Radboud University Nijmegen

Faculty of Social Science

A title

A project for the ages

Thesis MSc Artificial Intelligence

Student Information

Surname:	Hommelsheim
First name:	Ron
Student number:	s1000522
E-mail address:	ron.hommelsheim@ru.nl
Course code:	SOW-MKI94 (45 EC)
AI Specialisation:	Cognitive Computing

Supervisor Information

Role:	Supervisor 1
Surname:	Thill
First name:	Serge
Institute:	Radboud University
E-mail address:	serge.thill@donders.ru.nl
Supervision type:	Internal

Role:	Supervisor 2
Surname:	
First name:	
Institute:	Radboud University
E-mail address:	
Supervision type:	Internal

Abstract

Contents

1 Introduction **3**
1.1 Problem Statement . . . 5

2 Computation and the Computational Mind **6**

3 Language of Thought **6**

4 Grammar **6**

5 Syntax vs Semantics **6**

6 Program Synthesis **6**

7 DreamCoder **7**

8 DeepSynth **7**

9 GFlowNet **8**
9.1 Previous work on program synthesis. . 8

10 Search space **8**

11 Program space **8**

12 GFlowNet **8**

13 assumptions **8**

14 limitations **8**

15 biological plausibility **8**
15.1 Purpose of the Study . . . 9
15.2 Relevance of the Study . . . 9

16 Theoretical Framework **9**
16.1 Definitions . . . 9
16.2 The Need for Discernment in Cognitive Agents . . . 9
16.3 The Concept of Identity and Attribution of Cause and Effect, in relation to the self . . . 9
16.3.1 Karl Friston's perspective . . . 9
16.3.2 Michael Levin's perspective . . 9
16.4 The Necessity of World Modeling . . . 9
16.5 Alternative Approaches to Cognition . 9
16.5.1 Enactivism . . . 9
16.5.2 Embodiment . . . 9
16.6 Analogy as the Core of Cognition: Douglas Hofstadter's Ideas . . . 10
16.7 Virtualism and the "Game Engine in the Head": Perspectives of Joscha Bach and Josh Tenenbaum . . . 10

17 The Art and Science of Model-making **10**
17.1 Symbolic vs. Sub-symbolic Approaches 10
17.2 Analytic vs. Geometric Solutions . . . 10
17.3 System 1 & 2: Daniel Kahneman's theories and relation to Models . . . 10
17.4 Hemispheric Lateralization: Iain McGilchrist's Theory . . . 10

18 Spaces in Cognition **10**
18.1 Latent Spaces . . . 10
18.2 Cognitive and Conceptual Spaces . . . 10
18.2.1 Peter Gärdenfors's Perspective 10
18.3 The Geometry of Cognitive Spaces . . 10
18.3.1 Hyperbolic Spaces . . . 10

19 Computationalism and Computational Theories of the Universe **10**
19.1 Computational Mind . . . 10
19.2 Theories of Stephen Wolfram and David Deutsch . . . 10
19.3 Discussion on the Nature of Reality: Continuous vs. Discrete . . . 10

20 The Language of Thought **10**
20.1 Gödel's Incompleteness Theorem and Turing's Extensions . . . 10
20.2 Chomsky's Hierarchy . . . 10
20.3 Computational Irreducibility . . . 10
20.4 Constructivism . . . 10
20.5 Syntax vs. Semantics in Language . . 10
20.6 Semiotics . . . 10
20.7 The Notion of a Semantic Grammar . 10

21 The Nature of Meaning **10**
21.1 Meaning in the context of Modeling Approaches . . . 10

21.2 The Grounding Problem: Harnard's perspective 10
21.3 Concept of Meaningfulness 10

22 The Formation and Representation of Concepts 10
22.1 Approaches to Representation 10
22.2 The Theory of Memes: Richard Dawkins's perspective 10
22.3 Concepts as a Self-Organizing System 10
22.4 Connection to Cellular Automata . . . 10

23 Representation of Concepts as Probabilistic Programs 10
23.1 Overview of Probabilistic Programs . . 10
23.2 Role of Program Synthesis 10

24 Conclusion 10
24.1 Summary of Findings 10
24.2 Implications for the Field 10
24.3 Recommendations for Future Research 10
24.4 Background 11
24.4.1 Research Question 11
24.4.2 Motivation 11
24.4.3 Impact and Importance 11
24.4.4 Methods 11
24.4.5 Design 11
24.4.6 Analysis 11

25 Scientific and Societal Relevance 11

26 Conclusion 11

27 Discussion 11

1 Introduction

The human mind, in its complexity and adaptability, has long been a subject of fascination and inquiry for scientists and philosophers alike. The concept of 'self', a unique and individual consciousness that arises from cognitive processes, is a cornerstone of this exploration. The emergence of the 'self' is a phenomenon that remains elusive, despite the significant strides made in cognitive science and neuroscience. This thesis aims to delve into the intricacies of cognitive processes, specifically focusing on model-based and model-free approaches, and their potential role in the emergence of the 'self'.

Cognitive processes are the mechanisms through which we perceive, think, remember, and understand the world around us. Two primary paradigms have been proposed to explain these processes: model-based and model-free approaches. Model-based cognition posits that our minds construct internal models of the world, which we use to predict and interpret future events. This approach suggests a dynamic and adaptable cognitive process, where our understanding of the world is constantly updated based on new experiences and information.

On the other hand, model-free cognition proposes that our responses to the world are based on learned associations and habits, rather than internal models. This approach suggests a more static cognitive process, where behavior is driven by reinforcement learning and the repetition of actions that have previously led to rewarding outcomes.

The dichotomy between these two approaches raises intriguing questions about the nature of the 'self'. Does the 'self' emerge from the dynamic predictions and interpretations of the model-based approach, or does it arise from the static associations and habits of the model-free approach? Or perhaps, is the 'self' a product of an intricate interplay between these two cognitive processes?

This thesis will explore these questions, drawing on a wide range of research from cognitive science, neuroscience, psychology, and philosophy. It will critically analyze the strengths and weaknesses of both model-based and model-free approaches, and their implications for our understanding of the 'self'. Furthermore, it will investigate how these approaches can be integrated into a comprehensive framework for understanding the emergence of the 'self'.

The exploration of the 'self' through the lens of model-based and model-free cognition offers a novel perspective on one of the most profound questions of human existence. By unraveling the cognitive processes that underpin our sense of 'self', we can gain a deeper understanding of our minds, our behavior, and our place in the world. This thesis aims to contribute to this understanding, and to stimulate further research and discussion in this fascinating field.

- human intelligence seems to find the essence of situations. we are able to quickly condense information and extrapolate - is working memory like LLMs range?

The impressive capabilities of LLMs are undeniable, however they have limitations as well as differences to how humans understand the world.

The aim of the thesis is to investigate possible approaches of cognition and human thought. What's so cool about human thought? We are able to imagine, plan, reason, consider counterfactuals, induction, abduction, deduction, causal relationships, etc.

1.1 Problem Statement

2 Computation and the Computational Mind

3 Language of Thought

4 Grammar

- GOFAI - Connectionism - Hybrid

5 Syntax vs Semantics

Let's assume the premise that the mind is computational, in one form or another. and we are building a world model which can be thought of as probabilistic programs.

6 Program Synthesis

Probabilistic programs, in essence, represent a form of causal reasoning. By representing beliefs as probability distributions and reasoning patterns as programmatic structures, they offer a nuanced way of modeling complex real-world systems. Given the inherently uncertain nature of our environment and the myriad possible interpretations of sensory data, a probabilistic approach is naturally aligned with the cognitive demands of building accurate and adaptable world models.

3. Program Synthesis: A Key to Unlocking Cognition

Program synthesis refers to the automatic generation of programs from a higher-level specification. In the context of cognition, it implies a process where an agent—human or machine—generates a novel programmatic structure to represent or reason about its environment. Understanding this synthesis process becomes crucial for several reasons:

* *Generativity and Flexibility:* Human cognition is not merely reactive. It proactively creates, hypothesizes, and experiments. By studying how programs are synthesized, we might glean insights into the generative aspects of thought.

* *Abstraction:* At the heart of program synthesis is the ability to abstract away from particulars and generate general rules or patterns. This mirrors cognitive abilities like generalization and analogy-making.

* *Efficiency:* Just as efficient algorithms are prized in computing, efficient cognitive strategies are essential for survival. Program synthesis might hold the key to understanding how humans prune vast possibility spaces to arrive at functional solutions quickly.

Certainly! Here's the problem statement with LaTeX formatting for mathematical expressions:

—

Problem Statement: Navigating the Vast Program Search Space Within Context-Free Grammars

—

1. Background and Introduction

The field of programming has always been underpinned by the intricacies of formal grammars. Context-Free Grammars (CFGs), a subset of formal grammar, are essential in defining the syntactical structures of many programming languages. However, given the generative nature of CFGs, the potential program space defined by even a modestly complex grammar can be immensely vast. Searching for a specific program within this space, or ensuring that a particular space is sufficiently explored, poses significant computational challenges.

2. Problem Definition

2.1 CFG and Program Space

Let $G = (N, \Sigma, P, S)$ be a Context-Free Grammar, where:

- N is a finite set of non-terminal symbols.
- Σ is a finite set of terminal symbols with $N \cap \Sigma = \emptyset$
- P is a finite set of production rules, where each rule is of the form $N \to (N \cup \Sigma)^*$
- S is the start symbol, with $S \in N$

Given such a CFG, the derived program space $\Pi(G)$ is the set of all possible strings (or sequences of symbols) derivable from S.

2.2 Problem Statement

Given a Context-Free Grammar G and a defined objective function f that maps any program $p \in \Pi(G)$ to a real value representing its desirability or fitness:

Find p^* such that:

$$p^* = \arg\max_{p \in \Pi(G)} f(p)$$

In other words, the problem is to locate a program p^* within the vast program space $\Pi(G)$ defined by G that maximizes (or, alternatively, minimizes) the objective function f.

3. Challenges and Complications

3.1 Size of the Search Space

The generative capacity of CFGs means that even grammars of moderate complexity can define immensely vast program spaces. The sheer size of these spaces poses computational and search challenges.

3.2 Non-Linearity and Discontinuities

The mapping between programs and their fitness as defined by f might be non-linear with multiple local maxima, making search strategies based on gradient ascent or other linear heuristics suboptimal.

3.3 Generalization vs Specialization

While CFGs provide a generalized representation of possible programs, the objective function might lead to highly specialized solutions. Balancing between the two is non-trivial.

3.4 Syntactic vs Semantic Validity

A CFG ensures syntactic validity but does not guarantee semantic correctness. Ensuring that a program derived from a CFG is semantically meaningful or error-free in a given context is an additional layer of complexity.

4. Significance

Solving or even approximating solutions for this problem has far-reaching consequences. It touches on fields from program synthesis, where specific algorithms or code snippets are automatically generated to meet specific requirements, to genetic programming, where evolutionary methods are employed to 'evolve' optimal or near-optimal solutions.

Additionally, insights from this exploration can impact compiler design, optimization strategies in high-performance computing, and even areas like natural language processing, where CFGs have historically played a foundational role.

5. Conclusion

The exploration of the vast program space generated by CFGs and efficiently searching within it for optimal programs represents a formidable challenge. It is a nexus of computational theory, practical programming, and numerous applied domains. Addressing it promises not just solutions to specific computational problems but also deeper insights into the nature of computation, representation, and optimization.

4. Techniques in Program Synthesis

Several techniques have emerged as pivotal in the domain of program synthesis:

* *Deductive Synthesis:* Rooted in formal logic, this method transforms specifications into programs. The use of logic mirrors certain cognitive tasks, especially those demanding strict reasoning.

* *Stochastic Search:* By randomly exploring the space of possible programs, these methods mirror heuristic-based cognitive processes. Genetic algorithms, for instance, mimic evolutionary processes to evolve optimal or near-optimal solutions.

* *Neural Program Synthesis:* Neural networks, especially recurrent ones, have shown promise in generating programmatic structures. The parallels between neural networks and neural structures in the brain offer tantalizing possibilities for cognitive science.

* *Example-Based Synthesis:* Drawing inspiration from how humans often learn—from examples—these methods generate programs by generalizing from provided instances. This mirrors pedagogical processes and experiential learning.

7 DreamCoder

A Bayesian View –¿ What exactly happens?

8 DeepSynth

–¿ Introducing different search strategies.

9 GFlowNet

A generative policy

(Probabilistic) Context free grammar.

9.1 Previous work on program synthesis.

Check out microsoft paper. DreamCoder DeepSynth

10 Search space

11 Program space

- How can we calculate the size of that space? - Hierarchical structure - Abstract Syntax Trees

12 GFlowNet

-

In a probabilistic programming representation of world models and thoughts, as a type of Language of Thought, which aligns with constructivism we can conceive of the problem statement as We want to construct objects.

In a language of thought, regardless of most details, we tend to think in a paradigm in which thoughts are compositional. In one way or another, thoughts are hierarchical. Both in ontologies, i.e. the way concepts are structured, (animal -¿ bird -¿ fink) but also in the sequential nature of thought construction (as in natural language, reasoning tasks, etc.). We are creating parse trees, or abstract syntax trees.

In a probabilistic programming paradigm, we view this as compositional functions, with the many properties of functional programming analogous to currying, etc.

[Thoughts as trajectories]

We can formalise this as a hierarchical latent structure Z.

Framing the problem.

Combinatorial search problem.

- If we are using encoder + decoder, are we violating the Markovian Flow assumption? Look at the smiley example. If the NN would get [[left brow], [left brow, right brow], [left brow, right brow, smile]] as input, i.e. the sequence of the states, it would violate the assumption. but it only gets the current state, e.g. [left brow, right brow], and from that it has to infer the next step. When using a decoder, we would indeed give it the whole trajectory of states, so it would violate the assumption. But even now i am encoding the sequence and giving the whole trajectory as input. and since the CFG is essentially a tree, there is only one parent for each state.

It would probably be faster to do it bottom up like in HEAP search from Nathanael and from GFN-EM, then we could also use sub-trajectory balance, i.e. calculate intermediate Rewards.

Another thing we could do is predict a bunch of terminals at once, and then combine in each step.

Other ideas: wave function collapse

13 assumptions

14 limitations

15 biological plausibility

15.1 Purpose of the Study

15.2 Relevance of the Study

16 Theoretical Framework

16.1 Definitions

- What is a cognitive agent?
- What is the mind?
- What is the self?
- What is the "I"?
- What is a thought? (GFN trajectory, GPT trajectory, trajectory of some sort)

16.2 The Need for Discernment in Cognitive Agents

The concept of discernment in cognitive agents is fundamental to understanding cognition. In essence, discernment is the ability of an agent to differentiate and distinguish between various elements within its environment. Without discernment, an agent would be unable to make sense of its world or make decisions to optimize its performance within that world. Discernment, therefore, lies at the foundation of cognition, allowing the agent to separate itself from its surroundings and engage meaningfully with the world.

16.3 The Concept of Identity and Attribution of Cause and Effect, in relation to the self

16.3.1 Karl Friston's perspective

16.3.2 Michael Levin's perspective

16.4 The Necessity of World Modeling

Representationalism Representationalism suggests that cognitive agents create internal representations or models of the world to understand and interact effectively with their environment. These models, while never fully capturing the complexity of the external world, provide a means for the agent to anticipate and respond to future scenarios.

Contrast to Statistical Pattern Recognition Unlike statistical pattern recognition, which involves extracting patterns and regularities from data without necessarily understanding the underlying structures, model-based reasoning involves creating and manipulating symbolic models of the world. The advantage of this approach is that it allows for a more comprehensive understanding of the world and more flexible and adaptive behavior.

16.5 Alternative Approaches to Cognition

16.5.1 Enactivism

Enactivism posits that cognition arises from the dynamic interaction between an organism and its environment. Rather than relying on internal representations, cognition is seen as a process of embodied action, where knowledge is gained through physical interaction with the world.

16.5.2 Embodiment

The theory of embodiment emphasizes the role of the body in shaping the mind. According to this view, many features of cognition are shaped by aspects of the body beyond the brain, including sensory and motor systems. This perspective encourages a holistic view of cognition, bridging the mind-body divide.

16.6 Analogy as the Core of Cognition: Douglas Hofstadter's Ideas

16.7 Virtualism and the "Game Engine in the Head": Perspectives of Joscha Bach and Josh Tenenbaum

17 The Art and Science of Model-making

17.1 Symbolic vs. Sub-symbolic Approaches

17.2 Analytic vs. Geometric Solutions

17.3 System 1 & 2: Daniel Kahneman's theories and relation to Models

17.4 Hemispheric Lateralization: Iain McGilchrist's Theory

18 Spaces in Cognition

18.1 Latent Spaces

18.2 Cognitive and Conceptual Spaces

18.2.1 Peter Gärdenfors's Perspective

18.3 The Geometry of Cognitive Spaces

18.3.1 Hyperbolic Spaces

19 Computationalism and Computational Theories of the Universe

19.1 Computational Mind

19.2 Theories of Stephen Wolfram and David Deutsch

19.3 Discussion on the Nature of Reality: Continuous vs. Discrete

20 The Language of Thought

20.1 Gödel's Incompleteness Theorem and Turing's Extensions

20.2 Chomsky's Hierarchy

Current AI is focused on deep learning. difference in approaches: - statistical pattern recognition - models of the world

- we need very little data, - solve et coagula, we have concepts which we can use to dream, imagine, etc.

- who am i?

24.4 Background

- Problem statement - program synthesis, previously . . . - DreamCoder, DeepCoder, DeepSynth, etc. CFG, PCFG, etc.

- They are predicting weights for a PCFG.

24.4.1 Research Question

24.4.2 Motivation

- Modeling human thought

Test

- problem statement

24.4.3 Impact and Importance

- Understanding/ modeling cognition. - Better AI models, OOD generalization, reasoning, ...

24.4.4 Methods

- GFlowNet: definition, flow matching constraint,

24.4.5 Design

24.4.6 Analysis

25 Scientific and Societal Relevance

26 Conclusion

27 Discussion

References