

Finding Adam
Advancing Reasoning in Artificial Intelligence as a
Potential Step towards Artificial General Superintelligence

A Doctoral Thesis

Abhinav Madahar ◦ अभिनव मद्दहार

Department of Computer Science

University of TBD

abhinavmadahar@tbd.edu

Contents

Land Acknowledgement	iv
Dedication	v
Chapter I. Introduction	1
Chapter II. History	3
1. The History of Progress	4
2. The History of Conceptualizations of Change and Progress	5
3. The History of Decelerationism	6
4. The History of Scientific Research and Progress in Artificial Intelligence	7
5. The History of Conceptualizations of Artificial Intelligence	8
6. Historicizing Decelerationism in Artificial Intelligence	9
Chapter III. What Are Intelligence and Consciousness?	11
7. What Is Intelligence?	12
8. What Is Consciousness?	13
9. What Is the Relationship between Intelligence and Language?	14
10. What Is the Relationship between Intelligence and Consciousness?	15
11. Does There Exist a Meaningful Notion of Superintelligence?	16
12. Is There a Qualitative or Merely a Quantitative Difference between Artificial General Intelligence and Artificial General Superintelligence?	17
Chapter IV. Advancing Reasoning in Artificial Intelligence	19
13. Chain-of-Thought Reasoning	20
13.1. Understanding the Mechanism through which Chain-of-Thought Reasoning Works	20
13.2. Improving Chain-of-Thought Reasoning using Reinforcement Learning from Human Feedback	20
Chapter V. How Advancing Reasoning in Artificial Intelligence Can Achieve Artificial General Superintelligence	21
Chapter VI. Ethical Matters	23
14. Ethical Downsides	24
14.1. Job Losses	24
14.2. Malicious Use of Artificial General Superintelligence	24
14.3. Misalignment of Artificial General Superintelligence	24
15. Ethical Upsides	25
15.1. Increases in Total Factor Productivity	25
15.2. Answering Scientific and Academic Questions which Humans Struggle to Answer	25

15.3. Post-scarcity	25
16. Effective Accelerationism	26
17. Alignment	27
Chapter VII. Conclusion	29

(Land Acknowledgement based on the university which I attend)

Dedicated to the giants on whose shoulders I stand.

CHAPTER I

Introduction

(Placeholder for epigraph)

(Placeholder for epigraph) Author 1

Author 2

CHAPTER II

History

1. The History of Progress

2. The History of Conceptualizations of Change and Progress

3. The History of Decelerationism

4. The History of Scientific Research and Progress in Artificial Intelligence

5. The History of Conceptualizations of Artificial Intelligence

6. Historicizing Decelerationism in Artificial Intelligence

CHAPTER III

What Are Intelligence and Consciousness?

7. What Is Intelligence?

8. What Is Consciousness?

9. What Is the Relationship between Intelligence and Language?

10. What Is the Relationship between Intelligence and Consciousness?

11. Does There Exist a Meaningful Notion of Superintelligence?

12. IS THERE A QUALITATIVE OR MERELY A QUANTITATIVE DIFFERENCE BETWEEN ARTIFICIAL GENERAL INTELLIGENCE AND ARTIFICIAL GENERAL SUPERINTELLIGENCE?

**12. Is There a Qualitative or Merely a Quantitative Difference between
Artificial General Intelligence and Artificial General Superintelligence?**

CHAPTER IV

Advancing Reasoning in Artificial Intelligence

13. Chain-of-Thought Reasoning

13.1. Understanding the Mechanism through which Chain-of-Thought Reasoning Works.

13.2. Improving Chain-of-Thought Reasoning using Reinforcement Learning from Human Feedback.

CHAPTER V

How Advancing Reasoning in Artificial Intelligence Can Achieve Artificial General Superintelligence

CHAPTER VI

Ethical Matters

14. Ethical Downsides

14.1. Job Losses.

14.2. Malicious Use of Artificial General Superintelligence.

14.3. Misalignment of Artificial General Superintelligence.

15. Ethical Upsides

15.1. Increases in Total Factor Productivity.

15.2. Answering Scientific and Academic Questions which Humans Struggle to Answer.

15.3. Post-scarcity.

16. Effective Accelerationism

17. Alignment

CHAPTER VII

Conclusion