ALMA MATER STUDIORUM – UNIVERSITY OF BOLOGNA

Master's Degree in Computer Science and Engineering

Unity for Collektive: Reducing Reality Gap in the Simulation of Collective Adaptive Systems

Thesis in:
SOFTWARE PROCESS ENGINEERING

Supervisor
Prof. Danilo Pianini

Co-supervisors
Martina Baiardi
Angela Cortecchia

Candidate
Filippo Gurioli

Graduation Session: IV
Academic Year 2024-2025

Abstract

Max 2000 characters, strict.

Optional. Max a few lines.

Contents

List of Figures

List of Listings

Chapter 1

Introduction

Modern computing is moving away from the era of powerful and isolated machines toward one composed by massively interconnected ensembles of devices. We can observe this transition everywhere, from global cro:IoTInternet of Thing (IoT) sensor networks to smart city infrastructures. In such scenarios, the focus shifts from 'how to compute' to 'how to coordinate'.

As the number of devices in these systems grows into the thousands or millions, traditional centralized management becomes a bottleneck. The latency, bandwidth constraints, and single-point-of-failure risks of a 'command-and-control' architecture make it unsuitable for the dynamic, often unpredictable environments these systems inhabit. Instead, we must look toward decentralized coordination, where collective intelligence arises from local interactions rather than global oversight.

This thesis explores the intersection of high-level collective programming and high-fidelity simulation. Specifically, it addresses the engineering gap between abstract coordination models, such as cro:ACAggregate Computing (AC), and the practical requirements of developing, testing, and deploying these models within realistic 3D environments. By leveraging the power of modern game engines and automated development workflows, this work aims to provide a robust infrastructure for the next generation of collective system design.

1.1 Motivation: Swarm Behaviour

The natural world provides the strongest precedence for the goal of resilient decentralized coordination. From the coordinated flashing of fireflies to the intricate architectural achievements of termite mounds and the smooth collective motion of starling murmurings, biological systems exhibit an efficiency that is frequently difficult for classical engineering to match. These phenomena, which are collectively referred to as cro:SISwarm Intelligence (SI), arise from the interaction of many simple agents that follow localized rules rather than from a global supervisor.

In a natural swarm, intelligence is inherently distributed and emergent. Individual agents (be they ants, bees or birds) possess only a partial perception of their surroundings. The collective however can solve high-order problems such as finding the shortest path to a food source or executing rapid evasive maneuvers against predators. From an engineering perspective, these systems offer three indispensable properties:

- the absence of a central controller; the loss of individual units does not compromise the mission.
- The logic governing ten agents often remains functional for ten thousand, as interactions remain local regardless of total population size.
- Swarms autonomously adapt to dynamic environments, re-configuring their behaviour in response to external stimuli.

As we attempt to port these characteristics into the digital and physical domains (specifically through paradigms like AC) we face a significant translation gap. While the mathematical models for collective logic are maturing, the infrastructure to test them in realistic, high-fidelity environments remains fragmented. To truly harness the potential of swarm behaviour in human-made systems, we must develop tools that can simulate the complex interplay between decentralized algorithms and the physical world.

1.2 Problem Statement: Engineering Challanges in Simulation

Simulation has been widely explored in terms of scalability, but not many researches have been done regarding high-fidelity. This field brings into play hard constraints that mathematical rigor often does not consider. Physics collisions, gravity and friction are just examples of what a good high-fidelity simulator could add to a cooperative swarm simulation. Traditional simulators often prioritize the number of agents at the expense of environmental complexity, leading to a 'reality gap' that complicates the deployment of algorithms onto physical hardware. Fortunately, game engines do this work for us; they add physics engines capable of computing the result of physical interactions with rigor. The real problem now becomes only one: bridging these two worlds.

The challenge of bridging high-level coordination with game-engine-driven physics is not merely a matter of data transfer, but one of architectural alignment. In particular:

- synchronism: collective programming models rely on discrete logical steps whereas game engines operate on a continuous, high-frequency tick (e.g. 60Hz, 60 frames per second).
- Abstraction: collective models treat agents like points in space whereas high-fidelity environment represent them as complex entities with mass, inertia and physical bounds.
- Scalability: the simulator should still be able to compete with other collective programming simulators in terms of nodes represented inside the experiments and their interactions.

Chapter 2

Background and State of the Art

2.1 Distributed Systems and Organizational Complexity

2.2 Self-Organizing Frameworks

2.2.1 Aggregate Computing

2.3 Simulation Landscape

2.3.1 Paradigms

2.3.2 The Reality Gap

2.3.3 Reealism vs. Scalability

2.4 Game Engines as Simulators

Chapter 3

Unity-Package-Template: Automated Unity Development Infrastructure

3.1 Requirements

3.2 Features

Chapter 4

Collektive×Unity: Designing a 3D Simulator for Collective Systems

4.1 Goal

4.2 Requirements

4.3 Architecture

Chapter 5

Implementation of Collektive$\times$Unity

5.1 Design

5.2 Implementation Details

Chapter 6

Case Study: Environment-aware Gradient Ascent

Chapter 7

Results

7.1 Comparison with Socket-based Communication

Chapter 8

Conclusions and Future Work

Bibliography

[AB93] Abderrahmane Aggoun and Nicolas Beldiceanu. Extending chip in order to solve complex scheduling and placement problems. *Mathematical and computer modelling*, 17(7):57–73, 1993.

[ACO21] Andrea Agiollo, Giovanni Ciatto, and Andrea Omicini. *Shallow2Deep*: Restraining neural networks opacity through neural architecture search. In Davide Calvaresi, Amro Najjar, Michael Winikoff, and Kary Främling, editors, *Explainable and Transparent AI and Multi-Agent Systems. Third International Workshop, EXTRAAMAS 2021, Virtual Event, May 3–7, 2021, Revised Selected Papers*, volume 12688 of *Lecture Notes in Computer Science*, pages 63–82. Springer Nature, Basel, Switzerland, 2021.

[ADT95] Robert Andrews, Joachim Diederich, and Alan B. Tickle. Survey and critique of techniques for extracting rules from trained artificial neural networks. *Knowl. Based Syst.*, 8(6):373–389, 1995.

[AFWZ02] Alessandro Artale, Enrico Franconi, Frank Wolter, and Michael Zakharyaschev. A temporal description logic for reasoning over conceptual schemas and queries. In *European Workshop on Logics in Artificial Intelligence (JELIA 2002)*, pages 98–110. Springer, 2002.

[AK12a] M. Gethsiyal Augasta and T. Kathirvalavakumar. Reverse engineering the neural networks for rule extraction in classification problems. *Neural Processing Letters*, 35(2):131–150, April 2012.

[AK12b] M. Gethsiyal Augasta and T. Kathirvalavakumar. Reverse engineering the neural networks for rule extraction in classification problems. *Neural Process. Lett.*, 35(2):131–150, 2012.

[Apt90] Krzysztof R Apt. Logic programming. *Handbook of Theoretical Computer Science, Volume B: Formal Models and Sematics (B)*, 1990:493–574, 1990.

[Baa03] Franz Baader. Basic description logics. In *The Description Logic Handbook: Theory, Implementation, and Applications*, pages 43–95, USA, 2003. Cambridge University Press.

[BDKT97] Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial intelligence*, 93(1–2):63–101, 1997.

[BFOS84] Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone. *Classification and Regression Trees.* Chapman & Hall/CRC, 1984.

[BH16] Guido Bologna and Yoichi Hayashi. A rule extraction study on a neural network trained by deep learning. In *2016 International Joint Conference on Neural Networks, IJCNN 2016, Vancouver, BC, Canada, July 24-29, 2016*, pages 668–675. IEEE, 2016.

[BH18] Guido Bologna and Yoichi Hayashi. A comparison study on rule extraction from neural network ensembles, boosted shallow trees, and svms. *Appl. Comput. Intell. Soft Comput.*, 2018:4084850:1–4084850:20, 2018.

[BHS79] A. E. Bryson, Y. Ho, and G. M. Siouris. Applied optimal control: Optimization, estimation, and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(6):366–367, June 1979.

[BKB17] Osbert Bastani, Carolyn Kim, and Hamsa Bastani. Interpreting blackbox models via model extraction. *CoRR*, abs/1705.08504, 2017.

[BL04] Ronald J. Brachman and Hector J. Levesque. The tradeoff between expressiveness and tractability. In Ronald J. Brachman and Hector J. Levesque, editors, *Knowledge Representation and Reasoning*, The Morgan Kaufmann Series in Artificial Intelligence, pages 327–348. Morgan Kaufmann, San Francisco, 2004.

[BU18] Tarek R. Besold and Sara L. Uckelman. The what, the why, and the how of artificial explanations in automated decision-making. *CoRR*, abs/1808.07074:1–20, 2018.

[CBMO19] Giovanni Ciatto, Michael Bosello, Stefano Mariani, and Andrea Omicini. Comparative analysis of blockchain technologies under a coordination perspective. In Fernando De La Prieta, Alfonso González-Briones, Pawel Pawleski, Davide Calvaresi, Elena Del Val, Fernando Lopes, Vicente Julian, Eneko Osaba, and Ramón Sánchez-Iborra, editors, *Highlights of Practical Applications of Survivable Agents and Multi-Agent Systems. The PAAMS Collection*, volume 1047 of *Communications in Computer and Information Science*, chapter 7, pages 80–91. Springer, June 2019.

[CCDMSO20] Ashley Caselli, Giovanni Ciatto, Giovanna Di Marzo Serugendo, and Andrea Omicini. Engineering semantic self-composition of services through tuple-based coordination. In Tiziana Margaria and Bernhard Steffen, editors, *Leveraging Applications of Formal Methods, Verification and Validation: Engineering Principles*, volume 12477 of *Lecture Notes in Computer Science*, pages 205–223. Springer International Publishing, Cham, 2020.

[CCDO19a] Roberta Calegari, Giovanni Ciatto, Jason Dellaluce, and Andrea Omicini. Interpretable narrative explanation for ML predictors with LP: A case study for XAI. In Federico Bergenti and Stefania Monica, editors, *WOA 2019 – 20th Workshop "From Objects to Agents"*, volume 2404 of *CEUR Workshop Proceedings*, pages 105–112. Sun SITE Central Europe, RWTH Aachen University, 26–28 June 2019.

[CCDO19b] Roberta Calegari, Giovanni Ciatto, Enrico Denti, and Andrea Omicini. Engineering micro-intelligence at the edge of CPCS: Design guidelines. In *Internet and Distributed Computing Systems (IDCS 2019)*, volume 11874 of *Lecture Notes in Computer Science*, pages 260–270. Springer, 10–12 October 2019.

[CCDO20] Roberta Calegari, Giovanni Ciatto, Enrico Denti, and Andrea Omicini. Logic-based technologies for intelligent systems: State of the art and perspectives. *Information*, 11(3):1–29, March 2020. Special Issue "10th Anniversary of Information—Emerging Research Challenges".

[CCM+18a] Roberta Calegari, Giovanni Ciatto, Stefano Mariani, Enrico Denti, and Andrea Omicini. Logic programming in space-time: The case of situatedness in LPaaS. In Massimo Cossentino, Luca Sabatucci, and Valeria Seidita, editors, *WOA 2018 – 19th Workshop "From Objects to Agents"*, volume 2215 of *CEUR Workshop Proceedings*, pages 63–68. Sun SITE Central Europe, RWTH Aachen University, 29–30 June 2018.

[CCM+18b] Roberta Calegari, Giovanni Ciatto, Stefano Mariani, Enrico Denti, and Andrea Omicini. LPaaS as micro-intelligence: Enhancing IoT with symbolic reasoning. *Big Data and Cognitive Computing*, 2(3), 2018.

[CCM+18c] Roberta Calegari, Giovanni Ciatto, Stefano Mariani, Enrico Denti, and Andrea Omicini. Micro-intelligence for the IoT: SE challenges and practice in LPaaS. In *2018 IEEE International Conference on Cloud Engineering (IC2E 2018)*, pages 292–297. IEEE Computer Society, 17–20 April 2018.

[CCM+18d] Giovanni Ciatto, Roberta Calegari, Stefano Mariani, Enrico Denti, and Andrea Omicini. From the blockchain to logic programming and back: Research perspectives. In Massimo Cossentino, Luca Sabatucci, and Valeria Seidita, editors, *WOA 2018 – 19th Workshop*

"From Objects to Agents", volume 2215 of *CEUR Workshop Proceedings*, pages 69–74. Sun SITE Central Europe, RWTH Aachen University, June 2018.

[CCMO21a] Roberta Calegari, Giovanni Ciatto, Viviana Mascardi, and Andrea Omicini. Logic-based technologies for multi-agent systems: A systematic literature review. *Autonomous Agents and Multi-Agent Systems*, 35(1):1:1–1:67, 2021. Collection "Current Trends in Research on Software Agents and Agent-Based Software Development".

[CCMO21b] Roberta Calegari, Giovanni Ciatto, Viviana Mascardi, and Andrea Omicini. Logic-based technologies for multi-agent systems: Summary of a systematic literature review. In *20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2021)*, pages 1721–1723, May 2021.

[CCN+21] Davide Calvaresi, Giovanni Ciatto, Amro Najjar, Reyhan Aydoğan, Leon Van der Torre, Andrea Omicini, and Michael Schumacher. Expectation: Personalized explainable artificial intelligence for decentralized agents with heterogeneous knowledge. In Davide Calvaresi, Amro Najjar, Michael Winikoff, and Kary Främling, editors, *Explainable and Transparent AI and Multi-Agent Systems. Third International Workshop, EXTRAAMAS 2021, Virtual Event, May 3–7, 2021, Revised Selected Papers*, volume 12688 of *Lecture Notes in Computer Science*, pages 331–343. Springer Nature, Basel, Switzerland, 2021.

[CCO20] Roberta Calegari, Giovanni Ciatto, and Andrea Omicini. On the integration of symbolic and sub-symbolic techniques for XAI: A survey. *Intelligenza Artificiale*, 14(1):7–32, 2020.

[CCO21a] Giovanni Ciatto, Roberta Calegari, and Andrea Omicini. 2p-kt: A logic-based ecosystem for symbolic ai. *SoftwareX*, 2021.

[CCO21b] Giovanni Ciatto, Roberta Calegari, and Andrea Omicini. Lazy stream manipulation in Prolog via backtracking: The case of

2P-KT. In Wolfgang Faber, Gerhard Friedrich, Martin Gebser, and Michael Morak, editors, *Logics in Artificial Intelligence*, volume 12678 of *Lecture Notes in Computer Science*, pages 407–420. Springer, 2021. 17th European Conference, JELIA 2021, Virtual Event, May 17–20, 2021, Proceedings.

[CCS+20] Giovanni Ciatto, Roberta Calegari, Enrico Siboni, Enrico Denti, and Andrea Omicini. 2P-KT: logic programming with objects & functions in kotlin. In Roberta Calegari, Giovanni Ciatto, Enrico Denti, Andrea Omicini, and Giovanni Sartor, editors, *WOA 2020 – 21th Workshop "From Objects to Agents"*, volume 2706 of *CEUR Workshop Proceedings*, pages 219–236, Aachen, Germany, October 2020. Sun SITE Central Europe, RWTH Aachen University. 21st Workshop "From Objects to Agents" (WOA 2020), Bologna, Italy, 14–16 September 2020. Proceedings.

[CCSO20] Giovanni Ciatto, Davide Calvaresi, Michael I. Schumacher, and Andrea Omicini. An abstract framework for agent-based explanations in AI. In *19th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1816–1818. International Foundation for Autonomous Agents and Multiagent Systems, May 2020. Extended Abstract.

[CDD17] Giovanni Ciatto, Elisabetta De Maria, and Cinzia Di Giusto. Spiking neural networks as timed automata. In *Proc. of the Thematic Research School on Advances in Systems and Synthetic Biology (ASSB)*, pages 55–69. EDP Sciences, 2017.

[CDMSL+20] Giovanni Ciatto, Giovanna Di Marzo Serugendo, Maxime Louvel, Stefano Mariani, Andrea Omicini, and Franco Zambonelli. Twenty years of coordination technologies: COORDINATION contribution to the state of art. *Journal of Logical and Algebraic Methods in Programming*, 113:1–25, June 2020.

[CH94] William W. Cohen and Haym Hirsh. Learning the classic description logic: Theoretical and experimental results. In *Principles of Knowledge Representation and Reasoning*, pages 121–133. Elsevier, 1994.

[Cim06] Philipp Cimiano. *Ontology Learning and Population from Text*. Springer US, 2006.

[Cla77] Keith L. Clark. Negation as failure. In Hervé Gallaire and Jack Minker, editors, *Logic and Data Bases, Symposium on Logic and Data Bases, Centre d'études et de recherches de Toulouse, France, 1977*, Advances in Data Base Theory, pages 293–322, New York, 1977. Plemum Press.

[CMMO19] Giovanni Ciatto, Alfredo Maffi, Stefano Mariani, and Andrea Omicini. Towards agent-oriented blockchains: Autonomous smart contracts. In Yves Demazeau, Eric Matson, Juan Manuel Corchado, and Fernando De la Prieta, editors, *Advances in Practical Applications of Survivable Agents and Multi-Agent Systems: The PAAMS Collection*, volume 11523 of *Lecture Notes in Computer Science*, pages 29–41. Springer International Publishing, June 2019.

[CMMO20a] Giovanni Ciatto, Alfredo Maffi, Stefano Mariani, and Andrea Omicini. Smart contracts are more than objects: Pro-activeness on the blockchain. In Javier Prieto, Ashok Das Kumar, Stefano Ferretti, António Pinto, and Juan Manuel Corchado, editors, *Blockchain and Applications*, volume 1010 of *Advances in Intelligent Systems and Computing*, pages 45–53. Springer, 2020.

[CMMO20b] Giovanni Ciatto, Stefano Mariani, Alfredo Maffi, and Andrea Omicini. Blockchain-based coordination: Assessing the expressive power of smart contracts. *Information*, 11(1):1–20, January 2020. Special Issue "Blockchain Technologies for Multi-Agent Systems".

[CMO17] Giovanni Ciatto, Stefano Mariani, and Andrea Omicini. Programming the interaction space effectively with ReSpecTX. In Mirjana

Ivanović, Costin Bădică, Jürgen Dix, Zoran Jovanović, Michele Malgeri, and Miloš Savić, editors, *Intelligent Distributed Computing XI*, volume 737 of *Studies in Computational Intelligence*, pages 89–101. Springer, 2017.

[CMO18a] Giovanni Ciatto, Stefano Mariani, and Andrea Omicini. Blockchain for trustworthy coordination: A first study with Linda and Ethereum. In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 696–703, December 2018.

[CMO18b] Giovanni Ciatto, Stefano Mariani, and Andrea Omicini. ReSpecTX: Programming interaction made easy. *Computer Science and Information Systems*, 15(3):655–682, October 2018. Special Section: Contemporary Topics in Intelligent Distributed Computing.

[CMO+18c] Giovanni Ciatto, Stefano Mariani, Andrea Omicini, Franco Zambonelli, and Maxime Louvel. Twenty years of coordination technologies: State-of-the-art and perspectives. In Giovanna Di Marzo Serugendo and Michele Loreti, editors, *Coordination Models and Languages*, volume 10852 of *Lecture Notes in Computer Science*, pages 51–80. Springer, 2018. 20th IFIP WG 6.1 International Conference, COORDINATION 2018, Held as Part of the 13th International Federated Conference on Distributed Computing Techniques, DisCoTec 2018, Madrid, Spain, June 18-21, 2018. Proceedings.

[CMOZ20] Giovanni Ciatto, Stefano Mariani, Andrea Omicini, and Franco Zambonelli. From agents to blockchain: Stairway to integration. *Applied Sciences*, 10(21):7460:1–7460:22, 2020. Special Issue "Advances in Blockchain Technology and Applications 2020".

[CNCC21] Giovanni Ciatto, Amro Najjar, Jean-Paul Calbimonte, and Davide Calvaresi. Towards explainable visionary agents: License to dare and imagine. In Davide Calvaresi, Amro Najjar, Michael Winikoff, and Kary Främling, editors, *Explainable and Transparent AI and Multi-Agent Systems. Third International Workshop, EXTRAA-*

MAS 2021, Virtual Event, May 3–7, 2021, Revised Selected Papers, volume 12688 of *Lecture Notes in Computer Science*, pages 139–157. Springer Nature, Basel, Switzerland, 2021.

[Col86] Alain Colmerauer. Theoretical model of prolog ii. In M. van Canegham and D.-H.D. Warren, editors, *Logic Programming and its applications*, pages 3–31. Ablex Publishing Corporation, 1986.

[CR93] Alain Colmerauer and Philippe Roussel. The birth of prolog. In John A. N. Lee and Jean E. Sammet, editors, *History of Programming Languages Conference (HOPL-II)*, pages 37–52. ACM, April 1993.

[Cra16] Kate Crawford. Artificial intelligence's white guy problem. *The New York Times*, 25, 2016.

[CROM19] Giovanni Ciatto, Lorenzo Rizzato, Andrea Omicini, and Stefano Mariani. TuSoW: Tuple spaces for edge computing. In *The 28th International Conference on Computer Communications and Networks (ICCCN 2019)*, Valencia, Spain, 29 July–1 August 2019.

[CS95] Mark W. Craven and Jude W. Shavlik. Extracting tree-structured representations of trained networks. In David S. Touretzky, Michael Mozer, and Michael E. Hasselmo, editors, *Advances in Neural Information Processing Systems 8, NIPS, Denver, CO, USA, November 27-30, 1995*, pages 24–30. MIT Press, 1995.

[CSOC20] Giovanni Ciatto, Michael I. Schumacher, Andrea Omicini, and Davide Calvaresi. Agent-based explanations in ai: Towards an abstract framework. In Davide Calvaresi, Amro Najjar, Michael Winikoff, and Kary Främling, editors, *Explainable, Transparent Autonomous Agents and Multi-Agent Systems*, volume 12175 of *Lecture Notes in Computer Science*, pages 3–20. Springer, Cham, 2020. Second International Workshop, EXTRAAMAS 2020, Auckland, New Zealand, May 9–13, 2020, Revised Selected Papers.

[Cyb89] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, Dec 1989.

[DCB19] Federica Di Castro and Enrico Bertini. Surrogate decision tree visualization. In *Joint Proceedings of the ACM IUI 2019 Workshops (ACMIUI-WS 2019)*, volume 2327 of *CEUR Workshop Proceedings*, March 2019.

[dGBG01] Artur S. d'Avila Garcez, Krysia Broda, and Dov M. Gabbay. Symbolic knowledge extraction from trained neural networks: A sound approach. *Artif. Intell.*, 125(1-2):155–207, 2001.

[dGBR+15] Artur S. d'Avila Garcez, Tarek R. Besold, Luc De Raedt, Peter Földiák, Pascal Hitzler, Thomas Icard, Kai-Uwe Kühnberger, Luís C. Lamb, Risto Miikkulainen, and Daniel L. Silver. Neural-symbolic learning and reasoning: Contributions and challenges. In *2015 AAAI Spring Symposia, Stanford University, Palo Alto, California, USA, March 22-25, 2015*. AAAI Press, 2015.

[dK15] Luc de Raedt and Angelika Kimmig. Probabilistic (logic) programming concepts. *Machine Learning*, 100(1):5–47, 2015.

[dKL98] Mark d'Inverno, D. Kinney, and Michael Luck. Interaction protocols in agentis. In *Proceedings International Conference on Multi Agent Systems (Cat. No. 98EX160)*, pages 112–119. IEEE, 1998.

[DLR77] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, 39(1):1–38, 1977.

[DMDGC17] Elisabetta De Maria, Cinzia Di Giusto, and Giovanni Ciatto. Formal validation of neural networks as timed automata. In *Proceedings of the 8th International Conference on Computational Systems-Biology and Bioinformatics*, CSBio '17, pages 15–22, New York, NY, USA, 2017. ACM.

[DRW96] Steven Dawson, C. R. Ramakrishnan, and David S. Warren. Practical program analysis using general purpose logic programming systems—a case study. In *Proceedings of the ACM SIGPLAN 1996 Conference on Programming Language Design and Implementation*, PLDI '96, pages 117—126, New York, NY, USA, 1996. Association for Computing Machinery.

[DVK17] Finale Doshi-Velez and Been Kim. Towards a rigorous science of interpretable machine learning. *CoRR*, abs/1702.08608, 2017.

[Ell19] Anthony Elliott. *The Culture of AI: Everyday Life and the Digital Revolution*. Routledge, 2019.

[FH17a] Marion Fourcade and Kieran Healy. Categories all the way down. *Historical Social Research/Historische Sozialforschung*, pages 286–296, 2017.

[FH17b] Nicholas Frosst and Geoffrey E. Hinton. Distilling a neural network into a soft decision tree. In Tarek R. Besold and Oliver Kutz, editors, *Proceedings of the First International Workshop on Comprehensibility and Explanation in AI and ML 2017 co-located with 16th International Conference of the Italian Association for Artificial Intelligence (AI*IA 2017), Bari, Italy, November 16th and 17th, 2017*, volume 2071 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2017.

[FK97] Tze Ho Fung and Robert Kowalski. The IFF proof procedure for abductive logic programming. *The Journal of Logic Programming*, 33(2):151–165, 1997.

[FV17] Ruth C. Fong and Andrea Vedaldi. Interpretable explanations of black boxes by meaningful perturbation. pages 3449–3457, 2017.

[FW99] Jacques Ferber and Gerhard Weiss. *Multi-agent systems: an introduction to distributed artificial intelligence*, volume 1. Addison-Wesley Reading, 1999.

[GBC16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[GF17] Bryce Goodman and Seth Flaxman. European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine*, 38(3):50–57, 2017.

[GG12] Sanjeev Goyal and Sandeep Grover. Applying fuzzy grey relational analysis for ranking the advanced manufacturing systems. *Grey Systems: Theory and Application*, 2(2):284–298, 2012.

[GMR+19] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM Comput. Surv.*, 51(5):93:1–93:42, 2019.

[GR68] C. Cordell Green and Bertram Raphael. The use of theorem-proving techniques in question-answering systems. In *1968 23rd ACM National Conference*, pages 169–181, 1968.

[Gun16] David Gunning. Explainable artificial intelligence (XAI). Funding Program DARPA-BAA-16-53, DARPA, 2016.

[Han06] David J Hand. Data mining. *Encyclopedia of Environmetrics*, 2, 2006.

[HE06] Eduardo R. Hruschka and Nelson F.F. Ebecken. Extracting rules from multilayer perceptrons in classification problems: A clustering-based approach. *Neurocomputing*, 70(1-3):384–397, 2006.

[Hel19] Dirk Helbing. Societal, economic, ethical and legal challenges of the digital revolution: From big data to deep learning, artificial intelligence, and manipulative technologies. In *Towards Digital Enlightenment*, pages 47–72. Springer, 2019.

[Hen08] J. Hendler. Avoiding another ai winter. *IEEE Intelligent Systems*, 23(2):2–4, March 2008.

[Hor05] Ian Horrocks. OWL: A description logic based ontology language. In Peter Van Beek, editor, *Principles and Practice of Constraint Programming (CP 2005)*, pages 5–8. Springer, 2005. Extended Abstract.

[HQR17] Robert Hoehndorf and Núria Queralt-Rosinach. Data science and symbolic ai: Synergies, challenges and opportunities. *Data Science*, 2017.

[HRHL01] Nick Howden, Ralph Rönnquist, Andrew Hodgson, and Andrew Lucas. Intelligent agents-summary of an agent infrastructure. In *Proceedings of the 5th International conference on autonomous agents*, 2001.

[Hub99] Marcus J. Huber. Jam: A bdi-theoretic mobile agent architecture. In *Proceedings of the third annual conference on Autonomous Agents*, pages 236–243, 1999.

[JL87] Joxan Jaffar and J.-L. Lassez. Constraint logic programming. In *Proceedings of the 14th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*, pages 111–119, October 1987.

[JN09] Ulf Johansson and Lars Niklasson. Evolving decision trees using oracle guides. In *Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2009, part of the IEEE Symposium Series on Computational Intelligence 2009, Nashville, TN, USA, March 30, 2009 - April 2, 2009*, pages 238–244. IEEE, 2009.

[KA09] Humar Kahramanli and Novruz Allahverdi. Rule extraction from trained adaptive neural networks using artificial immune systems. *Expert Syst. Appl.*, 36(2):1513–1522, 2009.

[KBB+21] Philipp Körner, Michael Beuschel, João Barbosa, Vítor Santos Costa, Verónica Dahl, Manuel V. Hermenegildo, Jose F. Morales, Jan Wielemaker, Daniel Diaz, Salvador Abreu, and Giovanni

Ciatto. A multi-walk through the past, present and future of prolog. *Theory and Practice of Logic Programming*, 2021.

[Kot07] Sotiris Kotsiantis. Supervised machine learning: A review of classification techniques. In *Emerging Artificial Intelligence Applications in Computer Engineering*, volume 160 of *Frontiers in Artificial Intelligence and Applications*, pages 3–24. IOS Press, October 2007.

[Kow74] Robert A. Kowalski. Predicate logic as programming language. In Jack L. Rosenfeld, editor, *Information Processing, Proceedings of the 6th IFIP Congress*, pages 569–574. North-Holland, August 1974.

[KSB99] R. Krishnan, G. Sivakumar, and P. Bhattacharya. Extracting decision trees from trained neural networks. *Pattern Recognition*, 32(12):1999–2009, 1999.

[LB87] Hector J. Levesque and Ronald J. Brachman. Expressiveness and tractability in knowledge representation and reasoning. *Comput. Intell.*, 3:78–93, 1987.

[LD94] Jaeho Lee and Edmund H. Durfee. Structured circuit semantics for reactive plan execution systems. In Barbara Hayes-Roth and Richard E. Korf, editors, *Proceedings of the 12th National Conference on Artificial Intelligence*, volume 2, pages 1232–1237, Seattle, WA, USA, 31 July—4 August 1994. AAAI Press / The MIT Press.

[Lip18] Zachary C. Lipton. The mythos of model interpretability. *Commun. ACM*, 61(10):36–43, 2018.

[Llo90] John W Lloyd. *Computational logic*. Springer, 1990.

[Md94] Stephen Muggleton and Luc de Raedt. Inductive logic programming: Theory and methods. *The Journal of Logic Programming*, 19-20:629–679, 1994. Special Issue: Ten Years of Logic Programming.

[MH03] Ralf Moller and Volker Haarslev. *Description logic systems*, pages 282–305. Cambridge University Press, 2003.

[Mil56] George Abram Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97, March 1956.

[Min75] Marvin Minsky. A framework for representing knowledge representation. In *The Psychology of Computer Vision*. Mc Graw-Hill, New-York (NY, US), 1975.

[MM82] Alberto Martelli and Ugo Montanari. An efficient unification algorithm. *ACM Trans. Program. Lang. Syst.*, 4(2):258–282, April 1982.

[MOC17] Stefano Mariani, Andrea Omicini, and Giovanni Ciatto. Novel opportunities for tuple-based coordination: XPath, the Blockchain, and stream processing. In Pasquale De Meo, Maria Nadia Postorino, Domenico Rosaci, and Giuseppe M.L. Sarné, editors, *WOA 2017 – 18th Workshop "From Objects to Agents"*, volume 1867 of *CEUR Workshop Proceedings*, pages 61–64. Sun SITE Central Europe, RWTH Aachen University, June 2017.

[MP88] Marvin L. Minsky and Seymour A. Papert. *Perceptrons: Expanded Edition*. MIT Press, Cambridge, MA, USA, 1988.

[MS58] John Mccarthy and Claude Shannon. Automata studies. *Journal of Symbolic Logic*, 23(1):59–60, 1958.

[MTC+10] Marco Montali, Paolo Torroni, Federico Chesani, Paola Mello, Marco Alberti, and Evelina Lamma. Abductive logic programming as an effective technology for the static verification of declarative business processes. *Fundamenta Informaticae*, 102(3–4):325–361, 2010.

[NM96] Anil Nerode and G. Metakides. *Principles of Logic and Logic Programming*. Elsevier Science Inc., USA, 1996.

[Pau18] Lawrence C. Paulson. Computational logic: its origins and applications. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2210):20170872, 2018.

[PCC+18] Danilo Pianini, Giovanni Ciatto, Roberto Casadei, Stefano Mariani, Mirko Viroli, and Andrea Omicini. Transparent protection of aggregate computations from Byzantine behaviours via blockchain. In *GOODTECHS'18 – Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good*, pages 271–276, New Work, NY, USA, November 2018. ACM.

[PCCO20] Giuseppe Pisano, Giovanni Ciatto, Roberta Calegari, and Andrea Omicini. Neuro-symbolic computation for XAI: Towards a unified model. In Roberta Calegari, Giovanni Ciatto, Enrico Denti, Andrea Omicini, and Giovanni Sartor, editors, *WOA 2020 – 21th Workshop "From Objects to Agents"*, volume 2706 of *CEUR Workshop Proceedings*, pages 101–117, Aachen, Germany, October 2020. Sun SITE Central Europe, RWTH Aachen University. 21st Workshop "From Objects to Agents" (WOA 2020), Bologna, Italy, 14–16 September 2020. Proceedings.

[Pol87] John L. Pollock. Defeasible reasoning. *Cognitive science*, 11(4):481–518, 1987.

[PW78] David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, 1(4):515–526, December 1978.

[Rao96] Anand S. Rao. Agentspeak(l): BDI agents speak out in a logical computable language. In Walter Van de Velde and John W. Perram, editors, *Agents Breaking Away, 7th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, Eindhoven, The Netherlands, January 22-25, 1996, Proceedings*, volume 1038 of *Lecture Notes in Computer Science*, pages 42–55. Springer, Berlin, Heidelberg, 1996.

[Rei80] Raymond Reiter. A logic for default reasoning. *Artificial intelligence*, 13(1–2):81–132, 1980.

[RN16] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,, 2016.

[Rob65] John Alan Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–41, 1965.

[Ros57] Frank Rosenblatt. *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.

[Ros00] Francesca Rossi. Constraint (logic) programming: A survey on research and applications. In Krzysztof R. Apt, Eric Monfroy, Antonis C. Kakas, and Francesca Rossi, editors, *New Trends in Constraints*, pages 40–74. Springer Berlin Heidelberg, Berlin, Heidelberg, 2000.

[RR19] Avi Rosenfeld and Ariella Richardson. Explainability in human-agent systems. *Autonomous Agents and Multi-Agent Systems*, 33(6):673–705, November 2019.

[RSG16] Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should I trust you?": Explaining the predictions of any classifier. In Balaji Krishnapuram, Mohak Shah, Alexander J. Smola, Charu C. Aggarwal, Dou Shen, and Rajeev Rastogi, editors, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, pages 1135–1144. ACM, 2016.

[Rud19] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, 2019.

[RVBW08] Francesca Rossi, Peter Van Beek, and Toby Walsh. Constraint programming. *Foundations of Artificial Intelligence*, 3:181–211, 2008.

[SCO21] Federico Sabbatini, Giovanni Ciatto, and Andrea Omicini. GridEx: An algorithm for knowledge extraction from black-box regressors. In Davide Calvaresi, Amro Najjar, Michael Winikoff, and Kary Främling, editors, *Explainable and Transparent AI and Multi-Agent Systems. Third International Workshop, EXTRAAMAS 2021, Virtual Event, May 3–7, 2021, Revised Selected Papers*, volume 12688 of *Lecture Notes in Computer Science*, pages 18–38. Springer Nature, Basel, Switzerland, 2021.

[Sea80] John R. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3):417–424, 1980.

[Smo87] P. Smolensky. Connectionist ai, symbolic ai, and the brain. *Artificial Intelligence Review*, 1(2):95–109, Jun 1987.

[Sow91] John F Sowa, editor. *Principles of semantic networks: Explorations in the representation of knowledge.* Morgan Kaufmann Series in Representation and Reasoning. Morgan Kaufmann Pub, May 1991.

[SS04] Alex J. Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, August 2004.

[ST01] M. Sato and H. Tsukimoto. Rule extraction from neural networks via decision tree induction. In *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, volume 3, pages 1870–1875 vol.3, 2001.

[STY17] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pages 3319–3328. PMLR, 2017.

[Sun01] R. Sun. Artificial intelligence: Connectionist and symbolic approaches. In Neil J. Smelser and Paul B. Baltes, editors, *Inter-*

national Encyclopedia of the Social & Behavioral Sciences, page 783–789. Pergamon, Oxford, 2001.

[TSHL17] Gabriele Tolomei, Fabrizio Silvestri, Andrew Haines, and Mounia Lalmas. Interpretable predictions of tree-based ensembles via actionable feature tweaking. In *23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 465–474. ACM, 2017.

[Tur50] Alan M. Turing. Computing machinery and intelligence. *Mind*, 59(October):433–60, 1950.

[Twa10] Bhekisipho Twala. Multiple classifier application to credit risk assessment. *Expert Systems with Applications*, 37(4):3326–3336, 2010.

[Vit06] Andrew J. Viterbi. A personal history of the viterbi algorithm. *IEEE Signal Process. Mag.*, 23(4):120–142, 2006.

[VvdB17] Paul Voigt and Axel von dem Bussche. *The EU General Data Protection Regulation (GDPR). A Practical Guide*. Springer, 2017.

[YL99] John Yen and Reza Langari. *Fuzzy logic: intelligence, control, and information*, volume 1. Prentice Hall Press, Upper Saddle River, NJ, 1999.

[YWC+18] Quanming Yao, Mengshuo Wang, Yuqiang Chen, Wenyuan Dai, Hu Yi-Qi, Li Yu-Feng, Tu Wei-Wei, Yang Qiang, and Yu Yang. Taking human out of learning applications: A survey on automated machine learning. pages 1–26, 2018.

[ZJC83] Zhi-Hua Zhou, Yuan Jiang, and Shi-Fu Chen. Extracting symbolic rules from trained neural network ensembles. *Mineral Processing and Extractive Metallurgy Review*, 1(1-2):207–248, 1983.

Acknowledgements

Optional. Max 1 page.