A Review of the Introduction to Artificial Intelligence for the Analysis of Literary Works and Social Media

Dr. Keikhosrokiani

Dr. AR Costa- Ralha

Abstract

Artificial intelligence (AI) is the capacity of a digital computer or robot operated by a computer to carry out actions frequently performed by intelligent beings. The phrase is widely used in reference to the effort to create artificial intelligence (AI) systems that possess human-like cognitive abilities like the capacity for reasoning, meaning-finding, generalisation, and experience-based learning. It has been proven that computers can be programmed to perform extremely complicated tasks—like, for example, finding proofs for mathematical theorems or playing chess—with remarkable proficiency ever since the development of the digital computer in the 1940s. Nevertheless, despite ongoing improvements in computer processing speed and memory space, there are currently no programmes that can match human adaptability across a larger range of activities or those needing a substantial amount of background knowledge.

Introduction

Artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience. Since the development of the digital computer in the 1940s, it has been demonstrated that computers can be programmed to carry out very complex tasks—as, for example, discovering proofs for mathematical theorems or playing chess—with great proficiency. Still, despite continuing advances in computer processing speed and memory capacity, there are as yet no programs that can match human flexibility over wider domains or in tasks requiring much everyday knowledge. On the other hand, some programs have attained the performance levels of human experts and professionals in performing certain specific tasks, so that artificial intelligence in this limited sense is found in applications as diverse as medical diagnosis, computer search engines, and voice or handwriting recognition. *(Read Ray Kurzweil'sBritannica essay on the future of "Nonbiological Man.").*

The female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Artificial intelligence (AI), also known as machine intelligence, is a branch of computer science that focuses on building and managing technology that can learn to autonomously make decisions and carry out actions on behalf of a human being.

AI is not a single technology. Instead, it is an umbrella term that includes any type of software or hardware component that supports machine learning, computer vision, natural language understanding, natural language generation, natural language processing and robotics. Today's AI uses conventional CMOS hardware and the same basic algorithmic functions that drive traditional software. Future generations of AI are expected to inspire new types of brain-inspired circuits and architectures that can make data-driven decisions faster and more accurately than a human being can. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

Future generations of AI are expected to inspire new types of brain-inspired circuits and architectures that can make data-driven decisions faster and more accurately than a human being can. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

What is intelligence

All but the simplest human behaviour is ascribed to intelligence, while even the most complicated insect behaviour is never taken as an indication of intelligence. What is the difference? Consider the behaviour of the digger wasp, *Sphex ichneumoneus*. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

There are a number of different forms of learning as applied to artificial intelligence. The simplest is learning by trial and error. For example, a simple computer program for solving mate-in-one chess problems might try moves at random until mate is found. The program might then store the solution with the position so that the next time the computer encountered the same position it would recall the solution. This simple memorizing of individual items and procedures—known as rote learning—is relatively easy to implement on a computer. More challenging is the problem of implementing what is called generalization. Generalization involves applying past experience to analogous new situations. For example, a program that learns the past tense of regular English verbs by rote will not be able to produce the past tense of a word such as *jump* unless it previously had been presented with *jumped*, whereas a program that is able to generalize can learn the "add *ed*" rule and so form the past tense of *jump* based on experience with similar verbs.

The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances

There are a number of different forms of learning as applied to artificial intelligence. The simplest is learning by trial and error. For example, a simple computer program for solving mate-in-one chess problems might try moves at random until mate is found. The program might then store the solution with the position so that the next time the computer encountered the same position it would recall the solution. This simple memorizing of individual items and procedures—known as rote learning—is relatively easy to implement on a computer. More challenging is the problem of implementing what is called generalization. Generalization involves applying past experience to analogous new situations. For example, a program that learns the past tense of regular English verbs by rote will not be able to produce the past tense of a word such as *jump* unless it previously had been presented with *jumped*, whereas a program that is able to generalize can learn the "add *ed*" rule and so form the past tense of *jump* based on experience with similar verbs.

on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

There are a number of different forms of learning as applied to artificial intelligence. The simplest is learning by trial and error. For example, a simple computer program for solving mate-in-one chess problems might try moves at random until mate is found. The program might then store the solution with the position so that the next time the computer encountered the same position it would recall the solution. This simple memorizing of individual items and procedures—known as rote learning—is relatively easy to implement on a computer. More challenging is the problem of implementing what is called generalization. Generalization involves applying past experience to analogous new situations. For example, a program that learns the past tense of regular English verbs by rote will not be able to produce the past tense of a word such as *jump* unless it previously had been presented with *jumped*, whereas a program that is able to generalize can learn the "add *ed*" rule and so form the past tense of *jump* based on experience with similar verbs.

The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

To reason is to draw inferences appropriate to the situation. Inferences are classified as either deductive or inductive. An example of the former is, "Fred must be in either the museum or the café. He is not in the café; therefore he is in the museum," and of the latter, "Previous accidents of this sort were caused by instrument failure; therefore this accident was caused by instrument failure." The most significant difference between these forms of reasoning is that in the deductive case the truth of the premises guarantees the truth of the conclusion, whereas in the inductive case the truth of the premise lends support to the conclusion without giving absolute assurance. Inductive reasoning is common in science, where data are collected and tentative models are developed to describe and predict future behaviour—until the appearance of anomalous data forces the model to be revised. Deductive reasoning is common in mathematics and logic, where elaborate structures of irrefutable theorems are built up from a small set of basic axioms and rules.

The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp's instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

To reason is to draw inferences appropriate to the situation. Inferences are classified as either deductive or inductive. An example of the former is, “Fred must be in either the museum or the café. He is not in the café; therefore he is in the museum,” and of the latter, “Previous accidents of this sort were caused by instrument failure; therefore this accident was caused by instrument failure.” The most significant difference between these forms of reasoning is that in the deductive case the truth of the premises guarantees the truth of the conclusion, whereas in the inductive case the truth of the premise lends support to the conclusion without giving absolute assurance. Inductive reasoning is common in science, where data are collected and tentative models are developed to describe and predict future behaviour—until the appearance of anomalous data forces the model to be revised. Deductive reasoning is common in mathematics and logic, where elaborate structures of irrefutable theorems are built up from a small set of basic axioms and rules.

The real nature of the wasp’s instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp’s instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances. When the female wasp returns to her burrow with food, she first deposits it on the threshold, checks for intruders inside her burrow, and only then, if the coast is clear, carries her food inside. The real nature of the wasp’s instinctual behaviour is revealed if the food is moved a few inches away from the entrance to her burrow while she is inside: on emerging, she will repeat the whole procedure as often as the food is displaced. Intelligence—conspicuously absent in the case of *Sphex*—must include the ability to adapt to new circumstances.

Problem solving, particularly in artificial intelligence, may be characterized as a systematic search through a range of possible actions in order to reach some predefined goal or solution. Problem-solving methods divide into special purpose and general purpose. A special-purpose method is tailor-made for a particular problem and often exploits very specific features of the situation in which the problem is embedded. In contrast, a general-purpose method is applicable to a wide variety of problems. To reason is to draw inferences appropriate to the situation. Inferences are classified as either deductive or inductive. An example of the former is, "Fred must be in either the museum or the café. He is not in the café; therefore he is in the museum," and of the latter, "Previous accidents of this sort were caused by instrument failure; therefore this accident was caused by instrument failure." The most significant difference between these forms of reasoning is that in the deductive case the truth of the premises guarantees the truth of the conclusion, whereas in the inductive case the truth of the premise lends support to the conclusion without giving absolute assurance. Inductive reasoning is common in science, where data are collected and tentative models are developed to describe and predict future behaviour—until the appearance of anomalous data forces the model to be revised.

To reason is to draw inferences appropriate to the situation. Inferences are classified as either deductive or inductive. An example of the former is, "Fred must be in either the museum or the café. He is not in the café; therefore he is in the museum," and of the latter, "Previous accidents of this sort were caused by instrument failure; therefore this accident was caused by instrument failure." The most significant difference between these forms of reasoning is that in the deductive case the truth of the premises guarantees the truth of the conclusion, whereas in the inductive case the truth of the premise lends support to the conclusion without giving absolute assurance. Inductive reasoning is common in science, where data are collected and tentative models are developed to describe and predict future behaviour—until the appearance of anomalous data forces the model to be revised. Deductive reasoning is common in mathematics and logic, where elaborate structures of irrefutable theorems are built up from a small set of basic axioms and rules.

through a range of possible actions in order to reach some predefined goal or solution. Problem-solving methods divide into special purpose and general purpose. A special-purpose method is tailor-made for a particular problem and often exploits very specific features of the situation in which the problem is embedded. In contrast, a general-purpose method is applicable to a wide variety of problems. To reason is to draw inferences appropriate to the situation. Inferences are classified as either deductive or inductive. An example of the former is, “Fred must be in either the museum or the café. He is not in the café; therefore he is in the museum,” and of the latter, “Previous accidents of this sort were caused by instrument failure; therefore this accident was caused by instrument failure.” The most significant difference between these forms of reasoning is that in the deductive case the truth of the premises guarantees the truth of the conclusion, whereas in the inductive case the truth of the premise lends support to the conclusion without giving absolute assurance. Inductive reasoning is common in science, where data are collected and tentative models are developed to describe and predict future behaviour—until the appearance of anomalous data forces the model to be revised.

Reference

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., & Devin, M. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv:1603.04467.

Araujo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., … Campilho, A. (2017). Classification of breast cancer histology images using convolutional neural networks. PLoS One, 12, e0177544.

Babiker, H. K. B., & Goebel, R. (2017). An introduction to deep visual explanation. arXiv:1711.09482.

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLoS One, 10, e0130140.

Baumgartner, C. F., Koch, L. M., Tezcan, K. C., Ang, J. X., & Konukoglu, E. (2017). Visual feature attribution using wasserstein gans. Paper presented at Proceedings of the IEEE computer society conference on computer vision and pattern recognition.

Bengio, Y. (2009). Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2, 1–127.

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35, 1798–1828.

Berkes, P., & Wiskott, L. (2006). On the analysis and interpretation of inhomogeneous quadratic forms as receptive fields. Neural Computation, 18, 1868–1895.

Bevan, N. (1995). Measuring usability as quality of use. Software Quality Journal, 4, 115–130.

Biffi, C., Oktay, O., Tarroni, G., Bai, W., De Marvao, A., Doumou, G., ... Rueckert, D. (2018). Learning interpretable anatomical features through deep generative models: Application to cardiac remodeling. Paper presented at International conference on medical image computing and computer-assisted intervention (pp. 464–471). Springer.

Bologna, G., & Hayashi, Y. (2017). Characterization of symbolic rules embedded in deep dimlp networks: A challenge to transparency of deep learning. Journal of Artificial Intelligence and Soft Computing Research, 7, 265–286.

Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. Paper presented at 21th ACM SIGKDD international conference on knowledge discovery and data mining (KDD '15) (pp. 1721–1730). ACM.

Doran, D., Schulz, S., & Besold, T. R. (2017). What does explainable AI really mean? A new conceptualization of perspectives. arXiv:1710.00794.

19424795, 2019, 4, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/widm.1312 by INASP/HINARI - PAKISTAN, Wiley Online Library on [11/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Erhan, D., Bengio, Y., Courville, A., & Vincent, P. (2009). Visualizing higher-layer features of a deep network. University of Monetreal Technical Report Nr. 1341.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, 115–118.

Fellmeth, A. X., & Horwitz, M. (2009). Guide to Latin in international law. Oxford, England: Oxford University Press.

Gal, Y. (2016). Uncertainty in deep learning. Cambridge, England: University of Cambridge.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. Science,

349, 273–278.

Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P., & Holzinger, A. (2018). Explainable Ai: The new 42?. Paper presented at

Springer lecture notes in computer science LNCS 11015 (pp. 295–303). Springer.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., … Bengio, Y. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling,

C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), Advances in neural information processing systems (NIPS) (pp. 2672–2680). Montreal, Canada: Neural

Information Processing Systems Foundation.

Guillaume, S. (2001). Designing fuzzy inference systems from data: An interpretability-oriented review. IEEE Transactions on Fuzzy Systems, 9, 426–443.

Hamburg, M. A., & Collins, F. S. (2010). The path to personalized medicine. New England Journal of Medicine, 363, 301–304.

Hinton, G. E. (2012). A practical guide to training restricted boltzmann machines. In Neural networks: Tricks of the trade, lecture notes in computer science LNCS

(Vol. 7700, pp. 599–619). Heidelberg: Springer.

Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2018). Metrics for explainable AI: Challenges and prospects. arXiv:1812.04608.

Holzinger, A. (2005). Usability engineering methods for software developers. Communications of the ACM, 48, 71–74.

Holzinger, A. (2014). Trends in interactive knowledge discovery for personalized medicine: Cognitive science meets machine learning. IEEE Intelligent Informatics

Bulletin, 15, 6–14.

Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop?Brain Informatics, 3, 119–131.

Holzinger, A. (2017). Introduction to machine learning and knowledge extraction (MAKE). Machine Learning and Knowledge Extraction, 1, 1–20 http://www.mdpi.

com/2504-4990/1/1/1

Holzinger, A (2018). From machine learning to explainable AI. Paper presented at 2018 World symposium on digital intelligence for systems and machines (DISA).

Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2017). What do we need to build explainable AI systems for the medical domain? arXiv:1712.09923.

Holzinger, A., Dehmer, M., & Jurisica, I. (2014). Knowledge discovery and interactive data mining in bioinformatics—State-of-the-art, future challenges and research

directions. BMC Bioinformatics, 15, I1.

Holzinger, A., Plass, M., Holzinger, K., Crisan, G. C., Pintea, C.-M., & Palade, V. (2017). A glass-box interactive machine learning approach for solving np-hard problems with the human-in-the-loop. arXiv:1708.01104.

Holzinger, A., Plass, M., Kickmeier-Rust, M., Holzinger, K., Crian, G. C., Pintea, C.-M., & Palade, V. (2018). Interactive machine learning: Experimental evidence for

the human in the algorithmic loop. Applied Intelligence, 1–14.

Hudec, M., Bednrov, E., & Holzinger, A. (2018). Augmenting statistical data dissemination by short quantified sentences of natural language. Journal of Official Statistics (JOS), 34, 981–1010.

Johnson, W. L. (1994). Agents that learn to explain themselves. Paper presented at Twelfth national conference on artificial intelligence (AAAI '94) (pp. 1257–1263).

AAAI.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349, 255–260.

Kendall, A., & Gal, Y. (2017). What uncertainties do we need in bayesian deep learning for computer vision? In Advances in neural information processing systems

(pp. 5574–5584). Long Beach, CA: Neural Information Processing Systems Foundation.

Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. PLoS Computational Biology,

10, e1003915.

Kieseberg, P., Weippl, E., & Holzinger, A. (2016). Trust for the doctor-in-the-loop. In European Research Consortium for Informatics and Mathematics (ERCIM) news:

Tackling big data in the life sciences (Vol. 104, pp. 32–33). Sophia Antipolis, France: European Research Consortium for Informatics and Mathematics.

Komura, D., & Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. Computational and Structural Biotechnology Journal, 16, 34–42.

Lacave, C., & Diez, F. J. (2002). A review of explanation methods for Bayesian networks. The Knowledge Engineering Review, 17, 107–127.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. Behavioral and Brain Sciences, 40, e253.

Lakkaraju, H., Kamar, E., Caruana, R., & Leskovec, J. (2017). Interpretable and explorable approximations of black box models. arXiv:1707.01154.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521, 436–444.

Lee, H., Grosse, R., Ranganath, R., & Ng, A. Y. (2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. Paper

presented at 26th Annual international conference on machine learning (ICML '09) (pp. 609–616). ACM.

Lee, S., & Holzinger, A. (2016). Knowledge discovery from complex high dimensional data. In S. Michaelis, N. Piatkowski, & M. Stolpe (Eds.), Solving large scale learning tasks. Challenges and algorithms, lecture notes in artificial intelligence, LNAI 9580 (pp. 148–167). Heidelberg, Germany: Springer.

Lipton, Z. C. (2016). The mythos of model interpretability. arXiv:1606.03490.

Ma, J., Broda, K., Goebel, R., Hosobe, H., Russo, A., & Satoh, K. (2010) Speculative abductive reasoning for hierarchical agent systems. Paper presented at International workshop on computational logic in multi-agent systems (pp. 49–64). Springer.

McCarthy, J. (1960). Programs with common sense (pp. 75–91). RLE and MIT Computation Center.

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (1984). Machine learning: An artificial intelligence approach. Heidelberg, Germany: Springer.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv:1301.3781.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1–38.

Miller, T., Howe, P., & Sonenberg, L. (2017) Explainable AI: Beware of inmates running the asylum or: How i learnt to stop worrying and love the social and behavioural sciences. arXiv:1712.00547.

Montavon, G., Samek, W., & Müller, K.-R. (2017). Methods for interpreting and understanding deep neural networks. arXiv:1706.07979.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Chess-playing programs and the problem of complexity. IBM Journal of Research and Development, 2, 320–335.

Nguyen, A., Dosovitskiy, A., Yosinski, J., Brox, T., & Clune, J. (2016). Synthesizing the preferred inputs for neurons in neural networks via deep generator networks.

In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett (Eds.), Advances in neural information processing systems 29 (NIPS 2016) (pp. 3387–3395).

Barcelona, Spain: Neural Information Processing Systems Foundation.

Oord, A. v. d., Kalchbrenner, N., and Kavukcuoglu, K. (2016). Pixel recurrent neural networks. arXiv:1601.06759.

Pawlowski, N., Brock, A., Lee, M. C., Rajchl, M., & Glocker, B. (2017). Implicit weight uncertainty in neural networks. arXiv preprint arXiv:1711.01297.

Pearl, J. (2009). Causality: Models, reasoning, and inference (2nd ed.). Cambridge, England: Cambridge University Press.

Pearl, J. (2018). Theoretical impediments to machine learning with seven sparks from the causal revolution. arXiv:1801.04016.

Pearl, J., & Mackenzie, D. (2018). The book of why. New York, NY: Basic Books.

Peters, J., Janzing, D., & Schölkopf, B. (2017). Elements of causal inference: Foundations and learning algorithms. Cambridge, MA: MIT-Press.

Poole, D. L., Mackworth, A. K., & Goebel, R. (1998). Computational intelligence: A logical approach. New York, NY: Oxford University Press.

Popper, K. (1935). Die Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft. Wien, Austria: Springer-Verlag.

Poulin, B., Eisner, R., Szafron, D., Lu, P., Greiner, R., Wishart, D. S., … Anvik, J. (2006). Visual explanation of evidence with additive classifiers. In National conference on artificial intelligence (pp. 1822–1829). Cambridge, MA: MIT Press.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016a). Model-agnostic interpretability of machine learning. arXiv:1606.05386.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016b). Why should I trust you?: Explaining the predictions of any classifier. Paper presented at Proceedings of the 22nd

ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135–1144). ACM.

Rolnick, D., Veit, A., Belongie, S., & Shavit, N. (2017). Deep learning is robust to massive label noise. arXiv:1705.10694.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323, 533–536.

Russell, S. J., & Norvig, P. (2010). Artificial intelligence: A modern approach (3rd ed.). Upper Saddle River: Prentice Hall.

Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., & Langs, G. (2017). Unsupervised anomaly detection with generative adversarial networks to guide

marker discovery. In International conference on information processing in medical imaging (pp. 146–157). Heidelberg, Germany: Springer.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85–117.

Seeböck, P., Waldstein, S. M., Klimscha, S., Bogunovic, H., Schlegl, T., Gerendas, B. S., ... Langs, G. (2018). Unsupervised identification of disease marker candidates

in retinal oct imaging data. IEEE Transactions on Medical Imaging., 1.

Shortliffe, E. H., & Buchanan, B. G. (1975). A model of inexact reasoning in medicine. Mathematical Biosciences, 23, 351–379.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556.

Singh, D., Merdivan, E., Psychoula, I., Kropf, J., Hanke, S., Geist, M., & Holzinger, A. (2017). Human activity recognition using recurrent neural networks. In

A. Holzinger, P. Kieseberg, A. M. Tjoa, & E. Weippl (Eds.), Machine learning and knowledge extraction: Lecture notes in computer science LNCS 10410

(pp. 267–274). Cham: Springer.

Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. arXiv preprint arXiv:1703.01365.

Swartout, W., Paris, C., & Moore, J. (1991). Explanations in knowledge systems: Design for explainable expert systems. IEEE Expert, 6, 58–64.

Szafron, D., Lu, P., Greiner, R., Wishart, D. S., Poulin, B., Eisner, R., ... Fyshe, A. (2004). Proteome analyst: Custom predictions with explanations in a web-based tool

for high-throughput proteome annotations. Nucleic Acids Research, 32, W365–W371.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. Science, 331, 1279–1285.

Ting, D. S. W., Cheung, C. Y.-L., Lim, G., Tan, G. S. W., Quang, N. D., Gan, A., ... Lee, S. Y. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. Jama, 318, 2211–2223.

Turkay, C., Jeanquartier, F., Holzinger, A., & Hauser, H. (2014). On computationally-enhanced visual analysis of heterogeneous data and its application in biomedical

informatics. In A. Holzinger & I. Jurisica (Eds.), Interactive knowledge discovery and data mining: State-of-the-art and future challenges in biomedical informatics.

Lecture notes in computer science LNCS 8401 (pp. 117–140). Heidelberg, Germany: Springer.

Widrow, B., & Lehr, M. A. (1990). 30 years of adaptive neural networks: Perceptron, madaline, and backpropagation. Proceedings of the IEEE, 78, 1415–1442.

Xu, Y., Zhu, J.-Y., Chang, E. I. C., Lai, M., & Tu, Z. (2014). Weakly supervised histopathology cancer image segmentation and classification. Medical Image Analysis, 18, 591–604.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. Paper presented at Proceedings of the IEEE conference on computer vision and pattern recognition (2921–2929).