A note on the asymptotic expressiveness of ZF and ZFC

Maciej Bendkowski

CLA 2020 13 October 2020

Problem statement

Problem

Consider Zermelo-Fraenkel set theory zF and its extension zFc including the axiom of choice. What is the asymptotic density of theorems provable within zFc? Is it true that zFc is asymptotically more expressive than zF?

Problem statement

Problem

Consider Zermelo-Fraenkel set theory zF and its extension zFc including the axiom of choice. What is the asymptotic density of theorems provable within zFc? Is it true that zFc is asymptotically more expressive than zF?

Definition

We define the *asymptotic expressiveness* of a theory T as the asymptotic density of its theorems among all possible sentences:

$$\mu(\mathcal{T}) = \lim_{n \to \infty} \frac{|\{\varphi : |\varphi| = n \land \mathcal{T} \vdash \varphi\}|}{|\{\varphi : |\varphi| = n\}|}.$$
(1)

The language of Zermelo-Fraenkel set theory consists of a single binary *membership* predicate (\in) and no function symbols. More specifically, assume that \mathcal{F} is a functionally complete set of proposition connectives, *e.g.* { \land , \lor , \neg }. Then,

• $(x \in y)$ is a valid formula for any two variables x and y,

The language of Zermelo-Fraenkel set theory consists of a single binary *membership* predicate (\in) and no function symbols. More specifically, assume that \mathcal{F} is a functionally complete set of proposition connectives, *e.g.* { \land , \lor , \neg }. Then,

• $(x \in y)$ is a valid formula for any two variables x and y,

 if φ₁,..., φ_n are valid formulae, then so is ∘(φ₁,..., φ_n) for all connectives (∘) ∈ F,

The language of Zermelo-Fraenkel set theory consists of a single binary *membership* predicate (\in) and no function symbols. More specifically, assume that \mathcal{F} is a functionally complete set of proposition connectives, *e.g.* { \land , \lor , \neg }. Then,

• $(x \in y)$ is a valid formula for any two variables x and y,

- if φ₁,...,φ_n are valid formulae, then so is ∘(φ₁,...,φ_n) for all connectives (∘) ∈ F,
- if φ is a valid formula and x is a variable, then (∀x.φ) and (∃x.φ) are valid formulae.

The language of Zermelo-Fraenkel set theory consists of a single binary *membership* predicate (\in) and no function symbols. More specifically, assume that \mathcal{F} is a functionally complete set of proposition connectives, *e.g.* { \land , \lor , \neg }. Then,

• $(x \in y)$ is a valid formula for any two variables x and y,

- if φ₁,..., φ_n are valid formulae, then so is ∘(φ₁,..., φ_n) for all connectives (∘) ∈ F,
- if φ is a valid formula and x is a variable, then (∀x.φ) and (∃x.φ) are valid formulae.

Example

 $\exists x. \ [\forall y. (y \notin x)].$

In the current work, we adopt the De Bruijn notation for formulae.

Example $\exists x. \ [\forall y. (y \notin x)] \equiv \exists \ [\forall (\underline{0} \notin \underline{1})].$ Symbolically, the set Φ_{∞} of formulae is defined as

$$\Phi_{\infty} = V \in V \mid orall \Phi_{\infty} \mid \exists \Phi_{\infty} \mid igcup_{(\circ) \in \mathcal{F}} \circ (ec{\Phi}_{\infty})$$

where indices V are represented in unary, *i.e.* $\underline{n} = S^{(n)}0$. Consequently V is specified by V = 0 | SV.

Size model

We assume a *natural* size notion for formulae. In other words, the *size* $|\varphi|$ of a formula φ is the number of its building constructors. Example $|\exists \ [\forall (\underline{0} \notin \underline{1})]| = 7.$

$$\Phi_{\infty} = |V \in V| \, orall \Phi_{\infty} \, | \, \exists \Phi_{\infty} \, | \, igcup_{(\circ) \in \mathcal{F}} \circ (ilde{\Phi}_{\infty})$$

Therefore

$$\Phi_{\infty}(z) = z \left(\frac{z}{1-z}\right)^2 + 2z \Phi_{\infty}(z) + \sum_{(\circ) \in \mathcal{F}} z \Phi_{\infty}(z)^{\operatorname{arity}(\circ)}.$$

Counting formulae

Proposition

The generating function Φ_{∞} admits a Puiseux expansion in form of

$$\Phi_{\infty}(z) = a - b\sqrt{1 - \frac{z}{\rho}} + O\left(\left|1 - \frac{z}{\rho}\right|\right)$$

Moreover

 $[z^n]\Phi_{\infty}(z)\sim C\cdot\rho^n n^{-3/2}.$

Counting formulae

Proposition

The generating function Φ_{∞} admits a Puiseux expansion in form of

$$\Phi_{\infty}(z) = a - b\sqrt{1 - \frac{z}{\rho}} + O\left(\left|1 - \frac{z}{\rho}\right|\right)$$

Moreover

$$[z^n]\Phi_{\infty}(z)\sim C\cdot\rho^n n^{-3/2}.$$

Proof.

By Post's theorem \mathcal{F} must contain a connective of arity $k \geq 2$. In consequence the defining equation of Φ_{∞} is non-linear in Φ_{∞} . We can therefore apply the Drmota–Lalley–Woods theorem.

Recall that we are interested in *sentences* rather than formulae. Hence we want to count formulae in which each index is *bound*, *i.e.* has a corresponding quantifier.

Example

 $\exists [\underline{0} \land (\forall \underline{0} \rightarrow \underline{2})]$ is not a sentence as $\underline{2}$ is *free*.

Recall that we are interested in *sentences* rather than formulae. Hence we want to count formulae in which each index is *bound*, *i.e.* has a corresponding quantifier.

Example

 $\exists [\underline{0} \land (\forall \underline{0} \rightarrow \underline{2})]$ is not a sentence as $\underline{2}$ is *free*.

Definition

We call a formula φ *m*-open if prepending φ with *m* head quantifiers, be it universal or existential, turns φ into a sentence, *i.e.* a formula without free indices.

Note that sentences are 0-open (closed) formulae.

By symbolic methods, the definition of *m*-open formulae gives rise to an infinitely nested system involving all the classes of *m*-open formulae Φ_m :

By symbolic methods, the definition of *m*-open formulae gives rise to an infinitely nested system involving all the classes of *m*-open formulae Φ_m :

 $\Phi_0 := \forall \Phi_1 \mid \exists \Phi_1 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ (\vec{\Phi}_0)$

By symbolic methods, the definition of *m*-open formulae gives rise to an infinitely nested system involving all the classes of *m*-open formulae Φ_m :

$$\begin{split} \Phi_0 &:= \forall \Phi_1 \mid \exists \Phi_1 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ (\vec{\Phi}_0) \\ \Phi_1 &:= \forall \Phi_2 \mid \exists \Phi_2 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ (\vec{\Phi}_1) \mid V_{<1} \in V_{<1} \end{split}$$

By symbolic methods, the definition of *m*-open formulae gives rise to an infinitely nested system involving all the classes of *m*-open formulae Φ_m :

$$\begin{split} \Phi_0 &:= \forall \Phi_1 \mid \exists \Phi_1 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ(\vec{\Phi}_0) \\ \Phi_1 &:= \forall \Phi_2 \mid \exists \Phi_2 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ(\vec{\Phi}_1) \mid V_{<1} \in V_{<1} \\ \Phi_2 &:= \forall \Phi_3 \mid \exists \Phi_3 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ(\vec{\Phi}_2) \mid V_{<2} \in V_{<2} \end{split}$$

. . .

By symbolic methods, the definition of *m*-open formulae gives rise to an infinitely nested system involving all the classes of *m*-open formulae Φ_m :

$$\begin{split} \Phi_0 &:= \forall \Phi_1 \mid \exists \Phi_1 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ (\vec{\Phi}_0) \\ \Phi_1 &:= \forall \Phi_2 \mid \exists \Phi_2 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ (\vec{\Phi}_1) \mid V_{<1} \in V_{<1} \\ \Phi_2 &:= \forall \Phi_3 \mid \exists \Phi_3 \mid \bigcup_{(\circ) \in \mathcal{F}} \circ (\vec{\Phi}_2) \mid V_{<2} \in V_{<2} \end{split}$$

By symbolic methods, the definition of *m*-open formulae gives rise to an infinitely nested system involving all the classes of *m*-open formulae Φ_m :

> $\Phi_0 := \forall \Phi_1 \mid \exists \Phi_1 \mid \bigcup \circ (\vec{\Phi}_0)$ (°)∈*F* $\Phi_1 := \forall \Phi_2 \mid \exists \Phi_2 \mid \bigcup \circ (\vec{\Phi}_1) \mid V_{<1} \in V_{<1}$ (∘)∈*F* $\Phi_2 := orall \Phi_3 \mid \exists \Phi_3 \mid \bigcup \circ (ec \Phi_2) \mid V_{<2} \in V_{<2}$ (°)∈*F* . . . $\Phi_m := \forall \Phi_{m+1} \mid \exists \Phi_{m+1} \mid \bigcup \circ (\vec{\Phi}_m) \mid V_{< m} \in V_{< m}$ (°)∈*F*

Note that $\Phi_0 \subsetneq \Phi_1 \subsetneq \Phi_2 \subsetneq \cdots \subsetneq \Phi_m \cdots \subsetneq \Phi_\infty$ and so, intuitively,

$$\Phi_m \xrightarrow[m \to \infty]{} \Phi_{\infty}$$

Proposition

For all $m \ge 0$, the number of *m*-open formulae satisfies

$$[z^n]\Phi_m(z)\sim C_m\cdot\rho^n n^{-3/2}.$$

Idea: Re-use tools developed to count closed λ -terms¹. Note that

$$\Phi_m(z) = z \left(\frac{z(1-z^m)}{1-z}\right)^2 + 2z \Phi_{m+1}(z) + \sum_{(\circ)\in\mathcal{F}} z \Phi_m(z)^{\operatorname{arity}(\circ)}$$
$$\Phi_\infty(z) = z \left(\frac{z}{1-z}\right)^2 + 2z \Phi_\infty(z) + \sum_{(\circ)\in\mathcal{F}} z \Phi_\infty(z)^{\operatorname{arity}(\circ)}.$$

¹B., Bodini, Dovgal. "Statistical properties of lambda-terms". EJC. 2019.

Counting theorems

Plan

Since we can now estimate $[z^n]\Phi_0(z)$, *i.e.* the number of sentences of size *n*, we need to focus on the number of *theorems* of size *n*.

Counting theorems

Plan

Since we can now estimate $[z^n]\Phi_0(z)$, *i.e.* the number of sentences of size *n*, we need to focus on the number of *theorems* of size *n*.

Problem

Deciding whether a given sentence φ is a theorem of zF (or zFC) is *undecidable*. So is to compute $|\{\varphi : |\varphi| = n \land zF \vdash \varphi\}|$.

What does it mean that $\mathcal{T} \vdash \varphi$?

1. there exists a *proof tree* whose leaves are (instantiations of) axioms of \mathcal{T} or predicate logic, internal nodes represent valid *inference rules (e.g. modus ponens)*, and the root represents φ .

What does it mean that $\mathcal{T} \vdash \varphi$?

there exists a *proof tree* whose leaves are (instantiations of) axioms of *T* or predicate logic, internal nodes represent valid *inference rules (e.g. modus ponens)*, and the root represents φ.
 φ holds in all *models* of *T*.

What does it mean that $\mathcal{T} \vdash \varphi$?

there exists a *proof tree* whose leaves are (instantiations of) axioms of *T* or predicate logic, internal nodes represent valid *inference rules (e.g. modus ponens)*, and the root represents φ.
 φ holds in all *models* of *T*.

Definition

A model (structure) of a theory \mathcal{T} is a non-empty universe U of objects and an interpretation I of \mathcal{T} 's signature (*i.e.* predicates and function symbols).

What does it mean that $\mathcal{T} \vdash \varphi$?

there exists a *proof tree* whose leaves are (instantiations of) axioms of *T* or predicate logic, internal nodes represent valid *inference rules (e.g. modus ponens)*, and the root represents φ.
 φ holds in all *models* of *T*.

Definition

A model (structure) of a theory \mathcal{T} is a non-empty universe U of objects and an interpretation I of \mathcal{T} 's signature (*i.e.* predicates and function symbols).

1. $I(f): U^n \to U$ if f is an *n*-ary symbol.

What does it mean that $\mathcal{T} \vdash \varphi$?

there exists a *proof tree* whose leaves are (instantiations of) axioms of *T* or predicate logic, internal nodes represent valid *inference rules (e.g. modus ponens)*, and the root represents φ.
 φ holds in all *models* of *T*.

Definition

A model (structure) of a theory \mathcal{T} is a non-empty universe U of objects and an interpretation I of \mathcal{T} 's signature (*i.e.* predicates and function symbols).

- 1. $I(f): U^n \to U$ if f is an *n*-ary symbol.
- 2. $I(P): U^n \to \{\text{false, true}\} \text{ if } P \text{ is } n\text{-ary predicate.}$

Definition

A theory \mathcal{T} is said be *consistent* if $\mathcal{T} \not\vdash \varphi \land \neg \varphi$ for any sentence φ or, equivalently, if \mathcal{T} has a model. \mathcal{T} is said to be *complete* if for all sentences φ it holds $\mathcal{T} \vdash \varphi$ or $\mathcal{T} \vdash \neg \varphi$.

Definition

A theory \mathcal{T} is said be *consistent* if $\mathcal{T} \not\vdash \varphi \land \neg \varphi$ for any sentence φ or, equivalently, if \mathcal{T} has a model. \mathcal{T} is said to be *complete* if for all sentences φ it holds $\mathcal{T} \vdash \varphi$ or $\mathcal{T} \vdash \neg \varphi$.

Definition

A sentence φ is said to be *independent* of \mathcal{T} if $\mathcal{T} \not\vdash \varphi$ nor $\mathcal{T} \not\vdash \neg \varphi$.

Definition

A theory \mathcal{T} is said be *consistent* if $\mathcal{T} \not\vdash \varphi \land \neg \varphi$ for any sentence φ or, equivalently, if \mathcal{T} has a model. \mathcal{T} is said to be *complete* if for all sentences φ it holds $\mathcal{T} \vdash \varphi$ or $\mathcal{T} \vdash \neg \varphi$.

Definition

A sentence φ is said to be *independent* of \mathcal{T} if $\mathcal{T} \not\vdash \varphi$ nor $\mathcal{T} \not\vdash \neg \varphi$.

Gödel's incompleteness theorems

Sufficiently expressive theories (such as ZF or ZFC) cannot be both consistent and complete. Moreover, *sufficiently expressive* and consistent theories cannot prove their own consistency.

Formulae templates

Definition

A *template* C is a formula with a single hole $[\cdot]$ instead of some sub-formula in form of $(\underline{n} \in \underline{m})$. To denote the result of substituting a formula φ for $[\cdot]$ in C we write $C[\varphi]$.

Formulae templates

Definition

A *template* C is a formula with a single hole $[\cdot]$ instead of some sub-formula in form of $(\underline{n} \in \underline{m})$. To denote the result of substituting a formula φ for $[\cdot]$ in C we write $C[\varphi]$.

Definition

We call a template *m*-permissive if for each *m*-open formula φ the resulting $C[\varphi]$ is a sentence, *i.e.* is 0-open. By analogy to formulae, the *size* of a template is the total weight of its building constructors, assuming that $[\cdot]$ weights zero.

Formulae templates

Definition

A *template* C is a formula with a single hole $[\cdot]$ instead of some sub-formula in form of $(\underline{n} \in \underline{m})$. To denote the result of substituting a formula φ for $[\cdot]$ in C we write $C[\varphi]$.

Definition

We call a template *m*-permissive if for each *m*-open formula φ the resulting $C[\varphi]$ is a sentence, *i.e.* is 0-open. By analogy to formulae, the *size* of a template is the total weight of its building constructors, assuming that $[\cdot]$ weights zero.

Example

The template $C = \exists \forall [\cdot]$ is of size two. The result of substituting $(\underline{0} \notin \underline{1})$ into C is the formula $C[(\underline{0} \notin \underline{1})] = \exists \forall (\underline{0} \notin \underline{1})$. The hole $[\cdot]$ is proceeded by two quantifiers in C so C is 2-permissive.

Lemma

Let C be an *m*-permissive template and $\mathcal{L}(C)$ be the language it generates, *i.e.* $\mathcal{L}(C) = \{C[\varphi] : \varphi \text{ is } m\text{-open}\}$. Then, the set $\mathcal{L}(C)$ has positive asymptotic density in the set of all sentences.

Lemma

Let C be an *m*-permissive template and $\mathcal{L}(C)$ be the language it generates, *i.e.* $\mathcal{L}(C) = \{C[\varphi] : \varphi \text{ is } m\text{-open}\}$. Then, the set $\mathcal{L}(C)$ has positive asymptotic density in the set of all sentences.

Proposition

Let \mathcal{T} be a consistent set theory. Then, the set of \mathcal{T} -theorems cannot have a trivial asymptotic density, *i.e.* neither $\mu(\mathcal{T}) \neq 0$ nor $\mu(\mathcal{T}) \neq 1$. Moreover, \mathcal{T} is inconsistent if and only if $\mu(\mathcal{T}) = 1$.

Lemma

Let C be an *m*-permissive template and $\mathcal{L}(C)$ be the language it generates, *i.e.* $\mathcal{L}(C) = \{C[\varphi] : \varphi \text{ is } m\text{-open}\}$. Then, the set $\mathcal{L}(C)$ has positive asymptotic density in the set of all sentences.

Proposition

Let \mathcal{T} be a consistent set theory. Then, the set of \mathcal{T} -theorems cannot have a trivial asymptotic density, *i.e.* neither $\mu(\mathcal{T}) \neq 0$ nor $\mu(\mathcal{T}) \neq 1$. Moreover, \mathcal{T} is inconsistent if and only if $\mu(\mathcal{T}) = 1$.

Proof.

Let τ be an arbitrary tautology. Consider $C = ([\cdot] \lor \tau)$ and $\overline{C} = ([\cdot] \land \neg \tau)$. Note that these have positive asymptotic density and consist of tautologies and anti-tautologies, respectively.

Proposition

Let \mathcal{T} be a consistent theory and ϕ be an sentence independent of \mathcal{T} , *i.e.* $\mathcal{T} \not\vdash \phi$ nor $\mathcal{T} \not\vdash \neg \phi$. Then, there exists a set of theorems of the extended $\mathcal{T} + \phi$ which has positive asymptotic density.

Proposition

Let \mathcal{T} be a consistent theory and ϕ be an sentence independent of \mathcal{T} , *i.e.* $\mathcal{T} \not\vdash \phi$ nor $\mathcal{T} \not\vdash \neg \phi$. Then, there exists a set of theorems of the extended $\mathcal{T} + \phi$ which has positive asymptotic density.

Proof.

Let τ be an arbitrary tautology. Consider the context $C = (\tau \lor [\cdot]) \to \phi$ and the set $\mathcal{L}(C) = \{C[\varphi] : \varphi \in \Phi_0\}$ it generates. Note that $\mathcal{L}(C)$ has positive asymptotic density and consists of sentences which cannot be proven in the weaker theory \mathcal{T} .

Proposition

Let \mathcal{T} be a consistent theory and ϕ be an sentence independent of \mathcal{T} , *i.e.* $\mathcal{T} \not\vdash \phi$ nor $\mathcal{T} \not\vdash \neg \phi$. Then, there exists a set of theorems of the extended $\mathcal{T} + \phi$ which has positive asymptotic density.

Proof.

Let τ be an arbitrary tautology. Consider the context $C = (\tau \lor [\cdot]) \to \phi$ and the set $\mathcal{L}(C) = \{C[\varphi] : \varphi \in \Phi_0\}$ it generates. Note that $\mathcal{L}(C)$ has positive asymptotic density and consists of sentences which cannot be proven in the weaker theory \mathcal{T} .

Corollary

The axiom of choice is an example of a sentence independent of zF hence zFc and zF cannot share the same asymptotic expressiveness.

Proposition

Let μ be a predicate definable in ZFC such that ZFC $\vdash \mu(g)$ if and only if $\mu($ ZFC) exists and is equal to g. Let CONSISTENT be the (canonical) predicate encoding the consistency of ZFC. If ZFC is consistent and

 $zfc \vdash consistent \leftrightarrow \neg \mu(1),$

then zFC $\not\vdash \neg \exists g \colon \mu(g)$.

(2)

Proposition

Let μ be a predicate definable in ZFC such that ZFC $\vdash \mu(g)$ if and only if $\mu($ ZFC) exists and is equal to g. Let CONSISTENT be the (canonical) predicate encoding the consistency of ZFC. If ZFC is consistent and

$$z_{FC} \vdash consistent \longleftrightarrow \neg \mu(1),$$
 (2)

then zFC $\not\vdash \neg \exists g \colon \mu(g)$.

Proof.

Suppose that $z_{FC} \vdash \neg \exists g : \mu(g)$. Equivalently, $z_{FC} \vdash \forall g : \neg \mu(g)$. Hence in particular $z_{FC} \vdash \neg \mu(1)$. By (2) it holds $z_{FC} \vdash \text{consistent}$. Gödel's second incompleteness theorem, as instantiated for z_{FC} , provides the required contradiction.

1. ZFC $\not\vdash$ CONSISTENT along with the assumption that ZFC is consistent imply that ZFC has a model \mathcal{M} such that $\mathcal{M} \vDash \mu(1)$ which witnesses the final proposition.

- 1. ZFC $\not\vdash$ CONSISTENT along with the assumption that ZFC is consistent imply that ZFC has a model \mathcal{M} such that $\mathcal{M} \vDash \mu(1)$ which witnesses the final proposition.
- 2. It is, however, unclear if the same argumentation can be carried out for the weaker zF. Our proof of zFC \vdash CONSISTENT $\leftrightarrow \neg \mu(1)$ uses analytic combinatorics and, most likely, the axiom of choice.

- 1. ZFC $\not\vdash$ CONSISTENT along with the assumption that ZFC is consistent imply that ZFC has a model \mathcal{M} such that $\mathcal{M} \vDash \mu(1)$ which witnesses the final proposition.
- 2. It is, however, unclear if the same argumentation can be carried out for the weaker zF. Our proof of zFC \vdash CONSISTENT $\leftrightarrow \neg \mu(1)$ uses analytic combinatorics and, most likely, the axiom of choice.

Open problem

If zFC is consistent, then either:

- 1. ZFC $\not\vdash$ CONSISTENT along with the assumption that ZFC is consistent imply that ZFC has a model \mathcal{M} such that $\mathcal{M} \vDash \mu(1)$ which witnesses the final proposition.
- 2. It is, however, unclear if the same argumentation can be carried out for the weaker zF. Our proof of zFC \vdash CONSISTENT $\leftrightarrow \neg \mu(1)$ uses analytic combinatorics and, most likely, the axiom of choice.

Open problem

- If zFC is consistent, then either:
 - ▶ zFC has a model \mathcal{M} such that $\mathcal{M} \vDash \neg \exists g \colon \mu(g)$, or

- 1. ZFC $\not\vdash$ CONSISTENT along with the assumption that ZFC is consistent imply that ZFC has a model \mathcal{M} such that $\mathcal{M} \vDash \mu(1)$ which witnesses the final proposition.
- It is, however, unclear if the same argumentation can be carried out for the weaker zF. Our proof of zFC ⊢ CONSISTENT ↔ ¬µ(1) uses analytic combinatorics and, most likely, the axiom of choice.

Open problem

- If zFc is consistent, then either:
 - ► zFC has a model \mathcal{M} such that $\mathcal{M} \models \neg \exists g \colon \mu(g)$, or
 - ► $\mathcal{M} \models \exists g : \mu(g)$ holds for each model \mathcal{M} of zFC. A priori these witnesses do not have to be the same across all models.

Thank you!