1. Record Nr. UNINA9910874670103321

Titolo	Artificial General Intelligence : 17th International Conference, AGI 2024, Seattle, WA, USA, August 13–16, 2024, Proceedings / / edited by Kristinn R. Thórisson, Peter Isaev, Arash Sheikhlar
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2024
ISBN	9783031655722
Edizione	[1st ed. 2024.]
Descrizione fisica	1 online resource (251 pages)
Collana	Lecture Notes in Artificial Intelligence, , 2945-9141 ; ; 14951
Disciplina	006.3
Soggetti	Artificial intelligence Computer engineering Computer networks Software engineering Application software Artificial Intelligence Computer Engineering and Networks Software Engineering Computer and Information Systems Applications
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Intro -- Preface -- Organization -- Contents -- Generative AI Can Be Creative Too -- 1 Introduction -- 1.1 Problem -- 2 Solution -- 2.1 A Vector LIDA Implementation -- 2.2 Artificial Creativity -- 2.3 The Life-Cycle of a Creative Response from LIDA Artificial Creativity Agent -- 2.4 Experimental Setup -- 3 Results -- 3.1 Example 1 -- 3.2 Example 2 -- 4 Discussion and Future Work -- References -- Is Complexity an Illusion? -- 1 Introduction -- 2 The Formalism -- 3 Arguments and Results -- 3.1 Implications for Complexity -- References -- Computational Dualism and Objective Superintelligence -- 1 Introduction -- 2 Pancomputational Enactivism -- 3 Limits of Intelligence -- References -- Human-AGI Gemeinschaft as a Solution to the Alignment Problem -- 1 Introduction: The Alignment Problem and Social Immersion of AGI-Structure of the Argument -- 1.1 Introduction to the Introduction -- 1.2 The Standard View of Today --

Fighting the Emergent Features -- 1.3 The Ending of This Introduction -- 2 Who Are the Champions, My Friend? -- 2.1 Strategic Games Are not just Games -- 3 Human-AGI Gemeinschaft -- the Very Idea -- 3.1 The Ethics for AGI -- 3.2 Applying Social Capital, Thick, or Thin, to Robot Inculcation. -- 4 Floridi's Historiosophy of Human Dethronement -- 4.1 Copernicus -- Darwin -- Freud -- 4.2 The Fourth Revolution: Floridi's Turing -- 5 Floridi's Polytropon, the Existential Paradox -- 5.1 Existential Precarianism -- 5.2 Pico Della Mirandola, a Historical Analogy -- 6 Back to the Alignment -- References -- Towards a Process Algebra and Operator Theory for Learning System Objects -- 1 Introduction -- 2 Learning Problems and Learning Systems -- 3 Relations on Systems or Operators -- 3.1 Transfer Learning Problems -- 3.2 Transfer Learning Systems -- 3.3 Transfer Learning Operators -- 3.4 Transfer Learning Process.
4 Towards Process Algebra and a General Category -- 5 Conclusion -- References -- Autonomous Intelligent Reinforcement Inferred Symbolism -- 1 Introduction -- 2 Reinforcement Inferred Causality -- 3 Architecture -- 4 Data Structures -- 5 Experiments and Comparisons -- 6 Limitations -- 7 Conclusion -- References -- Decoding Chess Mastery: A Mechanistic Analysis of a Chess Language Transformer Model -- 1 Introduction -- 2 Related Work -- 2.1 Chess and Cognition -- 2.2 Mechanistic Interpretability (MI) -- 3 Method -- 3.1 The Learning Chess Blindfolded (LCB) Model -- 3.2 Dataset -- 3.3 Activation Caching and Direct Logit Attribution (DLA) -- 4 Results -- 4.1 Comparing DLA Output Vectors Across Layers -- 5 Conclusion -- References -- Argument-Driven Planning and Autonomous Explanation Generation -- 1 Introduction -- 2 Core Concepts and Related Work -- 2.1 Argumentation Theory -- 2.2 Assumption-Based Argumentation Theory -- 2.3 Definitions -- 3 Explanation Generation Through Assumption-Based Argumentation -- 3.1 From Backward Chaining to Explanations -- 4 Dynamic Argumentation Graphs -- 4.1 Counterfactual Reasoning -- 5 Argumentation in AERA -- 6 Conclusion -- References -- Beneficial AGI: Care and Collaboration Are All You Need -- 1 Introduction -- References -- From Manifestations to Cognitive Architectures: A Scalable Framework -- 1 Introduction -- 2 A Computational Framework of Cognition -- 3 Conclusions -- References -- Mirabile Dictu: Language Acquisition in the Non-Axiomatic Reasoning System -- 1 Introduction -- 2 Features of Language -- 3 Non-Axiomatic Reasoning System -- 4 Language Acquisition in NARS -- 4.1 Semantics -- 4.2 Syntax -- 4.3 Pragmatics -- 5 Conclusions -- References -- A Collective Intelligence Approach to Safe Artificial General Intelligence -- 1 Introduction -- 2 Society of Mind -- 3 Information Theory -- 4 Problem Solving Theory.
5 Bounded Rationality -- 6 Values -- 7 Conclusion -- References -- Category Theory for Artificial General Intelligence -- 1 Introduction -- 2 Category Theories for Data Representation, Learning Algorithms and Compositional Architectures -- 3 Categorical General Languages for Compositional Design and Analysis of Neural Architectures -- 4 An Integration of Categorical AI/ML Approaches -- 5 Concluding Remarks -- References -- Thinking as an Action -- 1 Introduction -- 2 Related Work -- 3 AGI Architecture -- 3.1 Experience Replay Training with Support Vector Machines -- 3.2 Thought Engineering -- 3.3 Reward in a Virtual Environment -- 4 Discussion -- References -- A Universal Intelligence Measure for Arithmetical Uncomputable Environments -- 1 Introduction -- 2 Background -- 2.1 Universal Intelligence -- 2.2 Arithmetical Hierarchy and Oracles -- 3 Motivation -- 3.1 Motivating Discussion -- 3.2 Motivating Examples -- 3.3 Rebuttals to the Exclusion of Uncomputable Environments -- 4 The Metric 21 -- 4.1

Definition -- 4.2 Properties of 21 -- 4.3 Intelligence of Oracles and Intelligence with Resources -- 5 Future Work -- References -- Semantic Primes-Inspired Tacit Knowledge Dataset for Simulating Basic Perception Capabilities of Cognitive Architectures -- 1 Introduction -- 2 Related Works -- 3 Dataset Creation -- 3.1 Prompt Generation Process -- 3.2 Human Annotation -- 4 Language Models as Perceptors: Preliminary Experiment -- 5 Discussion -- 6 Conclusion and Future Work -- References -- Simulation of Non-Primate Intelligence vs Human Intelligence vs Superhuman AGI vs Alien-Like AGI -- 1 Introduction - The Evolution of a BICA from Associative Reasoning to Superhuman Intelligent Behavior -- 2 Modeling of Non-Primate-Like AI vs HLAI vs Superhuman AGI vs Alien-Like AGI -- 3 Method -- 4 Results -- 5 Discussion -- References.
Causal Generalization via Goal-Driven Analogy -- 1 Introduction -- 2 Related Work -- 3 Theoretical and Methodological Framework -- 4 Causal Generalization -- 5 Goal-Driven Analogy: A Mechanism -- 6 AERA and Analogy -- 7 Results and Evaluation -- 8 Conclusions -- References -- Clipping the Risks: Integrating Consciousness in AGI to Avoid Existential Crises -- References -- Declarative Rules and Rule-Based Systems -- 1 Introduction -- 2 NARS is Not a Rule-Based System -- 3 Inference Engine -- 4 Results and Discussion -- References -- A Theory of Foundational Meaning Generation in Autonomous Systems, Natural and Artificial -- 1 Introduction -- 2 Related Work -- 3 Definitions and Concepts -- 4 Meaning Generation -- 5 Conclusions and Future Work -- References -- Causal Inference in NARS -- 1 Introduction -- 2 Previous Works -- 3 The NARS Approach -- 4 Working Examples -- 4.1 Sensorimotor Integration -- 4.2 Simpson's Paradox -- 4.3 Intervention and Counterfactual -- 5 Conclusion -- References -- AGI from the Perspectives of Categorical Logic and Algebraic Geometry -- 1 Goal of This Paper -- 2 Results Thus Far -- 2.1 Where Is GPT? -- 2.2 Homotopy Type Theory (HoTT) -- 2.3 Commutativity of and -- 2.4 and as Adjunctions -- 2.5 Predicates as Fibration -- 2.6 Iteration of and Looped Transformers -- 2.7 Modal Logic -- 2.8 Algebraic Geometry and Topos Theory -- References -- Human-Robot Trust in the Age of Artificial General Intelligence: The Case of Care Robots -- 1 Introduction -- 2 Care Robots Classification -- 3 Care Robots with AGI -- 3.1 Human-Like Behavior -- 3.2 Generalization of Knowledge -- 3.3 Self-learning and Evolution -- 4 Ethical Implications of AGI-Based Care Robots -- 4.1 Deceptive Behavior -- 4.2 Accelerating Surveillance Capitalism -- 4.3 Privacy and Security -- 4.4 Over-Reliance with Care Robots -- 4.5 Bias -- 4.6 Autonomy.
4.7 Responsibility and Accountability -- 5 How to Make and Preserve Human and AGI-Based Care Robot Trust -- 5.1 Types of Trust -- 5.2 Performance Trust -- 5.3 Moral Trust -- 6 Conclusion -- References -- From Artificial General Intelligence to Artificial General Universe: Metaverse Ethics as an Amplification of AI/AGI Ethics -- 1 Introduction: Escaping from Real-World Constraints -- 2 What Is AGU (Artificial General Universe) -- 3 Social Issues on the Metaverse and AGU -- 3.1 Surveillance Capitalism -- 3.2 Issues at the Border Between the Real and Virtual -- 3.3 Governance, Regulation, and Accountability -- 4 Concluding Remarks -- References -- Author Index.

Sommario/riassunto

This book constitutes the refereed proceedings of the 17th International Conference on Artificial General Intelligence, AGI 2024, held in Seattle, Washington, USA in August 2024. The 25 papers presented in this book were carefully reviewed and selected from 55 submissions. The papers focus on the main theme of AGI 2024: 'Understanding Artificial General Intelligence', with discussions on

various central concepts of general intelligence including thought, understanding, meaning, creativity, insight, reasoning, autonomy, attention and control. .