

Algebraic Logic for Deep Learning

by

YKY

A Thesis Submitted to
The Hong Kong University of Science and Technology
in Partial Fulfilment of the Requirements for
the Degree of Master of Philosophy
in Applied Mathematics

August 2024, Hong Kong

Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the Hong Kong University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

YKY

13 August 2024

Algebraic Logic for Deep Learning

by

YKY

This is to certify that I have examined the above MPhil thesis
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by
the thesis examination committee have been made.

Prof. Who, Thesis Supervisor

Prof. Who Else, Thesis Co-supervisor

Prof. Ikari Someone, Head of Department

Department of Maths

13 August 2024

ACKNOWLEDGEMENTS

Thank you, all the Evangelion.

TABLE OF CONTENTS

Title Page		i
Authorization		iii
Signature Page		v
Acknowledgements		vii
Table of Contents		ix
List of Figures		xi
List of Tables		xiii
List of Algorithms		xv
Abstract		xvii
Chapter 1	Introduction	1
	1.1 Background	1
	1.2 Paul Halmos’ algebraic logic	1
	1.3 Yuri Manin and Russians	2
	1.4 Topos and internal language	2
	1.5 Term rewriting and all that	2
	1.6 The set-up	2
	1.7 From abstract algebraic logic to concrete computations	3
	1.8 What does it mean to train the AI?	3
Chapter 2	Example chapter	7
	2.1 Background	7
	2.1.1 Cross references	7
	2.1.2 Citation and bibliography	7
	2.2 Math	8

2.2.1	Symbols	8
2.2.2	Equations	8
2.2.3	Theorem	8
2.3	Figure	9
2.4	Table	9
2.5	Code	10
2.5.1	Inline code	10
2.5.2	Code environment	10
Chapter 3	Conclusions	11
Appendix A	List of Publications	13
Appendix B	FYTGS Requirements	15
B.1	Components	15
B.1.1	Order	15
B.1.2	Authorization page	15
B.1.3	Signature page	16
B.1.4	Acknowledgments	16
B.1.5	Abstract	16
B.1.6	Bibliography	16
B.2	Language, Style and Format	16
B.2.1	Language	16
B.2.2	Pagination	17
B.2.3	Format	17
B.2.4	Footnotes	17
B.2.5	Appendices	18
B.2.6	Figures, Tables and Illustrations	18
B.2.7	Photographs/Images	18
B.2.8	Additional Materials	18

LIST OF FIGURES

Figure 2.1	An example tikz picture with a short caption.	9
Figure 2.2	An example tikz picture with breakline and a very long caption.	9

LIST OF TABLES

Table 2.1	A table with a short caption.	9
Table 2.2	Test result on different platforms with breakline and a very long caption.	9

LIST OF ALGORITHMS

Algorithm 1 Temp

8

Algebraic Logic for Deep Learning

by YKY

Department of Maths

The Hong Kong University of Science and Technology

Abstract

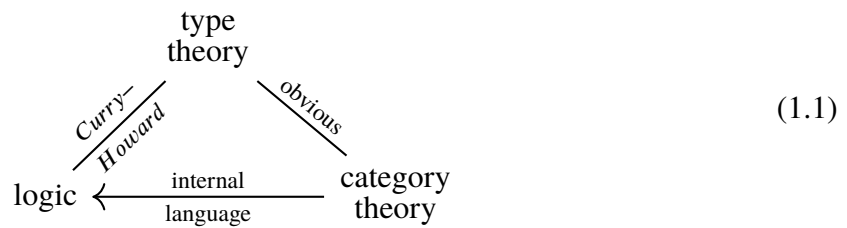
Some text.

CHAPTER 1

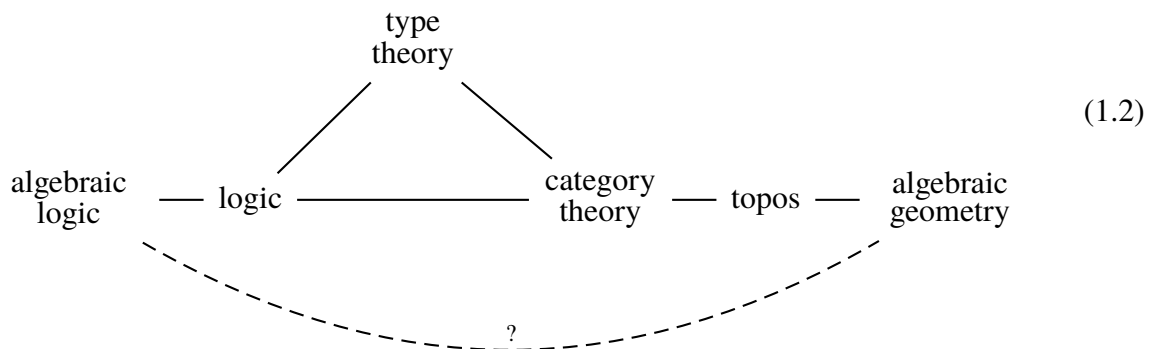
INTRODUCTION

1.1 Background

Lambek:



I am very curious to see if the two algebras below would coincide?



1.2 Paul Halmos' algebraic logic

Every Boolean algebra \mathbb{A} is isomorphic to the set of all continuous functions from X into \mathbb{O} , where X is the dual space of the algebra \mathbb{A} , and \mathbb{O} is the Boolean algebra with 2 elements. If there is a homomorphism f between Boolean algebras $\mathbb{A} \rightarrow \mathbb{B}$ then there is a dual morphism f^*

between their dual spaces $Y \rightarrow X$:

$$\begin{array}{ccc}
 \mathbb{A} & \xrightarrow{f} & \mathbb{B} \\
 \Downarrow & & \Downarrow \\
 \overline{X} & \xleftarrow{f^*} & \overline{Y} \\
 \downarrow & & \downarrow \\
 \mathbb{O} & & \mathbb{O}
 \end{array} \tag{1.3}$$

1.3 Yuri Manin and Russians

1.4 Topos and internal language

1.5 Term rewriting and all that

1.6 The set-up

The set of equations F defines an algebraic set = **the world**:

$$F(x) = 0. \tag{1.4}$$

The objective of an intelligent agent is to learn F .

We have the function f performing **prediction** of the immediate future:

$$\boxed{\text{current state}} \quad x_t \xrightarrow{f} x_{t+1} \quad \boxed{\text{next state}}. \tag{1.5}$$

In an infinitesimal sense, we can see f as a **differential equation** describing the **world trajectory**:

$$\dot{x} = f(x). \tag{1.6}$$

So F is the **solution** to this differential equation.

It seems that F and f are more or less equivalent ways to describe the world.

Logic can be turned into some form of algebra, and this algebra can be used to express either F or f . Perhaps both ways are feasible, or even mixing the two.

What does it mean to use logic to express F or f ?

The following table depicts the main correspondences relevant to our research:

LOGIC	facts human(socrates)	rules $\forall x.\text{human}(x) \rightarrow \text{mortal}(x)$
ALGEBRA	element $p \in \mathbb{A}$	element $(p \rightarrow q) \in \mathbb{A}$
WORLD	states x_t	state transitions $x_t \xrightarrow{f} x_{t+1}$

(1.7)

The relation between LOGIC and WORLD has been elucidated quite thoroughly in the AI literature. Note that the state x_t is made up of a set of facts (logic propositions). A single step of logic inference results in a new conclusion δx which is *added* (as a set element) to the current state x_t to form a new state x_{t+1} . Here t refers to “mental time” which does not necessarily coincide with real time.

1.7 From abstract algebraic logic to concrete computations

There are two main routes to make abstract algebraic logic concrete:

- Find **matrix representations** of the logical algebra
- Implement the logical algebra as the commutative algebra of (classical) **polynomials**

1.8 What does it mean to train the AI?

From the previous section,

$$F(x) = 0 \quad \text{is the solution to} \quad \dot{x} = f(x) \quad (1.8)$$

and the two descriptions (by F or by f) are equivalent.

The sensory data from the AI are a set of “world” points $\{x_i\}$ and we require either:

$$F(x_t) = 0 \quad \text{or} \quad f(x_t) = \delta x = x_{t+1} - x_t \quad (1.9)$$

and F or f can be trained by gradient descent to eliminate errors in the above conditions (equations).

- the x_t ’s are represented as **logic facts**

- F or f is represented as **logic rules**

and we need to **evaluate** $F(x_t)$ or $f(x_t)$.

Let's do some examples:

Logic formula	Algebraic form
human(socrates)	$h(s) = 1$
human(socrates) \wedge human(plato)	$h(s) \cdot h(p) = 1$
human(socrates) \rightarrow mortal(socrates)	$1 + h(s) + h(s) \cdot m(s) = 1$
$\forall x. \text{human}(x)$	$h(x)$ is a propositional function $\forall_x h(x)$ is a constant function mapping to 1 or 0
$\forall x. \text{human}(x) \rightarrow \text{mortal}(x)$	$\forall_x (1 + h(x) + h(x) \cdot m(x))$ is a constant function mapping to 1 or 0
$\forall x, y, z. \text{father}(x, y) \wedge \text{father}(y, z) \rightarrow$ $\text{grandfather}(x, z)$	$\forall_x \forall_y \forall_z (1 + f(x, y) \cdot f(y, z) + f(x, y) \cdot f(y, z) \cdot g(x, z))$ $\mapsto 0$ or 1
general Horn formula: $\forall_{x...} P \wedge Q \wedge R... \rightarrow Z$	$\forall_{x...} (1 + P \cdot Q \cdot R... + P \cdot Q \cdot R... \cdot Z)$ $\mapsto 0$ or 1

Now imagine there are millions of such rules. Number of predicates obviously increases.

Does each equation require new variables, or can variables be re-used? Seems yes, can be re-used.

The **loss function** would be the sum of squared errors over all equations:

$$\mathcal{L} = \sum_{\text{eqns}} \epsilon^2 = \sum_i (\phi_i(x...) - 1)^2. \quad (1.10)$$

Learning means to perform the **gradient descent** via $\nabla_{\Phi} \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \Phi}$ where Φ is the set of parameters for the equations.

How are new conclusions added to the state? What is the state? State = set of facts = set of **grounded** equations.

Inference: how to get from current state to next state? Big problem!! New ground facts have to be read off from satisfaction of all equations. Rather intractable...

Potential problems:

- How to represent the set of equations efficiently? Matrix of coefficients seems wasteful.
- We have lost the “**deepness**” of deep learning, but there is recent research showing that **shallow learning** may work well too.

- Need to iterate logical inference multiple times using the same set of equations.

CHAPTER 2

EXAMPLE CHAPTER

2.1 Background

2.1.1 Cross references

You can use `\cref{}` to automatically setup the cross reference name; instead, you can always use `\ref{}` to customize the appearance of the cross reference.

Chapter 1 tells you to read the latest PDF documentation.

Chapter 3 is a conclusion.

2.1.2 Citation and bibliography

Use `biber` as `BIBTEX` backend.

Cite a book[**LaTeX.Companion**], some papers[**KeshavACMSIGCOMMComput.Commun.Rev.2007**, **WhitesidesAdv.Mater.2004**], and a conference[**Babu2020IEEE33rdInt.Conf.MicroElectroMech.Syst.MEM**].

Citation database

Bibliography entries (database) are stored in `mythesis.bib`. You can use Zotero (or similar software) to generate `another.bib` file. You can add multiple database files by adding their filenames one-by-one in the following commands in `mythesis.tex`:

```
1 \addbibresource{mythesis.bib}
2 \addbibresource{another.bib}
```

Citation style

As described in the sample page from ECE department, the style is set to `ieee` by default. You can modify the style in the `hkustthesis.cls` file as you wish.

2.2 Math

2.2.1 Symbols

- Calligraphic letters: \mathcal{A}
- Mathbb letters: \mathbb{A}
- Mathfrak letters: \mathfrak{A}
- Math Sans serif letters: A
- Math bold letters: \mathbf{A}
- Math bold upright Greek letters: (Not displaying! Use the following one.)
- Math bold upright Greek letters: $\boldsymbol{\alpha}$
- Math bold italic Greek letters¹: $\boldsymbol{\alpha}$
- Math bold italic Greek letters in upper case: \boldsymbol{A}

2.2.2 Equations

$$E^2 = m^2 + p^2 \tag{2.1}$$

Equation (2.1) or Equation (2.1) gives the mass-energy relationship.

2.2.3 Theorem

Definition 1 LCL is orange juice.

Proof. They are both orange. □

Algorithm 1 Temp

1: Temp

Available theorem environments are listed below:

algorithm, assumption, axiom, conclusion, condition, corollary, definition, example, lemma, proof, property, proposition, remark, theorem.

¹Avoid using `bm` package as it conflicts with `unicode-math` and it is outdated for \LaTeX . You can alias some math commands by `\newcommand` or `\renewcommand` anyway.

2.3 Figure

An example image is shown in Figure (2.2) or Figure (2.2).



Figure 2.1: An example tikz picture with a short caption.

[illegible]

2.4 Table

An example table is shown in Table (2.2) or Table (2.2).

Table 2.1: A table with a short caption.

OS	TeX environment	Test
Overleaf	T _E X Live 2021~4	Pass
Windows 10	T _E X Live 2020	l_txhook problem
Ubuntu 20.04	T _E X Live 2021	Pass

[illegible]

OS	TeX environment	Test
Overleaf	T _E X Live 2021~4	Pass
Arch Linux (2024.10)	T _E X Live	Pass
Windows 10/11	T _E X Live 2021	Pass
macOS 10.15	T _E X Live 2021	Pass
Windows 10	T _E X Live 2020	l _t xhook problem
Ubuntu 20.04	T _E X Live 2021	Pass
Termux	T _E X Live 2021	Pass
Windows 11	MiK _T E _X 4.9	Pass
Windows 10	MiK _T E _X 4.4	Pass

2.5 Code

2.5.1 Inline code

Use `\lstinline|<code>|` to print code snippets. The `|` marks delimit the code and can be replaced by any character not in the code; *e.g.*, `\lstinline$<code>$` gives the same result.

2.5.2 Code environment

The code to draw the Figure (2.2) is listed below:

Listing 2.1: L^AT_EX code for inserting a figure

```
1 \begin{figure}[htb]
2   \begin{tikzpicture}
3     \draw (0,0) -- (1, 0) -- (1, 1) -- cycle;
4   \end{tikzpicture}
5   \caption{An example picture with long caption: \blindtext
6     \\\blindtext}
7   \label{fig:tikz example} % this is a comment
8 \end{figure}
```

CHAPTER 3

CONCLUSIONS

Some conclusion text.

APPENDIX A

LIST OF PUBLICATIONS

APPENDIX B

FYTGS REQUIREMENTS

The requirements are from the RPG Handbook.

B.1 Components

B.1.1 Order

A thesis should contain the following parts in the order shown:

1. Title page, containing in this order:
 - a. Thesis title
 - b. Full name of the candidate
 - c. Degree for which the thesis is submitted
 - d. Name of the University, *i.e.* The Hong Kong University of Science and Technology
 - e. Month and year of submission
2. Authorization page
3. Signature page
4. Acknowledgments
5. Table of contents
6. Lists of figures and tables
7. Abstract (≤ 300 words.)
8. Thesis body
9. Bibliography
10. Appendices and other addenda, if any.

B.1.2 Authorization page

On this page, students authorize the University to lend or reproduce the thesis.

1. The copyright of the thesis as a literary work vests in its author (the student).

2. The authorization gives HKUST Library a non-exclusive right to make it available for scholarly research.

B.1.3 Signature page

This page provides signatures of the thesis supervisor(s) and Department Head confirming that the thesis is satisfactory.

B.1.4 Acknowledgments

The student is required to declare, in this section, the extent to which assistance has been given by his/her faculty and staff, fellow students, external bodies or others in the collection of materials and data, the design and construction of apparatus, the performance of experiments, the analysis of data, and the preparation of the thesis (including editorial help). In addition, it is appropriate to recognize the supervision and advice given by the thesis supervisor(s) and members of TSC.

B.1.5 Abstract

Every copy of the thesis must have an English abstract, being a concise summary of the thesis, in 300 words or less.

B.1.6 Bibliography

The list of sources and references used should be presented in a standard format appropriate to the discipline; formatting should be consistent throughout.

Sample pages of both MPhil and PhD theses are provided here (MPhil / PhD), with specific instructions for formatting page content (centering, spacing, etc.).

B.2 Language, Style and Format

B.2.1 Language

Theses should be written in English.

Students in the School of Humanities and Social Science who are pursuing research work in the areas of Chinese Studies, and who can demonstrate a need to use Chinese to write their theses should seek prior approval from the School via their thesis supervisor and the divisional head.

If approval is granted, students are also required to produce a translation of the title page, authorization page, signature page, table of contents and the abstract in English.

B.2.2 Pagination

1. All pages, starting with the Title page should be numbered.
2. All page numbers should be centered, at the bottom of each page.
3. Page numbers of materials preceding the body of the text should be in small Roman numerals.
4. Page numbers of the text, beginning with the first page of the first chapter and continuing through the bibliography, including any pages with tables, maps, figures, photographs, etc., and any subsequent appendices, should be in Arabic numerals.
5. Start a new page after each chapter or section but not after a sub-section.

Note: That means the Title page will be page i; the first page of the first chapter will be page 1.

B.2.3 Format

1. A conventional font, size 12-point, 10 to 12 characters per inch must be used.
2. One-and-a-half line spacing should be used throughout the thesis, except for abstracts, indented quotations or footnotes where single line spacing may be used.
3. All margins—top, bottom, sides—should be consistently 25mm (or no more than 30mm) in width. The same margin should be used throughout a thesis. Exceptionally, margins of a different size may be used when the nature of the thesis requires it.

B.2.4 Footnotes

1. Footnotes may be placed at the bottom of the page, at the end of each chapter or after the end of the thesis body.
2. Like references, footnotes should be presented in a standard format appropriate to the discipline.
3. Both the position and format of footnotes should be consistent throughout the thesis.

B.2.5 Appendices

The format of each appended item should be consistent with the nature of that item, whether text, diagram, figure, etc., and should follow the guidelines for that item as listed here.

B.2.6 Figures, Tables and Illustrations

Figures, tables, graphs, etc., should be positioned according to the scientific publication conventions of the discipline, e.g., interspersed in text or collected at the end of chapters. Charts, graphs, maps, and tables that are larger than a standard page should be provided as appendices.

B.2.7 Photographs/Images

1. High contrast photos should be used because they reproduce well. Photographs with a glossy finish and those with dark backgrounds should be avoided.
2. Images should be dense enough to provide 300 ppi for printing and 72 dpi for viewing.

B.2.8 Additional Materials

Raw files, datasets, media files, and high resolution photographs/images of any format can be included.

Note: Students should get approval from their department head before deviating from any of the above requirements concerning paper size, font, margins, etc.