PERFECT PREDICTION
NOTRE DAME—FOOD FOR THOUGHT SEMINAR

JOEL DAVID HAMKINS

I. The axiom of choice

A. The axiom asserts that every family $\mathcal{A}$ of nonempty sets admits a choice function f, meaning $f(A) \in A$ for all $A \in \mathcal{A}$.

B. Introduced 1904 in Zermelo's proof of the well-order theorem.

C. Well-order theorem: every set admits a well order.

D. Included in the 1908 Zermelo theory.

E. Now a standard axiom in ZFC set theory.

F. Immediate controversy.

> When in 1904 Zermelo published his proof that every set can be well ordered, many questions lay unresolved. ... Was [the axiom of choice] true? Was it a law of logic? Should one postulate simultaneous, independent arbitrary choices in preference to successive, dependent ones? Did the cardinality of the set of choices affect the validity of the Axiom, so that the Denumerable Axiom was true but not the Axiom of Choice in general?... What methods were permissible in mathematics? Must such methods be constructive? If so, what constituted a construction? What did it mean to say that a mathematical object existed? Normally mathematicians avoided such quasi-philosophical questions, and addressed them only when they felt their discipline to face a crisis. That Zermelo's proof precipitated such a crisis was shown by the extent of the resulting controversy. From 1905 to 1908 eminent mathematicians in England, France, Germany, Holland, Hungary, Italy, and the United States debated the validity of his demonstration. Never in modern times have mathematicians argued so publicly and so vehemently over a proof.
>
> -Moore [Moo82, p.85]

G. Inadvertant use by critics, uncovered by Zermelo.

H. Russell's allegory

> A millionaire possesses an infinite number of pairs of shoes, and an infinite number of pairs of socks. One day, in a fit of eccentricity, the millionaire summons his valet and asks him to select one shoe from each pair. When the valet, accustomed to receiving precise instructions, asks for details as to how to perform the selection, the millionaire suggests that the left shoe be chosen from each pair. Next day the millionaire proposes to the valet that he select one sock from each pair. When asked as to how this operation is to be carried out, the millionaire is at a loss for a reply, since, unlike shoes, there is no intrinsic way of distinguishing one sock of a pair from the other. In other words, the selection of the socks must be truly arbitrary.
>
> -Bell [Bel21, fn 4]; also Russell (1907, §III, 1959, p. 78).

II. Nonmeasurability consequence
   A. Vitali 1905 shows the existence of nonmeasurable sets
   B. Banach-Tarski paradox 1924.

III. Mathematical justification
   A. Extrinsic justification
   B. Traditional discussion admits extrinsice justification, then uses Banach-Tarski to heap scorn on AC.
   C. Intrinsic justification
   D. A balanced discussion: AC rules out bizarre phenomenon.
      1. Sequential continuity of $f : \mathbb{R} \to \mathbb{R}$ versus ε/δ account.
      2. Non-isomorphic algebraic closures of the rational field $\mathbb{Q}$.
      3. Failure of dimension: vector spaces with no basis, or bases of different cardinality.
      4. Putting real numbers in buckets, but number of buckets is strictly greater than $\mathbb{R}$.
      5. Tree with no terminal nodes, but no infinite branch.

IV. Finally, let me present the nearly perfect prediction theorem
   A. Hardin and Taylor [HT08]
   B. My summary account [Ham25]
   C. The philosophical problem of nonconstructive prediction procedures

References

[Bel21] John L. Bell. "The Axiom of Choice". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2021. Metaphysics Research Lab, Stanford University, 2021.

[Ham25] Joel David Hamkins. "The Nearly Perfect Prediction Theorem". *Notices of the American Mathematical Society* 72.3 (2025). Short Stories column, pp. 308–309. DOI: 10.1090/noti3098.

[HT08] Christopher S. Hardin and Alan D. Taylor. "A Peculiar Connection Between the Axiom of Choice and Predicting the Future". *Amer. Math. Monthly* 115.2 (2008), pp. 91–96. ISSN: 0002-9890,1930-0972. DOI: 10.1080/00029890.2008.11920502.

[Moo82] Gregory H. Moore. *Zermelo's Axiom of Choice*. Vol. 8. Studies in the History of Mathematics and Physical Sciences. Its origins, development, and influence. Springer-Verlag, New York, 1982, xiv+410 pp. (1 plate). ISBN: 0-387-90670-3. DOI: 10.1007/978-1-4613-9478-5.

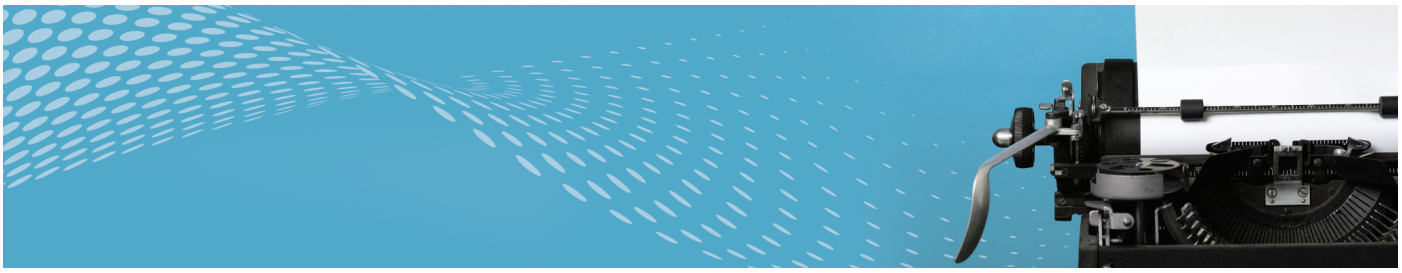
[Rus07] B. Russell. "On Some Difficulties in the Theory of Transfinite Numbers and Order Types". *Proceedings of the London Mathematical Society* 4 (1907), pp. 29–53. ISSN: 0024-6115. DOI: 10.1112/plms/s2-4.1.29.

[Rus59] Bertrand Russell. *My Philosophical Development*. Routledge Classics edition. Routledge, 1959.

(Joel David Hamkins) O'Hara Professor of Logic, University of Notre Dame, 100 Malloy Hall, Notre Dame, IN 46556 USA, & Guest Chair Professor, Peking University, Beijing, China

Email address: jdhamkins@nd.edu

URL: http://jdh.hamkins.org



The Nearly Perfect Prediction Theorem

Joel David Hamkins

Like stock-market traders trying to guess today's opening price and movements, we aim to predict the current and future values of an unknown function based on the observed history of previous values. The true function $f : \mathbb{R} \to \mathbb{R}$, not necessarily continuous, will be revealed only gradually.

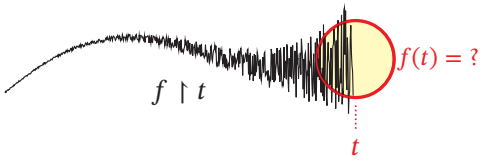


Figure 1.

The nearly perfect prediction theorem, a remarkable result of Christopher Hardin and Alan Taylor [HT08, HT13] shows that there is a nearly perfect prediction strategy, predicting the current and immediate future values of every function almost always correctly.

For predicting the present, we define a prediction strategy $P : \mathbb{R}^{<\mathbb{R}} \to \mathbb{R}$ on the set $\mathbb{R}^{<\mathbb{R}}$ of possible histories $f \restriction t$, each a real-valued function on domain $(-\infty, t)$ for some $t \in \mathbb{R}$, assigning them each to a corresponding present-value prediction $P(f \restriction t) \in \mathbb{R}$. For every function f, these are almost always correct:

$$f(t) = P(f \restriction t).$$

Specifically, the predictions will be correct except on a countable nowhere dense set of exceptional points. What is more, every real number s will be followed by an interval $(s, s + \varepsilon)$ on which the predictions are correct, and indeed there will be arbitrarily long intervals on which the predictions are perfectly correct. It is incredible.

To define the prediction strategy, simply fix a well order $\trianglelefteq$ on the set of all functions $f : \mathbb{R} \to \mathbb{R}$, and let P predict in accordance with the $\trianglelefteq$-least function standing in agreement with the observed data. That is, given historical information $f \restriction t$, let $P(f \restriction t) = f^*(t)$, where f^* is the $\trianglelefteq$-least function compatible with the prior data, meaning $f^* \restriction t = f \restriction t$.

To see that these predictions are almost always correct, suppose that the true function is $f : \mathbb{R} \to \mathbb{R}$. As time passes, more of the function is revealed and so the class of compatible functions becomes more restrictive. Consequently $s < t$ implies $f_s^* \trianglelefteq f_t^*$, where these are the $\trianglelefteq$-least functions in agreement with the data at those times. Further, if the prediction at time s is incorrect, then f_s^* would be incompatible with the data up to t, and so $f_s^* \neq f_t^*$. So the prediction functions climb strictly higher in the hierarchy $f_s^* \triangleleft f_t^*$ at every exceptional point, and since this hierarchy is well ordered, the exceptional points themselves must therefore be well ordered in the real numbers.

From this, all the amazing consequences follow. Above every real number s there is a next exceptional point, if any,

Joel David Hamkins is the O'Hara Professor of Logic at the University of Notre Dame and VRF at the Mathematical Institute of the University of Oxford. His email address is jdhamkins@nd.edu.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: https://doi.org/10.1090/noti3098

and so there is an open interval $(s, s+\varepsilon)$ on which all predictions are correct. The exceptional points are thus a countable nowhere dense set, countable since we can associate each exceptional s with a rational number in this interval. Since there must also be an absolutely smallest exceptional point r, if any, all predictions made at times $t \in (-\infty, r)$ are correct, and so there are arbitrarily long intervals on which all the predictions are correct. The real line is thus exhausted by a countable well-ordered sequence of intervals, which abut each other and their limits, with the predictions fully correct in the interior of each interval. Amazing!

Next, to predict the immediate future values of the function, we extend our earlier idea by defining a continuation prediction strategy $F : \mathbb{R}^{<\mathbb{R}} \to \mathbb{R}^{\mathbb{R}}$, which maps each history $f \restriction t$ to its predicted continuation $F(f \restriction t) = f^*$, the $\trianglelefteq$-least function standing in agreement with the observed data $f \restriction t = f^* \restriction t$.

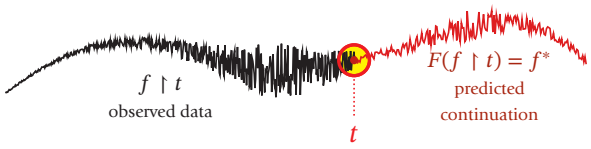


Figure 2.

Incredible as it may seem, these predicted continuations will almost always agree with the true function $f : \mathbb{R} \to \mathbb{R}$ in an interval extending strictly beyond t:

$$\exists \epsilon > 0 \quad \forall x < t + \epsilon \qquad f(x) = F(f \restriction t)(x).$$

We observed earlier in effect that if $s < t$, then $F(f \restriction s) \trianglelefteq F(f \restriction t)$, since any function compatible with the later data $f \restriction t$ is also compatible with the earlier data $f \restriction s$. As time passes, therefore, we are climbing in the $\trianglelefteq$ hierarchy, and we climb strictly higher whenever the prediction at s is not fully correct between s and t. So again the continuation prediction exceptional points will be well ordered in the real numbers, and consequently again they will form a countable nowhere dense set. So the continuation predictions are almost always correct. And again, following every point s will be an interval $(s, s + \epsilon)$ on which all predicted continuations are correct for the duration, and there will again be an absolutely least exceptional point r, meaning that the predicted continuations are fully correct for the entire interval $(-\infty, r)$, which includes arbitrarily long intervals on which the predictions are correct.

I find it interesting to notice a certain robust stability—whenever a correct continuation prediction is made, then the very same prediction will continue to be made for the duration of that correctness, since it will continue as the $\trianglelefteq$-least function in agreement with the observed data as long as it remains correct.

The theorem is interesting even just for continuous functions. Although predicting the present value $f(t)$ of a continuous function is trivial, of course, since it is just the limit of $f(x)$ as $x \to t$ from below, the nearly perfect prediction theorem enables us to predict how a continuous function will continue into the immediate future, which is not trivial. We can ensure that strategies P and F predict in accordance with continuity when possible simply by placing the continuous functions at the bottom of the $\trianglelefteq$ hierarchy.

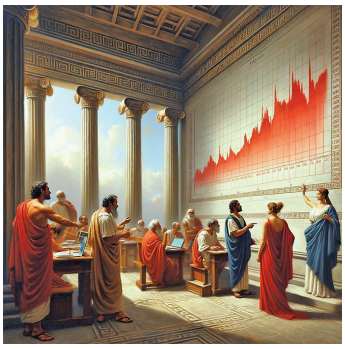


Figure 3. Stock traders in the platonic realm.

Meanwhile, the reader will have noticed that the proof of the nearly perfect prediction theorem is not constructive, since it relies on the axiom of choice via the well-order theorem. The proof does not provide a useful concrete prediction strategy—it reveals little about the actual predictions to be made. I regret to say that you had best keep your day job.

Rather, we proved the theorem as a pure-existence result, showing that there is a nearly perfect prediction strategy as a matter of abstract mathematical ontology. I am confident that the iridescent mathematical stock traders of the platonic realm are truly making bank.

References

[HT08] Christopher S. Hardin and Alan D. Taylor, *A peculiar connection between the axiom of choice and predicting the future*, Amer. Math. Monthly **115** (2008), no. 2, 91–96. MR2384262

[HT13] Christopher S. Hardin and Alan D. Taylor, *The mathematics of coordinated inference: A study of generalized hat problems*, Developments in Mathematics, vol. 33, Springer, Cham, 2013, https://doi.org/10.1007/978-3-319-01333-6. MR3100500



Joel David Hamkins

Credits

Figure 1 and Figure 2 were created by the author using TikZ.
Figure 3 was created by the author via DALL · E.
Author photo is courtesy of Joel David Hamkins.