Logic of Scientific Discovery: How Physical Induction Affects What Is Computable

Vladik Kreinovich and Olga Kosheleva
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu, olgak@utep.edu

http://www.cs.utep.edu/vladik
http://www.cs.utep.edu/vladik/olgavita.html

1. Outline

- Most of our knowledge about a physical world comes from *physical induction*:
  - if a hypothesis is confirmed by a sufficient number of observations,
  - we conclude that this hypothesis is universally true.
- We show that a natural formalization of this property affects what is computable.
- We explain how this formalization is related to Kolmogorov complexity and randomness.
- We also consider computational consequences of an alternative idea also coming from physics:
  - that no physical law is absolutely true,
  - that every physical law will sooner or later need to be corrected.

2. Physical Induction: Main Idea

- How do we come up with physical laws?
- Someone formulates a hypothesis.
- This hypothesis is tested, and if it confirmed sufficiently many times.
- Then we conclude that this hypothesis is indeed a universal physical law.
- This conclusion is known as *physical induction*.
- Different physicists may disagree on how many experiments we need to become certain.
- However, most physicists would agree that:
  - after sufficiently many confirmations,
  - the hypothesis should be accepted as a physical law.
- Example: normal distribution :-)

3. How to Describe Physical Induction in Precise Terms

- Let s denote the state of the world, and let $P(s,i)$ indicate that the property P holds in the i-th experiment.
- In these terms, physical induction means that for every property P, there is an integer N such that:
  - if the statements $P(s,1)$, ..., $P(s,N)$ are all true,
  - then the property P holds for all possible experiments – i.e., we have $\forall n\, P(s,n)$.
- This cannot be true for all *mathematically possible* states: we can have $P(s,1),\ldots,P(s,N)$ and $\neg P(s,N+1)$.
- Our understanding of the physicists' claims is that:
  - if we restrict ourselves to *physically meaningful* states,
  - then physical induction is true.

4. Resulting Definition

- Let S be a set; its elements will be called *states of the world.*
- Let $T \subseteq S$ be a subset of S. We say that T *consists of physically meaningful states* if:
  - for every property P, there exists an integer N_P such that
  - for each state $s \in T$ for which $P(s, i)$ holds for all $i \leq N_P$, we have $\forall n\, P(s, n)$.
- For this definition to be precise, we need to select a theory $\mathcal{L}$ which is:
  - rich enough to contain all physicists' arguments and
  - weak enough so that we will be able to formally talk about definability in $\mathcal{L}$.

5. Definition: Equivalent Form

- We can reformulate this definition in terms of *definable sets*, i.e.:
  - sets of the type $\{x : P(x)\}$
  - corresponding to definable properties $P(x)$.
- Let S be a set; its elements will be called *states of the world.*
- Let $T \subseteq S$ be a subset of S. We say that T *consists of physically meaningful states* if:
  - for every definable sequence of sets $\{A_n\}$, there exists an integer N_A
  - such that $T \cap \bigcap\limits_{n=1}^{N_A} A_n = T \cap \bigcap\limits_{n=1}^{\infty} A_n$.

6. Existence Proof

- There are no more than countably many words, so no more than countably many definable sequences.
- For real numbers, we can enumerate all definable sequence, as $\{A_n^1\}$, $\{A_n^2\}$, …Let us pick $\varepsilon \in (0,1)$.
- For each k, for the difference sets $D_n^k \stackrel{\text{def}}{=} \bigcap\limits_{i=1}^{n} A_n^k - \bigcap\limits_{i=1}^{\infty} A_n^k$, we have $D_{n+1}^k \subseteq D_n^k$ and $\bigcap\limits_{n=1}^{\infty} D_n^k = \emptyset$, thus, $\mu(D_n^k) \to 0$.
- Hence, there exists n_k for which $\mu\left(D_{n_k}^k\right) \le 2^{-k} \cdot \varepsilon$.
- We then take $T = S - \bigcup\limits_{k=1}^{\infty} D_{n_k}^k$.
- Here, $\mu\left(\bigcup\limits_{k=1}^{\infty} D_{n_k}^k\right) \le \sum\limits_{k=1}^{\infty} \mu\left(D_{n_k}^k\right) \le \sum\limits_{k=1}^{\infty} 2^{-k} \cdot \varepsilon = \varepsilon < 1$, and thus, the difference T is non-empty.
- For this set T, we can take $N_{A^k} = n_k$.

7. From States of the World to Specific Quantities

- Usually, we only have a partial information about a state: we have a definable f-n $f : S \to X$ which maps
  - every state of the world
  - into the corresponding partial information.
- Then the range $f(T)$ corresponding to all physically meaningful states has the same property as T:
- Let a set $T \subseteq S$ consist of physically meaningful states, and let $f : S \to X$ be a definable function.
- Then, for every definable sequence of subsets $B_n \subseteq X$, there exists an integer N_B such that

$$f(T) \cap \bigcap_{n=1}^{N_B} B_n = f(T) \cap \bigcap_{n=1}^{\infty} B_n.$$

8. Proof

- We want to prove that for some N_B,
  - if an element $x \in f(T)$ belongs to the sets $B_1, \ldots, B_{N_B}$,
  - then $x \in B_n$ for all n.
- An arbitrary element $x \in f(T)$ has the form $x = f(s)$ for some $s \in T$.
- Let us take $A_n \stackrel{\text{def}}{=} f^{-1}(B_n)$.
- Since T consists of physically meaningful states, there exists an appropriate integer N_A.
- Let us take $N_B \stackrel{\text{def}}{=} N_A$.
- By definition of A_n, the condition $x = f(s) \in B_i$ implies that $s \in A_i$; so we have $s \in A_i$ for all $i \le N_A$.
- Thus, we have $s \in A_n$ for all n, which implies that $x = f(s) \in B_n$. Q.E.D.

9. Computations with Real Numbers: Reminder

- From the physical viewpoint, real numbers x describe values of different quantities.
- We get values of real numbers by measurements.
- Measurements are never 100% accurate, so after a measurement, we get an approximate value r_k of x.
- In principle, we can measure x with higher and higher accuracy.
- So, from the computational viewpoint, a real number is a sequence of rational numbers r_k for which, e.g.,

$$|x - r_k| \le 2^{-k}.$$

- By an algorithm processing real numbers, we mean an algorithm using r_k as an "oracle" (subroutine).
- This is how computations with real numbers are defined in *computable analysis*.

10. Checking Equality of Real Numbers

- *Known:* equality of real numbers is undecidable.
- For physically meaningful real numbers, however, a deciding algorithm *is* possible:
  - *for every set* $T \subseteq \mathbb{R}^2$ *which consists of physically meaningful pairs* (x, y) *of real numbers,*
  - *there exists an algorithm deciding whether* $x = y$*.*
- *Proof:* We can take $A_n = \{(x, y) : 0 < |x - y| < 2^{-n}\}$. The intersection of all these sets is empty.
- Hence, T has no elements from $\bigcap\limits_{n=1}^{N_A} A_n = A_{N_A}$.
- Thus, for each $(x, y) \in T$, $x = y$ or $|x - y| \geq 2^{-N_A}$.
- We can detect this by taking $2^{-(N_A+3)}$-approximations x' and y' to x and y. Q.E.D.

11. Finding Roots

- In general, it is not possible, given a f-n $f(x)$ attaining negative and positive values, to compute its root.
- This becomes possible if we restrict ourselves to physically meaningful functions:
- *Let K be a computable compact.*
- *Let X be the set of all functions $f : K \to \mathbb{R}$ that attain 0 value somewhere on K. Then:*
  - *for every set $T \subseteq X$ consisting of physically meaningful functions and for every $\varepsilon > 0$,*
  - *there is an algorithm that, given a f-n $f \in T$, computes an ε-approximation to the set of roots*

$$R \stackrel{\text{def}}{=} \{x : f(x) = 0\}.$$

- In particular, we can compute an ε-approximation to one of the roots.

12. Finding Roots: Proof

- To compute the set $R = \{x : f(x) = 0\}$ with accuracy $\varepsilon > 0$, let us take an $(\varepsilon/2)$-net $\{x_1, \ldots, x_n\} \subseteq K$.
- For each i, we can compute $\varepsilon' \in (\varepsilon/2, \varepsilon)$ for which $B_i \stackrel{\text{def}}{=} \{x : d(x, x_i) \le \varepsilon'\}$ is a computable compact set.
- It is possible to algorithmically compute the minimum of a function on a computable compact set.
- Thus, we can compute $m_i \stackrel{\text{def}}{=} \min\{|f(x)| : x \in B_i\}$.
- Since $f \in T$, similarly to the previous proof, we can prove that $\exists N \, \forall f \in T \, \forall i \, (m_i = 0 \vee m_i \ge 2^{-N})$.
- Comp. m_i w/acc. $2^{-(N+2)}$, we check $m_i = 0$ or $m_i > 0$.
- Let's prove that $d_H(R, \{x_i : m_i = 0\}) \le \varepsilon$, i.e., that $\forall i \, (m_i = 0 \Rightarrow \exists x \, (f(x) = 0 \,\&\, d(x, x_i) \le \varepsilon))$ and $\forall x \, (f(x) = 0 \Rightarrow \exists i \, (m_i = 0 \,\&\, d(x, x_i) \le \varepsilon))$.

13. Finding Roots: Proof (cont-d)

- $m_i = 0$ means $\min\{|f(x)| : x \in B_i \stackrel{\text{def}}{=} B_{\varepsilon'}(x_i)\} = 0$.
- Since the set K is compact, this value 0 is attained, i.e., there exists a value $x \in B_i$ for which $f(x) = 0$.
- From $x \in B_i$, we conclude that $d(x, x_i) \le \varepsilon'$ and, since $\varepsilon' < \varepsilon$, that $d(x, x_i) < \varepsilon$.
- Thus, x_i is ε-close to the root x.
- Vice versa, let x be a root, i.e., let $f(x) = 0$.
- Since the points x_i form an $(\varepsilon/2)$-net, there exists an index i for which $d(x, x_i) \le \varepsilon/2$.
- Since $\varepsilon/2 < \varepsilon'$, this means that $d(x, x_i) \le \varepsilon'$ and thus, $x \in B_i$.
- Therefore, $m_i = \min\{|f(x)| : x \in B_i\} = 0$. So, the root x is ε-close to a point x_i for which $m_i = 0$.

14. Optimization

- In general, it is not algorithmically possible to find x where $f(x)$ attains maximum.
- Let K be a computable compact. Let X be the set of all functions $f : K \to \mathbb{R}$. Then:
  - *for every set $T \subseteq X$ consisting of physically meaningful functions and for every $\varepsilon > 0$,*
  - *there is an algorithm that, given a f-n $f \in T$, computes an ε-approx. to* $S = \left\{x : f(x) = \max\limits_y f(y)\right\}$.
- In particular, we can compute an approximation to an individual $x \in S$.
- *Reduction to roots:* $f(x) = \max\limits_y f(y)$ iff $g(x) = 0$, where $g(x) \stackrel{\text{def}}{=} f(x) - \max\limits_y f(y)$.

15. Computing Fixed Points

- In general, it is not possible to compute all the fixed points of a given computable function $f(x)$.
- Let K be a computable compact. Let X be the set of all functions $f : K \to K$. Then:
  - *for every set* $T \subseteq X$ *consisting of physically meaningful functions and for every* $\varepsilon > 0$,
  - *there is an algorithm that, given a f-n* $f \in T$, *computes an* ε*-approximation to the set* $\{x : f(x) = x\}$.
- In particular, we can compute an approximation to an individual fixed point.
- *Reduction to roots:* $f(x) = x$ iff $g(x) = 0$, where $g(x) \stackrel{\text{def}}{=} d(f(x), x)$.

16. Computing Limits

- *In general:* it is not algorithmically possible to find a limit $\lim a_n$ of a convergent computable sequence.
- Let K be a computable compact. Let X be the set of all convergent sequences $a = \{a_n\}$, $a_n \in K$. Then:
  - *for every set* $T \subseteq X$ *consisting of physically meaningful functions and for every* $\varepsilon > 0$,
  - *there exists an algorithm that, given a sequence* $a \in T$, *computes its limit with accuracy* ε.
- *Use:* this enables us to compute limits of iterations and sums of Taylor series (frequent in physics).
- *Main idea:* for every $\varepsilon > 0$ there exists $\delta > 0$ such that when $|a_n - a_{n-1}| \le \delta$, then $|a_n - \lim a_n| \le \varepsilon$.
- *Intuitively:* we stop when two consequent iterations are close to each other.

17. Random Sequences: Reminder

- If a sequence s is random, it satisfies all the probability laws such as the law of large numbers.
- If a sequence satisfies all probability laws, then for all practical purposes we can consider it random.
- Thus, we can define a sequence to be random if it satisfies all probability laws.
- A probability law is a statement S which is true with probability 1: $P(S) = 1$.
- So, a sequence is random if it belongs to all definable sets of measure 1.
- A sequence belongs to a set of measure 1 iff it does not belong to its complement $C = -S$ with $P(C) = 0$.
- So, *a sequence is* random *if it does not belong to any definable set of measure 0.*

Home Page
Title Page
◀◀ ▶▶
◀ ▶

Go Back
Full Screen
Close
Quit

18. Randomness and Kolmogorov Complexity

- Different definabilities lead to different randomness.
- When definable means computable, randomness can be described in terms of Kolmogorov complexity

$$K(x) \stackrel{\text{def}}{=} \min\{\text{len}(p) : p \text{ generates } x\}.$$

- Crudely speaking, an infinite string $s = s_1 s_2 \ldots$ is random if, for some constant $C > 0$, we have

$$\forall n\, (K(s_1 \ldots s_n) \geq n - C).$$

- Indeed, if a sequence $s_1 \ldots s_n$ is truly random, then the only way to generate it is to explicitly print it:

$$\texttt{print}(s_1 \ldots s_n).$$

- In contrast, a sequence like 0101...01 generated by a short program is clearly not random.

19. From Kolmogorov-Martin-Löf Theoretical Randomness to a More Physical One

- The above definition means that (definable) events with probability 0 cannot happen.
- In practice, physicists also assume that events with a *very small* probability cannot happen.
- For example, a kettle on a cold stove will not boil by itself – but the probability is non-zero.
- If a coin falls head 100 times in a row, any reasonable person will conclude that this coin is not fair.
- It is not possible to formalize this idea by simply setting a threshold $p_0 > 0$ below which events are not possible.
- Indeed, then, for N for which $2^{-N} < p_0$, no sequence of N heads or tails would be possible at all.

20. From Kolmogorov-Martin-Löf Theoretical Randomness to a More Physical One (cont-d)

- We cannot have a universal threshold p_0 such that events with probability $\le p_0$ cannot happen.
- However, we know that:
  - for each decreasing ($A_n \supseteq A_{n+1}$) sequence of properties A_n with $\lim p(A_n) = 0$,
  - there exists an N above which a truly random sequence cannot belong to A_N.
- *Resulting definition:* we say that $\mathcal{R}$ is a *set of random elements* if
  - for every definable decreasing sequence $\{A_n\}$ for which $\lim P(A_n) = 0$,
  - there exists an N for which $\mathcal{R} \cap A_N = \emptyset$.

21. Random Sequences and Physically Meaningful Sequences

- Let $\mathcal{R}_K$ denote the set of all elements which are random in Kolmorogov-Martin-Löf sense. Then:
- *Every set of random elements consists of physically meaningful elements.*
- *For every set T of physically meaningful elements, the intersection $T \cap \mathcal{R}_K$ is a set of random elements.*
- *Proof:* When A_n is definable, for $D_n \stackrel{\text{def}}{=} \bigcap\limits_{i=1}^{n} A_i - \bigcap\limits_{i=1}^{\infty} A_i$, we have $D_n \supseteq D_{n+1}$ and $\bigcap\limits_{n=1}^{\infty} D_n = \emptyset$, so $P(D_n) \to 0$.
- Therefore, there exists an N for which the set of random elements does not contain any elements from D_N.
- Thus, every set of random elements indeed consists of physically meaningful elements.

22. Proof (cont-d)

- Let T consist of physically meaningful elements. Let us prove that $\mathcal{T} \cap \mathcal{R}_K$ is a set of random elements.
- If $A_n \supseteq A_{n+1}$ and $P\left(\bigcap\limits_{n=1}^{\infty} A_n\right) = 0$, then for $B_m \stackrel{\text{def}}{=} A_m - \bigcap\limits_{n=1}^{\infty} A_n$, we have $B_m \supseteq B_{m+1}$ and $\bigcap\limits_{n=1}^{\infty} B_n = \emptyset$.
- Thus, by definition of a set consisting of physically meaningful elements, we conclude that $B_N \cap T = \emptyset$.
- Since $P\left(\bigcap\limits_{n=1}^{\infty} A_n\right) = 0$, we also know that $\left(\bigcap\limits_{n=1}^{\infty} A_n\right) \cap \mathcal{R}_K = \emptyset$.
- Thus, $A_N = B_N \cup \left(\bigcap\limits_{n=1}^{\infty} A_n\right)$ has no common elements with the intersection $T \cap \mathcal{R}_K$. Q.E.D.

23. Random Sequences: Conclusion

- Kolmogorov-Martin-Löf randomness means that events with probability 0 cannot occur.
- Physicists also argue that events with a *sufficiently small* probability cannot occur.
- Physical induction means that every sequence belongs to a set S of physically meaningful sequences.
- In particular, a physical Kolmogorov-Martin-Löf random sequence s must belong to the set S.
- The above result shows that this sequence s is random in the physical sense as well.
- In other words, physical induction implies that events with a sufficiently small probability cannot occur.

24. Additional Consequence

- Main *objectives* of science:
  - *guaranteed* estimates for physical quantities;
  - *guaranteed* predictions for these quantities.
- *Problem:* estimation and prediction are ill-posed.
- *Example:*
  - measurement devices are inertial;
  - hence suppress high frequencies ω;
  - so $\varphi(x)$ and $\varphi(x) + \sin(\omega \cdot t)$ are indistinguishable.
- *Existing approaches:*
  - statistical regularization (filtering);
  - Tikhonov regularization (e.g., $|\dot{x}| \le \Delta$);
  - expert-based regularization.
- *Main problem:* no guarantee.

25. On Physically Meaningful Solutions, Problems Become Well-Posed

- *State estimation – an ill-posed problem:*
  - *Measurement f:*
    state $s \in S \to$ observation $r = f(s) \in R$.
  - *In principle,* we can reconstruct $r \to s$:
    as $s = f^{-1}(r)$.
  - *Problem:* small changes in r can lead to huge changes in s (f^{-1} *not continuous*).
- *Theorem:*
  - Let S be a definably separable metric space.
  - Let $\mathcal{T}$ be a set of physically meaningful elements of S.
  - Let $f : S \to R$ be a continuous 1-1 function.
  - Then, the inverse mapping $f^{-1} : R \to S$ is *continuous* for every $r \in f(\mathcal{T})$.

26. Everything is Related – Einstein-Podolsky-Rosen (EPR) Paradox

- Due to *Relativity Theory*, two spatially separated simultaneous events cannot influence each other.
- *Einstein, Podolsky, and Rosen* intended to show that in quantum physics, such influence is possible.
- *In formal terms*, let x and x' be measured values at these two events.
- *Independence* means that possible values of x do not depend on x', i.e., $S = X \times X'$ for some X and X'.
- *Physical induction* implies that the pair (x, x') belongs to a set S of physically meaningful pairs.
- *Theorem: The set S cannot be represented as $X \times X'$.*
- Thus, everything *is* related – but we probably can't use this relation to pass information (S isn't computable).

27. Alternative Idea: No Physical Theory Is Perfect

- Physical induction implies that a physical law is universally valid.
- However, in the history of physics,
  - always new observations appear
  - which are not fully consistent with the original theory.
- Thus, many physicists believe that every physical theory is approximate.
- For each theory T, inevitably new observations will surface which require a modification of T.
- Let us analyze how this idea affects computations.

28. No Physical Theory Is Perfect: How to Formalize This Idea

- *Statement:* for every theory, eventually there will be observations which violate this theory.
- To formalize this statement, we need to formalize what are *observations* and what is a *theory*.
- Each *observation* can be represented, in the computer, as a sequence of 0s and 1s.
- Most sensors already produce the signal in the computer-readable form, as a sequence of 0s and 1s.
- Thus, all past and future observations form a (potentially) infinite sequence $\omega = \omega_1\omega_2\ldots$ of 0s and 1s.
- A physical *theory* may be very complex.
- All we care about is which sequences of observations ω are consistent with this theory and which are not.

29. What Is a Physical Theory? (cont-d)

- So, a physical theory T can be defined as the set of all sequences ω which are consistent with this theory.
- A physical theory must have at least one possible sequence of observations: $T \neq \emptyset$.
- A theory must be described by a finite sequence of symbols: the set T must be *definable.*
- How can we check that an infinite sequence $\omega = \omega_1\omega_2\ldots$ is consistent with the theory?
- The only way is check that for every n, the sequence $\omega_1 \ldots \omega_n$ is consistent with T; so:

$$\forall n\, \exists \omega^{(n)} \in T\, (\omega_1^{(n)} \ldots \omega_n^{(n)} = \omega_1 \ldots \omega_n) \Rightarrow \omega \in T.$$

- In mathematical terms, this means that T is *closed* in the Baire metric $d(\omega, \omega') \stackrel{\text{def}}{=} 2^{-N(\omega,\omega')}$, where

$$N(\omega, \omega') \stackrel{\text{def}}{=} \max\{k : \omega_1 \ldots \omega_k = \omega'_1 \ldots \omega'_k\}.$$

30. What Is a Physical Theory: Final Definition

- A theory must predict something new.
- So, for every sequence $\omega_1 \ldots \omega_n$ consistent with T, there is a continuation which does not belong to T.
- In mathematical terms, T is *nowhere dense.*
- *By a* physical theory, *we mean a non-empty closed nowhere dense definable set* T.
- *A sequence* ω *is* consistent with the no-perfect-theory principle *if it does not belong to any physical theory.*
- In precise terms, ω does not belong to the union of all definable closed nowhere dense set.
- There are countably many definable set, so this union is *meager* (= *Baire first category*).
- Thus, due to Baire Theorem, such sequences ω exist.

31. How to Describe General Computations

- Each computation is a solution to a well-defined problem.
- As a result, each bit in the resulting answer satisfies a well-defined mathematical property.
- All mathematical properties can be described, e.g., in terms of Zermelo-Fraenkel set theory ZF.
- So, each bit in each computation result can be viewed as the truth value of some statement formulated in ZF.
- Let α_n denote the truth value of the n-th ZF statement.
- In these terms, each computation partially compute the sequence $\alpha = \alpha_1 \ldots \alpha_n \ldots$

32. Relative Kolmogorov Complexity

- The usual notion of Kolmogorov complexity provides the complexity of computing x "from scratch".
- Suppose we have a (potentially infinite) sequence y.
- *Relative Kolmogorov complexity* $K(x\,|\,y)$ can be used to describe the relative complexity of computing x.
- This relative complexity is based on programs which are allowed to use y as a subroutine.
- When we compute the length of such programs, we do not count the auxiliary program computing y_n.
- $K(x\,|\,y)$ is then defined as the shortest length of such a y-using program which computes x.
- If x and y are unrelated, then $K(x\,|\,y) \approx K(x)$.
- If $K(x\,|\,y) \ll K(x)$, then y helps compute x.

33. Computations Under No-Perfect-Theory Principle: Main Result

- Let us show that under the no-perfect-theory principle, observations do indeed enhance computations.
- *Let α be a sequence of truth values of ZF statements.*
- *Let ω be an infinite binary sequence which is consistent with the no-perfect-theory principle.*
- *Then, for every integer $C > 0$, there exists an integer n for which $K(\alpha_1 \ldots \alpha_n \,|\, \omega) < K(\alpha_1 \ldots \alpha_n) - C$.*
- In other words, in principle, we can have an arbitrary large enhancement.

34. Proof: Main Ideas

- We need to prove: $K(\alpha_1 \ldots \alpha_n \,|\, \omega) < K(\alpha_1 \ldots \alpha_n) - C$.
- For that, we prove that the set T of all sequences for which $K(\alpha_1 \ldots \alpha_n \,|\, \omega) \ge K(\alpha_1 \ldots \alpha_n) - C$ is a theory.
- The set T is clearly non-empty: it contains, e.g., $\omega = 00 \ldots 0 \ldots$ which does not affect computations.
- The set T is also definable: we have just defined it.
- The fact that computations involve only finitely many bits of ω can be used to prove that T is closed.
- To prove that T is nowhere dense, we can extend each sequence $\omega_1 \ldots \omega_m$ with α_i's: $\omega' \stackrel{\text{def}}{=} \omega_1 \ldots \omega_n \alpha_1 \alpha_2 \ldots$.
- For this new sequence ω', computing $\alpha_1 \ldots \alpha_n$ is easy: just copy α_i, so $K(\alpha_1 \ldots \alpha_n \,|\, \omega') \ll K(\alpha_1 \ldots \alpha_n) - C$.

35. Acknowledgments

- This work was supported in part by the National Science Foundation grants:
  - HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and
  - DUE-0926721.
- The authors are thankful:
  - to Luc Longpré, Sergei Soloviev, and Michael Zakharevich for valuable discussions,
  - to the anonymous referees for useful suggestions, and
  - to Oleg Prosorov for his help.

36. A Formal Definition of Definable Sets

- Let $\mathcal{L}$ be a theory.
- Let $P(x)$ be a formula from $\mathcal{L}$ for which the set $\{x \mid P(x)\}$ exists.
- We will then call the set $\{x \mid P(x)\}$ $\mathcal{L}$-*definable*.
- Crudely speaking, a set is $\mathcal{L}$-definable if we can explicitly *define* it in $\mathcal{L}$.
- All usual sets are definable: $\mathbb{N}$, $\mathbb{R}$, etc.
- Not every set is $\mathcal{L}$-definable:
  - every $\mathcal{L}$-definable set is uniquely determined by a text $P(x)$ in the language of set theory;
  - there are only countably many texts and therefore, there are only countably many $\mathcal{L}$-definable sets;
  - so, some sets of natural numbers are not definable.

37. How to Prove Results About Definable Sets

- Our objective is to be able to make mathematical statements about $\mathcal{L}$-definable sets. Therefore:
  - in addition to the theory $\mathcal{L}$,
  - we must have a stronger theory $\mathcal{M}$ in which the class of all $\mathcal{L}$-definable sets is a countable set.
- For every formula F from the theory $\mathcal{L}$, we denote its Gödel number by $\lfloor F \rfloor$.
- We say that a theory $\mathcal{M}$ is *stronger* than $\mathcal{L}$ if:
  - $\mathcal{M}$ contains all formulas, all axioms, and all deduction rules from $\mathcal{L}$, and
  - $\mathcal{M}$ contains a predicate $\text{def}(n, x)$ such that for every formula $P(x)$ from $\mathcal{L}$ with one free variable,

$$\mathcal{M} \vdash \forall y \, (\text{def}(\lfloor P(x) \rfloor, y) \leftrightarrow P(y)).$$

38. Existence of a Stronger Theory

- As $\mathcal{M}$, we take $\mathcal{L}$ plus all above equivalence formulas.
- Is $\mathcal{M}$ consistent?
- Due to compactness, we prove that for any $P_1(x), \ldots, P_m(x)$, $\mathcal{L}$ is consistent with the equivalences corr. to $P_i(x)$.
- Indeed, we can take

$$\text{def}(n, y) \leftrightarrow (n = \lfloor P_1(x) \rfloor \,\&\, P_1(y)) \vee \ldots \vee (n = \lfloor P_m(x) \rfloor \,\&\, P_m(y)).$$

- This formula is definable in $\mathcal{L}$ and satisfies all m equivalence properties.
- Thus, the existence of a stronger theory is proven.
- The notion of an $\mathcal{L}$-definable set can be expressed in $\mathcal{M}$: S is $\mathcal{L}$-definable iff $\exists n \in \mathbb{N}\, \forall y\, (\text{def}(n, y) \leftrightarrow y \in S)$.
- So, all statements involving definability become statements from the $\mathcal{M}$ itself, *not* from metalanguage.