My path to AGI

Tom Rochette <tom.rochette@coreteks.org>

August 30, 2025 — 861fb9d0

0.1 Learned in this study

0.2 Things to explore

1 Preface

If you are interested in a more digestible path through the same content I've experienced, I suggest you look up my suggested path to AGI article. It basically contains most of the stuff I've gone through myself, but it is organized in a constructive manner to a new comer (at least, in my opinion).

If you have comments and suggestions, feel free to let me know through the comments!

Note: The dates in parenthesis indicate when I started and finished reading a given book/paper, not their publication date.

2 Overview

2.1 2014

2.1.1 Artificial Intelligence: A Modern Approach (2014-01)

Stuart Russell, Peter Norvig

Only reading the content without doing any exercises is about 60-65h of reading.

2.1.2 Augmenting Human Intellect (2014-02)

Douglas Engelbart

2.1.3 Playing Atari with Deep Reinforcement Learning (2014-02)

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, Martin Riedmiller

2.1.4 Genifer – an artificial general intelligence (2014-03 - ?)

YKY

2.1.5 How to Create a Mind: The Secret of Human Thought Revealed (2014-03)

Ray Kurzweil

2.1.6 Coursera - Machine Learning (2014-03)

Andrew Ng

2.2 2015

2.2.1 The Essential Turing (2015-06 - ?)

Alan M. Turing, B. Jack Copeland

- On Computable Numbers, with an Application to the Entscheidungsproblem (2015-06 - ?)
- Intelligent Machinery (2015-10)
- Computing Machinery and Intelligence (2015-10)
- Can Digital Computer Think? (2015-10)
- Solvable and Unsolvable Problems (2015-10)
- Systems of Logic Based on Ordinals (2016-01 - ?)
- Can Automatic Calculating Machines Be Said to Think?
- Lecture on the Automated Computing Engine
- The Chemical Basis of Morphogenesis
- Chess

2.2.2 The First Level of Super Mario Bros. is Easy with Lexicographic Orderings and Time Travel (2015-06)

Tom Murphy VII

2.2.3 Naive set theory (2015-08 - ?)

Paul R. Halmos

2.2.4 A Mathematical Theory of Communication (2015-08 - 2016-05)

Claude E. Shannon

2.2.5 A Model for Recursively Self Improving Programs (2015-09)

Matt Mahoney

2.2.6 Universal Artificial Intelligence (2015-09 - ?)

Marcus Hutter

2.2.7 A Collection of Definitions of Intelligence (2015-09)

Shane Legg

2.2.8 Gödel, Escher, Bach: an Eternal Golden Braid (2015-09 - ?)

Douglas R. Hofstadter

2.2.9 The Society of Mind (2015-11 - 2016-02)

Marvin Minsky

2.2.10 The Computer and the Brain (2015-12)

John von Neumann

2.3 2016

- Tensorflow 2016-01

2.3.1 PHP-Brain (2016-02 - ?)

See PHP-Brain

For about two weeks, I spent my time trying to find ways to do efficient string matching and document learning. The reason I wanted to do this was to attempt to learn language through examples. Provided with enough (hopefully valid) documents, it would be possible to assess if a given sentence is valid by matching it with previously seen sentences. It would also allow me to feed it a bunch of my chatlogs so that it may learn my vocabulary behavior and potentially predict what I would type next (basically autocompletion).

With a program that could do not much more than receive texts and match them against previously seen documents, I wanted to figure out a way to match generic data instead of words/sentences.

However, it wasn't only about matching generic data. It was also about being able to match data of different types with one another. In other words, if an image contained a sentence within one of its region, I'd like to be able to reuse that region to store the sentence as well as this region of the image. Thus, using only one run of the data I have two different informations: a sentence and a part of an image.

Furthermore, the idea is also to be able to relate concepts of varying degree of abstraction with one another.

In other to get an idea of how this could be accomplished, I started to reflect on the architecture of the human body and how the brain processes signal/information.

I began by designing a model-based brain, in other words, something that would look like a brain but from the inside would be something different. The main idea was to have a loop that would run forever, which would select tasks to execute. Input from the environment as well as actions to undertake (output) were considered to be tasks. The brain would have to process an input task, which would consist of a stream of data to be parsed and looked for patterns. An output task would be the emission of a stream of data (a stored pattern we expect to produce the desired result) to a target actuator. In other words, the brain would have to learn how to use its outputs through pattern emission, like it would try to learn to decipher the input streams it would receive.

After playing with the idea for a while, I came across an issue: what should I do when a task may take several seconds, even maybe minutes or hours? It doesn't make sense to stop processing everything else until said task is complete. That is not a real time approach. Thus, it meant we'd have to implement some sort of system to do task preemption. PHP being a single threaded language, this would prove to be more complicated than I wanted to deal with. Thus, I decided to move to a language that would support multithreading, C#.

2.3.2 Sharp-Brain (2016-03 - ?)

See Sharp-Brain

The first thing I did when moving to C# was to build a complete body in an object oriented fashion. Body, head, brain, eyes, ears, nerves and so on. The head had eyes, which were attached to their own optic nerve, which would then connect to the brain. The brain was connected to a set of nerves, such as the optic nerve, which could be used either to receive or transmit data/signal.

Since I was now using a multithreaded language, I made each input a thread of its own. Thus, there would be 2 threads for the eyes, one per eye, 2 threads for the ears, one per ear, a thread for the brain, and so on. This would be the model of a human.

However, it was possible to generalize the idea to a robot. If we replaced the eyes with a single webcam, then the architecture would be similar in the sense that the eye, now a webcam, would take a frame and transmit it over the optic nerve to the brain. The brain would receive this array of bytes and process it (whatever that would mean, at this point no processing was done within the brain other than receive the data).

The problem at this point though was that the brain would now build its own network of understanding/patterns, which would make it more opaque and difficult to understand compared to a system of tasks.

2.3.3 Artificial General Intelligence (2016-02 - ?)

Ben Goertzel, Cassio Pennachin

2.3.4 The Logic of Intelligence (2016-02)

Pei Wang

2.3.5 The Novamente Artificial Intelligence Engine (2016-02)

Ben Goertzel

2.3.6 Essentials of General Intelligence: The Direct Path to Artificial General Intelligence (2016-03)

Peter Voss

2.3.7 Artificial Brains (2016-03)

Hugo de Garis

2.3.8 The OpenNARS implementation of the Non-Axiomatic Reasoning System (2016-03)

Patrick Hammer, Tony Lofthouse, Pei Wang

2.3.9 The New AI: General & Sound & Relevant for Physics (2016-03)

Jürgen Schmidhuber

2.3.10 Gödel Machines: Fully Self-Referential Optimal Universal Self-improvers (2016-03)

Jürgen Schmidhuber

2.3.11 Playing Atari with Deep Reinforcement Learning (2016-03)

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, Martin Riedmiller

2.3.12 Mastering the Game of Go with Deep Neural Networks and Tree Search (2016-03)

David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, Demis Hassabis

2.3.13 Program Search as a Path to Artificial General Intelligence (2016-03)

Łukasz Kaiser

2.3.14 The Natural Way to Artificial Intelligence (2016-04)

Vladimir G. Red'ko

2.3.15 Neural Turing Machines (2016-04)

Alex Graves, Greg Wayne, Ivo Danihelka

2.3.16 3D Simulation: the Key to A.I. (2016-04)

Keith A. Hoyes

2.3.17 Cognitive Science - An Introduction to the Science of the Mind (2016-04 - 2016-07)

José Luis Bermúdez

2.3.18 Building Machines That Learn and Think Like People (2016-04 - 2016-08)

Breden Lake

2.3.19 EURISKO: A Program That Learns New Heuristics and Domain Concepts (2016-04)

Douglas Lenat

2.3.20 Why AM and Eurisko Appear to Work (2016-05)

Douglas Lenat

2.3.21 AM: An Artificial Intelligence Approach to Discovery in Mathematics as Heuristic Search (2016-05 - 2016-06)

Douglas Lenat

2.3.22 Discovery Systems: From AM to CYRANO (2016-05)

Ken Haase

2.3.23 AM: A Case Study in AI Methodology (2016-05)

G.D. Ritchie, F.K. Hanna

2.3.24 Report on a General Problem-Solving Program (2016-05)

Allen Newell, John Clifford Shaw, Herbert A. Simon

2.3.25 The Logic Theory Machine: A Complex Information Processing System (2016-05)

Allen Newell, Herbert A. Simon

2.3.26 The Copycat Project: A Model of Mental Fluidity and Analogy-Making (2016-05 - ?)

Douglas Hofstadter, Melanie Mitchell

2.3.27 Minds, Machines and Gödel (2016-05)

J.R. Lucas

2.3.28 Automated programming (2016-06 - ?)

Began research related to automated programming. The main goal is to determine what is doable at the moment (based on current knowledge of the field).

2.3.29 Automated refactoring (2016-06 - ?)

Began research related to automated refactoring. The main goal is to determine a list of refactoring tools that can be developed to simplify the job of refactoring of programmers.

2.3.30 Approaches to Automatic Programming (2016-06)

Alan Biermann

2.3.31 Duplicated Code Refactoring Advisor (DCRA) (2016-06)

Francesco Zanoni

2.3.32 How to Solve It: A New Aspect of Mathematical Method (2016-06 - ?)

George Pólya

2.3.33 The heuristic of George Polya and its relation to artificial intelligence (2016-06)

Allen Newell

2.3.34 Neuroscience: Exploring the Brain (2016-06 - 2016-11)

Mark F. Bear, Barry W. Connors, Michael A. Paradiso

2.3.35 A Concise Introduction to Logic (2016-06 - 2016-12)

Patrick J. Hurley

2.3.36 A Hypothetical Dialogue Exhibiting a Knowledge Base For a Program-Understanding System (2016-06 - 2016-07)

Cordell Green, David Barstow

2.3.37 Machine Learning (2016-06 - (2016-09))

Peter Flach

2.3.38 On Intelligence (2016-06 - 2016-08)

Jeff Hawkins

2.3.39 Programs with common sense (2016-06)

John McCarthy

2.3.40 Universal Value Function Approximators (2016-06 - 2016-07)

Tom Schaul, Dan Horgan, Karol Gregor, David Silver

2.3.41 The Programmer's Apprentice: A Research Overview (2016-07)

Charles Rich, Richard C. Waters

2.3.42 Towards a Programming Apprentice (2016-07)

Carl E. Hewitt, Brian Smith

2.3.43 Artificial Intelligence – a personal view (2016-07)

David Marr

2.3.44 Automated research (2016-08 - ?)

Began research related to automated research. The main goal is to determine a basic research flow and automate most of the process (researching papers, establishing important papers, determine a reading order, create a wordcloud of the most important terms, define which papers cite which papers, etc.).

2.3.45 Introduction to Automata Theory, Languages, and Computation (2016-08 - 2016-10)

John E. Hopcroft, Rajeev Motwani, Jeffrey D. Ullman

2.3.46 Theory of self-reproducing automata (2016-08 - 2016-10)

John von Neumann

2.3.47 From Seed AI to Technological Singularity via Recursively Self-Improving Software (2016-08)

Roman V. Yampolskiy

2.3.48 Self-improving AI: an Analysis (2016-08)

John Storrs Hall

2.3.49 Engineering Utopia (2016-08)

John Storrs Hall

2.3.50 Toward a Standard Metric of Machine Intelligence (2016-08)

Richard Yonck

2.3.51 From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0 (2016-08)

Masafumi Oizumi, Larissa Albantakis, Giulio Tononi

2.3.52 Automatic Learning of Context-Free Grammar (2016-08)

Tai-Hung Chen, Chun-Han Tseng, Chia-Ping Chen

2.3.53 TensorFlow for Machine Intelligence (2016-09)

Sam Abrahams, Danijar Hafner, Erik Erwitt, Ariel Scarpinelli

2.3.54 Language Identification in the Limit (2016-09)

E. Mark Gold

2.3.55 How to Invent Functions (2016-09)

J. Roland Olsson

2.3.56 Can Intelligence Explode? (2016-09)

Marcus Hutter

2.3.57 A Complete Theory of Everything (will be subjective) (2016-09)

Marcus Hutter

2.3.58 Automated test generator (2016-09 - ?)

Began research related to automated test generator. The main goal of this project is to use the knowledge I've acquired thinking about automated programming and refactoring and attempt to apply it to test generation. One of the goals is to develop a tool which will automate most of the process of testing PHP OOP code, specifically MVC-oriented websites using frameworks such as Laravel or Symfony.

2.3.59 Static analysis (2016-09 - ?)

My research into static analysis is related to my research on automated programming, automated refactoring and automated test generator. The main goal is to produce a piece of software that will be able to reason as a programmer would reason about a program in his head.

2.3.60 Automating Software Testing Using Program Analysis (2016-10)

Patrice Godefroid, Peli de Halleux, Aditya Nori, Sriram Rajamani, Wolfram Schulte, Nikolai Tillmann, Michael Y. Levin

2.3.61 Software Testing: A Craftsman's Approach (2016-10 - 2016-11)

Paul Jorgensen

2.3.62 Advanced Symbolic Analysis for Compilers (2016-10 - 2016-11)

Thomas Fahringer, Bernhard Scholz

2.3.63 Principles of Program Analysis (2016-10 - 2016-12)

Flemming Nielson, Hanne R. Nielson, Chris Hankin

2.4 2017

2.4.1 Programming the Semantic Web (2017-01)

Toby Segaran, Colin Evans, Jamie Taylor

2.4.2 Deep Learning (2017-02 - 2017-05)

Ian Goodfellow, Yoshua Bengio, Aaron Courville

2.4.3 Reinforcement Learning (2017-02 - 2017-03)

David Silver

2.4.4 A Framework for Searching for General Artificial Intelligence (2017-02)

Marek Rosa, Jan Feyereisl, The GoodAI Collective

2.4.5 Knowledge base (2017-02)

I wrote a polymer/javascript-based application which dealt with triples in order to manage knowledge. Everything is considered to be a node and nodes can be associated to one another through anonymous.

See Knowledge base.

2.4.6 Curriculum Learning (2017-04)

Yoshua Bengio, Jérôme Louradour, Ronan Collobert, Jason Weston

2.4.7 A Postulate on the Brain's Basic Wiring Logic (2017-04)

Joe Z. Tsien

2.4.8 Brain Computation Is Organized via Power-of-Two-Based Permutation Logic (2017-04)

Kun Xie, Grace E. Fox, Jun Liu, Cheng Lyu, Jason C. Lee, Hui Kuang, Stephanie Jacobs, Meng Li, Tianming Liu, Sen Song, Joe Z. Tsien

2.4.9 How to Make a Complete Map of Every Thought (2017-04)

Lion Kimbro

2.4.10 Induction of Decision Trees (2017-04)

J.R. Quinlan

2.4.11 Handwriting recognition (2017-04 - ?)

Began experimenting with the problem of handwriting recognition. The main goal is to be able to feed a page through a camera/webcam/scanner and have its text content extracted.

2.4.12 Traitement automatique des langues (2017-04)

Hugo Larochelle

2.4.13 RobustFill: Neural Program Learning under Noisy I/O (2017-04)

Jacob Devlin, Jonathan Uesato, Surya Bhupatiraju, Rishabh Singh, Abdel-rahman Mohamed, Pushmeet Kohli

2.4.14 Generative Adversarial Nets (2017-04)

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio

2.4.15 Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks (2017-04)

Alex Graves, Jürgen Schmidhuber

2.4.16 Auto-Encoding Variational Bayes (2017-04)

Diederik P. Kingma, Max Welling

2.4.17 Tacotron: Towards End-To-End Speech Synthesis (2017-05)

Yuxuan Wang, RJ Skerry-Ryan, Daisy Stanton, Yonghui Wu, Ron J. Weiss, Navdeep Jaitly, Zongheng Yang, Ying Xiao, Zhifeng Chen, Samy Bengio, Quoc Le, Yannis Agiomyrgiannakis, Rob Clark, Rif A. Saurous

2.4.18 DeepZip: Lossless Compression using Recurrent Networks (2017-05)

Kedar Tatwawadi

2.4.19 Scan, Attend and Read: End-to-End Handwritten Paragraph Recognition with MDLSTM Attention (2017-05)

Théodore Bluche, Jérôme Louradour, Ronaldo Messina

2.4.20 Neural machine translation by jointly learning to align and translate (2017-05)

Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio

2.4.21 Deepcoder: Learning to Write Programs (2017-05)

Matej Balog, Alexander L. Gaunt, Marc Brockschmidt, Sebastian Nowozin, Daniel Tarlow

2.4.22 Neural Architecture Search With Reinforcement Learning (2017-05)

Barret Zoph, Quoc V. Le

2.4.23 SMARTPASTE: Learning to Adapt Source Code (2017-05)

Miltiadis Allamanis, Marc Brockschmidt

2.4.24 Third-Person Imitation Learning (2017-05)

Bradly C. Stadie, Pieter Abbeel, Ilya Sutskever

2.4.25 Elements of Information Theory (2017-05 - (2017-07))

Thomas M. Cover, Joy A. Thomas

2.4.26 Deep Voice: Real-time Neural Text-to-Speech (2017-06)

Sercan O. Arik, Mike Chrzanowski, Adam Coates, Gregory Diamos, Andrew Gibiansky, Yongguo Kang, Xian Li, John Miller, Andrew Ng, Jonathan Raiman, Shubho Sengupta, Mohammad Shoeybi

2.4.27 Deep Voice 2: Multi-Speaker Neural Text-to-Speech (2017-06)

Sercan Arik, Gregory Diamos, Andrew Gibiansky, John Miller, Kainan Peng, Wei Ping, Jonathan Raiman, Yanqi Zhou

2.4.28 Toward an Architecture for Never-Ending Language Learning (2017-06)

Andrew Carlson, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R. Hruschka, Tom M. Mitchell

2.4.29 Human-level control through deep reinforcement learning (2017-06)

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin A. Riedmiller, Andreas Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, Demis Hassabis

2.4.30 SophoLab; experimental computational philosophy (2017-06 - (2017-07))

Vicent Wiegel

2.4.31 Multi-Dimensional Recurrent Neural Networks (2017-06)

Alex Graves, Santiago Fernandez, Juergen Schmidhuber

2.4.32 Grid Long Short-Term Memory (2017-06)

Nal Kalchbrenner, Ivo Danihelka, Alex Graves

2.4.33 One Model To Learn Them All (2017-06)

Lukasz Kaiser, Aidan N. Gomez, Noam Shazeer, Ashish Vaswani, Niki Parmar, Llion Jones, Jakob Uszkoreit

2.4.34 Schema Networks: Zero-shot Transfer with a Generative Causal Model of Intuitive Physics (2017-06)

Ken Kansky, Tom Silver, David A. Mély, Mohamed Eldawy, Miguel Lázaro-Gredilla, Xinghua Lou, Nimrod Dorfman, Szymon Sidor, Scott Phoenix, Dileep George

2.4.35 The Predictron: End-To-End Learning and Planning (2017-06)

David Silver, Hado van Hasselt, Matteo Hessel, Tom Schaul, Arthur Guez, Tim Harley, Gabriel Dulac-Arnold, David Reichert, Neil Rabinowitz, André Barreto, Thomas Degris

2.4.36 Automated Curriculum Learning for Neural Networks (2017-06)

Alex Graves, Marc G. Bellemare, Jacob Menick, Rémi Munos, Koray Kavukcuoglu

2.4.37 The Strategic Student Approach for Life-Long Exploration and Learning (2017-06)

Manuel Lopes, Pierre-Yves Oudeyer

2.4.38 Attention Is All You Need (2017-07)

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia

2.4.39 Show, Attend and Tell: Neural Image Caption Generation with Visual Attention (2017-07)

Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron C. Courville, Ruslan Salakhutdinov, Richard S. Zemel, Yoshua Bengio

2.4.40 Deep Reinforcement Learning from Human Preferences (2017-07)

Paul F. Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, Dario Amodei

2.4.41 The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World (2017-07 - 2017-08)

Pedro Domingos

2.4.42 Algorithms to Live By: The Computer Science of Human Decisions (2017-07 - 2017-08)

Brian Christian, Tom Griffiths

2.4.43 Programmable Agents (2017-08)

Misha Denil, Sergio Gomez Colmenarejo, Serkan Cabi, David Saxton, Nando de Freitas

2.4.44 Principles of Computational Modelling in Neuroscience (2017-07 - (2017-08))

David Sterratt, Bruce Graham, Andrew Gillies, David Willshaw

2.4.45 How to Build a Brain: A Neural Architecture for Biological Cognition (2017-08 - ?)

Chris Eliasmith

2.4.46 Why Philosophers Should Care About Computational Complexity (2017-08 - 2017-09)

Scott Aaronson

2.4.47 Superintelligence: Paths, Dangers, Strategies (2017-08 - 2017-11)

Nick Bostrom

2.4.48 The Brain: The Story of You (2017-08 - 2017-09)

David Eagleman

2.4.49 The Consciousness Prior (2017-09)

Yoshua Bengio

2.4.50 Learning with Opponent-Learning Awareness (2017-09)

Jakob N. Foerster, Richard Y. Chen, Maruan Al-Shedivat, Shimon Whiteson, Pieter Abbeel, Igor Mordatch

2.4.51 StarCraft II: A New Challenge for Reinforcement Learning (2017-09)

Oriol Vinyals, Timo Ewalds, Sergey Bartunov, Petko Georgiev, Alexander Sasha Vezhnevets, Michelle Yeo, Alireza Makhzani, Heinrich Küttler, John Agapiou, Julian Schrittwieser, John Quan, Stephen Gaffney, Stig Petersen, Karen Simonyan, Tom Schaul, Hado van Hasselt, David Silver, Timothy P. Lillicrap, Kevin Calderone, Paul Keet, Anthony Brunasso, David Lawrence, Anders Ekermo, Jacob Repp, Rodney Tsing

2.4.52 Is the theory of everything merely the ultimate ensemble theory? (2017-09)

Max Tegmark

2.4.53 The Allure of Machinic Life: Cybernetics, Artificial Life, and the New AI (2017-10 - ?)

John Johnston

2.4.54 Developmental Biology (2017-10 - ?)

Scott Gilbert

2.4.55 Mastering the game of Go without human knowledge (2017-10)

David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas R Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy P. Lillicrap, Fan Xin Hui, Laurent Sifre, George van den Driessche, Thore Graepel, Demis Hassabis

2.4.56 Fast and Slow with Deep Learning and Tree Search (2017-10)

Thomas Anthony, Zheng Tian, David Barber

2.4.57 Rainbow: Combining Improvements in Deep Reinforcement Learning (2017-11)

Matteo Hessel, Joseph Modayil, Hado van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Daniel Horgan, Bilal Piot, Mohammad Gheshlaghi Azar, David Silver

2.4.58 Reverse Curriculum Generation for Reinforcement Learning (2017-11)

Carlos Florensa, David Held, Markus Wulfmeier, Pieter Abbeel

2.4.59 The Consciousness Prior (2017-12)

Yoshua Bengio

2.4.60 Deep Voice 3: 2000-Speaker Neural Text-to-Speech (2017-12)

Wei Ping, Kainan Peng, Andrew Gibiansky, Sercan Ömer Arik, Ajay Kannan, Sharan Narang, Jonathan Raiman, John Miller

2.4.61 Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm (2017-12)

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, Demis Hassabis

2.4.62 Bandit Based Monte-Carlo Planning (2017-12)

Levente Kocsis, Csaba Szepesvári

2.4.63 R-MAX - A General Polynomial Time Algorithm for Near-Optimal Reinforcement Learning (2017-12)

Ronen I. Brafman, Moshe Tennenholtz

2.4.64 Life 3.0: Being Human in the Age of Artificial Intelligence (2017-12 - 2018-02)

Max Tegmark

2.5 2018

2.5.1 Speculations concerning the first ultraintelligent machine (2018-01)

Irving John Good

2.5.2 A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting (2018-01)

Yoav Freund, Robert E. Schapire

2.5.3 Voice detection (2017-09 - 2018-01)

I have recorded a ton of audio of myself while I use my computer at home. I wanted to be able to find the slices of time where I speak and say something intelligible in order to extract those regions and then potentially send them to a service such as Google Voice to have it converted into text.

2.5.4 A Trainable Spaced Repetition Model for Language Learning (2018-02)

Burr Settles, Brendan Meeder

2.5.5 When will computer hardware match the human brain? (2018-02)

Hans Moravec

2.5.6 Machine Super Intelligence (2015-09 - 2018-02)

Shane Legg

2.5.7 The free-energy principle: a unified brain theory? (2018-02)

Karl Friston

2.5.8 Precise Condition Synthesis for Program Repair (2018-02)

Yingfei Xiong, Jie Wang, Runfa Yan, Jiachen Zhang, Shi Han, Gang Huang, Lu Zhang

2.5.9 Intrinsically Motivated Learning of Hierarchical Collections of Skills (2018-02)

Satinder P. Singh, Andrew G. Barto, Nuttapong Chentanez

2.5.10 History Driven Program Repair (2018-02)

Xuan-Bach D. Le, David Lo, Claire Le Goues

2.5.11 Machine Theory of Mind (2018-02)

Neil C. Rabinowitz, Frank Perbet, H. Francis Song, Chiyuan Zhang, S. M. Ali Eslami, Matthew M Botvinick

2.5.12 One Big Net For Everything (2018-05 - ?)

Jurgen Schmidhuber

2.5.13 PowerPlay: training an increasingly general problem solver by continually searching for the simplest still unsolvable problem (2018-05 - ?)

Jurgen Schmidhuber