DEEP THREE MATCH: USING NEURAL NETWORKS TO LEARN AN EVALUATION FUNCTION

by

THOMAS BRERETON

A thesis submitted to
The University of Birmingham
for the degree of
MASTER OF SCIENCE

School of Computer Science
College of Engineering and Physical Sciences
The University of Birmingham
August 2017

Abstract

This document only aims to demonstrate how one can use bhamthesis, a LaTeX class written by me. Please see the source file for more comments. Instead of making something up myself, I quote several pieces that may be of interest to mathematicians, and include some of my comments. *The contents of this document should not be treated seriously.* In particular, this document is not meant to be a thesis intended for submission.

CONTENTS

CHAPTER 1

THESIS WRITING

1.1 Pictures

It is not surprising that the English proverb ‘a picture is worth a thousand words’ has analogues in many languages, including Chinese and Japanese. Pictures are a very useful tool in explaining mathematics. The following is from Hardy’s *A Mathematician’s Apology* [4, §23].

> Let us suppose that I am giving a lecture on some system of geometry, such as ordinary Euclidean geometry, and that I draw figures on the blackboard to stimulate the imagination of my audience, rough drawings of straight lines or circles or ellipses. It is plain, first, that the truth of the theorems which I prove is in no way affected by the quality of my drawings. Their function is merely to bring home my meaning to my hearers, and, if I can do that, there would be no gain in having them redrawn by the most skillful draughtsman. They are pedagogical illustrations, not part of the real subject-matter of the lecture.

TeX and LaTeX are not very good at drawing pictures. However, it may still worth the effort to include some more important diagrams.

1.2 The mathematical experience

In the following excerpt from *The Mathematical Experience* [3, pp. 34–37], Davis and Hersh described how the ideal mathematician writes.

> To talk about the ideal mathematician at all, we must have a name for his "field," his subject. Let's call it, for instance, "non-Riemannian hypersquares." [...]
>
> To his fellow experts, [the ideal mathematician] communicates [his] results in a casual shorthand. "If you apply a tangential mollifier to the left quasi-martingale, you can get an estimate better than quadratic, so the convergence in the Bergstein theorem turns out to be of the same order as the degree of approximation in the Steinberg theorem."
>
> This breezy style is not to be found in his published writings. There he piles up formalism on top of formalism. Three pages of definitions are followed by seven lemmas and, finally, a theorem whose hypotheses take half a page to state, while its proof reduces essentially to "Apply Lemmas 1–7 to definitions A–H."
>
> His writing follows an unbreakable convention: to conceal any sign that the author or the intended reader is a human being. It gives the impression that, from the stated definitions, the desired results follow infallibly by a purely mechanical procedure. In fact, [to] read his proofs, one must be privy to a whole subculture of motivations, standard arguments and examples, habits of thought and agreed-upon modes of reasoning. The intended readers (all twelve of them) can decode the formal presentation, detect the new idea hidden in lemma 4, ignore the routine and uninteresting calculations of lemmas 1, 2, 3, 5, 6, 7, and see what the author is doing and why he does it. But for the uninitiate, this is a cipher that will never yield its secret. If (heaven forbid) the fraternity of non-Riemannian hypersquarers should ever die out, our hero's writings would become less translatable than those of the Maya.

Perhaps mathematical writing, and especially thesis writing, should not be this cold? Why should one take away *all* the human elements from his/her three years of mathematical experience?

1.3 Respectable mathematics

How much is your thesis supervisor involved in your thesis writing? Crilly and Johnson tell the story of Brouwer in their chapter in *History of Topology* [2, Section 7].

> [Brouwer's] doctoral thesis of 1907, *On the Foundations of Mathematics* (*Over de Grondslagen der Wiskunde*), marked the real beginning of his mathematical career. The work revealed the twin interests in mathematics that dominated his entire career: his fundamental concern with critically assessing the foundations of mathematics, which led to his creation of Intuitionism, and his deep interest in geometry, which led to his seminal work in topology

> [. . .]. Brouwer quickly found that his philosophical ideas sparked controversy. D.J. Korteweg (1848–1941), his thesis supervisor, had not been pleased with the more philosophical aspects of the thesis and had even demanded that several parts of the original draft be cut from the final presentation [. . .]. Korteweg urged Brouwer to concentrate on more "respectable" mathematics, so that the young man might enhance his mathematical reputation and thus secure an academic career.

You may get the same 'advice' from your supervisor, 90 years after Brouwer's days, if you try to put some philosophy into your thesis (and if your supervisor reads it). I do not think one should be discouraged for putting his/her ideas into the thesis, as long as the majority of the work is still mathematics. (This is only my personal opinion.) Given Brouwer's strong personality, one may expect him to have insisted on what he wanted to do. For some reason, he did not. As Crilly and Johnson [2] write:

> Brouwer was fiercely independent and did not follow in anybody's footsteps, but he apparently took his teacher's advice and set out to solve some really hard problems of mathematics. Brouwer put in a prodigious effort in these early years and rapidly produced a flood of papers on continuous group theory and topology — more than forty major papers in less than five years [. . .].

Perhaps I would only have a choice if I were as good as Brouwer.

By the way, the 'rejected parts' of Brouwer's thesis is now published [7]. You can even see the big crosses his supervisor drew on his drafts.

CHAPTER 2

THE NATURE OF MATHEMATICS

2.1 Arithmetical splitting

I was recently in a conversation with Andrey Bovykin, a former graduate student of the University of Birmingham. He told me about the idea of *arithmetic splitting*, which essentially says that there is no 'absolute truth' about the natural numbers. Pettigrew [6, pp. 19–20] explains it in a better way.

> This list of facts [...] gives a glimpse of the varied zoology of natural number structures that it supports. This, I propose, should be our foundation for arithmetic. [...]
>
> Of course, at first sight this proposal will seem extremely radical. It will seem that it entails changing much number-theoretic practice. Most importantly, if this foundation were adopted, each of our number-theoretic statements would have to be relativized to a particular collection of natural number structures. For instance, if I were to say that there are infinitely many primes, I would have to say in which natural number structures I take this to holds. Does it hold only in those closed under exponentiation? Or also in weaker structures? Thus, on my proposal, arithmetic would come to resemble branches of algebra such as group theory, in which we prove theorems that hold of all Abelian groups, for instance, or all cyclic groups; or field theory, in which sometimes we prove theorems that hold of all finite fields. [...]
>
> But perhaps [this] feature is [not] as revolutionary as it seems at first. [...] Much current research in mathematical logic concerns the strength of the number-theoretic assumptions required to prove certain propositions. For instance, in the case of Euclid's theorem that there are infinitely many primes, it has recently been shown that this holds in more natural number structures than previously thought.[...] This result, along with many, many similar to it, belongs to the research project known as *Weak* or *Bounded Arithmetic*,

> which studies the deductive power of a certain sort of fragment of first-order arithmetic.[...] In a similar vein, the research project of *Reverse Mathematics*, inaugurated by Harvey Friedman and carried on by Stephen G. Simpson, aims to identify, for a given proposition, the weakest fragment of second-order arithmetic in which that proposition may be proved.[...] Both research projects occupy central positions in contemporary research in arithmetic (more usually called *number theory* by mathematicians). So, it seems that a foundation for arithmetic in which UNIQUENESS fails might be more appropriate to the concerns of contemporary arithmetic than the traditional foundation, which guarantees UNIQUENESS.
>
> Furthermore, arithmetic belongs to that part of mathematics that is used to model phenomena in other disciplines. And, in some of these disciplines, it may well help to be able to choose between natural number structures with different properties. For instance, consider the notion of feasibility in computer science. We might well wish to say that the class of natural numbers that measure the feasible computations is closed under successor, but not under exponentiation [...]. And, if so, it will be difficult to model the notion of feasibility in arithmetic with the traditional foundation. But it is straightforward on the foundation proposed in this paper since there are natural number structures closed under successor but not under exponent[i]ation. Such a proposal requires more work than I have space to carry out here, but it suggests that there may be advantages to the foundation described and advocated in the previous section. [...]
>
> I conclude that mathematics took a wrong turning when it accepted the uniqueness of natural number structures and built into its foundations presuppositions that guarantee that uniqueness. Had it not taken that wrong turning, the orthodox foundation for arithmetic might have been $\mathsf{BST}_2^{\mathsf{Bnd}}$, which permits many different sorts of natural number structure: structures closed under addition and structures closed under multiplication, structures shorter than a structure that is closed under exponentiation, and so on. Had this been the foundation for arithmetic, we might have begun the study of so-called Weak Arithmetic and Reverse Mathematics much earlier than we did. I propose we rectify our error now.

From my education and experience, I am convinced that the set of natural numbers exists objectively. Therefore, in Pettigrew terminology, there is only one true 'natural number structure' for me. Set theory is different. We thought we knew a lot about sets, and we thought we could 'visualize' the universe of sets. Then it came Russell's Paradox, and everyone, including me, gave up (some of) our intuitions and work with formal axioms instead. Therefore, I do not believe there is an objectively existing universe of sets, just as I do not believe set theory is consistent.[1]

This is a widow, probably caused by the footnote.

[1] I suppose it is now common consensus amongst set theorists that there is a variety of set theoretic

I tried to think as Bovykin and Pettigrew suggest, but I still cannot convince myself. Here is one of the arguments Bovykin told me: you were convinced that natural numbers exist in exactly the same way as you were convinced 'the' universe of sets exists; why should you believe in the first argument but not the second? There is also some objective evidence supporting arithmetical splitting, e.g., Gödel's Incompleteness Theorems (depending on how you view it), but *there has to be one and only one* $\mathbb{N}$*!* Perhaps I just have a psychological barrier?

On the other hand, I would never call what is commonly known as arithmetic 'number theory', nor would I call number theory 'arithmetic'. They are different.

Logicians may be biased. What do other mathematicians think?

2.2 Topological spaces

Definition. A *topological space* is a pair $(X, \mathcal{T})$ where X is a nonempty set and $\mathcal{T}$ is a collection of subsets of that contains the empty set and X, and is closed under finite intersections and arbitrary unions.

As mathematicians, we are familiar with the idea of a topological space. The notion of topology abstracts the idea of open sets in a metric space, and the axioms for a topological space come from the properties of open sets. I remember having some difficulties accepting the definition of a topological space in my second year. Why are the open sets a good candidate for abstracting the notion of continuity? What do they model? If they model open sets in a metric space, then why should we allow topological spaces that are not Hausdorff? As Vickers said [9], mathematicians sometimes choose to work with a notion just because it works. Perhaps mathematicians should look for better motivations. In the particular case of topological spaces, you may find the answers to my questions in Vicker's book [8].

universes, none of which is superior to the others. There are (only) one or two exceptions. For example, Woodin thinks $2^{\aleph_0} = \aleph_2$ with strong reasons.

This is another awkward looking widow.

2.3 Set theory as *the* foundation

I recently participated in a conversation between several mathematicians, most of which are, strictly speaking, not logicians. One of them, who happens to know quite a lot of logic, insisted that *all* mathematicians should know enough set theory to distinguish whether they are working with sets, (proper) classes, or families of class, etc. to avoid running into set theoretic paradoxes.

I disagree with him. As a logician, I like set theory being popularized too, but mathematics *is* not set theory. Conway devoted a whole appendix in his *On Numbers and Games* [1] to this topic. Here is an excerpt on pages 65–67.

> In this simpler formalisation, a number is still a pretty complicated thing, namely a certain function in ZF, which is of course a certain set of Kuratowskian ordered pairs. The first members of these ordered pairs will be ordinals in the sense of von Neumann, and the second members chosen from the particular two-element set we take to represent $\{+, -\}$.
>
> The curiously complicated nature of these constructions tells us more about the nature of formalisations within ZF than about our system of numbers, and it is partly for this reason that we did not present any such formalised theory in this book. But the main reason was that we regard it as almost self-evident that our theory is as consistent as ZF, and that formalisation in ZF destroys a lot of its symmetry. [...]
>
> It seems to us, however, that mathematics has now reached the stage where formalisation within some particular axiomatic set theory is irrelevant, even for foundational studies. It should be possible to specify conditions on a mathematical theory which would suffice for embeddability within ZF (supplemented by additional axioms of infinity if necessary), but which do not otherwise restrict the possible constructions in that theory. Of course the conditions would apply to ZF itself, and to other possible theories that have been proposed as suitable foundations for mathematics (certain theories of categories, etc.), but would not restrict us to any particular theory. This appendix is in fact a cry for a Mathematicians' Liberation Movement!
>
> Among the permissible kinds of construction we should have:
>
> (i) Objects may be created from earlier objects in any reasonably constructive fashion.
>
> (ii) Equality among the created objects can be any desired equivalence relation.
>
> In particular, set theory would be such a theory, sets being constructed from earlier ones by processes corresponding to the usual axioms, and the equality relation being that of having the same members. But we could also,

for instance, freely create a new object (x, y) and call it the ordered pair of x and y. We could also create an ordered pair $[x, y]$ different from (x, y) but co-existing with it, and neither of these need have any relation to the set $\{\{x\}, \{x, y\}\}$. If instead we wanted to make (x, y) into an unordered pair, we could define equality by means of the equivalence relation $(x, y) = (z, t)$ if and only if $x = z$, $y = t$ *or* $x = t$, $y = z$. I hope it is clear that this proposal is not of any particular theory as an alternative to ZF (such as a theory of categories, or of the numbers or games considered in this book). What is proposed is instead that we give ourselves the freedom to create arbitrary mathematical theories of these kinds, but prove a metatheorem which ensures once and for all that any such theory could be formalised in terms of any of the standard foundational theories. The situation is analogous to the theory of vector spaces. Once upon a time these were collections of n-tuples of numbers, and the interesting theorems were those that remained invariant under linear transformations of these numbers. Now even the initial definitions are invariant, and vector spaces are defined by axioms rather than as particular objects. However, it is proved that every vector space has a base, so that the new theory is much the same as the old. But now no particular base is distinguished, and usually arguments which use particular bases are cumbrous and inelegant compared to arguments directly in terms of the axioms.

We believe that mathematics itself can be founded in an invariant way, which would be equivalent to, but would not involve, formalisation within some theory like ZF. No particular axiomatic theory like ZF would be needed, and indeed attempts to force arbitrary theories into a single formal strait-jacket will probably continue to produce unnecessarily cumbrous and inelegant contortions.

For those who doubt the possibility of such a programme, it might be worthwhile to note that certainly principles (i) and (ii) of our Mathematicians' Lib movement can be expressed directly in terms of the predicate calculus without any mention of sets (for instance), and it can be shown that any theory satisfying the corresponding restrictions can be formalised in ZF together with sufficiently many axioms of infinity.

Finally, we note that we have adopted the modern habit of identifying ZF (which properly has only sets) with the equiconsistent theory NBG (which has proper Classes as well) in this appendix and elsewhere. The classification of objects as Big and small is not peculiar to this theory, but appears in many foundational theories, and also in our formalised versions of principles (i) and (ii).

Formalization is probably still quite important to logicians though.

APPENDIX

TWO

The following is a famous quote from Knuth [5, Section 3.1].

> In a sense there is no such thing as a random number; for example, is 2 a random number?

I strongly believe that *two is not a random number*. There are many ways to see this. For example, two is the first number that allows *branching*. So while unary trees are not so interesting and ternary trees are too complicated, binary trees are just good.

More importantly, if one is interested in the relationships between objects, then one naturally considers *binary* relations to start with. We can draw the diagram

$$\circ \xrightarrow{R} \circ$$

to mean the left object is in relation R with the right object. This leads us to the study of *arrows*, or in technical terms, *categories*. Arrows have two ends. So they have a special kind of symmetry: one can flip an arrow by changing its head to its tail, and its tail to its head. When you do this *twice*, you get back to where you started. Such kinds of symmetries are called *involutions*, and this phenomenon is known as *duality*. In conclusion, if mathematics is about relationships between objects, then *two* must be a very important number.

If the previous paragraph does not look convincing to you, then you may find the following well-known theorem easier to take in.

Theorem (Folklore)**.** *All natural numbers are interesting.*

Proof. Suppose not. Then there is a natural number that is not interesting. Since the natural numbers are well-ordered, we can find a smallest such number n. Nothing can be more interesting than being the smallest uninteresting natural number. Therefore, n is both interesting and uninteresting, which is a contradiction. $\square$

LIST OF REFERENCES

[1] John H. Conway. *On Numbers and Games*, volume 6 of *London Mathematical Society Monographs*. Academic Press, London, 1976.

[2] Tony Crilly and Dale Johnson. The emergence of topological dimension theory. In Ioan Mackenzie James, editor, *History of Topology*, chapter 1, pages 1–24. Elsevier, Amsterdam, first edition, 1999.

[3] Philip J. Davis and Reuben Hersh. *The Mathematical Experience*. Birkhäuser, Boston, 1981. With an introduction by Gian-Carlo Rota.

[4] Godfrey Harold Hardy. *A Mathematician's Apology*. Cambridge University Press, London: Bentley House, 1940.

[5] Donald E. Knuth. *The Art of Computer Programming*, volume 2: Seminumerical Algorithms. Addison-Wesley Publishing Company, Reading, Massachusetts, 1969.

[6] Richard Pettigrew. The non-uniqueness of the natural numbers. Available for download at `http://seis.bris.ac.uk/~rp3959/possnoniso5unblind.pdf`. Accessed on 24 December 2009. Preprint dated 13 February 2009.

[7] Walter P. van Stigt. The rejected parts of Brouwer's dissertation on the foundations of mathematics. *Historia Mathematica*, 6:385–404, 1979.

[8] Steve Vickers. *Topology via Logic*, volume 5 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, 1989.

[9] Steve Vickers. Aspects of topology. Departmental Seminar at the School of Computer Science, University of Birmingham. The slides are available for download at `http://www.cs.bham.ac.uk/~sjv/SchoolSeminar.pdf`, 30 October 2008.