From Machine Learning to Robotics: Challenges and Opportunities for Embodied Intelligence

Nicholas Roy NICKROY@CSAIL.MIT.EDU
Massachusetts Institute of Technology, Cambridge, MA USA

Ingmar Posner INGMAR@ROBOTS.OX.AC.UK
University of Oxford, Oxford, UK

Tim Barfoot TIM.BARFOOT@UTORONTO.CA
University of Toronto, Toronto, ON Canada

Philippe Beaudoin PHIL@MYWAVERLY.COM
Waverly, Montreal, QC Canada

Yoshua Bengio YOSHUA.BENGIO@MILA.QUEBEC
Mila, University of Montreal, Montreal, QC Canada

Jeannette Bohg BOHG@STANFORD.EDU
Stanford University, Stanford, CA USA

Oliver Brock OLIVER.BROCK@TU-BERLIN.DE
Technische Universität Berlin, Berlin, Germany

Isabelle Dépatie ISABELLE.DEPATIE@ELEMENTAI.COM
Element AI – A ServiceNow Company, Montreal, QC Canada

Dieter Fox FOX@CS.WASHINGTON.EDU
University of Washington, Seattle, WA USA

Dan Koditschek KOD@SEAS.UPENN.EDU
University of Pennsylvania, Philadelphia, PA USA

Tomás Lozano-Pérez TLP@CSAIL.MIT.EDU
Massachusetts Institute of Technology, Cambridge, MA USA

Vikash Mansinghka VKM@MIT.EDU
Massachusetts Institute of Technology, Cambridge, MA USA

Christopher Pal CHRISTOPHER.PAL@SERVICENOW.COM
Element AI – A ServiceNow Company, Montreal, QC Canada

Blake Richards BLAKE.RICHARDS@MILA.QUEBEC
Mila, McGill University, Montreal, QC Canada

Dorsa Sadigh DORSA@CS.STANFORD.EDU
Stanford University, Stanford, CA USA

Stefan Schaal STEFAN.K.SCHAAL@GMAIL.COM
[Google] X, Mountain View, CA USA

Gaurav Sukhatme GAURAV@USC.EDU
University of Southern California, Los Angeles, CA USA

Denis Thérien DENIS@SERVICENOW.COM
ServiceNow, Montreal, QC Canada

Marc Toussaint TOUSSAINT@TU-BERLIN.DE
Technische Universität Berlin, Berlin, Germany

Michiel Van de Panne VAN@CS.UBC.CA
University of British Columbia, Vancouver, BC Canada

Abstract

Machine learning has long since become a keystone technology, accelerating science and applications in a broad range of domains. Consequently, the notion of applying learning methods to a particular problem set has become an established and valuable modus operandi to advance a particular field. In this article we argue that such an approach does not straightforwardly extended to robotics — or to *embodied intelligence* more generally: systems which engage in a purposeful exchange of energy and information with a physical environment. In particular, the purview of embodied intelligent agents extends significantly beyond the typical considerations of main-stream machine learning approaches, which typically (i) do not consider operation under conditions significantly different from those encountered during training; (ii) do not consider the often substantial, long-lasting and potentially safety-critical nature of interactions during learning and deployment; (iii) do not require ready adaptation to novel tasks while at the same time (iv) effectively and efficiently curating and extending their models of the world through targeted and deliberate actions. In reality, therefore, these limitations result in learning-based systems which suffer from many of the same operational shortcomings as more traditional, engineering-based approaches when deployed on a robot outside a well defined, and often narrow operating envelope. Contrary to viewing embodied intelligence as another application domain for machine learning, here we argue that it is in fact a key driver for the advancement of machine learning technology. In this article our goal is to highlight challenges and opportunities that are specific to embodied intelligence and to propose research directions which may significantly advance the state-of-the-art in *robot* learning.

1. Introduction

Robots and autonomous vehicles are being deployed with increasing frequency in an ever-increasing number of applications, from fleets of self-driving cars, to increasingly unpopulated factory environments to drones making commercial deliveries to people's houses. The expectation of the general public is that, more and more, robots are able to operate robustly and effectively, carrying out a range of tasks in a variety of complex domains. However, the reality is that today's robots are far from as robust, efficient and intelligent as they may seem. Today's robots are still brittle in the face of the unexpected, lack versatility, and are able to perform only a specific set of tasks within a very narrow set of operating conditions. It is an open secret that the vast majority of today's autonomous robots rely heavily on human supervision and intervention when deployed out in the world.

In the last ten years, there has been rapid progress on certain kinds of tasks in computer vision and natural language processing, driven by machine learning. For specific problems such as face recognition and machine translation, especially in the context of the web, engineered models have been no match for learned systems. Robotics and embodied systems have derived some benefit from machine learning in machine vision, and there are specific capabilities which enable systems such as self-driving cars that would not exist without learning. However, machine learning appears to have run into many of the same problems of brittleness and lack of versatility as more traditional engineering approaches when deployed on a robot operating outside well-defined conditions such as the warehouse, factory floor, or carefully mapped streets in a city with an unvarying, sunny climate.

The challenge of developing embodied, physical robots that can learn reliably suggests that learning for robots and embodied agents may somehow be different from the domains

2

where there has been greater operational success in using learning. In this article our goal is to highlight the challenges and opportunities that are specific to embodied intelligence and to propose research directions that may significantly advance the current state-of-the-art in robot learning.

The Challenges of Embodied Intelligence

Although intelligent agents that inhabit a real physical body are often referred to as "embodied", there is not a consensus as to what precisely makes an intelligence "embodied" (Pfeifer and Bongard, 2006; Cangelosi et al., 2015; Howard et al., 2019; Savva et al., 2019). Whether embodiment requires a physical body or whether the term can also apply to agents acting purely in simulation is a matter of ongoing debate. We therefore define embodied intelligence as *the purposeful exchange of energy and information with a physical environment* (Koditschek, 2021). We use the term "physical environment", rather than "real world" environment, in that a physical environment, simulated or otherwise, is one that enforces physical constraints of different kinds. In particular, our motivation lies in the field of robotics, where we are concerned with embodied agents which are equipped to act and interact in the real world and within the physical constraints imposed thereby. It has long been conjectured that physical constraints present an embodied intelligent agent with a very different learning landscape than traditional AI agents, and the technical requirements for enabling these agents to learn are fundamentally different from the requirements of learning agents that do not interact with a physical world.

Firstly, the fact that embodied agents exchange energy with the world implies that the agent's actions can have substantial and long-lasting effects, creating challenges of learning safely. Often unbeknownst to the agent, parts of the world may carry an extremely high penalty of exploration. Embodied agents must be aware of potentially catastrophic events that can fundamentally end any further learning or agency. Embodied agents must also be aware of their own limits in terms of available energy and specific power, and be able to learn in real-time. Nature provides existence proofs of embodied agents that can learn to execute complex tasks from a tiny number of examples without endangering themselves or the environment, but artificial embodied intelligence has struggled to replicate this ability. Behavioral studies of animal sensorimotor intelligence (Gallistel, 1990) and tool use (Seed and Byrne, 2010) reveal capacities for spatial intelligence and problem solving that outstrip our best autonomous systems not only in robustness and accuracy but also data efficiency and power efficiency.

Secondly, the physical world provides a much larger and richer source of information for training than any dataset or even most simulated worlds can provide. However, the data from the physical world are poorly aligned with the assumptions commonly made by most learning techniques. For example, learned classifiers often assume that the data are independent and identically distributed (IID) between training and test regimes. The assumption of IID data is critical to much of modern machine learning theory, but is fundamentally untrue of the physical world. Embodied agents live in a non-stationary, partially observable world where the data might be correlated through an intractably large number of latent factors. The data distributions are constantly changing, requiring the agent to learn over time what variations in the data might simply be the result of perceptual noise (or aliasing),

versus changes resulting from its own actions or the interventions of other agents, versus variations that might represent a fundamental change to the environment.

Thirdly, embodied agents cannot assume that their goals, specifications and rewards are fixed for all time. Following Simon (1956, 1996), it is useful to consider an embodied intelligent agent in terms of the range of tasks the agent must accomplish, the environments it must operate in, and how the tasks and environments affect the architecture of the agent's intelligence. Another of our conjectures is that a good architecture for an embodied agent is one that generalizes across tasks and environments, which implies that the agent's representation of tasks and specifications must allow generalization and rapid adaptation to new tasks and environments. Reinforcement learning and decision theory have historically represented tasks and goals through rewards and loss functions. However, reward functions are not easily adapted to substantial changes in the environment, nor are they easily adapted to more complex changes to the task than simply varying the goal state, and in fact may be an inefficient representation for allowing generalization.

Fourthly, the fact that the environment and tasks can change has implications for how the agent learns. While embodied agents may have access to orders of magnitude more data, the distribution of data at any one time is very local, and drawn from a much smaller measure than the true distribution of data the agent can encounter over its lifetime. An embodied agent must recognize the need to act to acquire specific kinds of data, not only to perform well at the current task, but also to build a rich enough theory of the world to allow the agent to generalize to future tasks and future environments. It may be precisely the feedback loop between action and inference that allows an embodied agent to learn a model that can make reliable predictions of the effects of actions, rather than merely learning to approximate the functional relationship between input and output.

Finally, the morphology of an agent is itself a decision variable for embodied intelligence (Lakoff and Johnson, 1980; Shapiro, 2007). What senses are available to the agent, the degrees of freedom of its actuators, the specific power available to the agent, all have tremendous implications about what the agent *can* learn about the world and what actions the agent can decide to take. The morphology of an agent also influences what the agent must learn about its environment, as the concrete embodiment can encode inductive biases that greatly facilitate learning.

The success of many robotic systems has been driven by the ability to constrain the operational environment and the set of tasks, avoiding many of the challenges described above. The constraints of a given environment and task enable an *inductive bias* in the model and the learner, and we can consider a spectrum of inductive biases (see Fig. 1). At one end of the spectrum, the environment and task are so constrained that a very strong inductive bias is possible, allowing learning to happen quickly and reliably from small amounts of data. The price we pay for such success is a high degree of specialization in the resulting system, with little generalization. Towards the other end of the spectrum, the inductive biases are significantly weaker (or stronger inductive biases remain elusive), leading to greater flexibility of the learner but also requiring increasing amounts of training data. At this end, in the limit, nothing is known and everything has to be learned. In reality, roboticists commonly find themselves somewhere between these extrema. We may assume *some* prior knowledge, but we can consider a system truly autonomous when it can function long enough such that almost every prior assumption is eventually violated.
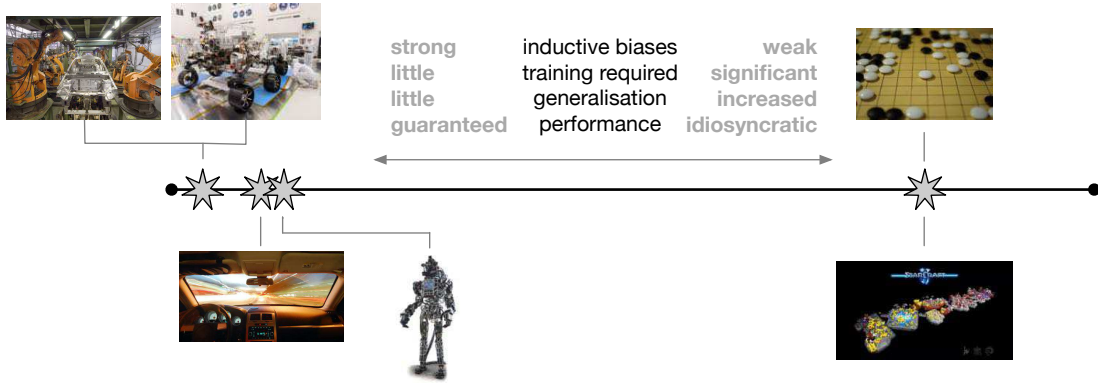
Figure 1: An *agent-based* view of the spectrum of *machine* learning. At one extreme (far left) everything is known. No learning is required, and performance is rigid but can be guaranteed via techniques from traditional systems engineering. At the other extreme (far right) nothing is assumed known and therefore everything must be learned from very large corpora of data. At this end, in theory, a system is able to generalize extremely well from one task to another. While approaches such as *MuZero* (Schrittwieser et al., 2020) have demonstrated the art of the possible with rather weak biases, to date such successes have mainly been confined to the context of game environments. In reality, AI agents commonly live in the continuum between these extrema. Current successes in real-world *embodied* agents remain focused on specific tasks such as factory automation and autonomous driving while general learning techniques have not yet been shown to succeed on real world robotics tasks to a similar degree. We posit that current advances in AI technology do not lend themselves readily to advancing embodied agents towards the right-hand side of the spectrum. As research pushes towards the safe deployment of fully autonomous vehicles as well as the development of more versatile articulated robots, bridging this divide poses both challenges and opportunities for *robot* learning.

In poorly understood, data-impoverished, highly dynamic, potentially adversarial domains such as disaster response, planetary exploration or even everyday life, the need for a mature theory to enable embodied agents to learn correctly, efficiently and safely is critical. A central challenge is therefore how to choose an architecture with an inductive bias of the learner that generalizes across a range of unknown and changing tasks and environments, and matches the requirements of physical constraints described above.

We begin with an examination of inductive biases for embodied intelligence and then move on to exploring some of them in more detail. In particular, we pose the following guiding questions for our discussion:

- What are the appropriate inductive biases that allow an embodied intelligence to learn most effectively while being robust to changes in task and environment? We conjecture that the necessary inductive biases for embodied learning are based in action and perception, and we address this conjecture in section 2.

5

- How do we design the architecture of an embodied intelligence such that it can learn effectively and act robustly? We conjecture that an essential ingredient of such an architecture is the ability to effectively amortize reasoning while maintaining the introspection and capability to perform more elaborate inference on demand. We address this conjecture in section 3.

- What is a representation appropriate for specifying both an embodied agent's model and the task to be performed? We conjecture that a specific form of compositionality is required of the representation that is absent from many architectures, and we address this conjecture in section 4.

- To what extent does the morphology of the agent affect its ability to learn? We conjecture that physical design is significantly complemented by effective learning methodology and should not be treated in isolation. We discuss this conjecture in section 5.

Finally, an open question in embodied intelligence is how these systems should be evaluated. At heart, how do we as a community know if progress is being made? We conclude with a discussion of the challenges of not only evaluating but verifying an embodied intelligence that can learn.

2. Inductive Biases for Embodied Intelligence

It is well understood that any learning system must have an inductive bias in order to make predictions outside the training set (Wolpert and Macready, 1997; Wolpert, 2002). An inductive bias is formally "any basis for choosing one generalization over another, other than strict consistency with the observed training instances" (Mitchell, 1980). Learning for embodied agents is no different in requiring inductive biases, however, we are not the first to observe that the nature of the bias may be very different for an embodied agent (Thrun and Mitchell, 1995; Burgard et al., 2020; Kaelbling, 2020). While traditional inductive biases in machine learning, such as minimum description length or weight sparsity, are useful for embodied intelligence, additional architectural, algorithmic or even learned biases are increasingly moving into focus for embodied agents.

2.1 Limitations and Challenges

It is useful to consider how the challenges of learning for embodied agents has implications for inductive biases. Our first challenge is to learn safely, given potential risks to the learning agent, which requires a theory of how to order models or hypotheses in a way that minimizes risk to the agent. The field of human cognitive development provides one possible such theory in the "zone of proximal development" (Vygotsky, 1978), which is often used to explain how children explore and learn. Vygotsky defined the zone of proximal development as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance, or in collaboration with more capable peers", essentially the set of tasks that are just beyond what a child can do, but the child can do with help. There are a number of reasons why learning in the zone of proximal development is most effective

for teaching children, but an important part of learning in the zone of proximal development is that it minimizes risk to the learner. The notion of scaffolding — a process by which a teacher helps the learner within the zone of proximal development — exploits this idea for effective learning.

To address our second challenge of learning when the world is non-stationary, the data are not identically distributed, and the agent encounters data that are very different from its training set, the agent must be able to generalize. Fodor and Pylyshyn (1988) defined the concept of a "systematic capacity" as one where "the ability to produce / understand some sentences is intrinsically connected to the ability to produce / understand certain others." This in itself is not a sufficient inductive bias but Bahdanau et al. (2018) broadened the idea to "systematic generalization", that is "the ability to learn general rules on how to compose words". The ability to learn general rules that lead to a causal theory of the world is important to an embodied agent because in a non-stationary, non-IID world, the training data will never be sufficient to represent all future queries. We conjecture that an inductive bias that encodes systematic generalization by encouraging a set of general rules to be learned may be the best way to enable agents that are robust to variation in task and environment. We will revisit this notion more explicitly in section 4.

Embodied agents must also be able to manage the computational growth that can result from learning in an open-ended environment. An effective approach is to reason at multiple levels of abstraction, making local decisions with highly precise models of the environment and the agent's dynamics when necessary, but equally being able to reason globally using much more abstracted representations (e.g., Toussaint, 2015; Toussaint et al., 2018; Garrett et al., 2017, 2021). The use of hierarchical abstractions in decision making and task execution has been an active area of research for decades. Fueled by advances in deep learning, abstraction continues to draw increased attention, predominantly in the context of reinforcement and imitation learning (e.g., Nachum et al., 2018; Le et al., 2018; Tirumala et al., 2019). However, hierarchical abstraction has not yet been used substantially in learning for embodied agents, with some modest exceptions (e.g., Sutton et al., 1999; Dietterich, 2000; Levy et al., 2017; Konidaris et al., 2018). The state of the art in most operational embodied agents is to hand-code an abstraction, and even the most carefully hand-coded representations are brittle in the face of a non-stationary world. A useful inductive bias for an embodied agent may be an ability to learn its own abstractions which balance the computational efficiency of approximation induced by the abstraction against the potential for loss in performance. Recent works on state-space learning and control serve as examples (e.g., Watter et al., 2015; Karl et al., 2016), as well as the increasing emergence of work on composable, object-centric deep generative models (e.g., Eslami et al., 2016; Greff et al., 2016; Kosiorek et al., 2018; Burgess et al., 2019; Greff et al., 2019; von Kügelgen et al., 2020; Jiang et al., 2020; Engelcke et al., 2020; Nguyen-Phuoc et al., 2020).

A final requirement of embodied learning is to recognize that the agent has a finite amount of available energy or specific power. While some inductive biases such as minimum description length will tend to favor models that are more energy efficient, there is a need for a more substantial investigation of how an agent's energy limitations can form inductive biases.

2.2 Opportunities and Future Directions

Given these challenges, there are a number of possible directions which may yield inductive biases that enable richer forms of learning for robots and other embodied agents. Three directions, in particular, would address the requirements and advance our understanding of choosing appropriate inductive biases for embodied agents that are learning.

A first possible form of inductive bias is to leverage the idea from cognitive science known as the "core knowledge hypothesis" (Spelke and Kinzler, 2007), that human intelligence rests on four systems that are hardwired for specific reasoning tasks, have innate representational capabilities, and innate limits. The four systems are designed to reason about objects, actions, number, and space[1]. More generally, there are sets of concepts that seem universally useful to an embodied intelligence, such as physical properties (e.g., masses, inertias, rigid lengths etc.) and physical laws (e.g., Newton's laws, the interaction of light and sound with matter, the interaction of electromagnetic radiation more generally with matter, etc.) that should guide the learning process. Having a strong inductive bias that prefers models that are consistent with physics would seem to be an extremely useful property. Implementing an inductive bias with a realistic model of physics may however be computationally burdensome and likely incomplete, which raises questions of how accurate a model needs to be to be useful, and how strong the bias should be. Behavioral studies of human physical reasoning suggest that people draw on a noisy, approximate ability to simulate interactions between physical objects, that qualitatively coheres with mechanics but departs from it in quantifiable ways (Sanborn et al., 2013; Battaglia et al., 2013). Prediction accuracy may also be a key bias in developing an intuitive understanding of the world and shaping the underlying model. Ha and Schmidhuber (2018) explore this bias in the context of training an agent to act in a learned world-model and find that the agent is able to learn in a model that is *good enough* — but not perfect.

A second form of inductive bias may be in the form of structural biases that drive abstraction. While some disciplines such as natural language processing have had success by reducing their explicit priors over internal representations and structure, there is not yet substantial evidence that embodied intelligence will succeed using an inductive bias that is purely "signal-to-symbol", where the internal representation is a quantization of the input (Bajcsy, 1995) — a structural bias that is more than merely quantizing the input is needed. Reasoning about action is imbued with the study of dynamics, and the basins of attraction arising from the consequent energy landscape offer at least one source of effectively grounded symbols — their systematic study launched the field of topology. Systematic generalisation suggests that model composition must be a key element for embodied intelligence, and the prospects for a compositional language of basins appears bright (Koditschek, 2021).

There is a lack of general consensus as to whether neural networks themselves are composable, at least in part because there is a lack of consensus on the definition of composition (Hupkes et al., 2019). Nevertheless, with relatively few exceptions, current learning approaches for learning composition functions do not perform well when tasked with learning the kind of composition that abstract reasoning systems excel at — for example, extrap-

---

1. Spelke and Kinzler (2007) speculate that a fifth system may exist for reasoning about "potential social partners and social group members", which would match the requirement that embodied agents recognise they exist in the presence of other dynamic agents with goals and intentions.

olation outside the training set or inference dependent on global context. We will revisit the question of how to introduce compositionality as a structural inductive bias within a learning agent in Section 4. The compositionality question is related to inductive biases about causality. Causal knowledge describes underlying mechanisms that can be composed to explain observations, such as the laws of physics, and is related to System 2 inductive biases (Goyal and Bengio, 2020) discussed in the next two sections.

A third and implicit form of inductive bias may result from how data are curated and presented to the agent. For example, an ordering over tasks and environments, such as a curriculum, induces an ordering over the concepts to be learned, which often smooths the optimization path during learning but can also create an implicit bias in what is learned. Curricular learning has been shown to be an effective form of meta-learning (Finn et al., 2017; Narvekar et al., 2017), and is very much related to the ideas from developmental cognitive psychology such as the zone of proximal development that enable efficient and safe learning. However, while there have been recent efforts in addressing this challenge, e.g., Botvinick et al. (2019), there is not yet a principled theory of the relationship between meta-learning for embodied agents and the corresponding inductive bias, nor how to choose a curriculum that supports the different requirements of an inductive bias for embodied intelligence.

Finally, there are dangers of having inductive biases that are too strong. For many conventional machine learning applications, there are principles that can provide guidance in choosing an inductive bias appropriately. For an embodied intelligence that is attempting to learn from small amounts of data that are neither independent nor identically distributed, and when the size and scope of the domain cannot be known ahead of time, it is important to understand the risks and tradeoffs of different inductive biases.

3. Robots Thinking Fast and Slow

As we have already argued, when it comes to agents that can act and interact, many of the advances in AI have played to the strengths of virtual environments: infinite training data is available, risk-free exploration is possible, and acting is essentially free. In contrast, we require our robots to robustly operate in real-time, to learn from a limited amount of data, take mission- and sometimes safety-critical decisions and even display a knack for creative problem solving. Cognitive science suggests that, while humans are faced with similar complexity, there are a number of mechanisms which allow us to safely act and interact in the real world. In addition to concepts such as the core knowledge hypothesis and the zone of proximal development that represent promising inductive biases, we focus in this section on a particular set of architectural biases inspired directly by Dual Process Theory (DPT). Popularized by Daniel Kahneman's book *Thinking Fast and Slow* (Kahneman, 2011), DPT postulates that human thought arises as a result of two interacting processes: System 1, an unconscious, intuitive response system, and System 2, much more deliberate reasoning. If we accept that Dual Process Theory plays a central role in our own successful interactions with the world, we can explore a similar approach towards realising robust, versatile and safe embodied intelligent systems.

A key observation of DPT is that faced with an every-day challenge like game-play, driving or stacking plates, we usually do not explicitly analyse the governing laws of the

particular process. Instead we tend to simply act according to what the situation demands, as informed by our senses, based on intuition or even an innate reflex-like behaviour. The advent of deep learning has afforded our agents principally two things: (i) an ability to learn arbitrarily complex mappings from inputs to outputs; and (ii) an ability to execute these mappings in constant time. These, together with an ability to learn structured, task-relevant embeddings in an unsupervised manner, afford researchers a different view on the computational architectures they employ. The ability to learn complex mappings endows our agents with an ability to perform very complex tasks at useful execution speeds. Direct human supervision, reinforcement learning, task demonstrations, complex learned models as well as the increasingly popular concept of system-level self-supervision all fit into this narrative. In game play, DeepMind's AlphaGo Zero (Silver et al., 2017) as well as the closely related *Expert Iteration* algorithm (Anthony et al., 2017) distill knowledge from Monte Carlo Tree Search (the oracle) and self-play into a model which predicts value and probability of next move given a particular board position. In robotics, OpenAI's Learning Dexterity project (Andrychowicz et al., 2020) distils knowledge gained in simulation through reinforcement learning and domain randomisation (the oracle) into a policy which can control a Shadow Hand to move an object into a target pose. In the context of autonomous driving, Barnes et al. (2017) distill, via the automatic generation of training data, hundreds of person-hours of systems engineering into a neural network model which predicts where a human might drive given a particular situation. Recent work on intuitive physics learning distills data arrived at through physical simulation into neural network models, see, for example, Wu et al. (2015); Lerer et al. (2016); Wu et al. (2017); Li et al. (2017); Groth et al. (2018); Janner et al. (2019). Owing to their ability to mimic the expertise of an oracle in a time- (or generally resource-) efficient manner, one might view the execution of a neural network model as analogous to an efficient, *intuitive* System 1 response.

At the same time, roboticists and AI researchers have spent decades developing System 2 equivalents. Symbolic reasoning, traditional planning approaches and even simplistic but time-intensive brute force methods, all constitute deliberate and often effortful task solvers. Robust, real-world performance thus seems to require computationally efficient policies empirically tuned to a particular task and environment *as well as* more computationally intensive approaches capable of systematic generalisation in that they are robust to variations in the task and environment. In the view offered here, we may be able to leverage machine learning to distill more resource-intensive, deliberate System 2 responses into learned models which mimic these experts in an efficient and effective way to form an intuitive System 1 response.

3.1 Limitations and Challenges

The narrative of distilling knowledge into rapidly executable neural network models allows us to achieve significant, often game-changing, computational gains. However, as roboticists we are still faced with a substantive and foundational challenge when it comes to applying machine learning systems in the real world: the routine violation of the various assumptions made by our systems. As discussed in section 1, an embodied intelligence cannot assume the data are independent and identically distributed. In contrast, as robots operate over ever longer time scales in increasingly unstructured environments, the data encountered

significantly deviate from the training distribution. Additionally, we expect the embodied intelligent systems to generalize to unforeseen tasks and environments. In practice, together with the approximate nature of our algorithms, these assumptions lead to learned models which are often over-confident and whose performance can only loosely be bounded (if at all) using traditional methods (e.g., Grimmett et al., 2016; Richter and Roy, 2017; Rao et al., 2019). The result is that our robots lack the ability to reliably know when they do not know and take appropriate remedial action. Despite many attempts over the years at remedying this shortcoming (e.g., Settles, 2009; Hsu, 2010; Li et al., 2011; Pentina and Lampert, 2014; Dann et al., 2017) we are still no closer to a practicable solution. Our conjecture is that a Dual Process Theory perspective may well provide a way forward to address this challenge.

A second limitation of existing approaches is demonstrated by the considerable evidence that humans represent and use estimates of uncertainty for neural computation in perception, learning and cognition (Deroy et al., 2016). However, how *metacognitive* uncertainties are derived and utilized is only gradually being discovered. Special, metacognitive circuitry in the human brain suggests knowledge integration above and beyond raw perceptual signals (Deroy et al., 2016). The *Feeling of Knowing* process, for example, enables humans to effectively choose a cognitive strategy (e.g., recall vs. reasoning) likely to succeed in a given circumstance (Reder and Ritter, 1992). Moreover, recent work on *multi-sensory* perception suggests that metacognition is instrumental in discovering *causal* structures in order to form a coherent percept from multi-modal inputs (Deroy et al., 2016).

Finally, while we may now have a technical blueprint for components on either side of the systemic divide, much uncertainty remains around the nature of the cognitive processes involved. Similarly, a strict categorization of the complex landscape of inter-operating neural processes into two types of systems as proposed in DPT is widely recognized to be a significant oversimplification. Nevertheless, it serves as a conceptual starting point. How to design an architecture that effectively combines the best of both worlds remains an open and potentially fruitful research question (e.g. Goyal and Bengio, 2020). It is not clear whether components of both systems run in parallel or run on demand with an explicit handover between deliberate planning and low-level intuitive policies. In analogy with metacognitive challenges faced by humans, *when* a System 1 response is appropriate over a more deliberate deployment of System 2 remains one of the open questions to be addressed. In humans, these two forms of processing interact in that System 2 can suppress, inform and even train System 1 responses (Kahneman, 2011). Furthermore, it is not clear if the architecture is two separate systems or is in fact a continuum or tight integration of processes capable of fulfilling either part.

3.2 Opportunities and Future Directions

These limitations and challenges immediately point at a set of now viable technical approaches in which the outcome of a downstream system (either in terms of success/failure or in terms of confidence in outcome) given a particular input is distilled into a machine learning model[2]. Predictive models of performance are now relatively common-place in the robotics literature. They have a long-standing track record in predicting task success in manipulation and complex planning tasks (e.g., Pastor et al., 2011; Kappler et al., 2015;

---

2. Statistical outlier detection also falls into this category.

Krug et al., 2016; Pinto and Gupta, 2016; Morrison et al., 2020) and are increasingly used, for example, to predict the performance of perception and vision-based navigation systems (e.g., Gurău et al., 2016; Daftry et al., 2016; Dequaire et al., 2016).

Human intelligence is far more robust and uncertainty-aware than our best learning systems and operates in significantly non-stationary (in the statistical sense) environments. We draw on a rich capacity for *metacognition* (Cox, 2005; Deroy et al., 2016): the process of making a decision, the ability to know whether we have enough information to make a decision and the ability to analyze the outcome of a decision once made. A predictive neural network can be injected with dropout noise (Gal and Ghahramani, 2016) or another trained to predict the magnitude of out-of-distribution errors in these predictions, and thus estimate epistemic uncertainty (Gurau et al., 2018; Jain et al., 2021), which can be used to guide decision-making and exploratory behavior. One of the interesting aspects of a Dual Process Theory for robots is the fact that metacognition may find a natural place in such a construct: the open question is whether a theory of metacognition can be used to bridge the two systems.

Another set of viable technical approaches draws on new ways of combining causally structured generative models and Monte Carlo inference with machine learning. Causally structured models and Monte Carlo inference have a long history in AI and robotics, precisely because, taken together, they offer a way to model the causal structure of non-stationary systems, perform data-efficient parameter estimation, and build inference processes for state estimation that "know when they do not know". Unfortunately, it has not been easy to scale up inference in causally structured models such that it is possible to track large, complex environments in real-time. One approach is to use slow, offline inference to fit the generative models to data, yielding realistic simulations that can then be used to generate fully labeled, synthetic training data for bottom-up neural networks. Additionally, one can retain a lower-resolution structured generative model "in the loop" at inference time, and combine fast, bottom-up learned Monte Carlo updates from neural networks with slower top-down, model-based Monte Carlo. Recently, new probabilistic programming languages such as Gen (Cusumano-Towner et al., 2019) have been developed that make it much easier to write structured generative models and carry out real-time, approximate Bayesian inference via custom hybrids of neural, symbolic, and Monte Carlo methods. These approaches have been applied to solve 3D scene perception problems (e.g., Kulkarni et al., 2015), and can run in real time while producing more robust results than approaches that only depend on learning. One benefit of "model in the loop" architectures is that they use the posterior probability density under the causal generative model to assign quantitative (relative) confidence levels to the output of machine learning algorithms. Similar architectures might be applied to navigation and planning tasks that build on the outputs of perception, using top-down model-based reasoning to quantify confidence in the outputs from fast, bottom-up learning.

A tantalising opportunity lies in System 1 computation helping to solve one of the key computational limitations of symbolic AI systems, i.e., that of the intractable cost of search: deep generative models (e.g. Kingma and Welling, 2013; Goodfellow et al., 2014; Bengio et al., 2021) may play a role similar to imagination in that they can be trained to produce good candidates for search or planning consistent with more expensive System 2 sequential reasoning steps (e.g. Ha and Schmidhuber, 2018).

A cognitive science theory related to the DPT is the Global Workspace Theory (GWT) from Baars (1993, 1997), which postulates a communication bottleneck between specialized brain modules, with these expert modules competing for being able to send content through this bottleneck to be broadcast to all the modules. The connection with the DPT is that this bottleneck is the consciously accessible working memory which humans can report verbally, i.e., corresponding to System 2 content. However, the details of the computations performed inside the competing expert modules is not consciously accessible (and probably too complex to be completely verbalizable) and corresponds more to System 1 machinery. An interesting question is the purpose of such a strong communication constraint. Bengio (2017) and Goyal and Bengio (2020) hypothesize that it induces an inductive bias that would make modeling the world at the abstract System 2 level good at capturing the kind of sparse dependencies and causal mechanisms that humans describe through language. The next section elaborates on the notion of multiple levels of abstraction and representation and the kind of discrete representations and sequential reasoning processes which humans tend to use at the higher System 2 level.

We close this section by noting that opportunities exist not only for addressing learning in embodied agents, but conversely also for advancing our knowledge in cognitive science. The latter often requires complex experimental procedures which, by design, need to disrupt the agent's learning process. Robotics, on the other hand, *allows* the design and close inspection of the mechanisms involved in the learning process — including control over individual components, the environment and modes of interaction. Much of this section makes the case that mechanisms already discovered in the cognitive sciences may cast existing robotics work in new light, with the aim of establishing a meaningful technological equivalent to the Dual Process Theory. In particular, they may provide a blueprint towards architecture components we are still missing in order to build more robust, versatile, interpretable and safe embodied agents. Conversely, as also noted by Botvinick et al. (2019), the discovery of intelligence architectures which successfully deliver such dual process functionality may *equally* provide fruitful research directions in the cognitive sciences.

4. The Role of Logic in Embodied Intelligent Systems that Learn

Having identified the need for systems for inference and decision making that can reason at different levels of abstraction, we turn our attention to the question of what representations enable learning at the different levels. Learning representations for sensorimotor perception and control (i.e., System 1) is currently a well-studied problem as demonstrated by the vast body of work in deep representations for perception and control. However, for deliberative inference and decision making, it is critical to consider different classes of representation languages. A hallmark of problems that require deliberative reasoning is that they involve a form of dynamic compositionality over abstract concepts, long horizon and high dimensionality that may be more difficult to implement using System 1 machinery (Lake et al., 2017; Bahdanau et al., 2018). The strategy, in System 2, is to render these problems tractable by exploiting structure in the problem, including factoring into weakly interacting sub-components (possibly involving different aspects of the state space or different temporal sub-regions), and "lifting," or abstracting over arbitrary sets of entities in the domain. Classically, the languages that have effectively enabled compact representation and

efficient reasoning in very large problems are *logics*, including propositional logic, first-order logic, AI planning domain description languages (PDDL), graphical models (which can be viewed as a probabilistic version of propositional logic), probabilistic planning languages, probabilistic programs, and temporal logics. Similarly, the tools of topology (and the computational efficacy of its algebras) hold a relationship to robotics analogous to that of logic relative to computer science.

Models represented in any of these languages are much closer to a specification provided by a human engineer compared to reflex-based models represented as neural networks. However, the inference and learning questions for all of these modeling paradigms are not inherently different: in every case, the models can be used to make predictions or to determine action choices, and they can be fit to data by machine learning methods that endeavor to make the models' outputs match those in a data set. Nonetheless, the algorithmic techniques for model fitting may be substantially different and the choice of what kinds of problems are best matched to what kinds of models is not well-understood. The choice of representation language therefore depends on the types of problems it is suited for and the underlying inductive bias — does it fit the problem at hand, and thus substantially decrease the sample complexity for learning? One important feature of many logical representations is *compositionality*, where parts of the model have independent semantics, allowing them to be learned independently and then composed in combinatorial ways to solve many different problems, providing a particularly aggressive and useful form of generalization. It is through this compositionality that we may hope to achieve the systematic generalization introduced in section 2.

4.1 Limitations and Challenges

Learning logical representations has been an active research area for a considerable time. One of the main focal points of learning in logical representations is inductive logic programming (ILP), with the goal of inferring a logical theory of the world that entails a given dataset. While the use of ILP has been examined previously for embodied intelligence (Bratko, 2010) including mobile navigation (Leban et al., 2008) and learned grasping (Antanas et al., 2015), these results have not demonstrated the same level of success as alternate approaches. How to learn compositional rules that represent abstract actions such as learned STRIPS or PDDL representations of actions has been a subject of investigation from the early days of Shakey (Fikes et al., 1972) to more recent results (Konidaris et al., 2014; Sammut et al., 2015). However, much of the prior work assumes deterministic and often propositional representations. The ability to learn stochastic, factored and lifted models of both perception (Nyga et al., 2014) and action models (Mugan and Kuipers, 2011; Krüger et al., 2011) is crucial for operating in a physical world. In particular, Pasula et al. (2007) argue that propositional, relational and deictic learning represent different inductive biases, and "deictic learning provides a strong bias that can improve generalization". Unfortunately, learning logical representations through ILP often suffers from an exponential growth in complexity with the size of the theory to be learned, and also can have difficulty with noisy or inconsistent data. Recent progress has shown the effect of neurally-inspired induction, such as differentiable ILP (Evans and Grefenstette, 2018; Payani, 2020), that can deal effectively with noisy data. Neural logic machines (Dong et al., 2019) have shown

how to recover a set of lifted rules in very simplified physical environments while avoiding problems of scalability. Nevertheless, the application of any kind of ILP or learning logic machines to embodied intelligence has been limited in scope and represents a substantial open challenge.

Another limitation of most existing work in learning logical representations is the reliance on an initial theory of the world, often referred to as "background knowledge", comprising an initial set of predicates and propositions with truth values that are either known *a priori* or can be inferred from observation. In very simplified domains such as family tree reasoning (a common benchmark in the ILP community, Wang et al., 2015) or blocks world, it may be reasonable to write down an initial background knowledge theory that is sufficient to support induction of a complete set of predicates and rules. However, the background knowledge is "similar to features used in most forms of ML" (Cropper et al., 2021). Just as deep learning has substantially reduced the need to handcode feature functions for many problems, it is possible that the current reliance on handcoded background knowledge theories is a substantial limitation in learning logical representations for embodied intelligent systems. Techniques that get closest to learning symbolic representations from completely continuous, sensorimotor representation still rely on background knowledge of motor primitives or skills (Konidaris et al., 2018). At the same time, as discussed in section 2, the "core knowledge hypothesis" described in section 2 would seem to suggest that biological systems themselves depend on background knowledge. It is an open question whether background knowledge is a limiting inductive bias for embodied intelligence, or whether more research is needed to identify the *correct* background knowledge for an embodied intelligence.

Our notion of a learning embodied agent is one that will eventually need to represent concepts that could not have been anticipated ahead of time. Being able to grow the primitive logical representation from the raw sensor data, a process sometimes referred to as "symbol emergence" (Taniguchi et al., 2018) is an important ability for an autonomous agent. There has been work on attempting to learn a discrete planning domain model from completely continuous, sensorimotor representations (Asai and Fukunaga, 2018; Asai, 2019; Ames et al., 2018). Additionally, some previous work has demonstrated that these approaches in learning discrete representations for actions can be coupled with a convolutional neural network used for processing visual input, to create an end-to-end system that learns both a symbolic representation of the visual input and a logical theory of actions for planning in toy domains such as Sudoku (e.g., Wang et al., 2019; Dong et al., 2019). Similarly, previous work has demonstrated that a partial representation can be learned ahead of time, and then recruited by a reactive system that deforms real-time sensory instances into their learned topological model via real-time change of coordinates (Vasilopoulos et al., 2018). However, it is not well-understood how to scale up these initial results to more complex domains encountered by an embodied intelligent system.

Furthermore, in order to obtain the advantages of logical representations such as composition, the inductive biases of the learning process must not only satisfy the physical constraints of an embodied intelligence, but also meet the needs of the logical inference process. For example, in factoring, a decomposition must be found that renders different aspects of the problem relatively independent. In lifting, objects must be "reified" in a way that allows useful abstractions. This search for the fundamental discrete structures requires an inductive bias towards factoring the world into independent and composable pieces,

for example as might be formalized by a type theory of energetically grounded symbols (Koditschek, 2021).

Lastly, even with a well-defined discrete representation, an embodied intelligence exists in a continuous environment, and there must be a connection between the abstract, usually discrete, representation to the concrete, usually continuous input and output signals. This connection between the abstract representation and the physical world is the "symbol grounding problem" (Harnad, 1990). However, existing models of symbol grounding fail to meet nearly all the challenges faced by embodied agents (Coradeschi et al., 2013); for example, learned models of symbol grounding again assume the data are IID, and that there is a fixed and known alphabet of symbols to be grounded to the input and output signals. The inductive biases in the current approaches to learning symbol grounding do not yet have principled techniques for determining when the world has changed and the corresponding learned model of symbol grounding has changed. In recent years, approaches such as unsupervised learning for option or skill discovery (Gregor et al., 2016; Eysenbach et al., 2018; Bagaria and Konidaris, 2019) have attempted to learn primitives and discrete-level skills to incorporate compositionality for tasks such as long-horizon planning. While still preliminary, these works put forward a promising direction for instituting some of the benefits of logic into neural models.

4.2 Opportunities and Future Directions

Developing new techniques for inferring the base or primitive representation that allows for compositionality is one of the primary open questions for a learning embodied agent. There is increasing evidence that the internal layers of many neural networks can encode compositional symbols in a way that relates to discrete logical representations. For example, "disentanglement" techniques can be used to force a learned representation that is factored, and results so far reveal surprisingly intuitive structure about the learned representation (Higgins et al., 2018). These results suggest a path forward to learning the underlying abstraction that is the foundation of logical reasoning, although this work has not yet led to substantially improved performance in inferring compositional representations.

Opportunities exist for new techniques in learning lifted models, and graph neural networks (GNNs) provide a possible mechanism for neurally-inspired lifted models, in the sense that a model of fixed dimensionality can be learned and then applied to domains with arbitrary numbers of objects. Neural networks embedded within graphical models (Krishna et al., 2017; Armeni et al., 2019) or full GNNs have been applied to scene perception and entity abstraction (Veerapaneni et al., 2020; Qu et al., 2019). However, GNNs have not been shown to capture the power of quantification, and with the exception of some promising preliminary work (Simonovsky and Komodakis, 2018; Franceschi et al., 2019; Alet et al., 2019), the assumption is that the graph structure is known ahead of time.

There does already exist an expressive and reasonably complete language that is composable and supports abstraction, in human natural language. There has been considerable work in symbol grounding efforts for embodied agents that relies on natural language to provide the set of symbols that are to be grounded (Tellex et al., 2011; Matuszek et al., 2013; Thomason et al., 2015; Paul et al., 2016; Patki et al., 2019), however, these approaches have not yet demonstrated the ability for a robot to acquire a large knowledge base without con-

siderable human intervention and effort. The difficulty in much of this work is the need for highly-annotated, aligned corpora of data; there is an opportunity for new techniques that allow an embodied intelligence to acquire new symbols from language in a self-supervised manner. Attempts to learn symbolic or logical knowledge bases by, for example, reading the internet (Matuszek et al., 2005; Mitchell et al., 2015; Olivares-Alarcos et al., 2019) have made more progress but remain incomplete and have not had substantial impact on embodied intelligence.

There is an open question as to whether or not background knowledge should act as a prior on the learner, as in the inductive biases discussed in section 2, or if in fact the learner should be attempting to derive its theory of the world from scratch using principles such as systematic generalization. There is increasing evidence that a very plausible approach is to fix a simplified model ahead of time, e.g., a physics-based model, and then learn to correct for errors induced by abstraction (Ajay et al., 2018; Zeng et al., 2020). Alternately, there is evidence that embedding an entire physical model in a neural network and training it end-to-end may lead to models that are more robust and appropriate for the problem distributions they are faced with (Whiteson, 2018; Amos et al., 2018; Karkus et al., 2019). More recent techniques heavily rely on pretraining a universal model on large offline datasets—following a similar paradigm as universal language models such as GPT-3—which although is often not compositional can act as the background knowledge (Levine et al., 2020).

An interesting direction is to incorporate System 2 inductive biases in neural networks (e.g. Goyal and Bengio, 2020). One starting point inspired by the GWT (see 3) is to construct a modular architecture where modules compete to be activated and communicate (Goyal et al., 2019). Using attention mechanisms to operate on sets of elements (rather than vectors) makes it possible to implement a working memory as the communication bottleneck (Goyal et al., 2021b), as in the GWT, yielding better out-of-distribution generalization. Forcing the messages exchanged between modules to be discretized using a shared vocabulary of abstract concepts yields further improvements to modular architectures (Liu et al., 2021), including to the very popular Transformers (Vaswani et al., 2017). This kind of modular architecture can also be forced to process information through the sequential application of simpler rule-like but learned operations implemented by small expert MLP modules, yielding "neural production systems" (Goyal et al., 2021a). While these are inspired by classical AI production systems, all the rules are *learnable* and can operate on distributed representations: end-to-end learning of an attention machinery can again be used to dynamically control, which rules are applied when, in what order, in order to minimize some training loss.

Finally, logical representations are often used in domains that require guarantees of correctness of the inference, which is especially useful in ensuring physical safety in many engineered systems and seem to have promise for ensuring the safety of an embodied agent (Kress-Gazit et al., 2009; Kloetzer and Belta, 2008; Raman et al., 2015). However, by construction, the learning process itself cannot provide anything other than statistical guarantees. As we shall discuss in section 6, for an embodied intelligence, statistical guarantees provide statements of robustness but not statements of safety; some external structure around the learner is required to ensure safety properties. There is an open question as to what kinds of learning can be applied to a formal representational language that preserves the ability to provide guarantees of correctness.

5. The Impact of Morphology on Embodied Intelligence

A particularly strong but often unacknowledged inductive bias is introduced by the morphology of the robot. What sensors the agent possesses, what degrees of freedom it can act with, the dynamics of its motion, its own rigidity in interacting with the environment, the extent to which it *can* interact with the environment all have a tremendous influence on what the agent can and cannot learn, and what computations can be "offloaded" to the body (Hogan, 1985; Pfeifer and Bongard, 2006; Müller and Hoffmann, 2017). However, physical agents are currently constructed from a fairly limited range of possibilities. The vast majority of vehicles rely on high-precision range sensing (typically based on lidar) and high-precision, stiff actuation. These design choices are motivated by the requirements of existing estimation, perception and control algorithms to have highly accurate models of the agent's sensors and actuators, and to have highly accurate knowledge of the agent's state and that of the environment around it.

5.1 Limitations and Challenges

While the space of robots and embodied agents that have precise sensing and actuation include many useful types of platforms such as manufacturing robots, self-driving cars, unmanned air vehicles, and others, there are some limiting consequences to how robots are most often designed today. Robots with precise, stiff actuation tend to be either extremely slow, or have high energy, making it difficult and potentially unsafe for people to work alongside them. High-precision actuation is also typically expensive in terms of size, weight, power and cost and such robots are difficult to operate, limiting how widely they can be adopted by people interested in the questions of learning for an embodied intelligence. Highly precise robots constructed from rigid, articulated links generally must avoid contact with the environment except in specific and carefully controlled ways, which dramatically limits their ability to learn by interacting with the environment. This is one example of how the morphology of the robot impacts what kind of concepts can be learned. An interesting counter-example is represented by the low-cost manufacturing robots Baxter and Sawyer produced by Rethink Robotics. By virtue of low-energy series-elastic actuators, these robots were safe enough for people to work alongside, which drove their adoption by the robot learning community. These robots were also relatively imprecise, but it has been shown (Li et al., 2014; Cremer et al., 2016; Guan et al., 2018) that an embodied intelligence can learn effective and accurate manipulation strategies despite the imprecision of the end effectors.

A similar situation exists for sensing, in that the vast majority of operational autonomous robots rely heavily on laser range finders, which perceive the world in ways that are radically different from most biological systems. While laser range finders are extremely reliable in detecting obstacles and building detailed and dense geometric maps of the environment in order to plan collision-free motion, these sensors provide limited information about the visual appearance of the world (even when intensity information is utilized). There is a tremendous amount of appearance information embedded in the world (signs, landmarks, etc.) that is inaccessible to a vehicle using only a laser range finder for navigation. While most robots now combine ranging with passive vision, each sensing modality is used to perceive very specific forms of information that are essentially orthogonal to each

18

other. There is very limited understanding of how to trade off the different sensors, especially in a learning context. As with high-precision actuation, high-precision sensing is also expensive in terms of size, weight, power and cost. A successful counter-example is the Skydio drone, which demonstrates a surprising degree of navigation autonomy using a large number of cameras arranged in stereo pairs. However, the internal representation used by the Skydio vehicle is very similar to the dense geometric maps used by laser range finders, which creates an inductive bias in not leveraging all semantic information at every representational level. Only recently have results begun to emerge that use so-called "semantic information" for navigation (Civera et al., 2011; Atanasov et al., 2014; Kostavelis et al., 2016) supported without a dense geometric model. The inductive bias induced by assuming that perception is provided by a range measurement system is another example of how the morphology of the robot impacts what kind of concepts can be learned.

However, despite the evidence that the morphology and sensory modalities of an agent induce a substantial inductive bias in learning, there is relatively little understanding of exactly how morphology and sensor modality impact the ability to learn. For example, the tradeoffs between power, mass, sensing and computing are poorly understood, and especially how they impact learning. There is no principled way to modify the morphology to address a specific inductive bias. Similarly, we have limited understanding of how to design an embodied agent that can learn robustly across a specific range of tasks and environments; too often our agents are designed for one environment and immediately fail when presented with a slightly different environment or task[3].

5.2 Opportunities and Future Directions

Given these challenges, there are a number of possible directions where learning can be used to leverage a wider range of agent morphologies. Learning may be a way to reduce the costs of constructing and operating real robots. More compliant actuators and compliant bodies are often cheaper to build and safer to operate around people, and precision is only sometimes needed. Relatedly, robots can be designed to be more robust to situations likely to be encountered during online learning (Bhatt et al., 2021). Opportunities exist to design robots that learn to be precise only when necessary, using very different control strategies than are currently used, leveraging kinematic and dynamic constraints such as contact or inertia. Modalities such as sound provide a potentially rich and low-cost sensory stream that remains largely unexploited (Zöller et al., 2020). Constructing components from active materials represents an area where learning may be very useful in developing highly-capable control systems. Similarly, learning may enable richer forms of sensor fusion. An excellent example of how expanding our sense of plausible robot sensors through sensor fusion is the GelSight sensor (Li et al., 2014), which uses visible deformations in gel at robot fingertips as a form of tactile sensing. The ability to interpret the deformations as force and contact information is enabled by modern machine learning techniques. Learning may also be the key to unlocking high-density, large-area tactile sensing using systems

---

3. Many DARPA robotics programs have attempted to overcome this problem by requiring evaluation on a range of environments, such as the Fast Lightweight Autonomy program, which standardized the flight vehicle hardware but allowed sensor variability. The range of environments did appear to lead to performers choosing different sensor modalities that had different operational characteristics. Unfortunately, no sensor morphology was robust to all environments, signaling a technical gap.

such as Maiolino et al. (2013), which are increasingly emerging. Similarly, varying robot morphology and design may allow us to better understand existing limits in terms of energy consumption, computational complexity, reachability and observability.

As discussed in the introduction, a substantial difference between conventional machine learning and learning for embodied agents is the availability of training data that matches the assumptions of the learner. Exploring the space of different morphologies, sensing and actuation paradigms is impractical when each new design must actually be physically fabricated and then used to collect data for learning. Simulators will be needed that are sufficiently high fidelity to enable learning to occur for robots that vary significantly from the current paradigms. Ideally, new classes of simulators could enable co-design of both the learning approach and the hardware itself. It has long been a central tenet of robot learning in particular that simulation is "doomed to succeed" (Brooks and Mataric, 1993). However, the state of simulation, and consequently the value of simulation, is changing rapidly and the apparent limits of simulation may not in fact be fundamental to all simulations. Photo-realistic game engines are increasingly common-place and can enable embodied agents to learn from images. Physical simulators are also increasingly capable, although some physical phenomena are challenging to simulate in real-time such as deformations, friction effects and fluid effects. Nevertheless, with the advent of better, more physically realistic and photo-realistic simulation, new opportunities exist for learning in the context of new robot types. A combination of advances in simulators, safe learning, and sim-to-real methodologies is likely to enable designs with much wider variations in morphology, and, commensurately, insights into how morphology helps enable intelligent embodied behaviors.

6. Assessing Robot Learning

Embodied agents can have a very direct and physically damaging effect on their environment as compared to other applications that rely on learning-based methods, creating concerns of safety. Verification, validation and benchmark-based performance evaluation are the cornerstones of safe deployment of active control systems and increasingly a core component of deployment of autonomous systems. However, how to validate and verify the performance of an embodied agent that learns and adapts to novel experiences is an open question and it is likely that new principles for evaluating our embodied agents will be required.

6.1 Verification and Validation of a Learning Embodied Agent

Conventional verification and validation techniques assume a predefined specification of the desired system behavior in an environment. *Validating* the safety of a system means that no failure has been found after testing it under environment disturbances or when the probability of failure is below some threshold. Safety *verification* means that a system is provably safe to all disturbances. Both validation and verification therefore require a model of the disturbances. For most realistic systems in robotics, such a specification and realistic disturbance model is nearly impossible to specify or sample from, for all but a very narrow set of tasks and environments. It may well be that the traditional requirements of verification and validation are in direct conflict with many of the goals of learning. Furthermore, the kinds of guarantees provided by most learning approaches are at best statistical, describing aggregate or asymptotic behavior rather than providing guarantees on

any instantaneous query of a learned model. These kinds of guarantees also often rest on the previously discussed assumption that the training and test distributions are identical, which is hardly true in a real scenario. Probabilistic guarantees may be extended when it is possible to determine how far a training set is from the current domain, task and environment (see section 3). Given these challenges, the successful verification and validation of a learning embodied agent will require novel approaches, based on innovative ways to characterize desired system behavior.

An interesting technical direction is the idea of adversarial testing. While adversarial examples have been used to demonstrate potential failures of learned models, it may be possible to automatically generate adversarial examples that test the response of an autonomous system. This is a very active area of research when it comes to adversarial images or other perceptual input (e.g., Szegedy et al., 2014; Ranjan et al., 2019; Tu et al., 2020) but raises the question of how adversarial disturbances of an environment can be automatically generated that change in response to closed-loop behaviour (Sinha et al., 2020). The nature of a proof of correctness of a learning system may be that no adversarial examples exist, at least in the context of a particular task environment. The challenge is not only to find such examples given the high sample complexity in realistic problems, but also to ensure the learning system is robust to these. Corso et al. (2020) provide a broad survey of safety validation of black-box systems through simulation. These systems search for disturbances that cause the evaluated system to behave improperly and may use optimization, planning or reinforcement learning techniques for this search.

6.2 Performance Evaluation of a Learning Embodied Agent

Evaluating the performance of, or "benchmarking", learning-based methods in robotics is similar in many ways to verification and validation. In benchmarking we are interested in how well such a system performs especially in comparison to other systems tested in similar tasks and environments. Benchmarking differs from the binary assessment of verification in that benchmarks can be used to measure progress during development, not just whether the final system meets a target level of performance.

Current benchmarking on *real robotics systems or data* can be approximately grouped into three kinds. First, there are anecdotal demonstrations of solutions on specific scenarios. These are often proofs of concept. Second, there are (offline) datasets, e.g., for pick and place (Levine et al., 2017; Mahler and Goldberg, 2017; Pinto and Gupta, 2016) or autonomous driving (Geiger et al., 2013; Caesar et al., 2019; Sun et al., 2019) that often test specific perceptual skills such as inferring a robust grasp from visual data, accurate object detection or semantic segmentation in a street scene. Third, there are robotics challenges (Correll et al., 2018; Atkeson et al., 2015; Righetti et al., 2014; Eppner et al., 2018), most prominently the DARPA challenges that often reveal how far we have come in a specific subfield of robotics: manipulation, navigation, driving, tasks in disaster relief scenarios, etc.

Robotics poses a particular challenge for benchmarking learning-based solutions. A robot is a closed-loop system that acts and thereby influences the state of the environment, which in turn influences its next steps. This feedback loop usually prevents the use of offline datasets for evaluation, as is commonly done for benchmarking of algorithms in

computer vision or natural language processing. For example, the success of datasets like KITTI (Geiger et al., 2013) and the more recent Argoverse dataset (Chang et al., 2019) have helped drive progress in estimation and perception for self-driving cars, but they have not yet had a similar impact on decision making. Major companies focusing on autonomous vehicles – such as Waymo, Tesla and others, have publicly discussed the importance of simulations and the significant amount of simulated driving that they are currently using for testing and technology development. Fully interactive simulations allow for both performance evaluation and learning. Furthermore, rare corner cases can be slightly modified to multiply interaction scenarios, allowing systems to be better tested and trained. In this context, the Open Source CARLA simulation environment (Dosovitskiy et al., 2017) and the challenges run by the CARLA team serve an important role for the broader academic and industrial research community.

Simulation is a mechanism for systematic evaluation that does not require expensive and tedious deployment of robots into the real world. However, simulations for robotics have historically been notoriously low quality with a significant domain gap between simulation and the real world. Techniques such as so-called "sim2real" that use learned models to close the gap between the distribution of simulated and real data have shown some recent success in allowing embodied intelligence to learn effective control policies (Sadeghi et al., 2017; Ramos et al., 2019; Zhao et al., 2020; Hofer et al., 2020); nevertheless, simulating realistic sensory data (e.g., Porav et al., 2018; Yang et al., 2020; Weston et al., 2021) is still a challenge and simulating contact is even harder. While many learning-based methods are developed and evaluated in simulation, it is unclear whether they would work in the real world. Examples of such benchmarks include Meta-world (Yu et al., 2020), Point Navigation in Habitat-AI (Savva et al., 2019) iGibson (Shen et al., 2020), and the OpenAI Gym (Brockman et al., 2016). Andrychowicz et al. (2020) presented a rare example of transferring a policy learned in this simulation environment to a real robot, but the success rate dropped by 50% despite very large amounts of simulated training iterations on a randomized domain. A possible way forward is to consider "real2sim" which attempts to identify a simulator that more closely resembles a few real world examples (Jiang et al., 2021).

There has always been a conflict between functionality and generalization in robotics. Learning-based systems promise to generalize to unseen situations in many robotics applications, but when compared to methods that are application-specific (even with learning-based components), full learning-based methods often exhibit orders of magnitude worse performance on a specific benchmark. Therefore the question arises of what exactly are the correct metrics for a learning embodied agent — it is an open question if the best metrics are performance on any given task, or if the metrics should characterize generalization over many tasks or to unexpected situations. Specifically, Goodhart's law is often at play, summarized by Strathern as "When a measure becomes a target, it ceases to be a good measure" (Strathern, 1997). This phenomenon can be seen in robotics challenges where the point is to win rather than to find solutions to problems that may generalize to larger variations of the given task. While some general insights have been gained through challenges (e.g., compliance for manipulation, suction cups for pick and place) these challenges often expose point solutions that address the specific goal of the contest but do not solve the actual problem of interest (Atkeson et al., 2015; Correll et al., 2018).

It is imperative to develop new theories and metrics for evaluating embodied intelligence that lead to general solutions, rather than point solutions that are subject to Goodhart's law. Extending the ideas presented with respect to verification and validation (Corso et al., 2020) such as adversarial testing for benchmarking is one promising direction. Another open question is how to systematically change evaluation metrics in a principled way to avoid overfitting and point solutions. While an ordered system of evaluations has typically not been considered by the robotics community, the area of meta-learning allows progress in this direction by re-using learned knowledge about the learning process itself to learn new tasks faster. Finally, it is important to note that all of these approaches are linked by the notion of online learning, which raises concerns about the stability in terms of representation and models.

7. Conclusion

Embodied agents that are able to learn new representations, new models of the world and new skills will enable a new generation of robots and autonomous machines that can carry out a vast range of tasks over long time- and length-scales. At some point, the world no longer presents data from a stationary, independent and identically distributed test set. The agent must recognize the potentially unsafe consequences both to itself and to the environment around it as it acquires new data, builds a new model and tests the decisions it takes as a result of that model. The agent must also act to acquire that data, in the context of its own size, weight, and most importantly, energy constraints.

Current systems that rely on the state-of-the-art in machine learning and artificial intelligence are reaching the limits of performance. We posit that these limits are imposed by the challenges described in this article. Remarkably, many of the same challenges of learning for embodied intelligence we describe here have not changed in 30 years (Brooks and Mataric, 1993). Overcoming these challenges will require another step-change in both our understanding of, and the capabilities of, embodied agents and robots. In this article, we have argued that conventional machine learning and artificial intelligence do not adequately address the needs of an embodied agent that learns. There are technical challenges that must be addressed by the different communities building embodied intelligence systems. Our first hypothesis is that the inductive biases must be made explicit for embodied intelligent systems that learn, and in many cases new inductive biases must be developed to address the limitations of the current approaches to learning in an embodied intelligence. Second, we hypothesize that an embodied intelligence that learns must do so at multiple levels of abstraction, and claim that cognitive science theories like the dual process theory and the global workspace theory may in fact provide a path forward to a mature set of inductive biases enjoyed by humans and that may allow learning at such multiple scales and levels of abstraction. The absence of an explicit hierarchical representation of time and length scale is one limitation that has prevented robots from learning to perform complex tasks in stochastic environments. Third, we hypothesize that new languages are required for representing embodied world models at multiple time and length scales. In fact, we claim that some form of symbolic logic grounded in physical perception and action is essential for scalable representations that enable robots to learn and generalize effectively in the real world. Fourth, we hypothesize that robotics in particular has been too committed to spe-

cific morphologies, not only in physical form and actuation but also in sensing. While these actuation and sensing modalities have very much enabled robotic technologies in specific domains such as manufacturing or self-driving vehicles, these same modalities have limited our understanding of what embodied intelligence can really accomplish. Investing in new sensing and actuation modalities may create opportunities for robots with a very different set of capabilities, but such robots may depend heavily on learning for robust perception and control in ways that are not yet well-understood. Finally, embodied intelligent agents that learn must do so safely, and we need ways to verify that safety. We need meaningful ways to compare different agents both in learning and in performance. Existing technologies, metrics and evaluation techniques are not adequate for a variety of reasons, including the risk of over-committing to specific standards and benchmarks. The entire point of learning is to be able to generalize to the unseen, even out-of-distribution, and it is currently very difficult to assess learning performance on novel tasks.

Assessing these hypotheses and progressing our understanding of embodied intelligence is not a trivial task. However, our hope is that these hypotheses are sufficiently concrete to inspire new research directions and new investigations of embodied intelligence that are not currently underway. We look forward to future generations of embodied intelligent agents that are robust to a changing world, are robust to the assumptions of their human designers, and can operate in a populated world with true autonomy.

Acknowledgments

References

Anurag Ajay, Jiajun Wu, Nima Fazeli, Maria Bauza, Leslie P Kaelbling, Joshua B Tenenbaum, and Alberto Rodriguez. Augmenting physical simulators with stochastic neural networks: Case study of planar pushing and bouncing. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3066–3073. IEEE, 2018.

Ferran Alet, Erica Weng, Tomás Lozano-Pérez, and Leslie Pack Kaelbling. Neural relational inference with fast modular meta-learning. In *Advances in Neural Information Processing Systems*, pages 11804–11815, 2019.

B. Ames, A. Thackston, and G.D. Konidaris. Learning symbolic representations for planning with parameterized skills. In *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018.

Brandon Amos, Ivan Dario Jimenez Rodriguez, Jacob Sacks, Byron Boots, and J Zico Kolter. Differentiable MPC for end-to-end planning and control. In *Advances in Neural Information Processing Systems*, 2018.

OpenAI: Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafal Jozefowicz, Bob McGrew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, Jonas Schneider, Szymon Sidor, Josh Tobin, Peter Welinder, Lilian Weng, and Wojciech Zaremba. Learning dexterous in-hand manipulation. *The International Journal of Robotics Research*, 39(1):3–20, 2020.

Laura Antanas, Plinio Moreno, and Luc De Raedt. Relational kernel-based grasping with numerical features. In *International Conference on Inductive Logic Programming*, pages 1–14. Springer, 2015.

Thomas Anthony, Zheng Tian, and David Barber. Thinking fast and slow with deep learning and tree search. In *Advances in Neural Information Processing Systems*, pages 5360–5370, 2017.

Iro Armeni, Zhi-Yang He, JunYoung Gwak, Amir R Zamir, Martin Fischer, Jitendra Malik, and Silvio Savarese. 3D scene graph: A structure for unified semantics, 3D space, and camera. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5664–5673, 2019.

Masataro Asai. Unsupervised grounding of plannable first-order logic representation from images. In *Proceedings of the Twenty-Ninth International Conference on Automated Planning and Scheduling (ICAPS)*, 2019.

Masataro Asai and Alex Fukunaga. Classical planning in deep latent space: Bridging the subsymbolic-symbolic boundary. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2018.

Nikolay Atanasov, Menglong Zhu, Kostas Daniilidis, and George J Pappas. Semantic localization via the matrix permanent. In *Robotics: Science and Systems*, volume 2, 2014.

C. G. Atkeson, B. P. W. Babu, N. Banerjee, D. Berenson, C. P. Bove, X. Cui, M. DeDonato, R. Du, S. Feng, P. Franklin, M. Gennert, J. P. Graff, P. He, A. Jaeger, J. Kim, K. Knoedler, L. Li, C. Liu, X. Long, T. Padir, F. Polido, G. G. Tighe, and X. Xinjilefu. No falls, no resets: Reliable humanoid behavior in the DARPA robotics challenge. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pages 623–630, 2015.

Bernard J Baars. *A cognitive theory of consciousness*. Cambridge University Press, 1993.

Bernard J Baars. In the theatre of consciousness. global workspace theory, a rigorous scientific theory of consciousness. *Journal of consciousness Studies*, 4(4):292–309, 1997.

Akhil Bagaria and George Konidaris. Option discovery using deep skill chaining. In *International Conference on Learning Representations*, 2019.

Dzmitry Bahdanau, Shikhar Murty, Michael Noukhovitch, Thien Huu Nguyen, Harm de Vries, and Aaron Courville. Systematic generalization: what is required and can it be learned? *ArXiv:1811.12889, ICLR'2019*, 2018.

Ruzena Bajcsy. Signal-to-symbol transformation and vice versa: From fundamental processes to representation. *ACM Computing Surveys (CSUR)*, 27(3):310–313, 1995.

Dan Barnes, Will Maddern, and Ingmar Posner. Find your own way: Weakly-supervised segmentation of path proposals for urban autonomy. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, June 2017.

Peter W Battaglia, Jessica B Hamrick, and Joshua B Tenenbaum. Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, 110 (45):18327–18332, 2013.

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. *ArXiv:2106.04399, NeurIPS'2021*, 2021.

Yoshua Bengio. The consciousness prior. *arXiv preprint arXiv:1709.08568*, 2017.

Aditya Bhatt, Adrian Sieler, Steffen Puhlmann, and Oliver Brock. Surprisingly robust in-hand manipulation: An empirical study. In *Proceeedings of Robotics: Science and Systems*, 2021.

Mathew Botvinick, Sam Ritter, Jane X Wang, Zeb Kurth-Nelson, Charles Blundell, and Demis Hassabis. Reinforcement learning, fast and slow. *Trends in Cognitive Sciences*, 2019.

Ivan Bratko. Comparison of machine learning for autonomous robot discovery. In *Advances in Machine Learning I*, pages 441–456. Springer, 2010.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI gym. *CoRR*, abs/1606.01540, 2016.

Rodney A. Brooks and Maja J. Mataric. Real robots, real learning problems. In Jonathan H. Connell and Sridhar Mahadevan, editors, *Robot Learning*, pages 193–213. Springer US, Boston, MA, 1993.

Wolfram Burgard, Abhinav Valada, Noha Radwan, Tayyab Naseer, Jingwei Zhang, Johan Vertens, Oier Mees, Andreas Eitel, and Gabriel Oliveira. Perspectives on deep multimodel robot learning. In *Robotics Research*, pages 17–24. Springer, 2020.

Christopher P Burgess, Loic Matthey, Nicholas Watters, Rishabh Kabra, Irina Higgins, Matt Botvinick, and Alexander Lerchner. MONet: Unsupervised Scene Decomposition and Representation. *arXiv preprint arXiv:1901.11390*, 2019.

Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuScenes: a multimodal dataset for autonomous driving. *arXiv preprint arXiv:1903.11027*, 2019.

Angelo Cangelosi, Josh Bongard, Martin H Fischer, and Stefano Nolfi. Embodied intelligence. In *Springer Handbook of Computational Intelligence*, pages 697–714. Springer, 2015.

Ming-Fang Chang, John Lambert, Patsorn Sangkloy, Jagjeet Singh, Slawomir Bak, Andrew Hartnett, De Wang, Peter Carr, Simon Lucey, Deva Ramanan, et al. Argoverse: 3D tracking and forecasting with rich maps. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8748–8757, 2019.

Javier Civera, Dorian Gálvez-López, Luis Riazuelo, Juan D Tardós, and JMM Montiel. Towards semantic slam using a monocular camera. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1277–1284. IEEE, 2011.

Silvia Coradeschi, Amy Loutfi, and Britta Wrede. A short review of symbol grounding in robotic and intelligent systems. *KI-Künstliche Intelligenz*, 27(2):129–136, 2013.

N. Correll, K. E. Bekris, D. Berenson, O. Brock, A. Causo, K. Hauser, K. Okada, A. Rodriguez, J. M. Romano, and P. R. Wurman. Analysis and observations from the first Amazon picking challenge. *IEEE Transactions on Automation Science and Engineering*, 15(1):172–188, 2018.

Anthony Corso, Robert J. Moss, Mark Koren, Ritchie Lee, and Mykel J. Kochenderfer. A survey of algorithms for black-box safety validation. *arXiv preprint arXiv:2005.02979*, 2020.

Michael T Cox. Metacognition in computation: A selected research review. *Artificial intelligence*, 169(2):104–141, 2005.

Sven Cremer, Lawrence Mastromoro, and Dan O Popa. On the performance of the Baxter research robot. In *2016 IEEE international symposium on assembly and manufacturing (ISAM)*, pages 106–111. IEEE, 2016.

Andrew Cropper, Sebastijan Dumančić, Richard Evans, and Stephen H Muggleton. Inductive logic programming at 30. *arXiv preprint arXiv:2102.10556*, 2021.

Marco F Cusumano-Towner, Feras A Saad, Alexander K Lew, and Vikash K Mansinghka. Gen: a general-purpose probabilistic programming system with programmable inference. In *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation*, pages 221–236, 2019.

Shreyansh Daftry, Sam Zeng, J Andrew Bagnell, and Martial Hebert. Introspective perception: Learning to predict failures in vision systems. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1743–1750. IEEE, 2016.

Christoph Dann, Tor Lattimore, and Emma Brunskill. Unifying PAC and regret: Uniform PAC bounds for episodic reinforcement learning. In *Advances in Neural Information Processing Systems 30*, 2017.

Julie Dequaire, Chi Hay Tong, Winston Churchill, and Ingmar Posner. Off the beaten track: Predicting localisation performance in visual teach and repeat. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 795–800. IEEE, 2016.

Ophelia Deroy, Charles Spence, and Uta Noppeney. Metacognition in multisensory perception. *Trends in cognitive sciences*, 20(10):736–747, 2016.

Thomas G Dietterich. Hierarchical reinforcement learning with the MAXQ value function decomposition. *Journal of Artificial Intelligence Research*, 13:227–303, 2000.

Honghua Dong, Jiayuan Mao, Tian Lin, Chong Wang, Lihong Li, and Denny Zhou. Neural logic machines. In *International Conference on Learning Representations*, 2019.

Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017.

Martin Engelcke, Adam R Kosiorek, Oiwi Parker Jones, and Ingmar Posner. GENESIS: Generative Scene Inference and Sampling with Object-Centric Latent Representations. *International Conference on Learning Representations*, 2020.

Clemens Eppner, Sebastian Höfer, Rico Jonschkowski, Roberto Martín-Martín, Arne Sieverling, Vincent Wall, and Oliver Brock. Four aspects of building robotic systems: lessons from the Amazon Picking Challenge 2015. *Autonomous Robots*, 42(7):1459–1475, 2018.

S.M. Ali Eslami, Nicolas Heess, Theophane Weber, Yuval Tassa, David Szepesvari, and Geoffrey E Hinton. Attend, Infer, Repeat: Fast Scene Understanding with Generative Models. *Advances in Neural Information Processing Systems*, 2016.

Richard Evans and Edward Grefenstette. Learning explanatory rules from noisy data. *Journal of Artificial Intelligence Research*, 61:1–64, 2018.

Benjamin Eysenbach, Abhishek Gupta, Julian Ibarz, and Sergey Levine. Diversity is all you need: Learning skills without a reward function. *arXiv preprint arXiv:1802.06070*, 2018.

Richard E Fikes, Peter E Hart, and Nils J Nilsson. Learning and executing generalized robot plans. *Artificial intelligence*, 3:251–288, 1972.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1126–1135. JMLR. org, 2017.

Jerry A Fodor and Zenon W Pylyshyn. Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2):3–71, 1988.

Luca Franceschi, Mathias Niepert, Massimiliano Pontil, and Xiao He. Learning discrete structures for graph neural networks. *arXiv preprint arXiv:1903.11960*, 2019.

Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR, 2016.

Charles R Gallistel. *The organization of learning.* The MIT Press, 1990.

Caelan Reed Garrett, Tomas Lozano-Perez, and Leslie Pack Kaelbling. FFRob: leveraging symbolic planning for efficient task and motion planning. *The International Journal of Robotics Research*, 2017.

Caelan Reed Garrett, Rohan Chitnis, Rachel Holladay, Beomjoon Kim, Tom Silver, Leslie Pack Kaelbling, and Tomas Lozano-Perez. Integrated task and motion planning. *Annual review of control, robotics, and autonomous systems*, 4, 2021.

Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research (IJRR)*, 2013.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

Anirudh Goyal and Yoshua Bengio. Inductive biases for deep learning of higher-level cognition. *arXiv preprint arXiv:2011.15091*, 2020.

Anirudh Goyal, Alex Lamb, Jordan Hoffmann, Shagun Sodhani, Sergey Levine, Yoshua Bengio, and Bernhard Schölkopf. Recurrent independent mechanisms. *ArXiv:1909.10893, ICLR'2021*, 2019.

Anirudh Goyal, Aniket Didolkar, Nan Rosemary Ke, Charles Blundell, Philippe Beaudoin, Nicolas Heess, Michael Mozer, and Yoshua Bengio. Neural production systems. *arXiv preprint arXiv:2103.01937*, 2021a.

Anirudh Goyal, Aniket Didolkar, Alex Lamb, Kartikeya Badola, Nan Rosemary Ke, Nasim Rahaman, Jonathan Binas, Charles Blundell, Michael Mozer, and Yoshua Bengio. Coordination among neural modules through a shared global workspace. *arXiv preprint arXiv:2103.01197*, 2021b.

Klaus Greff, Antti Rasmus, Mathias Berglund, Tele Hao, Harri Valpola, and Jürgen Schmidhuber. Tagger: Deep Unsupervised Perceptual Grouping. *Advances in Neural Information Processing Systems*, 2016.

Klaus Greff, Raphaël Lopez Kaufmann, Rishab Kabra, Nick Watters, Chris Burgess, Daniel Zoran, Loic Matthey, Matthew Botvinick, and Alexander Lerchner. Multi-Object Representation Learning with Iterative Variational Inference. *International Conference on Machine Learning*, 2019.

Karol Gregor, Danilo Jimenez Rezende, and Daan Wierstra. Variational intrinsic control. *arXiv preprint arXiv:1611.07507*, 2016.

Hugo Grimmett, Rudolph Triebel, Rohan Paul, and Ingmar Posner. Introspective classification for robot perception. *The International Journal of Robotics Research*, 35(7):743–762, 2016.

Oliver Groth, Fabian B Fuchs, Ingmar Posner, and Andrea Vedaldi. Shapestacks: Learning vision-based physical intuition for generalised object stacking. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 702–717, 2018.

Charlie Guan, William Vega-Brown, and Nicholas Roy. Efficient planning for near-optimal compliant manipulation leveraging environmental contact. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 215–222. IEEE, 2018.

Corina Gurău, Chi Hay Tong, and Ingmar Posner. Fit for purpose? Predicting perception performance based on past experience. In *International Symposium on Experimental Robotics*, pages 454–464. Springer, 2016.

Corina Gurau, Alex Bewley, and Ingmar Posner. Dropout distillation for efficiently estimating model confidence. *arXiv preprint arXiv:1809.10562*, 2018.

David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018.

Stevan Harnad. The symbol grounding problem. *Physica D*, 43:335–346, 1990.

Irina Higgins, David Amos, David Pfau, Sebastien Racaniere, Loic Matthey, Danilo Rezende, and Alexander Lerchner. Towards a definition of disentangled representations. *arXiv preprint arXiv:1812.02230*, 2018.

Sebastian Hofer, Kostas Bekris, Ankur Handa, Juan Camilo Gamboa, Florian Golemo, and Melissa Mozifian. 2nd workshop on closing the reality gap in sim2real transfer for robotics, 2020. URL https://sim2real.github.io/.

N. Hogan. Impedance control: An approach to manipulation: Part 1 – theory. *Journal of Dynamic Systems, Measurement, and Control*, 107(1):1–7, 1985.

David Howard, Agoston E. Eiben, Danielle Frances Kennedy, Jean-Baptiste Mouret, Philip Valencia, and Dave Winkler. Evolving embodied intelligence from materials to machines. *Nature Machine Intelligence*, 1:12–19, 2019.

Daniel Joseph Hsu. *Algorithms for active learning*. PhD thesis, UC San Diego, 2010.

Dieuwke Hupkes, Verna Dankers, Mathijs Mul, and Elia Bruni. The compositionality of neural networks: integrating symbolism and connectionism. *arXiv preprint arXiv:1908.08351*, 2019.

Moksh Jain, Salem Lahlou, Hadi Nekoei, Victor Butoi, Paul Bertin, Jarrid Rector-Brooks, Maksym Korablyov, and Yoshua Bengio. Deup: Direct epistemic uncertainty prediction. *arXiv preprint arXiv:2102.08501*, 2021.

Michael Janner, Sergey Levine, William T. Freeman, Joshua B. Tenenbaum, Chelsea Finn, and Jiajun Wu. Reasoning about physical interactions with object-oriented prediction and planning. In *International Conference on Learning Representations*, 2019.

Jindong Jiang, Sepehr Janghorbani, Gerard De Melo, and Sungjin Ahn. SCALOR: Generative World Models with Scalable Object Representations. In *International Conference on Learning Representations*, 2020.

Yifeng Jiang, Tingnan Zhang, Daniel Ho, Yunfei Bai, C. Karen Liu, Sergey Levine, and Jie Tan. Simgan: Hybrid simulator identification for domain adaptation via adversarial reinforcement learning. *CoRR*, abs/2101.06005, 2021.

Leslie Pack Kaelbling. The foundation of efficient robot learning. *Science*, 369(6506):915–916, 2020.

Daniel Kahneman. *Thinking Fast and Slow*. Farrar, Straus and Giroux, 2011.

Daniel Kappler, Jeannette Bohg, and Stefan Schaal. Leveraging big data for grasp planning. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4304–4311. IEEE, 2015.

Peter Karkus, Xiao Ma, David Hsu, Leslie Pack Kaelbling, Wee Sun Lee, and Tomas Lozano-Perez. Differentiable algorithm networks for composable robot learning. In *Robotics: Science and Systems (RSS)*, 2019.

Maximilian Karl, Maximilian Soelch, Justin Bayer, and Patrick Van der Smagt. Deep variational Bayes filters: Unsupervised learning of state space models from raw data. *arXiv preprint arXiv:1605.06432*, 2016.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *ArXiv:1312.6114, ICLR'2014*, 2013.

Marius Kloetzer and Calin Belta. A fully automated framework for control of linear systems from temporal logic specifications. *IEEE Transactions on Automatic Control*, 53(1):287–297, 2008.

Daniel E. Koditschek. What is robotics? why do we need it and how can we get it? *Annual Review of Control, Robotics, and Autonomous Systems*, 4(1):1–33, May 2021.

George Konidaris, Leslie Kaelbling, and Tomas Lozano-Perez. Constructing symbolic representations for high-level planning. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.

George Konidaris, Leslie Pack Kaelbling, and Tomas Lozano-Perez. From skills to symbols: Learning symbolic representations for abstract high-level planning. *Journal of Artificial Intelligence Research*, 61:215–289, 2018.

Adam Kosiorek, Hyunjik Kim, Yee Whye Teh, and Ingmar Posner. Sequential Attend, Infer, Repeat: Generative Modelling of Moving Objects. *Advances in Neural Information Processing Systems*, 2018.

Ioannis Kostavelis, Konstantinos Charalampous, Antonios Gasteratos, and John K Tsotsos. Robot navigation via spatial and temporal coherent semantic maps. *Engineering Applications of Artificial Intelligence*, 48:173–187, 2016.

Hadas Kress-Gazit, Georgios E Fainekos, and George J Pappas. Temporal-logic-based reactive mission and motion planning. *IEEE transactions on robotics*, 25(6):1370–1381, 2009.

Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A Shamma, Michael S. Bernstein, and Fei-Fei Li. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *International Journal of Computer Vision*, 123(1):32–73, 2017.

Robert Krug, Achim J Lilienthal, Danica Kragic, and Yasemin Bekiroglu. Analytic grasp success prediction with tactile feedback. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 165–171. IEEE, 2016.

Norbert Krüger, Christopher Geib, Justus Piater, Ronald Petrick, Mark Steedman, Florentin Wörgötter, Aleš Ude, Tamim Asfour, Dirk Kraft, Damir Omrčen, Alejandro Agostini, and Rüdiger Dillmann. Object–action complexes: Grounded abstractions of sensory–motor processes. *Robotics and Autonomous Systems*, 59(10):740–757, 2011.

Tejas D Kulkarni, Pushmeet Kohli, Joshua B Tenenbaum, and Vikash Mansinghka. Picture: A probabilistic programming language for scene perception. In *Computer Vision and Pattern Recognition*, pages 4390–4399, 2015.

Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman. Building machines that learn and think like people. *Behavioral and brain sciences*, 40, 2017.

G Lakoff and M Johnson. *Metaphors We Live By (Chicago, IL: U. of Chicago P.)*. Univ. of Chicago Press, Chicago, Illinois, USA, 1980.

Hoang Le, Nan Jiang, Alekh Agarwal, Miroslav Dudík, Yisong Yue, and Hal Daumé. Hierarchical imitation and reinforcement learning. In *International Conference on Machine Learning*, pages 2917–2926. PMLR, 2018.

Gregor Leban, Jure Žabkar, and Ivan Bratko. An experiment in robot discovery with ilp. In *International Conference on Inductive Logic Programming*, pages 77–90. Springer, 2008.

Adam Lerer, Sam Gross, and Rob Fergus. Learning physical intuition of block towers by example. *arXiv preprint arXiv:1603.01312*, 2016.

Sergey Levine, Peter Pastor, Alex Krizhevsky, and Deirdre Quillen. Learning hand-eye coordination for robotic grasping with large-scale data collection. In Dana Kulić, Yoshihiko Nakamura, Oussama Khatib, and Gentiane Venture, editors, *2016 International Symposium on Experimental Robotics*, pages 173–184. Springer International Publishing, 2017.

Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:2005.01643*, 2020.

Andrew Levy, George Konidaris, Robert Platt, and Kate Saenko. Learning multi-level hierarchies with hindsight. *arXiv preprint arXiv:1712.00948*, 2017.

Lihong Li, Michael L Littman, Thomas J Walsh, and Alexander L Strehl. Knows what it knows: a framework for self-aware learning. *Machine learning*, 82(3):399–443, 2011.

Rui Li, Robert Platt, Wenzhen Yuan, Andreas ten Pas, Nathan Roscup, Mandayam A Srinivasan, and Edward Adelson. Localization and manipulation of small parts using GelSight tactile sensing. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3988–3993. IEEE, 2014.

Wenbin Li, Aleš Leonardis, and Mario Fritz. Visual stability prediction for robotic manipulation. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2606–2613. IEEE, 2017.

Dianbo Liu, Alex Lamb, Kenji Kawaguchi, Anirudh Goyal, Chen Sun, Michael Curtis Mozer, and Yoshua Bengio. Discrete-valued neural communication. *arXiv preprint arXiv:2107.02367*, 2021.

Jeffrey Mahler and Ken Goldberg. Learning deep policies for robot bin picking by simulating robust grasping sequences. In Sergey Levine, Vincent Vanhoucke, and Ken Goldberg, editors, *Proceedings of the 1st Annual Conference on Robot Learning*, volume 78 of *Proceedings of Machine Learning Research*, pages 515–524. PMLR, 13–15 Nov 2017.

Perla Maiolino, Marco Maggiali, Giorgio Cannata, Giorgio Metta, and Lorenzo Natale. A flexible and robust large scale capacitive tactile system for robots. *IEEE Sensors Journal*, 13(10):3910–3917, 2013.

Cynthia Matuszek, Michael Witbrock, Robert C Kahlert, John Cabral, Dave Schneider, Purvesh Shah, and Doug Lenat. Searching for common sense: Populating CYC from the web. *UMBC Computer Science and Electrical Engineering Department Collection*, 2005.

Cynthia Matuszek, Evan Herbst, Luke Zettlemoyer, and Dieter Fox. Learning to parse natural language commands to a robot control system. In *Experimental robotics*, pages 403–415. Springer, 2013.

T. Mitchell, W. Cohen, E. Hruschka, P. Talukdar, J. Betteridge, A. Carlson, B. Dalvi, M. Gardner, B. Kisiel, J. Krishnamurthy, N. Lao, K. Mazaitis, T. Mohamed, N. Nakashole, E. Platanios, A. Ritter, M. Samadi, B. Settles, R. Wang, D. Wijaya, A. Gupta, X. Chen, A. Saparov, M. Greaves, and J. Welling. Never-ending learning. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15)*, 2015.

Tom M Mitchell. *The need for biases in learning generalizations*. Department of Computer Science, Laboratory for Computer Science Research, 1980.

Douglas Morrison, Peter Corke, and Jürgen Leitner. Learning robust, real-time, reactive robotic grasping. *The International Journal of Robotics Research*, 39(2-3):183–201, 2020.

Jonathan Mugan and Benjamin Kuipers. Autonomous learning of high-level states and actions in continuous environments. *IEEE Transactions on Autonomous Mental Development*, 4(1):70–86, 2011.

Vincent C Müller and Matej Hoffmann. What is morphological computation? On how the body contributes to cognition and control. *Artificial life*, 23(1):1–24, 2017.

Ofir Nachum, Honglak Lee, Shixiang Gu, and Sergey Levine. Data-efficient hierarchical reinforcement learning. In *Advances in Neural Information Processing Systems*, pages 3303–3313, 2018.

Sanmit Narvekar, Jivko Sinapov, and Peter Stone. Autonomous task sequencing for customized curriculum design in reinforcement learning. In *IJCAI*, pages 2536–2542, 2017.

Thu Nguyen-Phuoc, Christian Richardt, Long Mai, Yong-Liang Yang, and Niloy Mitra. BlockGAN: Learning 3D Object-aware Scene Representations from Unlabelled Images. *arXiv preprint arXiv:2002.08988*, 2020.

Daniel Nyga, Ferenc Balint-Benczedi, and Michael Beetz. PR2 Looking at Things: Ensemble Learning for Unstructured Information Processing with Markov Logic Networks. In *IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, May 31-June 7 2014.

Alberto Olivares-Alarcos, Daniel Beßler, Alaa Khamis, Paulo Gonçalves, Maki Habib, J. Bermejo, Marcos Barreto, Mohammed Diab, Jan Rosell, João Quintas, Joanna Olszewska, Hirenkumar Nakawala, Edison Pignaton de Freitas, Amelie Gyrard, Stefano Borgo, Guillem Alenyà, Michael Beetz, and Howard Li. A review and comparison of ontology-based approaches to robot autonomy. *The Knowledge Engineering Review*, 34, 12 2019.

Peter Pastor, Mrinal Kalakrishnan, Sachin Chitta, Evangelos Theodorou, and Stefan Schaal. Skill learning and task outcome prediction for manipulation. In *2011 IEEE International Conference on Robotics and Automation*, pages 3828–3834. IEEE, 2011.

Hanna M Pasula, Luke S Zettlemoyer, and Leslie Pack Kaelbling. Learning symbolic models of stochastic domains. *Journal of Artificial Intelligence Research*, 29:309–352, 2007.

Siddharth Patki, Andrea F Daniele, Matthew R Walter, and Thomas M Howard. Inferring compact representations for efficient natural language understanding of robot instructions. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 6926–6933. IEEE, 2019.

Rohan Paul, Jacob Arkin, Nicholas Roy, and Thomas M Howard. Efficient grounding of abstract spatial concepts for natural language interaction with robot manipulators. In *Robotics: Science and Systems Foundation*, 2016.

Ali Payani. *Differentiable Neural Logic Networks And Their Application Onto Inductive Logic Programming*. PhD thesis, George Institute of Technology, 2020.

Anastasia Pentina and Christoph Lampert. A PAC-Bayesian bound for lifelong learning. In *International Conference on Machine Learning*, pages 991–999, 2014.

Rolf Pfeifer and Josh Bongard. *How the body shapes the way we think: a new view of intelligence*. MIT press, 2006.

Lerrel Pinto and Abhinav Gupta. Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 3406–3413. IEEE, 2016.

Horia Porav, Will Maddern, and Paul Newman. Adversarial training for adverse conditions: Robust metric localisation using appearance transfer. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 1011–1018. IEEE, 2018.

Meng Qu, Yoshua Bengio, and Jian Tang. GMNN: graph Markov neural networks. In *Proceedings of ICML*, 2019.

Vasumathi Raman, Alexandre Donzé, Dorsa Sadigh, Richard M Murray, and Sanjit A Seshia. Reactive synthesis from signal temporal logic specifications. In *Proceedings of the 18th international conference on hybrid systems: Computation and control*, pages 239–248, 2015.

Fabio Ramos, Rafael Carvalhaes Possas, and Dieter Fox. BayesSim: adaptive domain randomization via probabilistic inference for robotics simulators. *arXiv preprint arXiv:1906.01728*, 2019.

Anurag Ranjan, Joel Janai, Andreas Geiger, and Michael J. Black. Attacking optical flow. In *International Conference on Computer Vision (ICCV)*, 2019.

Dushyant Rao, Francesco Visin, Andrei Rusu, Razvan Pascanu, Yee Whye Teh, and Raia Hadsell. Continual unsupervised representation learning. In *Advances in Neural Information Processing Systems 32*, 2019.

Lynne M Reder and Frank E Ritter. What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, memory, and cognition*, 18(3):435, 1992.

Charles Richter and Nicholas Roy. Safe visual navigation via deep learning and novelty detection. In *Robotics: Science and Systems Foundation*, 2017.

Ludovic Righetti, Mrinal Kalakrishnan, Peter Pastor, Jonathan Binney, Jonathan Kelly, Randolph C. Voorhies, Gaurav S. Sukhatme, and Stefan Schaal. An autonomous manipulation system based on force control and optimization. *Autonomous Robots*, 36(1-2): 11–30, 2014.

Fereshteh Sadeghi, Alexander Toshev, Eric Jang, and Sergey Levine. Sim2real view invariant visual servoing by recurrent control. *arXiv preprint arXiv:1712.07642*, 2017.

Claude Sammut, Raymond Sheh, Adam Haber, and Handy Wicaksono. The robot engineer. In *ILP (Late Breaking Papers)*, pages 101–106, 2015.

Adam N Sanborn, Vikash K Mansinghka, and Thomas L Griffiths. Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological review*, 120(2):411, 2013.

Manolis Savva, Abhishek Kadian, Oleksandr Maksymets, Yili Zhao, Erik Wijmans, Bhavana Jain, Julian Straub, Jia Liu, Vladlen Koltun, Jitendra Malik, Devi Parikh, and Dhruv Batra. Habitat: A Platform for Embodied AI Research. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, et al. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588(7839): 604–609, 2020.

Amanda Seed and Richard Byrne. Animal tool-use. *Current biology*, 20(23):R1032–R1039, 2010.

Burr Settles. Active learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences, 2009.

Larry Shapiro. The embodied cognition research programme. *Philosophy compass*, 2(2): 338–346, 2007.

Bokui Shen, Fei Xia, Chengshu Li, Roberto Martín-Martín, Linxi Fan, Guanzhi Wang, Shyamal Buch, Claudia D'Arpino, Sanjana Srivastava, Lyne P. Tchapmi, Micael E. Tchapmi, Kent Vainio, Li Fei-Fei, and Silvio Savarese. iGibson, a Simulation Environment for Interactive Tasks in Large Realistic Scenes, 2020.

David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Yutian Chen, Timothy Lillicrap, Fan Hui, George Sifre, Laurent andG van den Driessche, Thore Graepel, and Demis Hassabis. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

H.A. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138, 1956.

Herbert A. Simon. *The sciences of the artificial*. MIT press, 3 edition, 1996.

Martin Simonovsky and Nikos Komodakis. GraphVAE: towards generation of small graphs using variational autoencoders. In *International Conference on Artificial Neural Networks*, pages 412–422. Springer, 2018.

Aman Sinha, Matthew O'Kelly, Russ Tedrake, and John C. Duchi. Neural bridge sampling for evaluating safety-critical autonomous systems. In *NeurIPS*, 2020.

Elizabeth S Spelke and Katherine D Kinzler. Core knowledge. *Developmental science*, 10 (1):89–96, 2007.

Marilyn Strathern. 'Improving ratings': Audit in the British University system. *European Review*, 5(3):305–321, 1997.

Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, Vijay Vasudevan, Wei Han, Jiquan Ngiam, Hang Zhao, Aleksei Timofeev, Scott Ettinger, Maxim Krivokon, Amy Gao, Aditya Joshi, Yu Zhang, Jonathon Shlens, Zhifeng Chen, and Dragomir Anguelov. Scalability in perception for autonomous driving: Waymo open dataset, 2019.

Richard S Sutton, Doina Precup, and Satinder Singh. Between MDPs and semi-MDPs: a framework for temporal abstraction in reinforcement learning. *Artificial intelligence*, 112 (1-2):181–211, 1999.

Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks. In *International Conference on Learning Representations*, 2014.

Tadahiro Taniguchi, Emre Ugur, Matej Hoffmann, Lorenzo Jamone, Takayuki Nagai, Benjamin Rosman, Toshihiko Matsuka, Naoto Iwahashi, Erhan Oztop, Justus Piater, and Florentin Wörgötter. Symbol emergence in cognitive developmental systems: a survey. *IEEE transactions on Cognitive and Developmental Systems*, 11(4):494–516, 2018.

Stefanie Tellex, Thomas Kollar, Steven Dickerson, Matthew R Walter, Ashis Gopal Banerjee, Seth Teller, and Nicholas Roy. Approaching the symbol grounding problem with probabilistic graphical models. *AI magazine*, 32(4):64–76, 2011.

Jesse Thomason, Shiqi Zhang, Raymond J Mooney, and Peter Stone. Learning to interpret natural language commands through human-robot dialog. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*. Citeseer, 2015.

Sebastian Thrun and Tom M Mitchell. Lifelong robot learning. *Robotics and autonomous systems*, 15(1-2):25–46, 1995.

Dhruva Tirumala, Hyeonwoo Noh, Alexandre Galashov, Leonard Hasenclever, Arun Ahuja, Greg Wayne, Razvan Pascanu, Yee Whye Teh, and Nicolas Heess. Exploiting hierarchy for learning and transfer in kl-regularized rl. *arXiv preprint arXiv:1903.07438*, 2019.

Marc Toussaint. Logic-geometric programming: An optimization-based approach to combined task and motion planning. In *IJCAI*, pages 1930–1936, 2015.

Marc Toussaint, Kelsey R Allen, Kevin A Smith, and Josh B Tenenbaum. Differentiable physics and stable modes for tool-use and manipulation planning. In *Proc. of Robotics: Science and Systems (R:SS 2018)*, 2018.

James Tu, Mengye Ren, Siva Manivasagam, Ming Liang, Bin Yang, Richard Du, Frank Cheng, and Raquel Urtasun. Physically realizable adversarial examples for lidar object detection, 2020.

Vasileios Vasilopoulos, William Vega-Brown, Omur Arslan, Nicholas Roy, and Daniel Koditschek. Sensor-based reactive symbolic planning in partially known environments. In *International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, 2018. IEEE.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.

Rishi Veerapaneni, John D Co-Reyes, Michael Chang, Michael Janner, Chelsea Finn, Jiajun Wu, Joshua Tenenbaum, and Sergey Levine. Entity abstraction in visual model-based reinforcement learning. In *Conference on Robot Learning*, pages 1439–1456. PMLR, 2020.

Julius von Kügelgen, Ivan Ustyuzhaninov, Peter Gehler, Matthias Bethge, and Bernhard Schölkopf. Towards causal generative scene models via competition of experts. *ICLR Workshop on Causal Learning for Decision Making"*, 2020.

L. S. Vygotsky. *Mind in Society: the development of higher mental processes*. Harvard University Press, Cambridge, MA, 1978.

Po-Wei Wang, Priya Donti, Bryan Wilder, and Zico Kolter. SATnet: Bridging deep learning and logical reasoning using a differentiable satisfiability solver. In *International Conference on Machine Learning*, pages 6545–6554. PMLR, 2019.

William Yang Wang, Kathryn Mazaitis, and William W Cohen. A soft version of predicate invention based on structured sparsity. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

Manuel Watter, Jost Springenberg, Joschka Boedecker, and Martin Riedmiller. Embed to control: A locally linear latent dynamics model for control from raw images. In *Advances in neural information processing systems*, pages 2746–2754, 2015.

Rob Weston, Oiwi Parker Jones, and Ingmar Posner. There and back again: Learning to simulate radar data for real-world applications. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2021.

Shimon Whiteson. TreeQN and ATreeC: Differentiable tree planning for deep reinforcement learning. In *Proceedings of the International Conference on Learning Representations*, 2018.

David H Wolpert. The supervised learning no-free-lunch theorems. In *Soft computing and industry*, pages 25–42. Springer, 2002.

David H Wolpert and William G Macready. No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, 1(1):67–82, 1997.

Jiajun Wu, Ilker Yildirim, Joseph J Lim, Bill Freeman, and Josh Tenenbaum. Galileo: Perceiving physical object properties by integrating a physics engine with deep learning. In *Advances in neural information processing systems*, pages 127–135, 2015.

Jiajun Wu, Erika Lu, Pushmeet Kohli, Bill Freeman, and Josh Tenenbaum. Learning to see physics via visual de-animation. In *Advances in Neural Information Processing Systems*, pages 153–164, 2017.

Zhenpei Yang, Yuning Chai, Dragomir Anguelov, Yin Zhou, Pei Sun, Dumitru Erhan, Sean Rafferty, and Henrik Kretzschmar. Surfelgan: Synthesizing realistic sensor data for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11118–11127, 2020.

Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Karol Hausman, Chelsea Finn, and Sergey Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *Conference on Robot Learning*, pages 1094–1100. PMLR, 2020.

Andy Zeng, Shuran Song, Johnny Lee, Alberto Rodriguez, and Thomas Funkhouser. Tossingbot: Learning to throw arbitrary objects with residual physics. *IEEE Transactions on Robotics*, 36(4):1307–1319, 2020.

Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 737–744, 2020.

Gabriel Zöller, Vincent Wall, and Oliver Brock. Active acoustic contact sensing for soft pneumatic actuators. In *2020 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2020.