u(\mathbf{0}), \cdots_x, u(v))^{s(x|c)}.$$

Using the induction hypothesis repeatedly:

$$\begin{aligned} G(u(\mathbf{0}), \cdots_x, u(v))(x|c)^s &= G(u(\mathbf{0}), \cdots_x, u(v(x|c)))^s \\ &= G^{\mathscr{M}_f}\left(u(x|\mathbf{0})^s, \ldots, u\left(x \,\middle|\, \overline{v(x|c)^s}\right)^s\right) \\ &= G^{\mathscr{M}_f}\left(u(x|\mathbf{0})^s, \ldots, u\left(x \,\middle|\, \overline{v^{s(x|c)}}\right)^s\right) \\ &= G^{\mathscr{M}_f}\left(u(x|\mathbf{0})^{s(x|c)}, \ldots, u\left(x \,\middle|\, \overline{v^{s(x|c)}}\right)^{s(x|c)}\right) \\ &= G(u(\mathbf{0}), \cdots_x, u(v))^{s(x|c)}. \end{aligned}$$

The next to last equation is justified because the terms whose "exponents" are changed do not depend on $x$. $\square$

A full Substitution Lemma is also true, but it requires a nonclassical definition of *substitutable*, which would take us too far afield.

# 4 Guessability and Quantifiers

**Definition 6.** Let $S \subseteq \mathbb{N}^{\mathbb{N}}$ be a set of sequences. Let $\phi$ be a sentence in $\mathscr{L}_{\max}$. We say that $\phi$ *defines* $S$ if, for every $f : \mathbb{N} \to \mathbb{N}$, $\mathscr{M}_f \models \phi$ iff $f \in S$.

**Theorem 5.** A set $S \subseteq \mathbb{N}^{\mathbb{N}}$ is guessable if and only if it is defined by some sentence $\forall x\, \exists y\, \phi$ and also by some sentence $\exists x\, \forall y\, \psi$, where $\phi$ and $\psi$ are quantifier-free.

We divide the proof of the theorem above into a sequence of lemmata.

**Lemma 6.** Suppose $S \subseteq \mathbb{N}^{\mathbb{N}}$ is guessable. Then $S$ is defined by some sentence $\exists x\, \forall y\, \phi$ and also by some sentence $\forall x\, \exists y\, \psi$, where $\phi$ and $\psi$ are quantifier-free.

*Proof.* Let $G$ be a guesser for $S$. For any $f : \mathbb{N} \to \mathbb{N}$, $f \in S$ if and only if $G(f(0), \ldots, f(n)) = 1$ for all $n$ sufficiently large. Therefore $S$ is defined by

$$\exists x\, \forall y\, ((y > x) \to \tilde{G}(\mathbf{f}(z)(\mathbf{0}), \cdots_z, \mathbf{f}(z)(y)) = \mathbf{1}),$$

where "$y > x$" is shorthand for $\tilde{>}(y, x)$. Similarly, $S$ is also defined by

$$\forall x\, \exists y\, ((y > x) \wedge \tilde{G}(\mathbf{f}(z)(\mathbf{0}), \cdots_z, \mathbf{f}(z)(y)) = \mathbf{1}).$$

$\square$

We will prove the converse of Lemma 6 shortly. To that end, a piece of technical machinery is needed.

**Definition 7.** A set $S \subseteq \mathbb{N}^{\mathbb{N}}$ is *overguessable* if there is a function $\mu : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N} \cup \{\infty\}$ such that:

1. For every $f \in S$, $\mu(f(0), \ldots, f(n))$ is eventually bounded by a finite number.

2. For every $f \notin S$, $\mu(f(0), \ldots, f(n)) \to \infty$ as $n \to \infty$.

**Lemma 7.** Suppose $S \subseteq \mathbb{N}^{\mathbb{N}}$ is defined by the sentence $\exists x\, \forall y\, \phi$ where $\phi$ is quantifier-free. Then $S$ is overguessable.

*Proof.* Given a tuple $(n_0, \ldots, n_k)$, define $\mu(n_0, \ldots, n_k)$ as follows. Let $h : \mathbb{N} \to \mathbb{N}$ be defined by $h(i) = n_i$ if $i \leq k$, $h(i) = 0$ otherwise. Given a pair $(a, b) \in \mathbb{N}^2$, consider the sentence $\phi(x, y|\bar{a}, \bar{b})$. Attempt to check whether $\mathscr{M}_h \models \phi(x, y|\bar{a}, \bar{b})$, using the inductive definition of the semantics of $\mathscr{M}_h$. If, in so doing, you must query $h(i)$ for some $i > k$, say that *the attempt failed*. Otherwise, say *the attempt succeeded*. If the attempt failed, *or* if $\mathscr{M}_h \models \phi(x, y|\bar{a}, \bar{b})$, then say that $(a, b)$ is *nice*.

Call a number $a$ *very nice* if $(a, b)$ is nice for every $b$. If there is any very nice number, then let $\mu(n_0, \ldots, n_k)$ be the smallest very nice number. Otherwise let $\mu(n_0, \ldots, n_k) = \infty$.

I claim the above $\mu$ witnesses that $S$ is overguessable.

First, suppose $f \in S$. Since $S$ is defined by $\exists x\, \forall y\, \phi$, $\mathscr{M}_f \models \exists x\, \forall y\, \phi$. By the Weak Substitution Lemma, for some $a$, $\mathscr{M}_f \models \phi(x, y|\bar{a}, \bar{b})$ for every $b$. When we attempt to check whether $\mathscr{M}_h \models \phi(x, y|\bar{a}, \bar{b})$ in the definition of $\mu(f(0), \ldots, f(k))$, if the attempt succeeds, then $\mathscr{M}_h \models \phi(x, y|\bar{a}, \bar{b})$ because $\mathscr{M}_f \models \phi(x, y|\bar{a}, \bar{b})$ and we never had to look at the part of $h$ which disagrees with $f$. So $(a, b)$ is nice for every $b$, so $a$ is very nice, so $\mu(f(0), \ldots, f(k))$ is bounded by $a$.

Next, suppose $f \notin S$. Let $a \in \mathbb{N}$, I claim $\mu(f(0), ..., f(n)) \neq a$ for all $n$ sufficiently large. Since $f \notin S$, $\mathscr{M}_f \not\models \exists x\, \forall y\, \phi$. By the Weak Substitution Lemma, there is some $b$ such that $\mathscr{M}_f \not\models \phi(x, y|\bar{a}, \bar{b})$. Since $\phi$ is quantifier-free, we invoke Corollary 3 on $\neg\phi(x, y|\bar{a}, \bar{b})$ and find $k$ such that $\mathscr{M}_g \not\models \phi(x, y|\bar{a}, \bar{b})$ whenever $g$ extends $(f(0), \ldots, f(k))$, and, to check whether $\mathscr{M}_g \models \phi(x, y|\bar{a}, \bar{b})$, we do not need to query $g(i)$ for $i > k$. Then, in the definition of $\mu(f(0), \ldots, f(k))$, for pair $(a, b)$, the attempt succeeds and $\mathscr{M}_h \not\models \phi(x, y|\bar{a}, \bar{b})$, so $(a, b)$ is not nice, so $a$ is not very nice, so $\mu(f(0), \ldots, f(k)) \neq a$, in fact, $\mu(f(0), \ldots, f(j)) \neq a$ for all $j \geq k$. By arbitrariness of $a$, $\mu(f(0), ..., f(n)) \to \infty$. □

**Lemma 8.** Suppose a set $S \subseteq \mathbb{N}^{\mathbb{N}}$ is defined by some sentence $\forall x\, \exists y\, \phi$ and also by some sentence $\exists x\, \forall y\, \psi$, where $\phi$ and $\psi$ are quantifier-free. Then $S$ is guessable.

*Proof.* By Lemma 7, find $\mu : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N} \cup \{\infty\}$ which overguesses $S$. And since $S^c$ is defined by $\exists x\, \forall y\, \neg\phi$, use Lemma 7 again to find $\nu : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N} \cup \{\infty\}$ which overguesses $S^c$.

Define $G : \mathbb{N}^{<\mathbb{N}} \to \{0, 1\}$ by saying $G(n_0, \dots, n_k) = 1$ if $\mu(n_0, \dots, n_k) \leq \nu(n_0, \dots, n_k)$ and 0 otherwise. If $f \in S$ then $\mu(f(0), \dots, f(k))$ is eventually bounded by a finite number and $\nu(f(0), \dots, f(k)) \to \infty$, so $G(f(0), \dots, f(k))$ converges to 1. The other case is similar. □

Combining Lemmata 6 and 8 proves Theorem 5.

**Proposition 9.** If $S \subset \mathbb{N}^{\mathbb{N}}$ is overguessable, then it is defined by some sentence $\exists x \forall y \phi$ with $\phi$ quantifier-free.

*Proof.* Suppose $S$ is overguessed by $\mu : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N} \cup \{\infty\}$. Define $\mu' : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ by saying $\mu'(n) = \mu(n) + 1$ if $\mu(n) \neq \infty$, $\mu'(n) = 0$ if $\mu(n) = \infty$. If $f : \mathbb{N} \to \mathbb{N}$, then $f \in S$ if and only if the sequence $\mu(f(0), \dots, f(n))$ is eventually bounded by some finite number. This is true if and only if $\mu'(f(0), \dots, f(n))$ is eventually bounded by some finite number and eventually nonzero. This latter equivalence can be expressed by

$$\exists m_1 \, \exists m_2 \, \forall m_3 \, ((m_3 > m_2) \to (0 < \mu'(f(0), \dots, f(m_3)) < m_1)).$$

Let $d : \mathbb{N} \to \mathbb{N}^2$ be any onto map from $\mathbb{N}$ to $\mathbb{N}^2$. Write $d(n) = (d_1(n), d_2(n))$, thus defining two functions $d_1, d_2 : \mathbb{N} \to \mathbb{N}$. Then the above formula is equivalent to

$$\exists m \, \forall m_3 \, ((m_3 > d_2(m)) \to (0 < \mu'(f(0), \dots, f(m_3)) < d_1(m))).$$

This can be formalized in $\mathscr{L}_{\max}$, providing a sentence $\exists x \, \forall y \, \phi$ which defines $S$, with $\phi$ quantifier-free. □

**Example 10.** Every countable subset of $\mathbb{N}^{\mathbb{N}}$ is overguessable.

*Proof.* Let $S \subseteq \mathbb{N}^{\mathbb{N}}$ be countable. Define $g : \mathbb{N}^2 \to \mathbb{N}$ by saying $g(m, n) = h_m(n)$ where $h_m$ is the $m$th element of $S$. Then $S$ is defined by

$$\exists x \, \forall y \, \tilde{g}(x, y) = \mathbf{f}(y).$$

By Lemma 7, $S$ is overguessable. □

**Remark 2.** Guessable and overguessable sets of sequences are analogous to computable and computably enumerable sets of naturals, respectively. One shows that $\Delta_1$ sets (in a much weaker logical setting than $\mathscr{L}_{\max}$) of naturals are computable by showing that they and their complements are c.e. by using the characterization of c.e. sets as sets which are $\Sigma_1$-definable (in the weaker setting). By comparison, I have shown that $\Delta_2$ sets of sequences (in a very strong logical setting) are guessable by showing that they and their complements are overguessable by using the characterization of overguessable sets as $\Sigma_2$-definable (in the stronger setting). These analogous phenomena in computability theory have been written about by Rogers [3], Enderton [2], Bilaniuk [1], and many other authors.

We will elaborate more on Remark 2 in Section 5.

**Lemma 11.** Suppose $S \subseteq \mathbb{N}^{\mathbb{N}}$ is definable by a sentence $\forall x\, \exists y\, \phi$ where $\phi$ is quantifier-free. If $S$ is countable then $S$ is guessable.

*Proof.* Suppose $S$ is countable. In the proof of Example 10, we showed $S$ is definable by a sentence $\exists x\, \forall y\, \psi$ where $\psi$ is quantifier-free. By Theorem 5, $S$ is guessable. □

**Example 12.** There are uncountably many permutations of $\mathbb{N}$.

*Proof.* A function $f : \mathbb{N} \to \mathbb{N}$ is a permutation iff

$$\forall m_1\, \forall m_2\, \exists n\, ((f(m_1) = f(m_2) \to m_1 = m_2) \wedge f(n) = m_2).$$

By appropriately coding $\langle m_1, m_2 \rangle$, the set $S$ of permutations is defined by a sentence $\forall x\, \exists y\, \phi$ where $\phi$ is quantifier-free.

Permutations are not guessable. If $G$ were a permutation-guesser, it would diverge on the following sequence. Let $f(0) = 0$, $f(1) = 1$, and so on until $G(f(0), \ldots, f(k_1)) = 1$ (this must happen since $G$ would converge to 1 if we kept going forever). Then skip a number, $f(k_1 + 1) = k_1 + 2$, $f(k_1 + 2) = k_1 + 3$, and keep going until $G(f(0), \ldots, f(k_2)) = 0$. Then fill in the gap, $f(k_2 + 1) = k_1 + 1$, and resume where we left off, $f(k_2 + 2) = k_2 + 2$, and so on until $G(f(0), \ldots, f(k_3)) = 1$. This process shows permutations are unguessable.

By Lemma 11, $S$ is uncountable. □

**Example 13.** *(Cantor)* There are uncountably many real numbers.

*Proof.* Consider the set $A$ of numbers in the interval $(0, 1)$ which have infinitely many 5s in their decimal expansions. There is an obvious bijection between $A$ and the set $S$ of sequences $f : \mathbb{N} \to \{0, 9\}$ such that $f(n) = 5$ infinitely often. This set $S$ is defined by

$$\forall x\, \exists y\, ((y > x) \wedge \mathbf{f}(y) = \mathbf{5} \wedge \mathbf{f}(x) \geq \mathbf{0} \wedge \mathbf{f}(x) \leq \mathbf{9}).$$

By Lemma 11, if $S$ is countable then it is guessable. But it is not: if $G$ were a guesser for $S$, then we could define a sequence on whose initial segments $G$ does not converge. Namely, let $f(0) = \ldots = f(x_k) = 0$ where $x_k$ is big enough that $G(f(0), ..., f(x_k)) = 0$, and then let $f(x_k + 1) = \cdots = f(x_{k+1}) = 5$, where $x_{k+1} > x_k$ is big enough that $G(f(0), ..., f(x_{k+1})) = 1$. And so on, alternating, forever. This shows $S$ is not guessable, so $S$ is not countable, so $A$ is uncountable, so $\mathbb{R}$ is uncountable. □

**Lemma 14.** If $S \subseteq \mathbb{N}^{\mathbb{N}}$ is definable by a sentence $\phi$ without nested quantifiers (that is, no quantifier appearing in the scope of another), then $S$ is guessable.

*Proof.* If so, then $\phi$ is a propositional combination of quantifier-free sentences and sentences of the form $\forall x\, \phi_0$ and $\exists x\, \phi_1$ where $\phi_0$, $\phi_1$ are quantifier-free. The sets defined by these component sentences are guessable by Theorem 5. Clearly guessable sets are closed under union, intersection, and complement, so $S$ itself is guessable. □

**Example 15.** The definition of onto functions cannot be simplified to get rid of nested quantifiers, not even with the full power of $\mathscr{L}_{\max}$.

*Proof.* By Lemma 14 and the fact the set of onto functions is not guessable, see Theorem 1. □

# 5 Descriptive Set Theory

In this section we will elaborate further on Remark 2. In descriptive set theory, $\mathbb{N}^{\mathbb{N}}$ is endowed with the topology whose basic open sets are those sets of the form

$$\{f \in \mathbb{N}^{\mathbb{N}} : f \text{ extends } f_0\}$$

where $f_0 \in \mathbb{N}^{<\mathbb{N}}$. Since $\mathbb{N}^{<\mathbb{N}}$ is countable, $\mathbb{N}^{\mathbb{N}}$ is second countable in the sense of basic topology. A set is called $G_\delta$ if it is a countable intersection of open sets, and $F_\sigma$ if it is a countable union of closed sets. A set is $\mathbf{\Delta}^0_2$ if it is both $G_\delta$ and $F_\sigma$ (equivalently, if it and its complement are both $G_\delta$). This $\mathbf{\Delta}^0_2$ is one of the levels of the *Borel hierarchy* which many authors, including Moschovakis [4], have written about.

**Theorem 16.** Let $S \subseteq \mathbb{N}^\mathbb{N}$. Then $S$ is guessable if and only if $S$ is $\mathbf{\Delta}^0_2$.

*Proof.* ($\Rightarrow$) Suppose $S$ is guessable. By Lemma 6, $S$ is defined by a sentence $\forall x\, \exists y\, \phi$ where $\phi$ is quantifier-free. For every $i, j \in \mathbb{N}$, let $S_i \subseteq \mathbb{N}^\mathbb{N}$ be the set defined by $\exists y\, \phi(x|i)$ and let $T_{ij} \subseteq \mathbb{N}^\mathbb{N}$ be the set defined by $\phi(x, y|i, j)$. It follows from Corollary 3 that each $T_{ij}$ is open. Each $S_i = \cup_j T_{ij}$, so each $S_i$ is open. Since $S = \cap_i S_i$, $S$ is $G_\delta$. Since $S^c$ is also guessable, identical reasoning shows $S^c$ is $G_\delta$, so $S$ is $\mathbf{\Delta}^0_2$.

($\Leftarrow$) Suppose $S$ is $\mathbf{\Delta}^0_2$. Write $S = \cap_{i\in\mathbb{N}} S_i$ where each $S_i$ is open. By second countability, write $S_i = \cup_{j\in\mathbb{N}} T_{ij}$ where each $T_{ij}$ is basic open. By the nature of basic open sets of $\mathbb{N}^\mathbb{N}$, there are $T^0_{ij} \in \mathbb{N}^{<\mathbb{N}}$ such that each $T_{ij}$ is exactly the set of infinite extensions of $T^0_{ij}$. Let $\tau : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N}$ be defined by saying $\tau(i, j, x_0, \ldots, x_k) = 1$ if $(x_0, \ldots, x_k) = T^0_{ij}$, $\tau$ is 0 everywhere else. Define $\ell : \mathbb{N}^2 \to \mathbb{N}$ by letting $\ell(i, j)$ be the length of $T^0_{ij}$. Then for any $f : \mathbb{N} \to \mathbb{N}$, $f$ extends $T^0_{ij}$ if and only if $\tau(i, j, f(0), \ldots, f(\ell(i, j))) = 1$. Thus, $f \in S$ if and only if

$$\forall i\, \exists j\, \tau(i, j, f(0), \ldots, f(\ell(i, j))) = 1.$$

This can be formalized in $\mathscr{L}_{\max}$. By dual reasoning applied to $S^c$, $S$ can also be defined by some $\exists i\, \forall j\, \phi$ where $\phi$ is quantifier-free. By Theorem 5, $S$ is guessable. □

# 6 Addendum

In January 2012, we learned that the notion of guessability was introduced some time ago in the PhD dissertation of William W. Wadge [7]. Instead of considering guesser functions, Wadge considered *guesser sets*, calling a subset $S \subseteq \mathbb{N}^\mathbb{N}$ guessable if there are disjoint sets $U, W \subseteq \mathbb{N}^{<\mathbb{N}}$ such that for every sequence $f$, $f \in S$ iff $f|k \in U$ for all but finitely many $k$, and $f \notin S$ iff $f|k \in W$ for all but finitely many $k$. This is clearly equivalent to our definition. Wadge gave a game-theoretical proof that guessability is equivalent to being $\mathbf{\Delta}^0_2$ (our Theorem 16) and then used this fact to show a special case, for $\mathbf{\Delta}^0_2$ sets, of what is now known as *Wadge's lemma*, an important result about Wadge degrees studied by descriptive set theorists.

# References

[1] S. Bilaniuk, *A Problem Course in Mathematical Logic*, self-published, 1991. http://euclid.trentu.ca/math/sb/pcml/pcml-16.pdf

[2] H. Enderton, *A Mathematical Introduction to Logic*, Academic Press, 2nd Edition, 2001.

[3] H. Rogers, *Theory of Recursive Functions and Effective Computability*, The MIT Press, 1987.

[4] Y. N. Moschovakis, *Descriptive Set Theory*, AMS, 2nd Edition, 2009.

[5] N. J. A. Sloane, Online Encyclopedia of Integer Sequences (OEIS), http://www.oeis.org/.

[6] B. Tsaban and L. Zdomskyy, Combinatorial images of sets of reals and semifilter trichotomy, *J. of Symbolic Logic* **73** (2008), 1278-1288.

[7] W. Wadge, *Reducibility and Determinateness on the Baire Space*, PhD Thesis, U.C. Berkeley, 1983.