1

Artificial Intelligence

Key Technologies and Opportunities

Wolfram Burgard

I. INTRODUCTION

Artificial Intelligence (AI) is a discipline that is concerned with the generation of software systems that provide functions, the execution of which requires what is typically referred to by the word intelligence. Thereby, the corresponding tasks can be performed by pure software agents as well as by physical systems, such as robots or self-driving cars.

As the term 'intelligence' is already very difficult to define, the definition of AI is, of course, correspondingly difficult and numerous definitions can be found in the literature.[1] Among them are several approaches that are based on human behavior or thinking. For example, the *Turing* test[2] introduced by *Alan Turing* in 1950, in which the actions generated by the system or robot should not be distinguishable from those generated by humans, has to be mentioned in this context. Such a *Turing* test for systems interacting with humans would then mean, for example, that a human could no longer determine whether a conversation partner on the telephone is a human or software.

However, most current AI systems aim to generate agents that think or act rationally. To realize systems that think rationally, often logic-based representations and reasoning systems are used. The basic assumption here is that rational thinking entails rational action if the reasoning mechanisms used are correct. Another group of definitional approaches deals with the direct generation of rational actions. In such systems, the underlying representations often are not human-readable or easily understood by humans. They often use a goal function that describes the usefulness of states. The task of the system is then to maximize this objective function, that is, to determine the state that has the maximum usefulness or that, in case of uncertainties, maximizes the future expected reward. If, for example, one chooses the cleanliness of the work surface minus the costs for the executed actions as the objective function for a cleaning robot, then in the ideal case this leads to the robot selecting the optimal actions in order to keep the work surface as clean as possible. This already shows the strength of the approach to generate rational behavior compared to the approach to generate human behavior. A robot striving for rational behavior can simply become more effective than one that merely imitates human behavior, because humans, unfortunately, do not show the optimal behavior in all cases. The disadvantage lies in the fact that the interpretation of the representations or structures learned by

---

[1] NJ Nilsson, *Artificial Intelligence: A New Synthesis* (1998); S Russell and P Norvig, *Artificial Intelligence: A Modern Approach* (4th ed. 2016).
[2] A Turing, 'Computing Machinery and Intelligence' (1950) 59 *Mind* 433.

the system typically is not easy, which makes verification difficult. Especially in the case of safety-relevant systems, it is often necessary to provide evidence of the safety of, for example, the control software. However, this can be very difficult and generally even impossible to do analytically, so one has to rely on statistics. In the case of self-driving cars, for example, one has to resort to extensive field tests in order to be able to prove the required safety of the systems.

Historically, the term AI dates back to 1956, when at a summer workshop called the Dartmouth Summer Research Project on Artificial Intelligence,[3] renowned scientists met in the state of New Hampshire, USA, to discuss AI. The basic idea was that any aspect of learning or other properties of intelligence can be described so precisely that machines can be used to simulate them. In addition, the participants wanted to discuss how to get computers to use language and abstract concepts, or simply improve their own behavior. This meeting is still considered today to have been extremely successful and has led to a large number of activities in the field of AI. For example, in the 1980s, there was a remarkable upswing in AI in which questions of knowledge representation and knowledge processing played an important role. In this context, for example, expert systems became popular.[4] Such systems used a large corpus of knowledge, represented for example in terms of facts and rules, to draw conclusions and provide solutions to problems. Although there were initially quite promising successes with expert systems, these successes then waned quite a bit, leading to a so-called demystification of AI and ushering in the AI winter.[5] It was not until the 1990s when mathematical and probabilistic methods increasingly took hold and a new upswing could be recorded. A prominent representative of this group of methods is Bayesian networks.[6] The systems resulting from this technique were significantly more robust than those based on symbolic techniques. This period also started the advent of machine learning techniques based on probabilistic and mathematical concepts. For example, support vector machines[7] revolutionized machine learning. Until a few years ago, they were considered one of the best performing approaches to classification problems. This radiated to other areas, such as pattern recognition and image processing. Face recognition and also speech recognition algorithms found their way into products we use in our daily lives, such as cameras or even cell phones. Cameras can automatically recognize faces and cell phones can be controlled by speech. These methods have been applied in automobiles, for example when components can be controlled by speech. However, there are also fundamental results from the early days of AI that have a substantial influence on today's products. These include, for example, the ability of navigation systems to plan the shortest possible routes[8] and navigate us effectively to our destination based on given maps. Incidentally, the same approaches play a significant role in computer games, especially when it comes to simulating intelligent systems that can effectively navigate the virtual environment. At the same time, there was also a paradigm shift in robotics. The probabilistic methods had a significant impact, especially on the navigation of mobile robots, and today, thanks to this development, it is well understood how to build mobile systems that move autonomously in their environment. This currently has an

[3] J McCarthy and others, 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955' (2006) 27(4) *AI Magazine* 12.
[4] F Hayes-Roth, DA Waterman, and DB Lenat, *Building Expert Systems* (1983).
[5] E Fast and E Horvitz, 'Long-Term Trends in the Public Perception of Artificial Intelligence' (2017) Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI).
[6] J Pearl, *Causality: Models, Reasoning and Inference* (2009) (hereafter Pearl, *Causality*).
[7] VN Vapnik, *Statistical Learning Theory* (1998) (hereafter Vapnik, *Statistical Learning Theory*).
[8] PE Hart, NJ Nilsson, and B Raphael, 'A Formal Basis for the Heuristic Determination of Minimum Cost Paths' (1968) 4(2) *IEEE Transactions on Systems Science and Cybernetics* 100 (hereafter Hart and others, 'A Formal Basis for the Heuristic Determination of Minimum Cost Paths').

important influence on various areas, such as self-driving cars or transport systems in logistics, where extensive innovations can be expected in the coming years.

For a few years now, the areas of machine learning and robotics have been considered particularly promising, based especially on the key fields of big data, deep learning, and autonomous navigation and manipulation.

II. MACHINE LEARNING

Machine learning typically involves developing algorithms to improve the performance of procedures based on data or examples and without explicit programming.[9] One of the predominant applications of machine learning is that of classification. Here the system is presented with a set of examples and their corresponding classes. The system must now learn a function that maps the properties or attributes of the examples to the classes with the goal of minimizing the classification error. Of course, one could simply memorize all the examples, which would automatically minimize the classification error, but such a procedure would require a lot of space and, moreover, would not generalize to examples not seen before. In principle, such an approach can only guess. The goal of machine learning is rather to learn a compact function that performs well on the given data and also generalizes well to unseen examples. In the context of classification, examples include decision trees, random forests, a generalization thereof, support vector machines, or boosting. These approaches are considered supervised learning because the learner is always given examples including their classes.

Another popular supervised learning problem is regression. Here, the system is given a set of points of a function with the task of determining a function that approximates the given points as well as possible. Again, one is interested in functions that are as compact as possible and minimize the approximation error. In addition, there is also unsupervised learning, where one searches for a function that explains the given data as well as possible. A typical unsupervised learning problem is clustering, where one seeks centers for a set of points in the plane such that the sum of the squared distances of all points from their nearest center is minimized.

Supervised learning problems occur very frequently in practice. For example, consider the face classification problem. Here, for a face found in an image, the problem is to assign the name of the person. Such data is available in large masses to companies that provide social networks, such as Facebook. Users can not only mark faces on Facebook but also assign the names of their friends to these marked faces. In this way, a huge data set of images is created in which faces are marked and labelled. With this, supervised learning can now be used to (a) identify faces in images and (b) assign the identified faces to people. Because the classifiers generalize well, they can subsequently be applied to faces that have not been seen before, and nowadays they produce surprisingly good results.

In fact, the acquisition of large corpora of annotated data is one of the main problems in the context of big data and deep learning. Major internet companies are making large-scale efforts to obtain massive corpora of annotated data. So-called CAPTCHAs (Completely Automated Public *Turing* tests to tell Computers and Humans Apart) represent an example of this.[10] Almost everyone who has tried to create a user account on the Internet has encountered such CAPTCHAs. Typically, service providers want to ensure that user accounts are not registered *en*

---

[9] TM Mitchell, *Machine Learning* (1997).
[10] L Von Ahn and others, 'CAPTCHA: Using Hard AI Problems for Security' (2003) Proceedings of the 22nd International Conference on Theory and Applications of Cryptographic Techniques, EUROCRYPT'03, 294.

masse by computer programs. Therefore, the applicants are provided with images of distorted text that can hardly be recognized by scanners and optical character recognition. Because the images are now difficult to recognize by programs, they are ideal for distinguishing humans from computer programs or bots. Once humans have annotated the images, learning techniques can again be used to solve these hard problems and further improve optical character recognition. At the same time, this ensures that computer programs are always presented with the hardest problems that even the best methods cannot yet solve.

1. *Key Technology Big Data*

In 2018, the total amount of storage globally available was estimated to be about 20 zettabytes (1 zettabyte = 10^{21} byte = 10^9 terabytes).[11] Other sources estimate internet data transfer at approximately 26 terabytes per second.[12] Of course, predictions are always subject to large uncertainties. Estimates from the International Data Corporation assume that the total volume will grow to 160 zettabytes by 2025, an estimated tenfold increase. Other sources predict an annual doubling. The number of pages of the World Wide Web indexed by search engines is enormous. Google announced almost ten years ago that they have indexed 10^{12} different URLs (uniform resource locators, reference to a resource on the World Wide Web).[13] Even though these figures are partly based on estimates and should therefore be treated with caution, especially with regard to predictions for the future, they make it clear that huge amounts of data are available on the World Wide Web. This creates an enormous potential of data that is available not only to people but also to service providers such as Apple, Facebook, Amazon, Google, and many others, in order to offer services that are helpful to people in other contexts using appropriate AI methods. One of the main problems here, however, is the provision of data. Data is not helpful in all cases. As a rule, it only becomes so when people annotate it and assign a meaning to it. By using learning techniques, images that have not been seen before can be annotated. The techniques for doing so will be presented in the following sections. We will also discuss which methods can be used to generate this annotated data.

2. *Key Technology Deep Learning*

Deep learning[14] is a technique that emerged a few years ago and that can learn from massive amounts of data to provide effective solutions to a variety of machine learning problems. One of the most popular approaches is the so-called deep neural networks. They are based on the neural networks whose introduction dates back to *Warren McCulloch* and *Walter Pitts* in 1943.[15] At that time, they tried to reproduce the functioning of neurons of the brain by using electronic circuits, which led to the artificial neural networks. The basic idea was to build a network consisting of interconnected layers of nodes. Here, the bottom layer is considered the input layer, and the top

---

[11] D Reinsel, J Gantz, and J Rydning, 'Data Age 2025: The Evolution of Data to Life-Critical' (IDC White Paper, 2017) www.import.io/wp-content/uploads/2017/04/Seagate-WP-DataAge2025-March-2017.pdf.

[12] Ibid.

[13] J Alpert and N Hajaj, 'We knew the web was big. . .' (Google Blog, 2008) https://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html.

[14] I Arel, DC Rose, and TP Karnowski, 'Research Frontier: Deep Machine Learning – A New Frontier in Artificial Intelligence Research' (2010) 5(4) *IEEE Computational Intelligence Magazine* 1; Y LeCun, Y Bengio, and G Hinton, 'Deep Learning' (2015) 521 *Nature* 436.

[15] WS McCulloch and WH Pitts, 'A Logical Calculus of the Ideas Immanent in Nervous Activity' (1943) 5 *Bulletin of Mathematical Biophysics* 115.

layer is considered the output layer. Each node now executes a simple computational rule, such as a simple threshold decision. The outputs of each node in a layer are then passed to the nodes in the next layer using weighted sums. These networks were already extremely successful and produced impressive results, for example, in the field of optical character recognition. However, even then there were already pioneering successes from today's point of view, for example in the No Hands Across America project,[16] in which a minivan navigated to a large extent autonomously and controlled by a neural network from the east coast to the west coast of the United States. Until the mid-80s of the last century, artificial neural networks played a significant role in machine learning, until they were eventually replaced by probabilistic methods and, for example, Bayesian networks,[17] support vector machines,[18] or Gaussian processes.[19] These techniques have dominated machine learning for more than a decade and have also led to numerous applications, for example in image processing, speech recognition, or even human–machine interaction. However, they have recently been superseded by the deep neural networks, which are characterized by having a massive number of layers that can be effectively trained on modern hardware, such as graphics cards. These deep networks learn representations of the data at different levels of abstraction at each layer. Particularly in conjunction with large data sets (big data), these networks can use efficient algorithms such as backpropagation to optimize the parameters in a single layer based on the previous layer to identify structures in data. Deep neural networks have led to tremendous successes, for example in image, video, or speech processing. But they have also been used with great success in other tasks, such as in the context of object recognition or deep data interpretation. The deep neural networks could impressively demonstrate their ability in their application within AlphaGo, a computer program that defeated *Lee Sidol*, one of the best Go players in the world.[20] This is noteworthy because until a few years ago it was considered unlikely that Go programs would be able to play at such a level in the foreseeable future.

III. ROBOTICS

Robotics is a scientific discipline that deals with the design of physical agents (robotic systems) that effectively perform tasks in the real world. They can thus be regarded as physical AI systems. Application fields of robotics are manifold. In addition to classical topics such as motion planning for robot manipulators, other areas of robotics have gained increasing interest in the recent past, for example, position estimation, simultaneous localization and mapping, and navigation. The latter is particularly relevant for transportation tasks. If we now combine manipulators with navigating platforms, we obtain mobile manipulation systems that can play a substantial role in the future and offer various services to their users. For example, production processes can become more effective and also can be reconfigured flexibly with these robots. To build such systems, various key competencies are required, some of which are already available or are at a quality level sufficient for a production environment, which has significantly increased the attractiveness of this technology in recent years.

[16] C Thorpe and others, 'Toward Autonomous Driving: The CMU Navlab. I. Perception' (1991) 6(4) *IEEE Expert* 31.
[17] Pearl, *Causality* (n 6).
[18] Vapnik, *Statistical Learning Theory* (n 7).
[19] CE Rasmussen and CKI Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)* (2005).
[20] D Silver and others, 'Mastering the Game of Go with Deep Neural Networks and Tree Search' (2016) 529 *Nature* 484.

1. *Key Technology Navigation*

Mobile robots must be able to navigate their environments effectively in order to perform various tasks effectively. Consider, for example, a robotic vacuum cleaner or a robotic lawnmower. Most of today's systems do their work by essentially navigating randomly. As a result, as time progresses, the probability increases that the robot will have approached every point in its vicinity once so that the task is never guaranteed but very likely to be completed if one waits for a sufficiently long time. Obviously, such an approach is not optimal in the context of transport robots that are supposed to move an object from the pickup position to the destination as quickly as possible. Several components are needed to execute such a task as effectively as possible. First, the robot must have a path planning component that allows it to get from its current position to the destination point in the shortest possible path. Methods for this come from AI and are based, for example, on the well-known A[*] algorithm for the effective computation of shortest paths.[21] For path planning, robotic systems typically use maps, either directly in the form of roadmaps or by subdividing the environment of the robot into free and occupied space in order to derive roadmaps from this representation. However, a robot can only assume under very strong restrictions that the once planned path is actually free of obstacles. This is, in particular, the case if the robot operates in a dynamic environment, for example in one used by humans. In dynamic, real-world environments the robot has to face situations in which doors are closed, that there are obstacles on the planned path or that the environment has changed and the given map is, therefore, no longer valid. One of the most popular approaches to attack this problem is to equip the robot with sensors that allow it to measure the distance to obstacles and thus avoid obstacles. Additionally, an approach is used that avoids collisions and makes dynamic adjustments to the previously planned path. In order to navigate along a planned path, the robot must actually be able to accurately determine its position on the map and on the planned path (or distance from it). For this purpose, current navigation systems for robots use special algorithms based on probabilistic principles,[22] such as the *Kalman* filter[23] or the particle filter algorithm.[24] Both approaches and their variants have been shown to be extremely robust for determining a probability distribution about the position of the vehicle based on the distances to obstacles determined by the distance sensor and the given obstacle map. Given this distribution, the robot can choose its most likely position to make its navigation decisions. The majority of autonomously navigating robots that are not guided by induction loops, optical markers, or lines utilize probabilistic approaches for robot localization. A basic requirement for the components discussed thus far is the existence of a map. But how can a robot obtain such an obstacle map? In principle, there are two possible solutions for this. First, the user can measure the environment and use it to create a map with the exact positions of all objects in the robot's workspace. This map can then be used to calculate the position of the vehicle or to calculate paths in the environment. The alternative is to use a so-called SLAM (Simultaneous Localization and Mapping)[25] method. Here, the robot is steered through its environment and, based on the data gathered throughout this process, automatically computes the map. Incidentally, this SLAM

[21] Hart and others, 'A formal basis for the heuristic determination of minimum cost paths' (n 8).

[22] S Thrun, W Burgard, and D Fox, *Probabilistic Robotics* (2005) (hereafter Thrun and others, *Probabilistic Robotics*).

[23] RE Kalman, 'A New Approach to Linear Filtering and Prediction Problems' (1960) *ASME–Journal of Basic Engineering* 35.

[24] D Fox and others, 'Monte Carlo Localization: Efficient Position Estimation for Mobile Robots' (1999) Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI) 343.

[25] Thrun and others, *Probabilistic Robotics* (n 22).

technique is also known in photogrammetry where it is used for generating maps based on measurements.[26] These four components: path planning, collision avoidance and replanning, localization, and SLAM for map generation are key to today's navigation robots and also self-driving cars.

2. *Key Technology Autonomous Manipulation*

Manipulation has been successfully used in production processes in the past. The majority of these robots had fixed programmed actions and, furthermore, a cage around them to prevent humans from entering the action spaces of the robots. The future, however, lies in robots that are able to robustly grasp arbitrary objects even from cluttered scenes and that are intrinsically safe and cannot harm people. In particular, the development of lightweight systems[27] will be a key enabler for human–robot collaboration. On the other hand, this requires novel approaches to robust manipulation. In this context, again, AI technology based on deep learning has played a key role over the past years and is envisioned to provide innovative solutions for the future. Recently, researchers presented an approach to apply deep learning to robustly grasp objects from cluttered scenes.[28] Both approaches will enable us in the future to build robots that coexist with humans, learn from them, and improve over time.

IV. CURRENT AND FUTURE FIELDS OF APPLICATION AND CHALLENGES

As already indicated, AI is currently more and more becoming a part of our daily lives. This affects both our personal and professional lives. Important transporters of AI technology are smartphones, as numerous functions on them are based on AI. For example, we can already control them by voice, they recognize faces in pictures, they automatically store important information for us, such as where our car is parked, and they play music we like after analyzing our music library or learning what we like from our ratings of music tracks. By analyzing these preferences in conjunction with those of other users, the predictions of tracks we like get better and better. This can, of course, be applied to other activities, such as shopping, where shopping platforms suggest possible products we might be interested in. This has long been known from search engines, which try to present us with answers that correspond as closely as possible to the Web pages for which we are actually looking. In robotics, the current key areas are logistics and flexible production (Industry 4.0). To remain competitive, companies must continue to optimize production processes. Here, mobile robots and flexible manipulation systems that can cooperate with humans will play a decisive role. This will result in significantly more flexible production processes, which will be of enormous importance for all countries with large manufacturing sectors. However, robots are also envisioned to perform various tasks in our homes.

By 2030, AI will penetrate further areas: Not only will we see robots performing ever more demanding tasks in production, but also AI techniques will find their way into areas performed

---

[26] P Agarwal, W Burgard, and C Stachniss, 'Survey of Geodetic Mapping Methods: Geodetic Approaches to Mapping and the Relationship to Graph-Based SLAM' (2014) 21(3) *IEEE Robotics & Automation Magazine* 63.

[27] G Hirzinger and others, 'On a New Generation of Torque Controlled Light-Weight Robots' (2001) 4 *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* 3356.

[28] J Mahler and others, 'Dex-net 1.0: A Cloud-Based Network of 3d Objects for Robust Grasp Planning Using a Multi-Armed Bandit Model with Correlated Rewards' (2016) Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) 1957.

by people with highly qualified training. For example, there was a paper in *Nature* that presented a system that could diagnose skin cancer based on an image of the skin taken with a cell phone.[29] The interesting aspect of this work is that the authors were actually able to achieve the detection rate of dermatologists with their deep neural networks-based system. This clearly indicates that there is enormous potential in AI to further optimize processes that require a high level of expertise.

With the increasing number of applications of systems relying on AI technology, there is also a growing need for the responsibility or the responsible governance of such systems. In particular, when they can impose risks for individuals, for example in the context of service robots that collaborate with humans or self-driving cars that co-exist with human traffic participants, where mistakes of the physical agent might substantially harm a person, the demands for systems whose behavior can be explained to, or understood by, humans are high. Even in the context of risk-free applications, there can be such a demand, for example, to better identify biases in recommender systems. A further relevant issue is that of privacy. In particular, AI systems based on machine learning require a large amount of data, which imposes the question of how these systems can be trained so that the privacy of the users can be maintained while at the same time providing all the necessary benefits. A further interesting tool for advancing the capabilities of such systems is fleet learning, learning in which all systems jointly learn from their users how to perform specific tasks. In this context, the question arises of how to guarantee that no system is taught inappropriate or even dangerous behavior. How can we build such systems so that they conform with values, norms, and regulations? Answers to these questions are by themselves challenging research problems and many chapters in this book address them.

---

[29] A Esteva and others, 'Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks' (2017) 542(7639) *Nature* 115.