Trends in Cognitive Sciences

CellPress REVIEWS

Series: Machine Behavior

Opinion

Artificial Intelligence and the Common Sense of Animals

Murray Shanahan,[1,2,*] Matthew Crosby,[2,*] Benjamin Beyret,[1,*] and Lucy Cheke[3,*]

The problem of common sense remains a major obstacle to progress in artificial intelligence. Here, we argue that common sense in humans is founded on a set of basic capacities that are possessed by many other animals, capacities pertaining to the understanding of objects, space, and causality. The field of animal cognition has developed numerous experimental protocols for studying these capacities and, thanks to progress in deep reinforcement learning (RL), it is now possible to apply these methods directly to evaluate RL agents in 3D environments. Besides evaluation, the animal cognition literature offers a rich source of behavioural data, which can serve as inspiration for RL tasks and curricula.

Highlights

Endowing computers with common sense remains one of the biggest challenges in the field of artificial intelligence (AI).

Most treatments of the topic foreground language, yet an understanding of everyday concepts such as objecthood, containers, obstructions, paths, etc. is arguably: (i) a prerequisite for language, and (ii) evident to some degree in non-human animals.

The recent advent of deep reinforcement learning (RL) in 3D simulated environments allows AI researchers to train and test (virtually) embodied agents in conditions analogous to the life of an animal.

With the right architecture, an RL agent inhabiting a simulated 3D world has the potential to acquire a repertoire of fundamental common sense concepts and principles, given suitable environments, tasks, and curricula.

Experimental protocols from the field of animal cognition can be repurposed for evaluating the extent to which an agent, after training, 'understands' a common sense concept or principle, in particular in a transfer setting.

Common Sense before Language

The challenge of endowing computers with common sense has been seen as a major obstacle to achieving the boldest aims of artificial intelligence (AI) since the field's earliest days [1] and it remains a significant problem today [2–6]. There is no universally accepted definition of common sense. However, most authors use language as a touchstone, following the example of [1], who stated that '[a] program has common sense if it automatically deduces for itself a sufficiently wide class of immediate consequences of anything it is told and what it already knows'. Consequently, tests for common sense are typically language based. For example, one such test uses so-called 'Winograd schemas' [7–9]. These are pairs of sentences that differ by a single word and contain an ambiguous pronoun whose resolution depends on understanding some aspect of common sense. Consider the sentences 'The falling rock smashed the bottle because it was heavy' and 'The falling rock smashed the bottle because it was fragile'. The pronoun 'it' refers to the rock in the first sentence, but to the bottle in the second. We are able to resolve the pronoun correctly in each case because of our common sense understanding of falling and fragility. In this paper, by contrast, we will set language temporarily to one side and focus on common sense capacities that are also found in non-human animals. Our rationale is that these capacities are also the foundation for human common sense. They are, so to speak, conceptually prior to language and human language rests on the foundation they provide [10].

Consider the phrase 'The falling rock smashed the bottle'. To understand this sentence, you have to know what a rock is and what it means for something to fall. But to understand what a rock is, you have to know what an object is. To grasp what falling is, you have to understand motion and space. And to understand the relationship between falling and smashing, you have to understand causality. Indeed, an understanding of objects, motion, space, and causality is a prerequisite for understanding any aspect of the everyday world, not just falling and rocks (cf. [11] and [12], not to mention Kant [13]). Unfortunately, this foundational layer of common sense, which is a prerequisite for human-level intelligence, is lacking in today's AI systems. Yet, thanks to a combination of evolution and learning, it is manifest in many non-human animals, to a greater or lesser degree [14–20]. For this reason, as we aim to show in this paper, the field of animal cognition [21] has a lot to offer AI. This is especially true in a reinforcement learning (RL) context, where, thanks to progress in deep learning [22], it is now possible to bring the methods of comparative cognition

[1]DeepMind, London, UK
[2]Imperial College London, London, UK
[3]University of Cambridge, Cambridge, UK

*Correspondence:
m.shanahan@imperial.ac.uk (M. Shanahan),
m.crosby@imperial.ac.uk (M. Crosby),
beyret@google.com (B. Beyret), and
lgc23@cam.ac.uk (L. Cheke).

 https://doi.org/10.1016/j.tics.2020.09.002

directly to bear [23,24]. In particular, animal cognition supplies a compendium of well-understood, nonlinguistic, intelligent behaviour; it suggests experimental methods for evaluation and benchmarking; and it can guide environment and task design [25].

Deep RL

Until the mid-2010s, it barely even made sense to think of assessing the cognitive abilities of a real AI system (as opposed to a hypothetical system of the future) using the same methods that are used to assess the cognitive abilities of animals. Students of animal cognition can take for granted a number of background assumptions that do not necessarily apply for an AI system. These include facts that are so obvious they go entirely unnoticed, such as the embodiment of an animal and its situatedness in a 3D spatial environment within which it can move and that contains objects with which it can interact [10]. An equally obvious assumption that animal researchers can safely (and unconsciously) make is that their subjects are motivated by various basic needs and will therefore exhibit purposeful behaviour, rather than, say, simply doing nothing. None of these things is inherently true of an AI system. A disembodied digital assistant, such as Siri or Alexa, cannot be placed in a maze or presented with a box containing food. In the context of a robot, an embodied system that interacts with the real world, at least such a prospect makes sense. But most robots are programmed to carry out predefined tasks in highly constrained circumstances and to present one with a novel situation is more likely to result in inaction or catastrophe than to elicit interesting behaviour.

With the advent of deep RL, however, these background assumptions can be satisfied and the cognitive prowess of an AI system can be evaluated using methods that were designed for animals. The RL setting, wherein an agent learns by trial-and-error to maximise its expected reward over time (Box 1), precludes inactivity and permits any cognitive challenge to be presented by means of a suitably designed environment and reward function. Until recently, RL systems with high-dimensional input (such as vision) were impractical. But this changed in the mid-2010s, when RL was paired with deep neural networks [26,27], inaugurating a new subfield. Among other successes, this led to the development of AlphaGo, the first program to defeat a top-

Box 1. Model-Free and Model-Based Reinforcement Learning (RL)

Suppose an agent acts according to a policy that maps its current input (possibly along with its internal state) to a recommendation for action. The job of RL is to improve the agent's policy, through trial and error, so as to maximise expected reward over time. In computer science, the study of RL has produced a substantial body of computational techniques for solving this problem [68]. Recently, this field has been dominated by deep RL, wherein a deep neural network is trained to compute the function that maps inputs to actions [22]. This has led to success in domains with high-dimensional input spaces, for example, the stream of images from a camera.

RL methods (independently of whether they use deep neural networks) can be broadly categorised as either model-based or model-free ([68], Chapter 8). The models in question are transition models that map states and actions to successor states and rewards (or to distributions of successor states and rewards). That is to say, they allow the agent to predict the outcome of its (prospective) actions. In model-free RL, the agent learns and enacts its policy without reference to an explicit transition model. Model-free RL can be extremely effective and is the basis for many of the most impressive recent results in the field. However, if a transition model is available, or can be learned, then an agent can use it to simulate interaction with the environment (a form of inner rehearsal), without having to interact with the environment directly, enabling it to improve its policy offline [69] and/or to plan a course of actions prior to their execution [70].

Additionally, a good model of how the world works has general application, which enables transfer learning and promotes data efficiency [30,71]. For example, if an agent understands that a long, thin, rigid object affords moving a reward item that is otherwise inaccessible, then it can apply that understanding to sticks and tubes, even if it has never encountered such objects before. However, current model-based deep RL methods, when applied to visual input, typically predict only the next few frames and do so at the pixel level. To realise the full promise of a model-based approach, RL methods will need to operate on a more abstract level.

ranked player at the game of Go [28]. But the original breakthrough was DeepMind's DQN, a deep RL system that mastered a suite of Atari video games, playing many at superhuman level [29] (Box 2).

Notwithstanding its impressive performance on Atari games, DQN inherited a number of shortcomings from deep learning [2,4]. First, it is not data efficient. It has to play a much larger number of games to reach human-level performance than a typical human. Second, it is brittle. A trained network is not robust to small changes in the game that a human would barely notice, such as background colour or the sizes of objects. Third, it is inflexible. Nothing of what it has learned on one game can be transferred to another similar game. (As many commentators have argued, human prowess in these respects can, in part, be attributed to the ability to bring common sense priors to bear when learning a new game, such as our everyday understanding of objects, motion, collision, gravity, etc. [2,4,30,31].) Despite progress on all these issues, none of them is fully resolved.

Nevertheless, the arrival of deep RL has opened up the possibility of training an agent in a 3D virtual environment with (somewhat) realistic physics, whose input is the scene rendered from the agent's point of view, and whose output is a set of actions enabling the agent to move within the environment and interact with the objects it contains (Figure 1) [32–34]. These objects can include universal reward items analogous to food in the natural world, such as green spheres that yield positive reward when touched, then disappear as if consumed [23,24]. At a fundamental level, the predicament of such an agent can be considered analogous to that of an animal. Although animals also act on various forms of intrinsic motivation (including curiosity, which we are certainly not ruling out for our agents), we contend that any cognitive challenge can be presented to a situated, (virtually) embodied RL agent in the guise of obtaining an external reward of a single type.

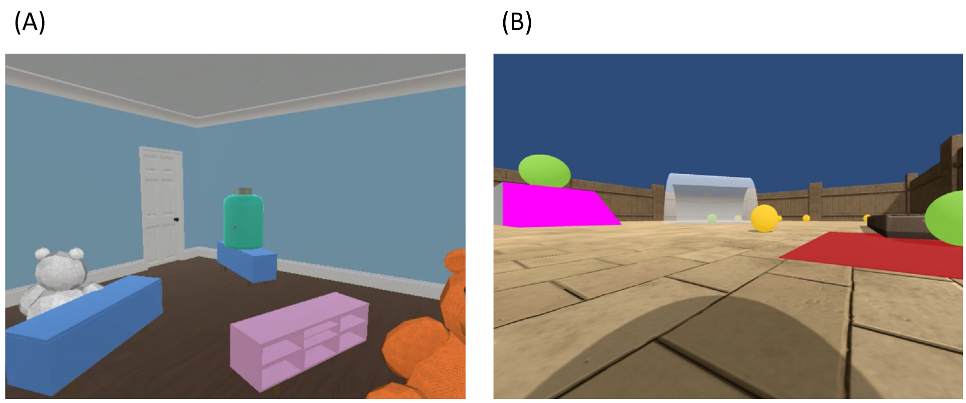
Moreover, to the extent that such an agent acquires an 'understanding' of any aspect of common sense, this can be thought of as grounded in its interaction with its world, even though that world is virtual not physical. (In the literature, the concept of grounding is typically associated with

Box 2. Training Protocols for Reinforcement Learning (RL)

A variety of protocols are used to train deep RL agents [22]. The simplest is the single task setting. The agent is presented with one task many times, such as the Atari game Space Invaders, and its performance slowly improves [29]. If the agent matches human performance, it is often said to have 'solved' the task. To solve a new task, such as the game Breakout, the agent is then re-initialised and learns it from scratch, losing its skill at Space Invaders in the process. In a multitask setting, by contrast, the agent learns many tasks together. If training is successful, the resulting agent will perform well on any of them. In the multitask setting, training tasks are often presented concurrently. Typically, episodes from different tasks are chopped up, interleaved, and stored in a replay buffer, and then presented to the learning component of the agent in a random order.

In a sequential multitask setting (an example of continual learning [57]), the agent learns tasks one at a time rather than concurrently. However, in contrast to the single task setting, the agent is not re-initialised after each task, and the final trained agent is expected to perform well on all the tasks [56]. From an engineering point of view, sequential multitask training is most difficult to get right, partly because of the phenomenon of catastrophic forgetting, wherein an agent trained on a new task loses its ability to perform well on tasks it was previously trained on [72,73].

Of course, the life of an animal is not divided into tasks (except perhaps in the laboratory). Rather, it is one long seamless episode. In this respect, the most realistic continual learning setting is one where there are no task boundaries, and this is arguably the most promising training protocol for acquiring the foundations of common sense. In RL terms, the objective is the same, to maximise expected reward over time, but the agent has a single, indefinitely extended 'life' rather than experiencing a sequence of discrete episodes. Despite the lack of task boundaries in this setting, it may be appropriate to structure the agent's life as a curriculum, where the agent has to become proficient on simpler tasks before it is confronted with more complex ones [74,75].



Trends in Cognitive Sciences

Figure 1. 3D Environments for Deep Reinforcement Learning Agents. (A) A screenshot from the DeepMind Playroom environment [43]. (B) A screenshot from the Animal-AI environment [23,24]. The basic setup is similar in both cases. What we see is the scene from the agent's point of view, that is to say, its virtual camera input. An agent can move about in the scene and push objects around. In the Playroom environment, the agent can also pick objects up and put them down and it obtains reward by successfully carrying out a natural language instruction, such as ' Put a teddy bear on a blue block '. In the Animal-AI environment, the agent receives reward for 'consuming' (moving over) green spheres, which function as a universal reward item.

symbols [35,36], so we are stretching the term somewhat here.) Even the most powerful contemporary language models trained purely on text, such as OpenAI's GPT-3, struggle with common sense inferences that a human would find straightforward [9]. Arguably, the lack of grounding for systems trained on text alone is an insuperable barrier to their ever fully matching humans in this respect.

Objects and Their Affordances

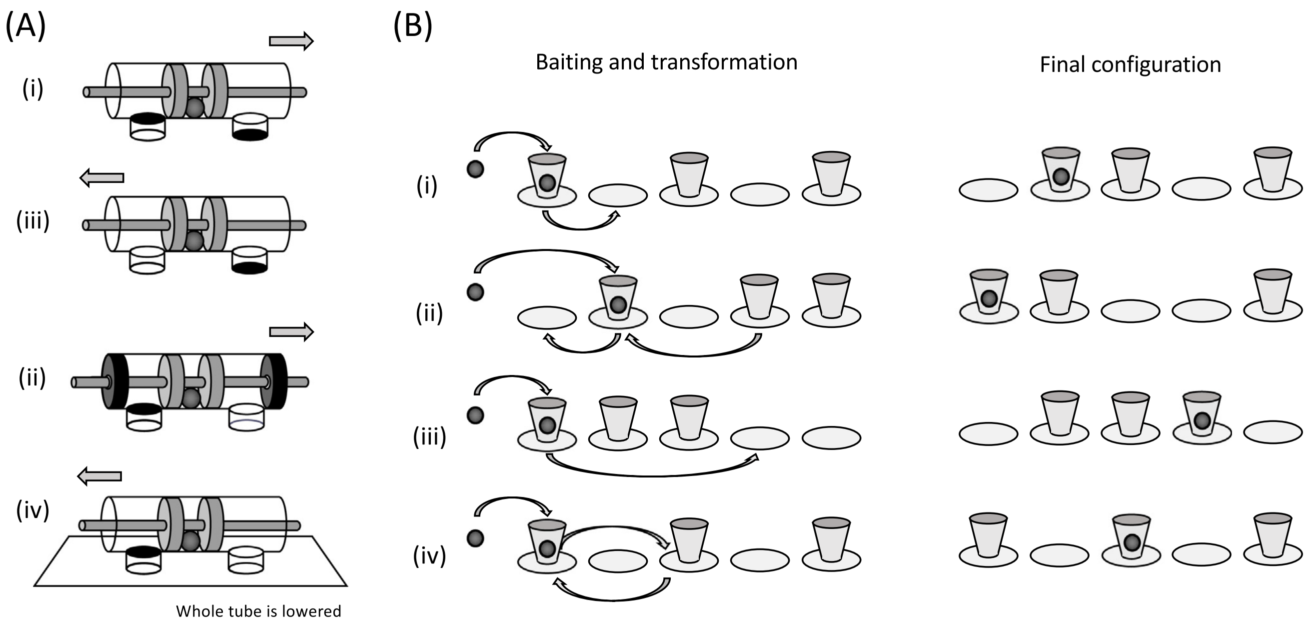
For many species, another assumption a researcher can safely make, particularly in the context of vision, is that an animal apprehends the world in terms of objects and surfaces. Gibson, for example, in advancing his well-known theory of affordances, contrasts orthodox psychology, according to which 'we perceive ... objects insofar as we discriminate their properties or qualities' with his own view that 'what we perceive when we look at objects are their affordances, not their qualities' ([37], Chapter 8). Even Gibson, whose approach to perception was so radical, takes it for granted that an animal perceives objects, that it will notice objects in its environment and respond to them *as objects*, rather than, say, as a superficial play of light with certain statistical regularities. What does it mean to respond to an object as an object? It means to be aware of the pattern of potential affordances presented by an object on account of the basic properties of objecthood, such as occupancy of a region of 3D space and persistence through time (or so-called 'object permanence'). For example, an object affords being moved by being pushed or pulled, or by being picked up, or, if already picked up, by being dropped again. An object affords investigation. An animal can look behind it or under it or inside it, potentially revealing other objects that were previously obscured or concealed.

Behaviourally speaking, an animal that is aware of this pattern of basic affordances will be able to exploit them directly, without having to discover them by accident or through exploration. Suppose we observe an animal that, on first encountering a new object, typically stares at it, then approaches it, then looks behind it or picks it up or pushes it. This should count as evidence that the animal perceives objects as objects. To be immediately aware of these basic affordances can be highly advantageous to an animal when presented with unfamiliar objects, as the literature on neophilia and object play will attest [38–41]. For example, in the so-called Aesop's fable task, birds are required to drop objects into a perspex tube containing water and a floating food item, in order to raise the water level and bring the food item within reach [17,18]. The overarching aim of the experiment is to establish whether the birds are able to distinguish between functional and nonfunctional objects with respect to the task, and to explore the extent to which they understand the causal relationships involved. But a prerequisite for success at any such task is the ability to recognise new objects as objects and to exploit the fact that they afford being picked up, carried to another location, and dropped. The experimental setup of the Aesop's fable task is typical in this regard. The same prerequisite must be met by the subjects of every experiment on animal tool-use.

Now, what about today's RL agents? Unlike many animals, it cannot be taken for granted that an RL agent *perceives* objects, properly speaking, at all. For sure, contemporary RL agents are capable of learning to perform complex tasks in 3D environments containing persistent objects (e.g., [42,43]). However, they can (and often do) accomplish this by exploiting correlations at the pixel level, rather than building on the common sense prior that the world contains persistent 3D objects that occupy space. The upshot is that, when faced with an unfamiliar object, such agents are not immediately able to exploit its basic affordances, the affordances it presents simply in virtue of being an object. In machine learning terms, such agents do not demonstrate effective out-of-distribution generalisation. In cognitive science terms, they exhibit poor transfer learning. By contrast, an RL agent that tends, not only to seek out and approach unfamiliar objects, but also to push and prod them, to turn them over, to pick them up, to carry them elsewhere, and so on, exhibits a more complete understanding of the basic affordances of objecthood and is consequently in a better position to transfer expertise acquired on one set of tasks to tasks and environments it has not previously seen ([44,45]; see also [46]).

To address this issue, it is not enough merely to learn perceptual grouping or scene segmentation [47,48]. This is certainly a step in the right direction, as it facilitates RL at a higher level of abstraction than the pixel, which can greatly reduce the amount of experience an agent needs to learn a new task [2,49]. (Although terms like 'object-level representation' and 'object-based RL' are sometimes used in this context [50], we are using the word 'object' in a more restricted sense here.) But to perceive an object as an object is much more than perceptual grouping. It is to engage a repertoire of behavioural responses that is systematically sensitive to objecthood, with all that entails [6]. It is to 'understand' that objects continue to exist when they are out of sight, that objects have 'backs' and 'underneaths'(even though these surfaces are not visible), that objects can be hidden behind other objects, that objects can be broken into parts, that objects have 'insides', and so on.

In animal cognition research, the claim that an animal 'understands' something is operationalised through carefully designed experimental protocols. For example, Seed *et al.* [15] investigated the extent to which rooks understood the causal properties of a trap-tube (Figure 2A). This is a transparent tube containing a food item that can be removed only by pushing (or pulling) a stick in the right direction; if the wrong direction is chosen, the food falls irretrievably into a trap. The authors wanted to establish whether birds that learned the task had done so through associative learning, or whether they understood the causal properties of the apparatus. To do this they gave the birds a number of transfer tasks, involving variations of the trap-tube with the same functional elements but differently arranged. An animal that solved the initial task by associative learning would fail the transfer condition. Some of the subjects, however, were successful in the transfer condition (within a few trials), which was taken as evidence of causal understanding.



Trends in Cognitive Sciences

Figure 2. Experimental Protocols from Animal Cognition That Test for Foundational Aspects of Common Sense That We Would Like Reinforcement Learning Agents to Acquire. (A) Testing for physical cognition using four variations of a trap-tube. If the stick is pulled out from the wrong end, the food item is lost. Having learned to solve variation (i) by trial-and-error, an animal that has only learned superficial associations will tend to perform poorly on the transfer tasks [variations (ii) to (iv))], while an animal that has acquired a causal understanding will tend to do well on first trial. (Adapted from [15].) (B) Testing for object permanence with an invisible displacement task. After baiting with a food item, the cups are moved as indicated into the final configuration on the right. The animal then has to select the cup containing the food. The cups are, of course, opaque, but the locations of the food item are shown for clarity. An animal that understands invisible displacement should do well at all four variations, even if it has never seen the objects involved before. (Adapted from [52].)

Evidence of 'understanding' in an RL agent (e.g., that objects continue to exist when they are out of sight, that objects have backs, that they have insides, and so on) must similarly be obtained under transfer conditions. So the challenge of endowing agents with such common sense principles can be cast as the problem of finding tasks and curricula that, given the right architecture, will result in trained agents that can pass suitably designed transfer tasks. These should be tasks that have not been encountered during training and that the agent will fail on if it has merely learned to exploit superficial correlations rather than grasping the way the world is structured at a deeper level. Moreover, for an agent that *has* grasped things at a deeper level, these should be tasks that can be solved zero- or few-shot, that is to say on first trial or after just a few trials. If the agent requires further training to solve a transfer task, this should not count as success.

Object Permanence

Behaviour such as poking at objects or flipping them over or pulling them apart, whether playful or not, often reveals other objects that can be interacted with in turn. But an animal or RL agent that exhibits such behaviour might have little or no grasp of the fundamental properties of object permanence and spatial occupancy that underpin this possibility. These properties are connected with one of the most fundamental aspects of common sense in humans: each person's understanding that there exists a world that is independent from them, a world that is spatially organised and contains enduring objects, one of which is their own body. Let us call this understanding common sense objectivity. A person does not have to be a philosopher to have common sense objectivity. It is manifest in their behaviour and it can be manifest to a greater or lesser degree in the behaviour of other animals. As such, it can be assessed, in humans, in other animals, and in RL agents. One hallmark of common sense objectivity is so-called 'object permanence', the understanding that objects continue to exist even when they are no longer perceived.

Drawing on work in developmental psychology from Piaget onwards, animal cognition researchers employ a number of experimental paradigms, to test for different levels of object permanence [51,52]. To test an animal's understanding of visible displacement, a rewarding object is simply moved out of sight while the subject is watching, pushed behind a screen, for example, or placed in a cup, and the subject passes the test if it searches for the object in the location where it disappeared. However, it is possible for an animal to learn through trial-and-error to pass such tests by searching in the place the object was last seen, which requires no conception of the object's continuing existence. A more demanding test involves invisible displacement, where the object is moved again after it has gone out of sight [51,52] (Figure 2B). For example, a rewarding object might be placed in one of three empty containers, while the subject is watching but prevented from approaching them. The container with the reward is then swapped with one of the empty containers, again while the subject is watching. The subject is then allowed to approach the containers. If the subject searches for the reward in the correct container, it suggests not only that they understand that the object still exists when out of sight, but also that they understand the physical constraints that govern its hidden trajectory.

Experimental paradigms designed to study object permanence in animals can readily be adapted to assess the same capacity in an RL agent that inhabits a 3D virtual world with universal reward items. However, as with all RL work, it is essential to differentiate training from testing. As Shettleworth notes, ' immediate accurate performance in a novel situation is necessary to rule out stimulus generalisation of previously reinforced behaviours. ... Here this means animals that have reached a given stage of object permanence should display evidence of it with novel objects and occluders '([21], p . 406). Moreover, when exporting these experimental paradigms from animal cognition to RL, it is also important to respect the fact that they were originally designed to probe pre-established cognitive capacities and not as training regimes in their own right. Training an RL agent to pass an object permanence test on one style of container, then testing it on another, would be unconvincing. A far more persuasive demonstration would be to train it on an altogether different task involving hidden objects and then, without further training, to test it using a standard paradigm from the animal cognition literature [23,24].

Containers and Enclosures

A great deal that matters in everyday life follows from the simple fact that the surface of one solid object is a barrier to the passage of another. There are abundant examples connected with obtaining food. The shell of a hazelnut precludes the immediate consumption of the kernel, obliging a squirrel to prise it apart first; a cavity in the bark of a tree affords enough protection to a grub to force a crow to use a twig to extract it; and to access the contents of their refrigerator, a person must first open the door. Each of these is an example of extractive foraging, 'the process by which an animal extracts food that is not directly perceptible from a substrate or shell' ([53], p. 77) and each features some sort of container, the nut-shell, the cavity in the bark, the refrigerator.

A container is an object whose interior surfaces present a barrier to the passage of its contents, resulting in a distinctive pattern of affordances [11,54]. They afford the confinement of their contents, preventing their escape, while simultaneously affording restricted access to those contents. This may necessitate moving them through an aperture (as with the grub in the bark) or the partial destruction of the container itself (as with the nut-shell). An RL agent that understood the general concept of a container would be able to apply it in a wide variety of situations, even if those situations superficially bore little resemblance to anything it had experienced before [55]. Since containers in various forms are ubiquitous, this would be a very useful fragment of common sense to have.

A viable approach to endowing an RL agent with such a concept is to train it to solve a large suite of tasks that involve using a variety of containers in a number of different ways. Although contemporary deep RL agents can learn to solve multiple tasks very effectively (e.g., [56]), and some architectures show rudimentary forms of transfer (e.g., [57]), it is far from clear that any current RL architecture is capable of acquiring such an abstract concept. But suppose we had a candidate agent, how would we test whether it had acquired the concept of a container?

Once again, the animal cognition literature is a valuable source of inspiration here, as well as a useful guide to experimental practice. Let us consider the squirrel again. The fact that a squirrel is adept at cracking nuts provides little support for the claim that it understands that nut-shells are containers, let alone that it grasps the concept of a container in a more abstract sense. Successful extractive foraging behaviour can be innate, or learned by trial-and-error, or some combination of the two. Analogously, good performance from an RL agent on a task that it was trained on does not support the hypothesis that it has acquired a general understanding of any aspect of the task domain. Evidence for such a claim has to come from performance on new tasks with unfamiliar objects, in other words in a transfer setting.

In an animal context, successful extractive foraging with novel objects provides more convincing support for the claim that an animal understands containers [58]. For example, keas (a species of parrot endemic to New Zealand) are especially competent at extracting food from novel objects, in both laboratory and urban settings [59]. But in such cases, an animal might be responding to cues, such as cracks, seams, holes, or ruptures, that are also found in familiar, naturally occurring container-like objects. An animal without any understanding of object interiority might still have a propensity to probe or work at these features, leading to success in extracting food. So further experimental assays would be required before a plausible case could be made that a species understood the high-level concept of a container or the general principle that objects have insides.

Now, what are the lessons for AI here? First, as already suggested, RL agents should be trained on large suites of varied tasks that involve the opening of container-like objects of various sorts to obtain a reward item. A prerequisite for this, of course, is the availability of 3D simulated worlds with sufficiently realistic physics to include, for example, objects with shells that can be cracked or prised apart, lids that can be unscrewed or levered off, packets that can be torn open, and so on. This is within the technological capabilities of today's physics engines, but such rich and realistic environments have yet to be deployed at scale for the training of RL agents. Second, trained agents need to be tested in a transfer setting. Performance on the training tasks indicates very little. Rather, transfer tasks need to be devised that are solvable for an agent in possession of the target common sense concept, but that differ from the training tasks as much as possible in other respects, and that control for alternative explanations for an agent's competence. The more they differ, the greater the transfer gap, so to speak, the more credit is due to a successful agent.

However, to properly get to grips with even this small fragment of common sense, we need to go much further. Extractive foraging is a response to just one aspect of containerhood. A container is also good for transporting things, for example [60]. (The invisible displacement tasks in Figure 2B exemplify this.) Ideally, an agent should understand that a container's usefulness for transporting things is a consequence of the same underlying principle of common sense that accounts for the difficulty of extractive foraging: the surface of one solid object is a barrier to the passage of another. The very same property of solid surfaces that prevents the contents of a container from falling out while it is being carried also makes it challenging to get at the contents of a container when it is closed. Does a bird that fills its crop with seeds understand that its crop is a container, analogous on one level to a nut-shell? This seems unlikely. But humans understand this and eventually we want our agents to as well.

In a similar vein, an enclosure is a kind of container, in which the role of contents is taken by the animal or agent itself. Just as objects have insides and surfaces have other sides, therefore enclosures have outsides. Animals have a notorious propensity to try to escape from any form of captivity, as Thorndike described in one of the earliest works on animal intelligence ([61]; see also [62], Chapter 7, and [63]). But does this mean they have a grasp of the common sense principle that surfaces have other sides? Combining the concept of container-as-carrier with the concept of container-as-enclosure, does a dog riding in a car understand that the car is a container and that it (the dog) is part of its contents? Whatever the case in respect of animals, we eventually want the answer to be affirmative for our trained agents, as it is for humans, which means the considerations just articulated in the context of extractive foraging apply. The key is a large and varied curriculum of relevant training tasks and a set of held-out transfer tasks for testing.

Concluding Remarks

In this review we have discussed a few fragments of common sense physics (objectivity and object permanence, containers and enclosures) so as to make a number of methodological points about training and evaluating RL agents. But the overall picture is, of course, much larger. Physics is just one common sense domain. We have neglected the whole domain of common sense psychological concepts (such as believing something or being unhappy) and the equally large (and related) domain of common sense social concepts (such as being with another agent, or giving something to someone). Within the domain of physics, our emphasis has been on solid objects, which are a special kind of thing. A more complete account would include liquids (puddles, streams, wine in a bottle), gaseous substances (smoke, mist, flames), and particulate matter (soil, sand) and would place more emphasis on deformable objects (sponges, paper,

Outstanding Questions

How do we build common sense objectivity, the understanding that there exists an external world that is independent of the agent, into an RL agent's cognitive make-up?

What RL architectures support the sorts of powerful generalisation needed to acquire an abstract common sense concept, such as that of a container?

How do we design environments, tasks, and curricula that will enable RL agents with the right architecture to acquire a repertoire of foundational common sense concepts and principles?

How can we precisely quantify a transfer gap, the difference between the environments and tasks an agent has been exposed to during training and those it can cope with after training?

Box 3. Aspects of Common Sense Physics

Let us call the fact that the surface of one solid object is a barrier to the passage of another the obstruction principle. Common sense concepts and principles like this can be viewed in terms of a hierarchy of generality. The obstruction principle, like the principle of object permanence, is foundational and has extremely wide application. The concept of a container is more specific and is one of many common sense concepts that derive from the obstruction principle.

Another important principle of common sense is that things tend to fall unless prevented from doing so. From this, in combination with the obstruction principle, we get the concept of support, which enables humans and other animals to make sense of tables, stairs, and stacks of blocks. As with other foundational common sense concepts, we would like to equip our RL agents with a general understanding of the concept of support to facilitate transfer of expertise from one domain to another [76,77]. For example, the concept of support is crucial for a proper understanding of the causal properties of a trap-tube (see Figure 2A in main text). It is the sudden absence of support (a hole in the tube) that allows the food item to fall when it is pulled from the wrong end and the presence of support (the base of the trap) that causes it to come to rest in an inaccessible place.

The role of contact in mediating motion, which is crucial to understanding the causal mechanisms implicated in tool-use, is another consequence of the obstruction principle. In a trap-tube, contact between the plunger and the food item allows it to be pulled out because the food item and the plunger cannot pass through each other. These little sequences of events in the trap-tube also illustrate another very general concept, one that complements the principle of obstruction: motion along a path. This concept is associated with the fundamental common sense principle that an object that moves from A to B (such as the food item moving along the tube) typically has to pass through a continuous series of points (a path) that connects A and B. The topology of 3D space implies that paths have a number of useful-to-know properties. For example, if an animal walks all the way around an object (such as a tree), it will end up back where it started.

string, clothing, leaves, twigs, animals' bodies) and empty spaces (holes, doorways, portals) [11,64,65].

Common sense can be thought of as an inter-related set of fundamental principles and abstract concepts (Box 3). Ideally, we would like to build AI technology that can grasp these inter-related principles and concepts as a systematic whole and that manifests this grasp in a human-level ability to generalise and innovate (see Outstanding Questions). How to build such AI technology remains an open question. But we advocate an approach wherein RL agents, perhaps with as-yet undeveloped architectures, acquire what is needed through extended interaction with rich virtual environments [4,24,49,66]. With the right architectures, environments, tasks, and curricula in place, we might at last be in a position to tackle language, where, thanks to analogy and metaphor, the foundational concepts of common sense find application at an altogether higher level of abstraction [67].

Acknowledgements

Thanks to Christos Kaplanis, Shane Legg, Drew Purves, David Reichert, and Bojan Vujatovic. This work was partly funded by the Leverhulme Centre for the Future of Intelligence.

References

1. McCarthy, J. (1959) Programs with common sense. In *Proceedings of the Teddington Conference on the Mechanization of Thought Processes*, pp. 75–91, Her Majesty's Stationary Office
2. Garnelo, M. *et al.* (2016) Towards deep symbolic reinforcement learning. *arXiv* 1609.05518
3. Davis, E. and Marcus, G. (2015) Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence. *Commun. ACM* 58, 92–103
4. Lake, B.M. *et al.* (2017) Building machines that learn and think like people. *Behav. Brain Sci.* 40
5. Marcus, G. and Davis, E. (2019) *Rebooting AI: Building Artificial Intelligence We Can Trust*, Ballantine Books Inc.
6. Smith, B.C. (2019) *The Promise of Artificial Intelligence: Reckoning and Judgment*, MIT Press
7. Levesque, H.J. *et al.* (2012) The Winograd schema challenge. In *KR'12: Proceedings of the Thirteenth International Conference on Principles of Knowledge Representation and Reasoning*, pp. 552–561
8. Sakaguchi, K. *et al.* (2020) WINOGRANDE: An Adversarial Winograd Schema Challenge at Scale. In *AAAI Conference on Artificial Intelligence*
9. Brown, T.B. *et al.* (2020) Language models are few-shot learners. *arXiv* 2005.14165
10. Shanahan, M. (2004) An attempt to formalise a non-trivial benchmark problem in common sense reasoning. *Artif. Intell.* 153, 141–165
11. Hayes, P.J. (1979) The Naive Physics Manifesto. In *Expert Systems in the Electronic Age* (Michie, D., ed.), pp. 242–270, Edinburgh University Press
12. Spelke, E.S. (2000) Core Knowledge. *Am. Psychol*. 2000, 1233–1243
13. Strawson, P.F. (1966) *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*, Methuen & Co.
14. Call, J. (2001) Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and children (*Homo sapiens*). *J. Comp. Psychol.* 115, 159–171
15. Seed, A.M. *et al.* (2006) Investigating Physical Cognition in Rooks, *Corvus frugilegus*. *Curr. Biol.* 16, 697–701
16. Taylor, A.H. *et al.* (2009) Do New Caledonian crows solve physical problems through causal reasoning? *Proc. R. Soc. B Biol. Sci.* 276, 247–254
17. Bird, C.D. and Emery, N.J. (2009) Rooks use stones to raise the water level to reach a floating worm. *Curr. Biol.* 19, 1410–1414
18. Cheke, L.G. *et al.* (2011) Tool-use and instrumental learning in the Eurasian jay (*Garrulus glandarius*). *Anim. Cogn.* 14, 441–455
19. Vallortigara, G. (2012) Core knowledge of object, number, and geometry: A comparative and neural approach. *Cogn. Neuropsychol.* 29, 213–236
20. Takagi, S. *et al.* (2016) There's no ball without noise: cats' prediction of an object from noise. *Anim. Cogn.* 19, 1043–1047
21. Shettleworth, S.J. (2010) *Cognition, Evolution, and Behavior*, Oxford University Press
22. Arulkumaran, K. *et al.* (2017) Deep reinforcement learning: a brief survey. *IEEE Signal Process. Mag.* 34, 26–38
23. Beyret, B. *et al.* (2019) The Animal-AI environment: training and testing animal-like artificial cognition. *arXiv* 1909.07483
24. Crosby, M. (2020) Building Thinking Machines by Solving Animal Cognition Tasks. *Mind. Mach.*
25. Versace, E. *et al.* (2018) Priors in animal and artificial intelligence: where does learning begin? *Trends Cogn. Sci.* 22, 963–965
26. LeCun, Y. *et al.* (2015) Deep learning. *Nature* 521, 436–444
27. Schmidhuber, J. (2015) Deep learning in neural networks: an overview. *Neural Netw.* 61, 85–117
28. Silver, D. *et al.* (2016) Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature* 529, 484–489
29. Mnih, V. *et al.* (2015) Human-level control through deep reinforcement learning. *Nature* 518, 529–533
30. Kansky, K. *et al.* (2017) Schema networks: zero-shot transfer with a generative causal model of intuitive physics. In *Proceedings International Conference on Machine Learning*, pp. 1809–1818
31. Dubey, R. *et al.* (2018) Investigating human priors for playing video games. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 1349–1357
32. Beattie, C. *et al.* (2016) DeepMind Lab. *arXiv* 1612.03801
33. Brockman, G. *et al.* (2016) OpenAI Gym. *arXiv* 1606.01540
34. Kempka, M. *et al.* (2016) ViZDoom: A Doom-based AI research platform for visual reinforcement learning. In *Proceedings IEEE Conference on Computational Intelligence and Games*
35. Harnad, S. (1990) The symbol grounding problem. *Phy. D Nonlinear Phenom.* 42, 335–346
36. Higgins, I. *et al.* (2018) SCAN: Learning Hierarchical Compositional Visual Concepts. In *Proceedings International Conference on Learning Representations*
37. Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*, Houghton Mifflin
38. Greenberg, R. and Mettke-Hofmann, C. (2001) Ecological aspects of neophobia and neophilia in birds. *Curr. Ornithol.* 16, 119–178
39. Paulos, R.D. *et al.* (2010) Play in wild and captive cetaceans. *Int. J. Comp. Psychol.* 23, 701–722

40. Bateson, P. (2014) Play, playfulness, creativity and innovation. *Anim. Behav. Cogn.* 1, 99–112
41. O'Hara, M. and Auersperg, A.M.I. (2017) Object play in parrots and corvids. *Curr. Opin. Behav. Sci.* 16, 119–125
42. Mirowski, P. *et al.* (2017) Learning to Navigate in Complex Environments. In *Proceedings International Conference on Learning Representations*
43. Hill, F. *et al.* (2020) Environmental drivers of systematicity and generalization in a situated agent. In *International Conference on Learning Representations. ICLR*
44. Haber, N. *et al.* (2018) Emergence of Structured Behaviors from Curiosity-Based Intrinsic Motivation. In *CogSci*, pp. 439–444
45. Watters, N. *et al.* (2019) COBRA: data-efficient model-based RL through unsupervised object discovery and curiosity-driven exploration. *arXiv* 1905.09275
46. Khetarpal, K. *et al.* (2020) What can I do here? A Theory of Affordances in Reinforcement Learning. In *Proceedings International Conference on Machine Learning*
47. Burgess, C.P. *et al.* (2019) MONet: unsupervised scene decomposition and representation. *arXiv* 1901.11390
48. Greff, K. *et al.* (2019) Multi-object representation learning with iterative variational inference. In *Proceedings International Conference on Machine Learning*
49. Lázaro-Gredilla, M. *et al.* (2019) Beyond imitation: Zero-shot task transfer on robots by learning concepts as cognitive programs. *Sci. Robot.* 4
50. Diuk, C. *et al.* (2008) An object-oriented representation for efficient reinforcement learning. In *Proceedings International Conference on Machine Learning*
51. Doré, F.Y. and Goulet, S. (1998) The comparative analysis of object knowledge. In *Piaget, Evolution, and Development* (Langer, J. and Killen, M., eds), pp. 55–72, Lawrence Erlbaum Associates Publishers
52. Jaakkola, K. (2014) Do animals understand invisible displacement? A critical review. *J. Comp. Psychol.* 128, 225–239
53. Tomasello, M. and Call, J. (1997) *Primate Cognition*, Oxford University Press
54. Davis, E. *et al.* (2017) Commonsense reasoning about containers using radically incomplete information. *Artif. Intell.* 248, 46–84
55. Hay, N. *et al.* (2018) Behavior Is Everything: towards representing concepts with sensorimotor contingencies. In *AAAI Conference on Artificial Intelligence*
56. Badia, A.P. *et al.* (2020) Agent57: outperforming the Atari human benchmark. In *Proceedings International Conference on Machine Learning*
57. Schwarz, J. *et al.* (2018) Progress & compress: a scalable framework for continual learning. In *Proceedings International Conference on Machine Learning*, pp. 4528–4537
58. Webster, S.J. and Lefebvre, L. (2001) Problem solving and neophobia in a columbiform-passeriform assemblage in Barbados. *Anim. Behav.* 62, 23–32
59. Huber, L. and Gajdon, G.K. (2006) Technical intelligence in animals: the kea model. *Anim. Cogn.* 9, 295–305
60. Davis, E. (2011) How does a box work? A study in the qualitative dynamics of solid objects. *Artif. Intell.* 175, 299–345
61. Thorndike, E.L. (1911) *Animal Intelligence: Experimental Studies*, Macmillan
62. Hediger, H. (1950) *Wild Animals in Captivity*, Butterworth-Heinemann
63. Wood, J.B. and Anderson, R.C. (2004) Interspecific evaluation of octopus escape behavior. *J. Appl. Anim. Welf. Sci.* 7, 95–106
64. P.J. Hayes, Naive Physics I: Ontology for Liquids, Morgan Kaufmann Publishers, 1989 pp. 484–502.
65. Davis, E. (2008) Pouring liquids: a study in commonsense physical reasoning. *Artif. Intell.* 172, 1540–1578
66. Garnelo, M. and Shanahan, M. (2019) Reconciling deep learning with symbolic artificial intelligence: representing objects and relations. *Curr. Opin. Behav. Sci.* 29, 17–23
67. Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*, University of Chicago Press
68. Sutton, R.S. and Barto, A.G. (2018) *Reinforcement Learning: An Introduction* (2nd edition), MIT Press
69. Kaiser, L. *et al.* (2020) Model Based Reinforcement Learning for Atari. In *International Conference on Learning Representations*
70. Racanière, S. *et al.* (2017) Imagination-augmented agents for deep reinforcement learning. In *Advances in Neural Information Processing Systems 30*, pp. 5690–5701
71. Hamrick, J.B. (2019) Analogues of mental simulation and imagination in deep learning. *Curr. Opin. Behav. Sci.* 29, 8–16
72. McCloskey, M. and Cohen, N.J. (1989) Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of Learning and Motivation* (Vol. 24), pp. 109–165, Elsevier
73. Kirkpatrick, J. *et al.* (2017) Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. U.S.A.* 114, 3521–3526
74. Elman, J.L. (1993) Learning and development in neural networks: the importance of starting small. *Cognition* 48, 71–99
75. Bengio, Y. *et al.* (2009) Curriculum Learning. In *Proceedings International Conference on Machine Learning*, pp. 41–48
76. Ullman, T.D. *et al.* (2017) Mind Games: game engines as an architecture for intuitive physics. *Trends Cogn. Sci.* 21, 649–665
77. Bapst, V. *et al.* (2019) Structured agents for physical construction. In *Proceedings International Conference on Machine Learning*, pp. 464–474