1 Introduction

The ability to act autonomously in the environment is a key feature of intelligence. An artificial intelligence (AI) acting system, for short an *actor*, is a computational artifact capable of autonomous operation in its environment. It can be a software system, such as a web-based service agent, or a robot embodied with sensory-motor devices. Actors require essential cognitive functions without which intelligence is hardly conceivable, and this book focuses on the functions of acting, planning, and learning:

- *Acting* is more than just the sensory-motor execution of low-level commands. There is always a need to decide *how* to perform each task, given the context, and to adapt online to changes in the environment.
- *Planning* involves choosing and organizing actions that can achieve a task or a goal. It usually involves reasoning on abstract models of a repertoire of actions the actor may perform.
- *Learning* is critical for acquiring knowledge about actions' actual effects, which actions to perform when, and how to perform and plan them. Conversely, acting and planning can be used to aid learning.

Combining these cognitive functions will be very important for the future of AI. To explain why, we briefly summarize some recent developments in AI research.

During the past few decades, AI has produced numerous success stories. However, most of them were costly, requiring huge development, modeling, and adaptation to their respective domains. They also tended to be brittle and narrow, with capabilities that were difficult to extend.[1] For many years, AI learning systems lacked a capability to adapt, generalize, and transfer to other domains. These adaptation capabilities, essential to intelligence, are beginning to be reached in two primary areas: data interpretation and data generation.[2]

- *Data interpretation.* Multi-layered neural networks have extended known principles to provide robust universal approximation classifiers. Moreover, they have incidentally provided, at several abstraction levels, representation features adapted to specific training data. For decades, the field of pattern recognition has devoted significant effort to designing representation features characterizing the data at hand. These features are now given for free as latent variables in the successive hidden layers of a neural net. They result from scalable training procedures, thanks to improvements in hardware performance and architectures. AI for data interpretation is no longer costly and narrow. It is

[1] An example is the Watson system [355], the impressive champion of the Jeopardy Q/A game, which was transposed to the medical domain but not successfully deployed despite huge investments.

[2] This oversimplifies a rich story. See, for example, [723, 756].

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

widely deployed for the analysis of all kinds of multi-modal data in numerous demanding applications, from astronomy to health and education.

- *Data generation.* Here also the principles have been known for a while: learn an adequate distribution for a domain, and sample from it for a given context. Generative sampling and prediction of the next term in a sequence have benefited from the progress in the performance and architectures of multi-layered networks. The recent multi-head attention transformer architecture of large language models, and their extensions in multimodal foundation models, have led to impressive performance in natural language processing and image generation tasks. They also demonstrate emergent but still fragile capabilities in other unexpected common-sense and reasoning tasks. Scalable AI tools for generating texts, images, videos, and sounds are now widely deployed.

From Data to Actions

This, we believe, is the next big, two-sided challenge for the field. On the one hand, AI has to pursue and leverage its successful achievements to transform current techniques for acting and planning into easily learned and scalable approaches. An actor should be able to extensively and efficiently learn how to act and how to plan. It should also be able to act and plan in order to better learn and adapt to its environment and mission. On the other hand, the challenge is to "put acting into AI." For example, the successful data interpretation and generation methods require numerous actions, such as to gather and select training data, choose meta-parameters, and so on. These should be part of the actions learned, planned for, and performed by the autonomous agent.

In two previous books, we wrote about automated planning [409] and about combining planning and acting [410]. The interactions between the acting, planning, and learning functions open essential perspectives for addressing the next big AI challenge. We hope through this book to contribute to the education and training of researchers and practitioners tackling this challenge.

The rest of this chapter is organized as follows. Section 1.1 presents a conceptual view of an AI actor, its architecture, and its main components. Section 1.2 introduces the types of models needed for the design of an actor. Section 1.3 expresses our concerns and recommendations about important ethical issues associated with autonomous actors. The outline and organization of the book are detailed in the Preface.

1.1 Architecture and Components of an Actor

This section introduces the main components and organization of a deliberative actor. It first presents a simplified, conceptual of view an actor's architecture. It then discusses the acting, planning, and learning functions and their interplay.

1.1.1 Architecture

The methods discussed in this book are relevant both for software actors and for actors embodied with sensory-motor devices. The latter are further detailed in Part VII on

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

motion and manipulation. A simplified view of an actor distinguishes two main parts: a deliberation part and an execution platform (see Figure 1.1).

The *execution platform* informs the actor about its environment and its current state. It transforms its commands into actuations that perform its actions (e.g., movements of a limb or of a virtual character). The platform of an embodied actor assembles sensors, actuators and signal processing functions. The actor has to control its platform (e.g., where to put and how to use its sensors and actuators). Hence, it needs a model of the platform's capabilities and limitations.

The *deliberation functions* are used to choose what to do and how to do it to achieve the actor's mission, how to react to changes in the environment, and how to interact with humans and other actors. We focus the book on the acting, planning, and learning functions. Other functions, namely perceiving, monitoring and goal reasoning, are briefly covered in Chapter 24. Communication, adaptation to, and interaction with other actors are also important. They are not developed *per se*, but Chapter 23 introduces large language models and discusses their possible use as deliberation functions.

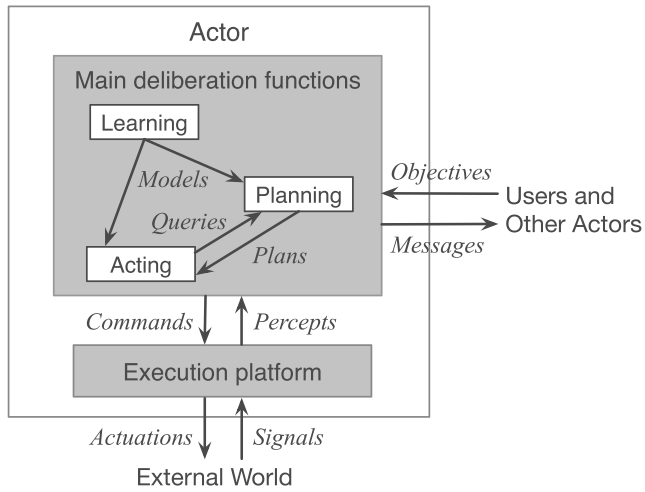

Figure 1.1 Conceptual architecture of an actor.

The architecture depicted in Figure 1.1 is a simplified conceptual schema that can be adapted to different classes of environments and actors. It presents the actor as a centralized system, while it can also be distributed. More importantly, there are two essential features, implicit in this figure:

- *Hierarchical processing within and across functions.* From abstract tasks to detailed actuations, a hierarchy of methods reduce the complexity of deliberation and integrate heterogeneous representations and models.
- *Continual online closed-loop adaptation.* The actor predicts what is expected; monitors what is taking place; reacts to events; extends, updates, and repairs its plan; and possibly revises its goals on the basis of its perception and deliberation.

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

These organizational principles provide a guideline to be adapted to different classes of environments and actors, about which the various parts of the book make different assumptions. Let us now discuss the main components of this architecture.

1.1.2 Planning

Planning is about *what* to do. It relies on a predictive model to foresee what may happen if some actions are performed, and a search over alternative options. It seeks to synthesize a plan, i.e., an organized set of actions that may lead, according to predictions, to a desired goal.

Planning problems vary in the kinds of actions, predictive models, and desired plans involved. In some cases, specialized planning methods can be used with specific problem representations. For instance, motion planning synthesizes a kinematic and dynamic trajectory for moving a device; perception planning generates sensing and interpretation actions to sense the world, or recognize or model an object or a scene.

In many cases, there are commonalities to various planning problems. Domain-independent planning tries to grasp these commonalities with abstract models. Domain-independent planners reason about actions by representing them uniformly as state-transformation operators over widely applicable representations of states as relations among objects.

Domain-independent and specialized planning are complementary. In a hierarchically organized actor, an abstract level can be tackled with domain-independent techniques, whereas lower levels may require specialized techniques. The integration of domain-independent and specialized planners raises several challenges, which are exemplified by the integrated task and motion planning problems in Part VII.

1.1.3 Acting

Acting is about *how* to do chosen actions while reacting, in a closed loop, to the observed context in which the activity takes place. An action is considered as a task to be progressively refined, given the current context, into more primitive actions and concrete commands. Whereas planning is a search over *predicted* states, acting is a continual assessment of the current *observed* state, and a consequent adaptation. Acting requires reacting to unexpected changes and exogenous events, which are independent from the actor's activity. It also requires a correct mapping between what is perceived and actuated and what is reasoned about for acting.

The techniques used in planning and acting can be compared as follows. Planning is organized as an *open-loop* search – a look-ahead process based on predictions. Acting is a *closed-loop* process, with feedback from observed effects and events used as input for subsequent decisions. Domain-independent planners can be developed to take advantage of commonalities among different forms of planning problems, but this is less true for acting systems, which require specific methods.

1.1.4 Interleaving Acting and Planning

Relationships between acting and planning are more complex than a simple linear sequence $\langle plan, act \rangle$. Seeking a complete plan before starting to act is not always

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

feasible, desirable, or needed. It is feasible when the environment is predictable and well modeled, as in a manufacturing production line. It is needed in domains with high costs or risks, or when actions are not reversible. In such domains, one often has to engineer the environment to reduce diversity as much as possible beyond what is modeled and can be predicted.

In open, dynamic domains with exogenous events that are difficult to model and fully predict, plans are expected to fail. They cannot be carried out blindly until the end. Plan modification and replanning are part of a global closed-loop process for acting. Replanning is normal and should be embedded in the design of an actor. Metaphorically, planning sheds light on the road ahead but does not lay an iron rail all the way to the goal.

The interplay between acting and planning can be organized in many ways, depending on how easy it is to plan, how predictable and dynamic the environment is, and how costly or risky the actions are. A general paradigm is the *receding-horizon* model of interleaved planning and acting. It consists of repeating the following two steps until the goal is reached:

1. Plan from the current state toward the goal, but not necessarily all the way to the goal, stopping at an arbitrary cutoff point called the *planning horizon*.
2. Perform one or more actions of the synthesized plan. Observe the current state and decide whether further planning is needed.

A receding-horizon scheme can have various instantiations. Options depend, for example, on the planning horizon, on what triggers replanning, on the number of actions performed after a planning stage, and on whether planning can be interrupted. Furthermore, the planning and acting procedures can be run either sequentially or in parallel with synchronization. A receding-horizon approach can scale up to large state spaces and can redirect the planning in a closed loop according to the results of acting. But it may also lead to situations from which the goal cannot be reached.

Depending on the planning horizon, the actor may execute each action as soon as it is planned or wait until a dynamically chosen planning horizon is reached. One should expect the observed state to differ from the predicted one and to evolve even if no action is executed. This may invalidate a plan and require replanning.

Interleaving acting and planning remains relevant if the planner synthesizes alternative courses of action for different contingencies (see Parts III and IV). It may not be worthwhile or even feasible to plan for all possible contingencies, or the planner may not know in advance what all of them are.

Several instances of the receding-horizon scheme will be illustrated throughout the book, including anytime approaches.

1.1.5 Learning

Learning is a very broad notion that includes many cognitive capabilities. An actor learns if it improves its performance with more autonomy and versatility, including ways to perform new tasks and adaptation to new or changing environments. Learning may rely on the actor's experiences, instructions from a tutor, and/or data and knowledge gathered from external sources.

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

Learning alleviates the costly efforts of programming an actor and specifying its environment. Even when such programming can be performed, it can hardly cover all the situations the actor may face, so adaptation by learning provides a significant advantage. Furthermore, learning allows an actor to acquire skills for which the designer may not have formalized knowledge or are difficult to program.[3]

An actor may want to learn a reactive function giving how to act in each situation and context, without further need of reasoning. Alternatively, it may want to learn models with which to reason for acting and planning. The former, called *end-to-end learning*, produces a reactive program that can be effective and efficient, and possibly amenable to continual adaptation; but it is usually a "black box" function, difficult to explain, verify, or validate. The latter, in contrast, aims at acquiring explicit models that are predictive but not executable; they can support analysis and explanation.

For example, a robot collaborating with a human should be proven safe to its users. To be accepted as a co-worker, it should also be able to explain what it is doing and remain intelligible. End-to-end learning may be less adequate in that regard. However, it can be very useful for acquiring low-level reactive sensory-motor skills, e.g., for grasping and manipulation, with additional mechanisms for verification and validation. It can also be very useful for acquiring domain-dependent search heuristics for more efficient planning and acting.

1.1.6 Integrating Acting, Planning, and Learning

Acting, planning, and learning are connected in many different ways, seldom limited to a simple sequence $\langle learn, plan, act \rangle$. There is learning to plan and learning to act, but there are also acting to learn and planning to learn. Let us mention a few possible interplays between these three functions.

An actor learns by acting. It may have the leisure to act for the sole purpose of learning. Possibly it may simulate its training actions to learn at an affordable cost. However, it is always desirable for an actor to keep learning while pursuing its activities, so it can improve and better adapt to a changing environment whose learned models need to be updated. Learning can be done when the actor fails or when it can benefit from additional advice or knowledge.

An actor or its user may reason about better ways to learn – for example, by planning how to find states and activities that may be useful for learning. For example, *curriculum learning* targets a progressive and rationally organized learning program, or a well-organized training database [111], as would be elaborated by an educator. Learning to learn, or *meta-learning*, seeks to improve learning.

Often, an actor engaged in its tasks as well as in learning will have to find a tradeoff between learning more versus advancing in its task. This is the *exploration versus exploitation* tradeoff. An actor without much knowledge may favor exploration, while an expert actor may prefer to exploit known behaviors.

The *planning-to-learn* paradigm is important in this book. A learner can provide models and control knowledge, such as heuristics, to an online planning–acting duo.

[3] These are related to the notion of *tacit knowledge*, e.g., how to recognize a face or ride a bicycle, as opposed to *explicit knowledge*, such as scientific facts and models [574].

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

Conversely, a planner can synthesize a number of random cases of problems and solutions to feed a learner's training database. Planning can be used to create curricula for curriculum learning. In a continual-learning scheme, the actor's experiences are fed back to the original planner for use in additional training to improve what has been learned. These interactions, partially depicted in Figure 1.2, may possibly require different planners and interactions with a simulator as well as with the real world.

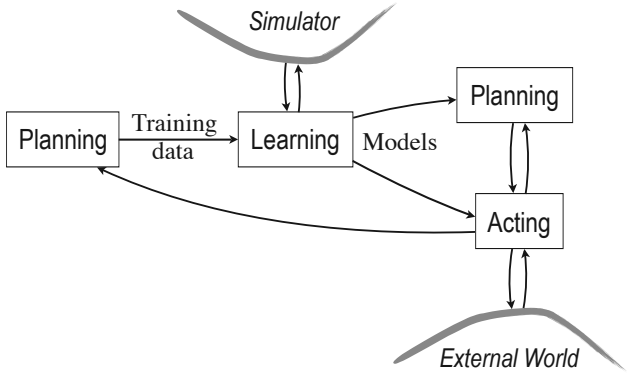

Figure 1.2 Interactions among acting, planning, and learning.

In some cases, a learner's output can be directly used for acting without additional planning. In these cases, the learner may synthesize a policy for reactive acting from a training database. This can be effective for focused and specialized functions, such as the sensory-motor control of a device. However, adaptation to a broad diversity of tasks and environments requires planning; hence it also requires learning for better planning, and possibly planning for learning as in the previous paragraph.

1.2 Descriptive and Operational Models of Actions

The book presents different models for acting, planning, and learning, starting from the simplest deterministic state-transition systems and proceeding to temporal, probabilistic, and nondeterministic cases. The formal representations used for expressing these models will be introduced when needed. Most of the chapters use discrete models, except for Part VII which uses continuous models of motion and manipulation.

Actors' models of actions can be classified into two types:

- *Descriptive models* specify *what* effects an action may have and *when* it is feasible. Descriptive models, also called causal models, are relations from the precondition to the effects of an action. The actor uses these models during planning, to reason about what actions may achieve the actor's objectives.
- *Operational models* specify *how* to perform an action: what commands to execute in the current context and how to organize them to achieve the action's intended effects. The actor uses these models during acting, to perform the actions that it has decided to perform.

Descriptive models are more abstract than operational models. They tend to ignore details and focus on the main effects needed to decide about the eventual use of an

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

action. For example, if you plan to take a book from a bookshelf, at planning time you usually are not concerned with the available space around the book to insert your fingers and extract the book. A descriptive model of an action abstracts away these details to focus on higher-level concerns, such as which shelf the book is on, whether it is within your reach, and whether you have a free hand with which to take it.

There are several reasons why these idealized abstract models are useful for planning. First, it is difficult to develop very detailed descriptive models. Second, these models may require information that is unknown during planning. Third, reasoning with detailed models is computationally very complex. Planners often need to search over many different combinations of actions, and if such a planner were to use operational (rather than descriptive) models for this search, it may run very slowly.

Operational models of how to perform actions cannot work with the simplifications allowed in descriptive models. To pick up a book on a shelf, you will need to determine precisely where the book is located, whether you need to remove an obstacle to reach the book, which positions of your hand and fingers give a feasible grasp, and which sequences of precise motions and manipulations will allow you to perform the action.

Furthermore, operational models may need to include ways to respond to *exogenous* events, that is, events that occur because of external factors beyond the actor's control. For example, someone might be standing in front of the bookshelf, or the stool you intended to use to reach the book on a high shelf might be missing, or a potentially huge number of other possibilities might interfere with your plan.

In principle, descriptive models can take into account the uncertainty caused by exogenous events (see Parts III and IV). However, exogenous events are often ignored in descriptive models because it is impractical to try to model all of the possible joint effects of actions and exogenous events or to plan in advance for all of the contingencies. In operational models, however, the need to handle exogenous events is much more compelling. Operational models must have ways to respond to such events if they happen, because they can interfere with the achievement of an action. In the bookshelf example, you might need to ask someone to move out of the way, or you might have to stand on a chair instead of the missing stool.

Finally, an actor's hierarchical organization and continual online processing can be integrated in these two types of models. We may have a hierarchy of operational models, sketching how to perform abstract tasks and giving more detailed recipes for primitive actions. Similarly, we may have a hierarchy of descriptive models, from abstract tasks down to the effects of commands executable by the platform. Furthermore, deliberation may perform a continual and interleaved processing of operational models and descriptive models at different levels of the hierarchy. The book illustrates instances of these hierarchical models.

1.3 Responsible Research on Autonomous Actors

Autonomous deliberative actors are scientifically and technically challenging for AI. They are also ethically very challenging. We, and all contributors to AI, hold a particular responsibility regarding ethical issues. However, since no chapter of this

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

book is devoted to ethics, we felt important to clarify our position and concerns here, particularly regarding actor-centered AI.

Discussions of the ethics of AI are very active, with numerous publications, committees, and recommendations (see for example [174, 334, 404, 1107]). Most of these discussions deal with data-centered ethical concerns, such as biases, privacy, fairness, transparency, trustworthiness, or ownership. They have been triggered by the significant AI advances in data interpretation and data generation. They are certainly very important. They need to be pursued and implemented into regulations (beyond the GDPR[4]), institutions (e.g., data trusts), and active monitoring processes.

These data-centered ethical concerns are more focused on individuals than on embracing broader social considerations, such as social cohesion, values, and democratic organization, which are becoming even more critical with the development of autonomous acting systems. Actor-related ethical issues may have more vital impacts on humanity – but they have not been as widely studied, possibly because of a less advanced state of the art.

Some of the actor-centered ethical issues are related to the possible automation of many human activities, including rewarding qualified professional and creative jobs. Such a trend, in particular if fast and widespread, would create economic problems regarding employment, inequalities, and social wealth sharing. It would entail a questioning of our role in and value to society, and hence to ourselves. Feeling socially superfluous, because machines might do most of what many people can do, may lead to significant human and social turmoil. It may cause infringements on human dignity.

Human interactions have already changed with social networks. They are fast changing with conversational agents becoming language-fluent and apparently knowledgeable. They will further change with the advent of autonomous actors that have not only the capabilities described earlier but also capable sensory-motor skills, detailed knowledge of a person, and the ability nudge or prod them with respect to dubious utility criteria. This prospect raises the risk of reduced autonomy and infringements on human freedom and agency.

Autonomous actors may possibly amplify inequalities and further tilt the power imbalance between human groups and nations. Leaders may be more likely to engage in conflicts if they can do so with no risks to their soldiers. Weaponized actors are a very serious concern. Despite a call from many scientists to ban lethal autonomous weapons [377], now supported by the UN and other organizations, there is unfortunately for the moment no international agreement on these matters. Strong opposition from most powerful nations remains.

Autonomous actors may also be beneficial to our well-being and health, for example as long-life empathic, serviceable, and trustable companions. We need to remain proactively engaged toward these ends, but we must also keep in mind that the individual acceptance of a technology (even as a widespread market) is not equivalent to its social acceptance or acceptability. The latter must include, among other things, long-term effects, social cohesion and values, and environmental impacts.

Neither the best outcome nor the worst are the most probable. However, our current social organization, and the profit-and-power motive for much of its development, do

[4] EU General Data Protection Regulation.

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press

not lean naturally toward the best. To avoid the worst, we need to be well aware of the risks and be proactive in mitigating them.

A possible ambition is to seek machines aligned with human values [230, 964]. However, it is unclear whether it is feasible to have machines behaving with and enforcing our values, if their understanding of those values comes from our specifications or from observation of our inconsistent behaviors.[5] It is even more questionable whether we could put the risks of fast deployment on hold until we are able to have all of our AI machines human-aligned.

A more questionable option is to seek machines capable of moral choices. Machines do not have intrinsic motivations, desires, no feelings with respect to which moral choices are meaningful. So-called "ethics by design" can be quite misleading: techniques cannot solve everything, including our ethical choices and responsibilities. We certainly must improve and implement verification and validation methods toward provable trustworthiness, under appropriate assumptions. However, the responsibilities for designing, using, and allowing the deployment of AI actors remain ours. Researchers should not only be concerned with how AI should not be used for harmful purposes but also with how it can be used to promote positive values and counteract antidemocratic and deceptive practices.

It is well known that technology is ambivalent, with both good and ugly faces.[6] Everyone in society is, to some degree, responsible for harmful technical deployments. Scientists hold particular responsibilities because they can investigate and foresee long-term risks and search for mitigating means. They can disseminate knowledge and be active in social debates about these risks. For that, we believe, they have to remain cautiously optimistic. This optimism is justified by the numerous expressions of risk-related concerns published by AI scientists and developers, and their calls for effective oversight and open independent verifications. It is also justified by some more advanced regulations (for example, the recently approved *European AI Act*). We urge our responsible reader to remain actively vigilant.

[5] After centuries of moral effort, we have been able to state some of these values in documents such as the Universal Declaration of Human Rights. However, these rights are routinely violated, and we are still unable to enforce them.

[6] Hephaestus, the Greek god of technology, is described as a limping deity.

https://doi.org/10.1017/9781009579346.003 Published online by Cambridge University Press