

## 从机器学习到机器人技术。实体智能的挑战和机遇

尼古拉斯-罗伊

nickroy@csail.mit.edu

edu

美国马萨诸塞州剑桥市, 麻省理工学院

英格玛-波斯纳

ingmar@robots.ox.ac.uk

uk

牛津大学, 英国, 牛津

蒂姆-巴富特

tim.barfoot@utoronto.ca

ca

多伦多大学, 加拿大安大略省多伦多市

Philippe Beaudoin

phil@mywaverly.com

com

加拿大魁北克省蒙特利尔市韦弗利

Yoshua Bengio

yoshua.Bengio@mila.quebec

Quebec

米拉, 蒙特利尔大学, 蒙特利尔, QC加拿大

詹妮特-博格

bohgan@stanford.edu

斯坦福大学, 美国加州斯坦福

奥利弗-布洛克

oliver.brock@tu-berlin.berlin

柏林工业大学, 德国, 柏林

Isabelle Dépatie

isabelle.depatie@elementai.com

Element AI - A ServiceNow Company, 蒙特利尔, QC 加拿大

Dieter Fox

fox@cs.washington.edu

华盛顿大学, 西雅图, 美国

Dan Koditschek@seas.upenn.edu

美国宾夕法尼亚大学, 宾夕法尼亚州费城, 美国

Tomás Lozano-Pérez

tlp@csail.mit.edu

edu

美国马萨诸塞州剑桥市, 麻省理工学院

Vikash Mansinghka

vkm@mit.edu

edu

美国马萨诸塞州剑桥市, 麻省理工学院

克里斯托弗-帕尔

christopher.pal@servicenow.com

Element AI - A ServiceNow Company, 蒙特利尔, QC 加拿大

布莱克-理查兹

Blake.Richards@mila.quebec

米拉, 麦吉尔大学, 加拿大魁北克省蒙特利尔市

Dorsa Sadigh

dorsa@cs.stanford.edu

斯坦福大学, 美国加州斯坦福

Stefan Schaal

stefan.k.schaal@gmail.com

[谷歌]X, 美国加州山景城

高拉夫-苏哈特梅

gaurav@usc.edu

南加州大学, 洛杉矶, 美国

Denis Thérien

denis@servicenow.com

ServiceNow, 蒙特利尔, QC 加拿大

马克-图桑

toussaint@tu-berlin.berlin

柏林工业大学, 德国, 柏林

## 摘要

机器学习早已成为一项关键技术，在广泛的领域内加速了科学和应用。因此，将学习方法应用于特定问题集的概念已经成为推动特定领域发展的一种既定的、有价值的操作方式。在这篇文章中，我们认为这种方法并没有直接扩展到机器人学--或者更广泛地扩展到*具身智能*：与物理环境进行有目的的能量和信息交换的系统。特别是，具身智能代理的范围大大超出了主流机器学习方法的典型考虑，这些方法通常 (i) 不考虑在与训练期间遇到的条件明显不同的情况下的操作。(ii) 不考虑学习和部署过程中的互动往往是实质性的、长期的和潜在的安全关键性的；(iii) 不要求随时适应新的任务，同时(iv) 通过有针对性的和故意的行动，有效和有意识地策划和扩展他们的世界模型。因此，在现实中，这些限制导致了基于学习的系统在部署在机器人上时，与更传统的、基于工程的方法一样，存在许多操作上的缺陷，而这些缺陷往往是在一个明确的、狭窄的操作范围之外。与将具身智能视为机器学习的另一个应用领域相反，我们认为它实际上是推动机器学习技术发展的一个关键因素。在这篇文章中，我们的目标是突出体现智能的挑战和机遇，并提出研究方向，这可能会极大地推动机器人学习的发展。

## 1. 简介

机器人和自主车辆被越来越频繁地部署在越来越多的应用中，从自动驾驶汽车，到越来越多的无人驾驶的工厂环境，再到无人机为人们提供商业送货服务。公众的期望是，越来越多的机器人能够稳健、有效地运行，在各种复杂的领域执行一系列任务。然而，现实情况是，今天的机器人远没有想象中的那么坚固、智能和聪明。今天的机器人在面对突发事件时仍然很脆弱，缺乏通用性，只能在非常狭窄的操作条件下执行一套特定的任务。一个公开的秘密是，今天的绝大多数自主机器人在部署到外面的世界时，在很大程度上依赖于人类的监督和干预。

在过去的十年里，在机器学习的推动下，计算机视觉和自然语言处理的某些任务取得了快速进展。对于诸如人脸识别和机器翻译这样的特殊问题，特别是在网络背景下，工程模型已经无法与学习系统相比。机器人和具身系统已经从机器视觉的机器学习中获得了一些好处，而且有一些特定的能力使诸如自动驾驶汽车这样的系统没有学习就不会存在。然而，机器学习似乎遇到了许多与更传统的工程方法相同的问题，即在仓库、工厂车间或气候不变、阳光充足的城市中精心绘制的街道等既定条件之外操作的机器人上，机器学习会变得脆弱和缺乏通用性。

开发能够可靠学习的有形机器人的挑战表明，机器人和有形代理的学习可能在某种程度上与其他领域不同。

在使用学习方面已经取得了更大的成功。在这篇文章中，我们的目标是强调具体体现在智能方面的挑战和机遇，并提出研究方向，以极大地推动目前机器人学习的最新进展。

## 实体化智能的挑战

尽管居住在真实的物理身体上的智能体经常被称为 "有形的"，但对于什么是智能体 "有形的" 并没有一个共识 (Pfeifer和Bongard, 2006 ; Cangelosi等人。 2015霍华德等人。 2019 ; Savva等人。 2019)。体现是否需要一个身体，或者这个术语是否也可以适用于纯粹的模拟行为的代理人，这是一个正在进行的辩论。因此，我们将具身智能定义为有目的的*能量和信息与物理环境的交换* (Koditschek, 2021)。我们使用 "物理环境" 这一术语，而不是 "现实世界" 的环境，因为物理环境，不管是模拟的还是其他的，都是一个执行不同种类的物理约束的环境。特别是，我们的动机是在机器人领域，我们关注的是在现实世界中和在由此产生的物理约束下行动和互动的具身代理。长期以来，人们一直猜测物理约束给具身智能代理带来了与传统人工智能代理截然不同的学习环境，而使这些代理能够学习的技术要求与不与物理世界互动的学习代理的要求有根本性的不同。

首先，具身的代理人与世界交换能量的事实意味着代理人的行动可能会产生实质性的、持久的影响，这给安全学习带来挑战。代理人往往不知道，世界的某些部分可能带有极高的探索惩罚。实体代理必须意识到潜在的灾难性事件，这些事件可以从根本上结束任何进一步的学习或代理。实体代理还必须意识到他们自己在可用能量和规格功率方面的极限，并能够实时学习。自然界提供了体现型代理的存在证明，它们可以从极少的例子中学习执行复杂的任务，而不危及自己或环境，但人工体现型智能一直在努力复制这种能力。对动物感觉运动智能的行为学研究 (Gallistel, 1990) 和工具的使用 (Seed和Byrne, 2010) 揭示了空间智能和解决问题的能力，这些能力不仅在鲁棒性和准确性方面超过了我们最好的自主系统，而且在数据*efficiency*和功率*efficiency*方面也是如此。

其次，物理世界为训练提供了比任何数据集甚至大多数模拟世界所能提供的更大、更丰富的信息来源。然而，来自物理世界的数据与大多数学习技术通常做出的假设并不一致。例如，学习型分类器通常假设数据在训练和测试过程中是独立和相同的分布 (IID)。IID数据的假设对于现代机器学习理论来说至关重要，但对于物理世界来说却根本不真实。具身的代理人生活在一个非稳态的、部分可观察的世界中，其中的数据可能通过大量的潜在因素而相互关联。数据分布是不断变化的，要求代理人随着时间的推移了解数据中的哪些变化可能仅仅是感知噪声 (或混叠) 的结果。

与其自身行为或其他代理人的干预所导致的变化相比，与可能代表环境根本变化的变化相比。

第三，具身的代理人不能假设他们的目标、规格和奖励在任何时候都是固定不变的。按照西蒙的说法(1956, 1996)，从代理必须完成的任务范围、必须运行的环境，以及任务和环境如何影响代理的智能架构等方面来考虑具身智能代理是很有用的。我们的另一个猜想是，对于一个具身的代理来说，一个好的架构是可以跨越任务和环境的，这意味着代理对任务和规格的表述必须允许泛化和快速适应新的任务和环境。强化学习和决策理论历来通过奖励和损失函数来表示任务和目标。然而，奖励函数不容易适应环境的实质性变化，也不容易适应比简单改变目标状态更复杂的任务变化，事实上，它可能是一种不允许泛化的古老表示。

第四，环境和任务可以改变的事实对代理的学习方式有影响。虽然具身的代理可以获得数量级更高的数据，但任何时候的数据分布都是非常局部的，与代理在其一生中可能遇到的数据的真实分布相比，数据分布的尺度小得多。一个具身的代理必须认识到需要采取行动来获取特定种类的数据，这不仅是为了在当前的任务中表现良好，也是为了建立一个足够丰富的世界理论，使代理能够归纳到未来的任务和未来的环境。也许正是行动和推理之间的反馈循环，使具身的代理人能够学习一个模型，对行动的后果作出可靠的预测，而不仅仅是学习近似输入和输出之间的功能关系。

最后，代理人的形态本身就是体现智力的一个决策变量（Lakoff和Johnson, 1980;Shapiro, 2007).代理人可利用的感官，其执行器的自由度，代理人可利用的特定权力，都对代理人可了解的世界和代理人可决定采取的行动有巨大的影响。代理人的形态也会影响到代理人必须了解其环境的内容，因为具体的体现可以编码归纳偏见，从而大大促进学习。

许多机器人系统的成功是由限制操作环境和任务集的能力驱动的，避免了上述的许多挑战。给定环境和任务的约束使模型和学习者能够有一个归纳性的偏见，我们可以考虑一个归纳性偏见的谱系（见图。 1).在谱系的一端，环境和任务受到如此大的限制，以至于有可能出现非常强烈的归纳性偏差，允许从少量数据中快速可靠地学习。我们为这种成功所付出的代价是所产生的系统的高度专业化，而很少有概括性。在光谱的另一端，归纳性偏差明显较弱（或更强的归纳性偏差仍然难以捉摸），导致学习者有更大的灵活性，但也需要越来越多的训练数据。在这一点上，在极限情况下，没有什么已是已知的，一切都必须学习。在现实中，机器人学家通常发现自己处于这两个极值之间。我们可能会假设一些先验知识，但当系统能够运行足够长的时间，以至于几乎所有的先验假设最终都被违反时

，我们可以认为它是真正的自主。

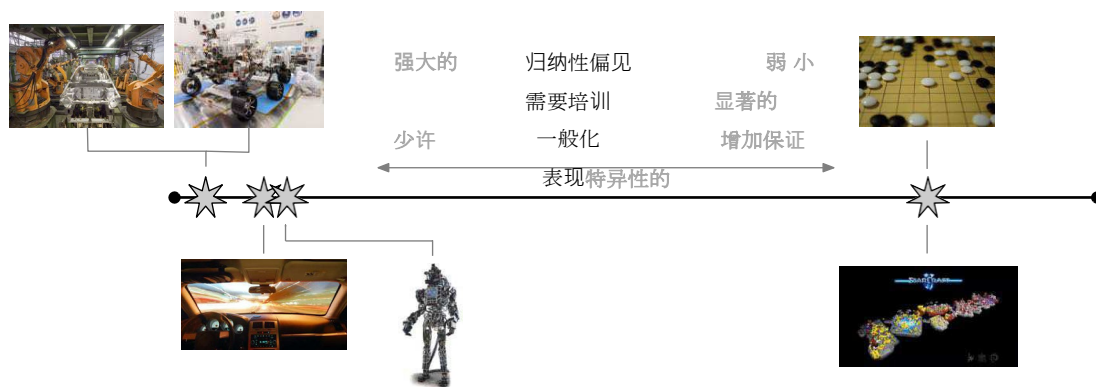


图1：基于代理的机器学习光谱的观点。在一个极端（最左边），一切都是已知的。不需要学习，性能是刚性的，但可以通过传统的系统工程技术来保证。在另一个极端（最右边），没有任何东西是已知的，因此一切都必须从非常大的数据体中学习。在这一端，从理论上讲，一个系统能够很好地从一个任务概括到另一个任务。虽然像MuZero (Schrittwieser等人。 2020)已经展示了具有相当弱的偏见的可能的艺术，但到目前为止，这种成功主要是在游戏环境中取得的。在现实中，人工智能代理通常生活在这些极值之间的真空地带。目前在现实世界中体现的代理的成功仍然集中在特定的任务上，如工厂自动化和自动驾驶，而一般的学习技术还没有被证明可以在现实世界的机器人任务上取得类似的成功。我们认为，目前人工智能技术的进步并不容易将具身代理推向光谱的右侧。随着研究推动完全自主车辆的安全部署，以及更多功能的关节型机器人的发展，弥合这一鸿沟为机器人学习带来了挑战和机遇。

在理解不足、数据匮乏、高度动态、潜在对抗性的领域，如灾难响应、行星探索甚至日常生活中，需要一个成熟的理论来使具身代理正确、efficiently和安全地学习，这一点至关重要。因此，一个核心的挑战是如何选择一个具有学习者的归纳偏见的架构，在一系列未知的和不断变化的任务和环境进行归纳，并符合上述物理约束的要求。

我们首先对具身智能的归纳偏见进行了考察，然后继续对其中的一些偏见进行更详细的探讨。特别是，我们为我们的讨论提出了以下指导性问题的。

- 什么是适当的归纳偏见，使具身智能能够最有效地学习，同时对任务和环境的变化具有鲁棒性？我们猜想，体现式学习的必要归纳偏见是基于行动和感知的，我们将在本节讨论这一猜想。 2.

- 我们如何设计一个具身智能的架构，使其能够有效地学习和稳健地行动？我们猜想，这种架构的一个基本要素是能够有效地摊销推理，同时保持在需求时进行更复杂推理的自省和能力。我们在本节中讨论这一猜想。 3.
- 什么是适合指定具身代理的模型和要执行的任务的表示？我们猜想，许多架构中都不存在一种特定形式的组合性，我们将在本节讨论这一猜想。 4.
- 代理人的形态在多大程度上影响其学习能力？我们猜想，物理设计在很大程度上是对有效的学习方法的补充，不应该被孤立地对待。我们在第二节讨论这个猜想。 5.

最后，体现式智能的一个开放性问题是评估这些系统。从根本上说，作为一个社区，我们如何知道是否正在取得进展？最后，我们讨论了不仅要评估而且要验证能够学习的具身智能的挑战。

## 2. 体现智能的归纳偏见

众所周知，任何学习系统都必须有一个归纳偏见，以便在训练集之外做出预测 (Wolpert 和 Macready, 1997 ; Wolpert, 2002). 归纳性偏见在形式上是 "除了与观察到的训练实例严格一致之外，选择一种归纳的任何依据" (Mitchell, 1980). 体现型代理的学习在要求归纳偏见方面没有什么不同，然而，我们并不是第一个观察到偏见的性质对体现型代理可能非常不同的人 (Thrun 和 Mitchell, 1995 伯加德等人。 2020 ; Kaelbling, 2020). 虽然机器学习中传统的诱导性偏见，如最小描述长度或权重稀疏性，对具身智能是有用的，但额外的架构、算法甚至学习性偏见正日益成为具身代理的焦点。

### 2.1 限制和挑战

考虑对具身代理的学习的挑战如何对归纳偏见产生影响是有用的。我们的第一个挑战是安全地学习，考虑到对学习代理的潜在风险，这需要一种理论，即如何以一种使代理的风险最小化的方式排列模型或假设。人类认知发展领域提供了一种可能的理论，即 "近似发展区" (Vygotsky, 1978)，它经常被用来解释儿童如何探索和学习。维果茨基将近似发展区定义为 "由独立解决问题所决定的实际发展水平与由在成人指导下或与更有能力的同伴合作解决问题所决定的潜在发展水平之间的距离"，实质上就是超出儿童能力的任务集，但儿童在帮助下可以完成。在近似发展区学习是最有效的，原因有很多

但在近似发展区学习的一个重要部分是，它将学习者的风险降到最低。支撑（scaffolding）的概念--教师在近似发展区帮助学习者的过程--利用这一理念进行有效的学习。

为了解决我们的第二个挑战，即当世界是非稳定的，数据不完全分布的，并且代理遇到的数据与它的训练集有很大区别的时候，代理必须能够概括。Fodor和Pylyshyn (1988)将 "系统能力" 的概念定义为："产生/理解某些句子的能力与产生/理解某些其他句子的能力有着内在的联系。"这本身并不是一种古老的归纳性偏见，但Bahdanau等人（）将这一概念扩大为 "系统性能力"。2018)将这一观点扩大为 "系统性概括"，即 "学习如何组成词语的一般规则的能力"。学习导致世界因果理论的一般规则的能力对一个具身的代理很重要，因为在一个非稳态的、非IID的世界中，训练数据永远不能代表所有未来的查询。我们猜想，通过鼓励学习一套一般规则来编码系统性概括的归纳偏见，可能是使代理人对任务和环境的变化具有鲁棒性的最好方法。我们将在第4节中更明确地重新审视这个概念。4.

体现的代理也必须能够管理计算量的增长，这种增长可能来自于在一个开放的环境中的学习。一个有效的方法是在多个抽象层次上进行推理，必要时用高度精确的环境和代理的动态模型进行局部决策，但同样能够使用更抽象的表征进行全局推理（例如，Toussaint, 2015；Toussaint等人。2018；Garrett等人。2017, 2021).几十年来，在决策和任务执行中使用分层抽象一直是一个活跃的研究领域。在深度学习进展的推动下，抽象继续引起越来越多的关注，主要是在强化和模仿学习方面（例如，Nachum等人。2018Le等人。2018；Tirumala等人。2019).然而，除了一些小的例外（如Sutton等人，；Dietterich等人，；Tirumala等人，），分层抽象还没有被大量用于具身代理的学习中。1999; Dietterich, 2000；Levy等人。2017；Konidaris等人。2018).大多数可操作的体现型代理的技术状态是手工编码的抽象，即使是最仔细的手工编码的代表在面对非平稳的世界时也是脆弱的。体现型代理的一个有用的归纳偏向可能是学习自己的抽象的能力，这种能力平衡了由抽象引起的近似的computational efficiency与潜在的性能损失。最近关于状态空间学习和控制的工作可以作为例子（例如，Watter等人。2015卡尔等人。2016），以及越来越多的关于可组合的、以对象为中心的深度生成模型的工作的出现（例如，Eslami等人。2016；Greff等人。2016；Kosiorrek等人。2018；Burgess等人。2019；Greff等人。2019; von Kgelgen et al, 2020; Jiang et al, 2020；Engelcke等人，2020；Nguyen-Phuoc等人。2020).

体现学习的最后一个要求是认识到代理人有一定量的可用能量或特定能力。虽然一些归纳性的偏见，如最小描述长度，会倾向于支持更节能的模型，但有必要对代理人的能量限制如何形成归纳性偏见进行更深入的研究。



## 2.2 机会和未来方向

鉴于这些挑战，有一些可能的方向可以产生归纳偏见，使机器人和其他具身代理的学习形式更加丰富。特别是有三个方向可以满足这些要求，并促进我们对为正在学习的具身代理选择适当的归纳偏见的理解。

归纳偏见的第一种可能形式是利用认知科学中被称为“核心知识假说”(Spelke和Kinzler, 2007)，即人类的智力依赖于四个系统，这四个系统是为特定的推理任务而硬连接的，具有先天的表示能力和先天的限制。这四个系统被设计用来推理物体、行动、数字和空间。<sup>1</sup>更广泛地说，有一些概念似乎对具身智能普遍有用，如物理属性（如质量、惯性、刚性长度等）和物理定律（如牛顿定律、光和声音与物质的相互作用、更普遍的电磁辐射与物质的相互作用等），应该指导学习过程。拥有强烈的归纳偏见，偏爱与物理学一致的模型，似乎是一个非常有用的属性。然而，用一个现实的物理学模型来实现归纳偏见可能是计算上的负担，而且很可能是不完整的，这就提出了这样的问题：一个模型需要多精确才有用，以及偏见应该有多强。对人类物理推理的行为研究表明，人们利用一种嘈杂的、近似的能力来模拟物理对象之间的相互作用，这种能力在质量上与力学一致，但以可量化的方式偏离它 (Sanborn等人。2013 Battaglia等人。2013)。预言的准确性也可能是发展对世界的直观理解和塑造基础模型的一个关键偏向。Ha和Schmidhuber (2018)在训练代理人在所学的世界模型中行动的背景下探讨了这种偏见，并发现代理人能够在一个足够好的模型中学习，但并不完美。

第二种形式的归纳偏见可能是驱动抽象化的结构偏见。虽然一些学科，如自然语言处理，通过减少其对内部表征和结构的明确预设而获得了成功，但还没有实质性的证据表明，体现性智能将使用纯粹的“信号到符号”的归纳偏见获得成功，其中内部表征是输入的量化 (Bajcsy, 1995)--需要一种结构性的偏向，而不仅仅是对输入的量化。对行动的推理充满了对动力学的研究，而由此产生的能量景观所产生的吸引盆地至少有一个 **effectively grounded symbols** 的来源--它们的系统研究启动了拓扑学的领域。系统化的概括表明，模型的构成必须是体现智能的一个关键因素，而盆地的构成语言的前景似乎很光明 (Koditschek, 2021)。

对于神经网络本身是否可以构成，目前还缺乏普遍共识，至少部分原因是对其构成的定义缺乏共识 (Hupkes et al, 2019)。然而，除了相对较少的例外，目前学习组成函数的方法在学习抽象推理系统擅长的那种组成时表现不佳--例如，推断----。

---

1. Spelke和Kinzler (2007)推测，第五个系统可能存在，用于推理“潜在的社会伙伴和社会群体

成员”，这将符合这样的要求，即具身的代理人认识到他们存在于其他具有目标和意图的动态代理人之中。

训练集之外的关联或依赖于全局背景的推理。我们将在本节中重新讨论如何在学习代理中引入构成性作为结构性归纳偏见的问题。4.构成性问题与关于因果关系的归纳性偏见有关。因果知识描述了可以被组成来解释观察结果的基本机制，例如物理定律，并与接下来两节讨论的系统2归纳偏见（Goyal and Bengio, 2020），在接下来的两节中讨论。

第三种隐性形式的归纳偏见可能来自于数据的策划和呈现给代理人的方式。例如，对任务和环境的排序，如课程，会引起对要学习的概念的排序，这通常会使学习过程中的优化路径变得平滑，但也会在学习的内容中产生隐性偏见。课程学习已经被证明是元学习的一种有效形式（Finn et al, 2017纳维卡等人。2017），并且与发展认知心理学的观点有很大的关系，比如近似发展区，它能使学习更有效和安全。然而，虽然最近在解决这一挑战方面有一些努力，例如Botvinick等人（）。2019），但对于具身代理的元学习和相应的归纳偏见之间的关系，以及如何选择支持具身智能的归纳偏见的不同要求的课程，目前还没有一个原则性的理论。

最后，归纳性偏见太强也有危险。对于许多传统的机器学习应用，有一些原则可以为适当地选择归纳偏见提供指导。对于试图从既不独立也不完全分布的少量数据中学习的具身智能来说，当领域的大小和范围无法提前知道时，了解不同的归纳偏见的风险和权衡是很重要的。

### 3. 机器人思考的快与慢

正如我们已经争论过的，当涉及到可以行动和互动的代理时，人工智能的许多进展都发挥了虚拟环境的优势：不确定的训练数据是可用的，无风险的探索是可能的，而且行动基本上是免费的。相比之下，我们要求我们的机器人在实时情况下稳健地运行，从有限的数据中学习，做出对任务、有时对安全至关重要的决定，甚至显示出创造性解决问题的诀窍。认知科学表明，虽然人类也面临着类似的复杂性，但有一些机制使我们能够在现实世界中安全地行动和互动。除了核心知识假说和近似发展区等代表有前途的归纳性偏见的概念外，我们在本节中重点讨论直接受双重过程理论（DPT）启发的一套特殊的架构性偏见。由丹尼尔-卡尼曼的《快与慢的思考》（Kahneman, 2011），DPT推测人类的思维是由两个相互作用的过程产生的。系统一是1,无意识的、直观的反应系统，系统二是更慎重的推理。如果我们接受双重过程理论在我们自己与世界的成功互动中起着核心作用，我们就可以探索一种类似的方法来实现强大的、多功能的和安全的具身智能系统。

DPT的一个关键观察点是，面对每天的挑战，如游戏、驾驶或堆叠盘子，我们通常不

会明确分析其管理法则。

特定的过程。相反，我们倾向于简单地根据情况的需要来行动，就像我们的感官所告知的那样，基于直觉，甚至是一种先天的类似于REFLEX的行为。深度学习的出现主要为我们的代理提供了两件事。(i) 学习从输入到输出的任意复杂映射的能力；(ii) 在恒定时间内执行这些映射的能力。这些，加上以无监督的方式学习结构化的、与任务相关的嵌入的能力，使研究人员对他们采用的计算架构有了不同的看法。学习复杂映射的能力使我们的代理有能力以有用的执行速度完成非常复杂的任务。人类的直接监督、强化学习、任务演示、复杂的学习模型以及日益流行的系统级自我监督的概念都符合这种叙述方式。在游戏中，DeepMind的AlphaGo Zero (Silver et al, 2017) 以及与之密切相关的专家迭代算法 (Anthony et al, 2017)从蒙特卡洛树搜索(神谕)和自我发挥中提炼出知识，并将其转化为一个模型，预测特定棋盘位置的价值和下一步棋的概率。在机器人领域，OpenAI的Learn- ing Dexterity项目 (Andrychowicz等人, 2020)通过强化学习和领域随机化(神谕)将模拟中获得的知识提炼成一个策略，该策略可以控制影子手将一个物体移动到目标姿势中。在自动驾驶的背景下，Barnes等人(2017)通过自动生成训练数据，将数百人小时的系统工程提炼成一个神经网络模型，预测人类在特定情况下可能的驾驶位置。最近关于直观物理学习的工作将通过物理模拟得到的数据提炼成神经网络模型，例如，见Wu等人 ( ) ；Lerer等人 ( ) 。2015) ；Lerer等人(2016) ；Wu等人(2017) ；Li et al.2017) ；Groth等人(2018) ；Janner等人(2019)。由于它们能够以一种时间上(或一般资源上)有效的方式模仿神谕的专业知识，我们可以把神经网络模型的执行视为类似于一种有效的、*直观的*系统反应<sup>1</sup>。

同时，机器人学家和人工智能研究人员已经花了几十年时间来开发Sys- tem 2的等价物。符号推理、传统的计划方法，甚至是简单但耗时的蛮力方法，都构成了深思熟虑的、往往是不可靠的任务解决方法。因此，强大的现实世界的性能似乎需要根据经验对特定的任务和环境进行调整的计算上的efficient策略，以及更多的计算密集型方法，这些方法能够进行系统的概括，对任务和环境的变化具有鲁棒性。根据这里的观点，我们也许能够利用机器学习，将更多的资源密集型、深思熟虑的系统2反应提炼成学习模型，这些模型以一种有效的方式模仿这些专家，形成直观的系统反应<sup>1</sup>。

### 3.1 限制和挑战

将知识提炼成可快速执行的神经网络模型的叙述使我们能够实现重大的、通常是改变游戏规则的计算收益。然而，作为机器人学家，当涉及到在现实世界中应用机器学习系统时，我们

仍然面临着一个实质性的基础挑战：我们系统所做的各种假设的常规违反。正如本节中所讨论的 1 中所讨论的，一个具身的智能体不能假设数据是独立和相同的分布。相反，随着机器人在越来越多的非结构化环境中在越来越长的时间尺度上运行，遇到的数据

显著地偏离了训练分布。此外，我们还期望体现的智能系统能够普及到不可预见的任务和环境。在实践中，加上我们算法的近似性，这些假设导致所学的模型往往过于复杂，其性能只能用传统的方法勉强约束（如果有的话）（例如，Grimmett et al, 2016; Richter和Roy, 2017 ; Rao等人。2019）。其结果是，我们的机器人缺乏可靠地知道它们何时不知道并采取适当补救措施的能力。尽管多年来在弥补这一缺陷方面做了许多尝试（例如，Settles, 2009; Hsu, 2010; Li et al, 2011; Pentina and Lampert, 2014 ; Dann等人。2017），我们仍然没有接近一个切实可行的解决方案。我们的猜想是，双重过程理论的观点很可能为解决这一挑战提供了一条出路。现有方法的第二个局限性体现在大量的证据上，即人类在感知、学习和认知中代表并使用不确定性的估计进行神经计算（Deroy等人, 2016）。然而，元认知的不确定性是如何产生和利用的，只是逐渐被发现。特别的是，人脑中的元认知回路表明，知识的整合超过了原始的感知信号（Deroy et al, 2016）。例如，“了解的感觉”过程使人类能够有效地选择一种认知策略（例如，回忆与推理），以便在特定情况下取得成功（Reder和Ritter, 1992）。

此外，最近关于多感官知觉的工作表明，元认知有助于发现因果结构，以形成从多模态输入中获得连贯的感知（Deroy et al, 2016）。

最后，虽然我们现在可能已经有了一个关于系统分界线两边组件的技术蓝图，但在所涉及的认知过程的性质方面仍有很多不确定性。同样，DPT中提出的将相互作用的神经过程的复杂景观严格归类为两种类型的系统，被广泛认为是一种明显的夸大。然而，它可以作为一个概念性的起点。如何设计一个能有效结合这两个世界的架构，仍然是一个开放的、可能富有成效的研究问题（例如Goyal和Bengio, 2020）。目前还不清楚这两个系统的组件是平行运行还是按需运行，在深思熟虑的计划和低层次的直觉政策之间有明确的交接。与人类面临的元认知挑战相类似，什么时候系统1的反应比系统2的更慎重的部署更合适，仍然是有待解决的问题之一。在人类中，这两种处理形式相互作用，系统可以2抑制、告知甚至训练系统1的反应（Kahneman, 2011）。此外，目前还不清楚该架构是两个独立的系统，还是实际上是一个连续体，或者是能够满足任何一部分的进程的紧密结合。

### 3.2 机会和未来方向

这些限制和挑战立即指向一套现在可行的技术方法，其中下游系统的结果（无论是在成功/失败方面还是在对结果的信心方面）给定的特定输入被提炼成机器学习模型。<sup>2</sup>性能的预测模型现在在机器人文献中已经比较常见了。它们在预测操纵和复杂规划任务的成功率方面有着长期的记录（例如，Pastor等人。2011 ; Kappler等人。2015；

---

2.统计学上的离群点检测也属于这一类。

Krug等人, 2016; Pinto and Gupta, 2016 ; 莫里森等人。 2020), 并越来越多地使用, 例如, 预测感知和基于视觉的导航系统的性能 (例如, Gurău等人。 2016 ; Daftry等人。 2016; Dequaire等人。 2016)。

人类智能比我们最好的学习系统要强大得多, 对不确定性的认识也更深刻, 并且在明显的非平稳 (统计意义上的) 环境中运行。我们利用丰富的元认知能力 (Cox, 2005 德鲁伊等人 2016): 做决定的过程, 知道我们是否有足够的信息来做决定的能力, 以及一旦做出决定后分析结果的能力。一个预测性神经网络可以注入辍学噪声 (Gal和Ghahramani。 2016)或另一个训练有素的网络, 以预测这些预测中的分布外误差的大小, 从而估计认识上的不确定性 (Gurau等人, 2018 ; Jain等人。 2021), 这可以用来指导决策和探索行为。机器人的双重过程理论的一个有趣的方面是, 元认知可能在这样的结构中找到一个自然的位置: 开放的问题是, 元认知的理论是否可以用来连接这两个系统。

另一套可行的技术方法是将因果结构生成模型和蒙特卡洛推理与机器学习相结合的新方法。因果结构模型和蒙特卡洛推理在人工智能和机器人领域有着悠久的历史, 这主要是因为, 它们共同提供了一种方法来对非静止系统的因果结构进行建模, 进行数据精确的参数估计, 并为状态估计建立推理程序, "在它们不知道的时候知道"。不幸的是, 要在因果结构模型中扩大推理的规模, 使之能够实时跟踪大型复杂环境, 并不容易。一种方法是使用缓慢的、精确的推理来使生成模型与数据相吻合, 产生现实的模拟, 然后可以用来为自下而上的神经网络产生完全标记的、合成的训练数据。此外, 我们可以在推理时保留一个较低分辨率的结构化生成模型, 并将神经网络的快速、自下而上的蒙特卡洛更新与较慢的、基于模型的蒙特卡洛相结合。最近, 新的概率编程语言, 如Gen (Cusumano-Towner et al, 2019)已经被开发出来, 使得编写结构化生成模型和通过自定义的神经、符号和蒙特卡洛方法的混合体进行实时、近似的贝叶斯推理变得更加容易。这些方法已被应用于解决三维场景感知问题 (例如, Kulkarni等人, 2006年), 并可以实时运行。 2015), 并且可以实时运行, 同时比只依赖学习的方法产生更稳健的结果。循环模型 "架构的一个好处是, 它们使用因果生成模型下的后验概率密度来表示机器学习算法输出的定量 (相对) 置信度。类似的架构可以应用于基于感知输出的导航和规划任务, 使用自上而下的基于模型的推理来量化快速、自下而上的学习输出的确定性。

一个诱人的机会在于系统计算<sup>1</sup>有助于解决符号人工智能系统的关键计算限制之一, 即难以解决的搜索成本问题: 深度生成模型 (例如Kingma和Welling, 2013 ; Goodfellow等人。 2014 ; Bengio等人。 2021)可能发挥类似于想象力的作用, 因为它们可以被训练来产生与更昂贵的系统<sup>2</sup>顺序推理步骤相一致的搜索或计划的良好候选者 (例如Ha和



Schmidhuber。 2018).

与DPT相关的一个认知科学理论是来自Baars的全球工作空间理论（GWT）。1993, 1997), 该理论假设在专门的大脑模块之间存在一个通信瓶颈, 这些专家模块竞争能够通过这个瓶颈发送内容, 并向所有模块广播。与DPT的联系是, 这个瓶颈是人类可以口头报告的有意识的工作记忆, 即对应于系统内容2。然而, 在相互竞争的专家模块内进行的计算的细节并不是有意识地可以获得的（而且可能太复杂, 不能完全用语言表达）, 而更多地是对应于系统机制1。一个有趣的问题是这样一个强大的交流约束的目的。Bengio (2017)以及Goyal和Bengio (2020)假设, 它诱发了一种归纳性的偏见, 使抽象的系统2层面的世界建模能够很好地捕捉到人类通过语言描述的那种稀疏的依赖关系和因果机制。下一节将阐述多层次的抽象和表示的概念, 以及人类在更高的系统层次2上倾向于使用的那种离散表示和顺序推理过程。

在本节的最后, 我们指出, 不仅存在解决具身代理中的学习的机会, 而且反过来也存在推进我们认知科学知识的机会。后者往往需要复杂的实验程序, 从设计上来说, 需要破坏代理人的学习过程。另一方面, 机器人学允许设计和密切监测参与学习过程的机制--包括对个别部件、环境和互动模式的控制。本节的大部分内容表明, 在认知科学中已经发现的机制可能会给现有的机器人工作带来新的启示, 目的是建立一个与双重过程理论相当的有意义的技术。特别是, 它们可以提供一个蓝图, 以构建我们仍然缺少的组件, 从而建立更强大的、多功能的、可解释的和安全的具身代理。相反, 正如Botvinick等人()所指出的那样, 智能架构的发现也是如此。2019)所指出的, 成功地提供这种双重过程功能的智能架构的发现, 同样可以为认知科学提供富有成效的研究方向。

#### 4. 逻辑在体现学习的智能系统中的作用

在确定了推理和决策系统需要能够在不同的抽象层次上进行推理后, 我们把注意力转向了什么表征能够在不同的层次上进行学习的问题。学习感觉运动感知和控制的表征（即系统1）是目前研究得很好的问题, 这一点在感知和控制的深度表征方面的大量工作中得到了证明。然而, 对于深思熟虑的推理和决策来说, 考虑不同类别的表征语言是至关重要的。需要深思熟虑推理的问题的一个特点是, 它们涉及到抽象概念的动态组合形式、长范围和高维度, 使用系统1机器可能更难实现（Lake et al, 2017 ; Bahdanau等人。2018).系统的策略是2,通过利用问题中的结构来使这些问题具有可操作性, 包括将问题分解成弱相互作用的子组件（可能涉及状态空间的不同方面或不同的时间子区域）, 以及“提升”, 或对领域中任意的实体集进行抽象。传统上, 那些能够有效地实现紧凑表示的语言和

在非常大的问题中进行科学推理的是逻辑学，包括命题逻辑、一阶逻辑、人工智能规划领域描述语言（PDDL）、图形模型（可被视为命题逻辑的概率版本）、概率规划语言、概率程序和时间逻辑。同样地，拓扑学的工具（以及其代数的com-putational efficacy）与机器人学的关系类似于逻辑学与计算机科学的关系。

与以神经网络为基础的模型相比，以任何这些语言表示的模型都更接近人类工程师提供的规格说明。然而，所有这些建模模式的推理和学习问题在本质上并没有什么不同：在每一种情况下，模型都可以用来进行预测或阻止行动选择，它们可以通过机器学习的方法对数据进行调整，努力使模型的输出与数据集相匹配。然而，模型测试的算法技术可能有很大的不同，而且什么样的问题与什么样的模型最匹配的选择还没有被充分理解。因此，对表示语言的选择取决于它所适合的问题类型和基本的归纳偏见--它是否适合手头的问题，从而大大降低学习的样本复杂性？许多逻辑表征的一个重要特征是组合性，即模型的各个部分具有独立的语义，允许它们独立学习，然后以组合的方式来解决许多不同的问题，提供一种特别积极和有用的概括形式。正是通过这种组合性，我们才有希望实现本节介绍的系统化泛化。

2.

#### 4.1 限制和挑战

在相当长的时间里，学习逻辑表征一直是一个活跃的研究领域。逻辑表征学习的主要焦点之一是归纳逻辑编程（ILP），其目标是推断出一个包含给定数据集的世界逻辑理论。虽然ILP的使用以前已经被研究过，用于体现智力（Bratko, 2010），包括移动导航（Leban et al, 2008）和学习抓取（Antanas等人, 2015），这些结果并没有显示出与其他方法相同的成功水平。如何学习代表抽象动作的组合规则，如学习STRIPS或PDDL的动作表示，从Shakey的早期（Fikes et al, 1972）到最近的成果（Konidaris等人, 2014; Sammut等人, 2015）。然而，之前的大部分工作都假设是确定性的，而且往往是命题性的表示。学习随机的、事实性的和提升的感知模型的能力（Nyga et al, 2014）和行动模型（Mugan和Kuipers, 2011；Köger等人, 2011）对于在物理世界中操作是至关重要的。特别是，Pasula等人(2007)认为，命题学习、关系学习和口述学习代表了不同的归纳偏见，而“口述学习提供了一个强大的偏见，可以提高概括能力”。不幸的是，通过ILP学习逻辑表征，其复杂度往往随着要学习的理论的大小而呈指数级增长，并且在处理嘈杂或不一致的数据时也会有困难。最近的进展显示了神经启发的归纳法的效果，如可重复的ILP（Evans和Grefenstette, 2018; Payani, 2020），它可以有效地处理噪声数据。神经逻辑机（Dong et al, 2019）已经表明

如何在非常简单的物理环境中恢复一套解除的规则，同时避免可扩展性问题。然而，任何类型的ILP或学习逻辑机在具身智能中的应用都是有限的，是一个巨大的开放性挑战。

大多数现有的学习逻辑表征的工作的另一个局限性是依赖于世界的初始理论，通常被称为"背景知识"，包括一组初始的谓词和命题，其真值要么是先验已知的，要么可以从观察中推断出来。在非常简单的领域中，如家系推理（ILP界的一个常见基准，Wang等人，2006）或积木世界，它可能是一个非常简单的问题。2015)或块状世界，写下一个初始背景知识理论可能是合理的，它足以支持归纳出一套完整的谓词和规则。然而，背景知识是"类似于大多数形式的ML中使用的特征"（Cropper等人，2021）。正如深度学习大大减少了许多问题对手写特征函数的需求一样，目前对手写背景知识理论的依赖有可能是体现式智能系统学习逻辑表征的一个实质性限制。最接近于从完全连续的感觉运动表征中学习符号表征的技术仍然依赖于运动基元或技能的背景知识（Konidaris et al, 2018）。同时，正如本节中所讨论的2第2节中描述的"核心知识假说"似乎表明，生物系统本身依赖于背景知识。背景知识是否是体现性智能的限制性归纳偏见，或者是否需要更多的研究来确定体现性智能的正确背景知识，这是一个开放的问题。

我们的学习型具身代理的概念是最终需要代表那些无法提前预期的概念。能够从原始的传感器数据中生长出原始的逻辑表示，这个过程有时被称为"符号的出现"（Taniguchi等人，2006），对于一个自主的代理来说是很重要的能力。2018)是自主代理的一项重要能力。已经有工作试图从完全连续的传感器运动表征中学习离散的规划领域模型（Asai和Fukunaga, 2018；Asai, 2019；Ames等人，2018）。此外，以前的一些工作表明，这些学习行动的离散表征的方法可以与用于处理视觉输入的卷积神经网络结合起来，创建一个端到端的系统，既学习视觉输入的符号表征，又学习用于计划玩具领域（如数独）的行动逻辑理论（例如，Wang等人，2019董建华等人，2019）。同样，以前的工作表明，部分表征可以提前学习，然后被一个反应式系统招募，该系统通过实时改变坐标将实时感觉实例转变为其学习的拓扑模型（Vasilopoulos et al, 2018）。然而，如何将这些初步的结果扩展到具身智能系统所遇到的更复杂的领域，还没有得到充分的理解。

此外，为了获得逻辑表征的优势，如composition，学习过程的归纳偏见不仅要满足具身智能的物理约束，还要满足逻辑推理过程的需要。例如，在因式分解中，必须找到一个分解，使问题的不同方面相对独立。在提升过程中，对象必须以一种允许有用的抽象的方式被"重新定义"。这种对基本离散结构的寻找需要一种归纳性的倾向，即把世界分解成独立的和可组合的部分。

例如，可以通过一个以能量为基础的符号的类型理论来正式确定 (Koditschek, 2021)。

最后，即使有一个完善的离散表征，一个具身的智能也存在于一个连续的环境中，抽象的、通常是离散的表征与具体的、通常是连续的输入和输出信号之间必须有一个联系。抽象表征和物理世界之间的这种联系就是 "符号接地问题" (Harnad, 1990)。然而，现有的符号接地模型几乎不能满足具身代理所面临的所有挑战 (Coradeschi et al, 2013)；例如，符号接地的学习模型再次假设数据是IID的，并且有一个固定的、已知的符号字母表，要与输入和输出信号接地。目前学习符号接地的方法中的归纳偏见还没有原则性的技术来确定什么时候世界已经改变，相应的学习的符号接地模型已经改变。近年来，诸如用于选项或技能发现的无序学习的方法 (Gregor et al, 2016；Eysenbach等人。 2018; Bagaria and Konidaris, 2019) 试图学习基元和离散水平的技能，以便为长线规划等任务纳入组合性。虽然仍是初步的，但这些工作提出了一个很有希望的方向，即把逻辑的一些好处纳入神经模型。

## 4.2 机会和未来方向

开发新的技术来推断允许构成性的基础或原始表征，是学习型具身代理的主要开放问题之一。越来越多的证据表明，许多神经网络的内部层可以以一种与离散逻辑表征有关的方式对组合性符号进行编码。例如，"解除纠缠"技术可以用来强迫学习到的表征是因子化的，到目前为止，结果显示了关于学习到的表征的令人惊讶的直观结构 (Higgins等人, 2018)。这些结果表明了学习作为逻辑推理基础的底层抽象的途径，尽管这项工作还没有使推断组合式表征的性能得到大幅提高。

学习提升模型的新技术存在机会，图神经网络 (GNNs) 为神经启发的提升模型提供了一个可能的机制，即一个固定维度的模型可以被学习，然后应用于具有任意数量对象的领域。嵌入在图形模型中的神经网络 (Krishna等人, 2017Armeni等人。 2019) 或完整的GNN已经被应用于场景感知和实体抽象 (Veerapaneni等人, 2020Qu等人。 2019)。然而，GNN还没有被证明能够捕捉到量化的力量，除了一些有希望的初步工作 (Simonovsky and Komodakis, 2018；Franceschi等人。 2019Alet等人。 2019)，其假设是图的结构是提前知道的。

在人类自然语言中，确实已经存在一种可表达的、合理的、完整的语言，它是可行的，并支持抽象化。在为具身代理提供符号基础方面，已经有相当多的工作，这些工作依赖于自然语言来提供需要基础的符号集 (Tellex et al, 2011；Matuszek等人。 2013; Thomason等人。 2015; Paul等人。 2016; Patki等人。 2019)，然而，这些方法还没有证明机器人有能力获得一个大型的知识库，而不需要CON-RAP。

这需要大量的人为干预和努力。这项工作的困难之处在于需要高度注释的、统一的数据体；现在有机会采用新技术，使具身智能以自我监督的方式从语言中获得新的符号。例如，通过阅读互联网来学习符号或逻辑知识库的尝试（Matuszek等人。2005米切尔等人。2015；Olivares-Arlcos等人。2019）已经取得了更多的进展，但是仍然不完整，对具身智能没有实质性的影响。

关于背景知识是否应该作为学习者的先验，就像本节中讨论的归纳性偏见一样，或者说学习者是否应该利用系统性概括等原则从头开始推导其世界理论，这是一个开放的问题。2中所讨论的那样，或者说，学习者是否应该尝试使用系统概括等原则，从头开始推导其世界理论。越来越多的证据表明，一个非常合理的方法是提前确定一个简化的模型，例如一个基于物理学的模型，然后学习纠正由抽象引起的错误（Ajay等人。2018曾志强等人。2020）。另外，有证据表明，在神经网络中嵌入整个物理模型并对其进行端到端的训练，可能会导致模型更加稳健并适合它们所面临的问题分布（Whiteson, 2018；Amos等人。2018；Karkus等人。2019）。最近的技术在很大程度上依赖于在大型offline数据集上预训练一个通用模型--遵循类似于通用语言模型（如GPT-3）的范式--尽管它通常不是构成性的，但可以作为背景知识（Levine等人，2020）。

一个有趣的方向是在神经网络工程中加入系统2的归纳偏见（例如Goyal和Bengio, 2020）。受GWT启发的一个出发点（见3）的一个出发点是构建一个模块化的架构，在这个架构中，模块竞争被激活并进行交流（Goyal et al, 2019）。使用注意力机制来操作元素集（而不是矢量），就有可能实现一个工作存储器作为通信瓶颈（Goyal等人，2021b），就像在GWT中那样，产生更好的分布外泛化。迫使模块之间交换的信息使用抽象概念的共享词汇进行离散化，可以进一步改善模块化架构（Liu等人，2021年），包括非常流行的变形金刚。2021），包括非常流行的变形金刚（Vaswani et al, 2017）。这种模块化架构也可以通过连续应用更简单的类似于规则但由小型前MLP模块实现的学习操作来强制处理信息，产生“神经生产系统”（Goyal等人，2021a）。虽然这些灵感来自于经典的人工智能生产系统，但所有的规则都是可学习的，并且可以在分布式表征上操作：注意力机械的端到端学习可以再次被用来动态控制，哪些规则在什么时候，以什么顺序应用，以最小化一些训练损失。

最后，逻辑表征经常被用于需要保证推理的正确性的领域，这在确保许多工程系统的物理安全方面特别有用，并且似乎有希望确保具身代理的安全（Kress-Gazit et al, 2009；Kloetzer和Belta, 2008；Raman等人。2015）。然而，从结构上看，学习过程本身不能提供任何其他的统计保证。正如我们将在本节中讨论的6对于具身智能来说，统计保证提供了鲁棒

性的声明，但不是安全声明；学习者周围的一些外部结构需要确保安全属性。对于什么样的学习可以应用于正式的表征语言，以保持提供正确性保证的能力，这是一个开放的问题。

## 5. 形态学对体现智能的影响

机器人的形态引入了一个特别强烈但往往不被认可的归纳偏见。代理人拥有什么样的传感器，它可以用什么样的自由度行动，它的运动动力学，它自己与环境互动的刚性，它可以与环境互动的程度，都对代理人可以学习和不能学习什么，以及什么计算可以 "加载 "到身体上产生巨大的影响 (Hogan, 1985Pfeifer和Bongard。 2006Muller和Hoffmann。 2017)。然而，物理代理目前是在一个相当有限的范围内构建的。绝大多数车辆都依赖于高精度的测距传感（通常基于激光雷达）和高精度的驱动装置。这些设计选择的动机是现有的估计、感知和控制算法的要求，即拥有代理人的传感器和执行器的高度精确模型，以及对代理人的状态和周围环境的高度精确的了解。

### 5.1 限制和挑战

虽然具有精确传感和驱动的机器人和具身代理的空间包括许多有用的平台类型，如制造机器人、自动驾驶汽车、无人驾驶航空器等，但今天最常见的机器人设计方式有一些限制性后果。具有精确的、僵硬的驱动的机器人往往要么非常慢，要么具有高能量，这使得人们与它们一起工作很困难，而且可能不安全。高精度驱动通常在尺寸、重量、功率和成本方面都很昂贵，而且这种机器人操作起来很困难，限制了对具身智能的学习问题感兴趣的人对它们的广泛采用。由刚性、铰接式链接构成的高度精确的机器人通常必须避免与环境接触，除非是以特定的、仔细控制的方式，这极大地限制了它们通过与环境互动来学习的能力。这是一个例子，说明机器人的形态如何影响到可以学习什么样的概念。一个有趣的反例是由 **Rethink Robotics**生产的低成本制造机器人**Baxter**和**Sawyer**。由于采用了低能量的串联弹性致动器，这些机器人足够安全，人们可以在旁边工作，这推动了机器人学习社区对它们的采用。这些机器人也是相对不精确的，但事实证明 (Li et al, 2014克雷默等人。 2016 ; Guan等人。 2018)，尽管终端检测器不精确，但具身智能可以学习有效和准确的操纵策略。

在传感方面也存在着类似的情况，因为绝大多数运行中的通用机器人在很大程度上依赖于激光测距仪，其感知世界的方式与大多数生物系统截然不同。虽然激光测距仪在探测障碍物和建立详细而密集的环境几何图以规划无碰撞运动方面非常可靠，但这些传感器提供的关于世界视觉外观的信息是有限的（即使在利用强度信息时）。世界上大量的外观信息（标志、地标等）是只用激光测距仪导航的车辆所无法获得的。虽然现在大多数机器人将测距和被动视觉结合起来，但每种传感方式都是用来感知非常具体的信息形式，这些信息基本上是正交的。



其他。对于如何交易不同的传感器，特别是在学习背景下，人们的理解非常有限。与高精度驱动一样，高精度传感在尺寸、重量、功率和成本方面也很昂贵。一个成功的反面例子是Skydio无人机，它使用大量立体排列的相机展示了惊人的导航自主性。然而，Skydio飞行器所使用的内部表征与激光测距仪所使用的密集几何图非常相似，这就造成了在每个表征层面上没有利用所有语义信息的归纳性偏差。直到最近，才开始出现使用所谓的“语义信息”进行导航的结果（Civera等人。2011；Atanasov等人。2014；Kostavelis等人。2016）没有密集的几何模型支持。假设感知是由一个范围测量系统提供的，所引起的归纳偏见是另一个例子，说明机器人的形态如何影响到什么样的概念可以被学习。

然而，尽管有证据表明代理人的形态和感觉模式在学习引起了很大的归纳性偏差，但对形态和传感器模式究竟如何影响学习能力的理解相对较少。例如，人们对功率、质量、传感和计算之间的权衡很不了解，特别是它们如何影响学习。没有原则性的方法来修改形态学以解决特定的归纳偏见。同样，我们对如何设计一个能在特定范围的任务和环境稳健学习的具身代理理解有限；我们的代理往往是为一个环境设计的，当遇到一个稍微不同的环境或任务时，就会立即失败。<sup>3</sup>。

## 5.2 机会和未来方向

鉴于这些挑战，有一些可能的方向，学习可以用来利用更广泛的代理形态。学习可能是降低建造和操作真实机器人的成本的一种方式。更顺从的执行器和顺从的身体通常建造起来更便宜，在人周围操作也更安全，而且精度只是有时需要。与此相关，机器人可以被设计得更加坚固，以应对在线学习中可能遇到的情况（Bhatt et al, 2021）。现在有机会设计出只在必要时才学习精确的机器人，使用与目前使用的非常不同的控制策略，利用运动学和动态约束，如接触或惯性。声音等模式提供了一个潜在的丰富而低成本的感觉流，但在很大程度上仍未被利用（Zler et al, 2020）。用活性材料构建部件代表了一个领域，在这个领域中，学习可能对开发高能力的控制系统非常有用。同样，学习可以使传感器的融合形式更加丰富。一个很好的例子是GelSight传感器（Li et al. ），它通过传感器融合扩大了对合理的机器人传感器的认识。2014），它使用机器人指尖的凝胶的可见变形作为触觉感应的一种形式。现代机器学习技术使其能够将变形解释为力和接触信息。学习也可能是解开高密度、大面积触觉感应系统的关键。

---

3.许多DARPA机器人项目试图通过要求在一系列环境中进行评估来克服这个问题，如快速轻量级自主项目，该项目将飞行器硬件标准化，但允许传感器的变化。一系列的环境似乎确实导致了参与者选择具有不同操作特性的不同传感器模式。不幸的是，没有一种传感器形态对所有的环境都是稳健的，这表明了一种技术差距。

如Maiolino等人(2013), 这些都在不断出现。同样, 改变机器人的形态和设计可以让我们更好地理解在能源消耗、计算复杂性、可及性和可观察性方面的现有限制。

正如引言中所讨论的, 传统的机器学习和具身代理的学习之间的一个重大区别是, 是否符合学习者假设的训练数据。当每个新的设计都必须实际制造, 然后用来收集学习数据时, 探索不同的形态、传感和驱动模式的空间是不切实际的。仿真器需要有足够高的成熟度, 以使机器人的学习发生在与目前的范式有很大的不同。理想情况下, 新类别的模拟器可以实现学习方法和硬件本身的共同设计。长期以来, 机器人学习的核心原则是模拟 "注定要成功" (Brooks and Mataric, 1993).然而, 模拟的状态, 以及由此产生的模拟价值, 正在迅速变化, 模拟的明显局限性事实上可能不是所有模拟的根本。照片逼真的游戏引擎越来越普遍, 可以使具身的代理从图像中学习。物理模拟器的能力也越来越强, 尽管一些物理现象在实时模拟方面具有挑战性, 如变形、摩擦效应和流体效应。然而, 随着更好的、更具有物理真实性和照片真实性的模拟的出现, 在新的机器人类型的背景下存在着新的学习机会。模拟器的进步、安全学习和模拟到真实的方法的结合, 可能会使设计在形态上有更广泛的变化, 并相应地对形态如何帮助实现智能体现行为有深入的了解。

## 6. 评估机器人学习

与其他依靠学习方法的应用相比, 具身代理可以对他们的环境产生非常直接和物理上的破坏性影响, 从而引起人们对安全的关注。检验、验证和基于基准的性能评估是安全部署主动控制系统的基石, 并且越来越成为部署自主系统的核心组成部分。然而, 如何验证和核实能学习和适应新经验的具身代理的性能是一个开放的问题, 很可能需要新的原则来评估我们的具身代理。

### 6.1 学习型具象代理的验证和确认

传统的验证和确认技术假定在一个环境中对所需的系统行为有一个预先的规定。验证系统的安全性意味着在环境干扰下测试后没有发现故障, 或者故障概率低于某个阈值。安全验证意味着系统在所有干扰下都是可证明的安全。因此, 验证和核实都需要一个干扰模型。对于大多数现实的机器人系统来说, 除了非常狭窄的任务和环境之外, 这样的规格化和现实的干扰模型几乎是不可能指定或取样的。很可能, 传统的验证要求与许多学习的目标直接冲突。此外, 大多数学习方法所提供的保证充其量是统计学上的, 描述的是总体或渐进行为, 而不是提供对学习的保证。

对所学模型的任何瞬时查询。这些类型的保证也经常建立在前面讨论的假设上，即训练和测试分布是相同的，这在实际情况下很难实现。当有可能确定训练集与当前的领域、任务和环境有多大差距时，概率保证可以得到扩展（见章节 3）。鉴于这些挑战，学习型具身代理的成功验证和确认将需要新颖的方法，基于创新的方式来描述所需的系统行为。

一个有趣的技术方向是对抗性测试的想法。虽然对抗性例子已经被用来证明学习模型的潜在失败，但有可能自动生成对抗性例子来测试一个自主系统的反应。当涉及到对抗性图像或其他知觉输入时，这是一个非常活跃的研究领域（例如，Szegedy等人。2014；Ranjan等人。2019；Tu等人。2020），但也提出了一个问题，即环境的对抗性干扰如何能够自动生成，并对闭环行为作出反应而改变(Sinha等人，2020)。学习系统正确性证明的本质可能是不存在对抗性例子，至少在特定的任务环境中是这样。考虑到现实问题中的高样本复杂性，挑战不仅在于找到这样的例子，而且还在于确保学习系统对这些例子具有鲁棒性。Corso等人(2020)对通过仿真对黑盒系统进行安全验证进行了广泛的调查。这些系统搜索导致被评估系统行为为不正常的缺陷，并可能使用优化、规划或强化学习技术进行搜索。

## 6.2 学习型具身代理的性能评估

评估机器人学中基于学习的方法的性能，或称“基准”，在许多方面与验证和确认类似。在基准测试中，我们对这样一个系统的表现感兴趣，特别是与其他在类似任务和环境中的测试的系统相比。基准测试与验证的二元评估不同，基准测试可以用来衡量开发过程中的进展，而不仅仅是衡量最终系统是否达到目标性能水平。

目前对真实机器人系统或数据的基准测试可以大致分为三种。首先，在特定的情况下，有一些解决方案的轶事演示。这些通常是概念的证明。第二，有一些数据集，例如，拣选和放置(Levine et al, 2017; Mahler和Goldberg, 2017; Pinto and Gupta, 2016)或自动驾驶(Geiger等人, 2013; Caesar等人。2019；Sun等人。2019)，这些测试通常是对特定的感知技能的测试，如从视觉数据中推断出一个稳健的把握，准确的物体检测或街道场景中的语义分割。第三，有一些机器人的挑战(Correll等人，2009)。2018；Atkeson等人。2015；Righetti等人。2014；Eppner等人。2018)，最突出的是DARPA的挑战，这些挑战往往揭示了我们在机器人技术的特定子领域所取得的成就：操纵、导航、驾驶、救灾场景下的任务等。

机器人技术对基于学习的解决方案的基准提出了一个特殊的挑战。机器人是一个闭环

系统，它的行动会影响到环境的状态，而环境的状态又会影响到它的下一步行动。这种反馈回路通常会阻止使用特定的数据集进行评估，就像通常用于评估算法的基准那样。

计算机视觉或自然语言处理。例如，像KITTI (Geiger等人) 这样的数据集的成功。 2013) 和最近的Argoverse数据集 (Chang等人, 2019)已经帮助推动了自动驾驶汽车的估计和感知方面的进展，但它们还没有对决策产生类似的影响。关注自动驾驶汽车的主要公司--如Waymo、特斯拉和其他公司，已经公开讨论了模拟的重要性，以及他们目前用于测试和技术开发的大量模拟驾驶。完全互动的模拟可以进行性能评估和学习。此外，罕见的角落案例可以稍作修改，以增加互动场景，使系统得到更好的测试和培训。在这种情况下，开源的CARLA模拟环境 (Dosovitskiy et al, 2017)和CARLA团队开展的挑战赛对更广泛的学术和工业研究界起到了重要作用。

仿真是一种系统评估机制，不需要在现实世界中部署昂贵而繁琐的机器人。然而，历史上机器人的模拟是出了名的低质量，在模拟和现实世界之间有很大的领域差距。诸如所谓的"sim2real"技术，使用学习模型来缩小模拟和真实数据分布之间的差距，最近在允许具身智能学习有效的控制政策方面取得了一些成功 (Sadeghi et al, 2017; Ramos等人, 2019; Zhao等人, 2020; Hofer等人, 2020); 尽管如此，模拟现实的感觉数据 (例如，Porav等人, 2018; Yang等人, 2020; Weston等人, 2021)仍然是一个挑战，而模拟接触则更加困难。虽然许多基于学习的方法是在模拟中开发和评估的，但目前还不清楚它们是否能在现实世界中发挥作用。这类基准的例子包括Meta-world (Yu等人, 2020)、Habitat-AI中的点导航 (Savva等人, 2019) iGibson (Shen等人, 2020)，以及OpenAI Gym (Brockman等人, 2016)。Andrychowicz等人(2020)提出了一个罕见的例子，将在这种模拟环境中学习到的策略转移到真正的机器人上，但是尽管在一个随机领域中进行了大量的模拟训练迭代，成功率还是下降了50%。一个可能的方法是考虑"real2sim"，它试图确定一个更接近真实世界例子的模拟器 (Jiang et al 2021)。

在机器人技术中，功能和通用性之间一直存在着矛盾。基于学习的系统有望在许多机器人应用中普及到未曾见过的情况，但当与特定应用的方法相比 (即使有基于学习的组件)，完全基于学习的方法往往在特定的基准上表现出数量级的糟糕表现。因此，出现了这样一个问题：对于一个学习型具身代理来说，什么才是正确的衡量标准--如果最好的衡量标准是在任何给定任务上的表现，或者衡量标准应该是在许多任务上或意外情况下的概括性，这是一个开放的问题。具体来说，古德哈特定律经常在起作用，斯特拉瑟恩将其总结为"当一个指标成为一个目标时，它就不再是一个好的指标" (斯特拉瑟恩, 1997)。这种现象可以在机器人挑战赛中看到，其重点是赢得比赛，而不是找到可能普及到给定任务的更大变化的问题解决方案。虽然通过挑战赛获得了一些一般性的见解 (例如，操纵的顺应性，取放的

吸盘)，但这些挑战赛往往暴露了解决比赛特定目标的点解决方案，但并没有解决实际问题（Atkeson等人。 2015阿特克森等人；科雷尔等人.....）。 2018).

当务之急是开发新的理论和衡量标准来评估具身智能，从而找到一般的解决方案，而不是受制于古德哈特定律的点解决方案。扩展所提出的关于验证和确认的想法 (Corso等人, 2020)，如用于基准测试的对抗性测试是一个有希望的方向。另一个开放的问题是如何以一种有原则的方式系统地改变评价指标，以避免过度评价和点解决方案。虽然机器人界通常不考虑有序的评价系统，但元学习领域通过重新利用学习过程本身的知识来更快地学习新的任务，从而在这个方向上取得进展。最后，需要注意的是，所有这些方法都是由在线学习的概念联系起来的，这引起了人们对表示和模型方面的稳定性的关注。

## 7. 总结

能够学习新的表征、新的世界模型和新的技能的具身代理将使新一代的机器人和自主机器能够在很长的时间和长度范围内执行广泛的任務。在某些时候，世界不再呈现来自固定的、独立的和相同分布的测试集的数据。代理人必须认识到对其自身和周围环境的潜在不安全后果，因为它获得了新的数据，建立了一个新的模型，并测试了它作为该模型的结果所采取的决定。代理人还必须在其自身的尺寸、重量和最重要的能量限制的背景下，采取行动来获取这些数据。

目前依靠最先进的机器学习和人工智能的系统已经达到了性能的极限。我们认为，这些极限是由本文所描述的挑战造成的。值得注意的是，我们在这里描述的许多关于具身智能学习的挑战30年来都没有改变 (Brooks and Mataric, 1993).要克服这些挑战，需要我们在对具身代理和机器人的认识和能力方面有新的改变。在这篇文章中，我们认为传统的机器学习和人工智能并不能充分满足具身代理学习的需要。构建具身智能系统的不同团体必须解决一些技术挑战。我们的第一个假设是，对于学习的具身智能系统来说，必须明确归纳偏见，而且在许多情况下，必须开发新的归纳偏见，以解决目前具身智能系统中学习方法的局限。第二，我们假设一个学习的具身智能必须在多个抽象层次上进行学习，并声称像双重过程理论和全局工作空间理论这样的认知科学理论事实上可能提供了一条通往人类享有的一套成熟的归纳偏见的道路，并可能允许在这样的多个规模和抽象层次上学习。缺乏明确的时间和长度尺度的层次表示是阻碍机器人学习在随机环境中执行复杂任务的一个限制。第三，我们假设需要新的语言来代表多个时间和长度尺度的具身世界模型。事实上，我们声称某种形式的以物理感知和行动为基础的符号逻辑对于可扩展的表征是至关重要的，它可以使机器人在现实世界中有效地学习和概括。第四，我们假设，特别是机器人学已经太过执着于spe-in的研究了。

体现智能的形态，不仅在物理形态和驱动方面，而且在感知方面。虽然这些驱动和传感方式在很大程度上使机器人技术在特定领域（如制造业或自动驾驶汽车）得以实现，但这些方式也限制了我们对具身智能的真正成就的理解。对新的传感和执行模式的投资可能会为具有不同能力的机器人创造机会，但这样的机器人可能在很大程度上依赖于学习，以实现强大的感知和控制，而这种方式还没有被充分理解。最后，学习的具身智能代理必须安全地进行，我们需要有方法来验证这种安全性。我们需要有意义的方法来比较不同代理的学习和性能。由于各种原因，现有的技术、衡量标准和评估技术是不够的，包括对特定标准和基准的过度承诺的风险。学习的全部意义在于能够归纳出未见过的东西，甚至是分布之外的东西，而目前评估新任务的学习性能是非常困难的。

评估这些假设并推进我们对具身智能的理解并不是一项琐碎的任务。然而，我们希望这些假设是足够具体的，以激发新的研究方向和目前尚未进行的对具身智能的新调查。我们期待着未来几代的具身智能代理，它们对不断变化的世界有很强的适应性，对人类设计者的假设有很强的适应性，并能在人口密集的世界中真正自主地运作。

## 鸣谢

本文是由Element AI赞助并举行的为期三天的研讨会的成果，我们非常感谢他们的支持。Leslie Kaelbling对本文中的一些想法提供了帮助，我们也非常感谢她的大量时间和帮助。

## 参考文献

Anurag Ajay, Jiajun Wu, Nima Fazeli, Maria Bauza, Leslie P Kaelbling, Joshua B Tenenbaum, and Alberto Rodriguez.用随机神经网络增强物理模拟器。平面推挤和弹跳的案例研究。在*IEEE/2018RSJ智能机器人和系统国际会议 (IROS)* 上，第3066-3073页。IEEE, 2018.

Ferran Alet, Erica Weng, ~~Th~~ Lozano-Pérez, and Leslie Pack Kaelbling.带有快速模块化元学习的神经关系推理。在*神经信息处理系统的进展中*，第1180-11815页。2019.

B.Ames, A. Thackston, and G.D. Konidaris.为具有参数化技能的规划学习符号表征。在2018年*IEEE/RSJ智能机器人和系统国际会议*的论文集中。2018.

Brandon Amos, Ivan Dario Jimenez Rodriguez, Jacob Sacks, Byron Boots, and J Zico Kolter.用于端到端计划和控制的重复的MPC。在*神经信息处理系统的进展中*。2018.



OpenAI: Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafal Jozefowicz, Bob Mc- Grew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, Jonas Schneider, Szymon Sidor, Josh Tobin, Peter Welinder, Lilian Weng, and Wojciech Zaremba.学习灵巧的手部操作。《国际机器人研究杂志》, 39 (1) : 3-20。 2020.

Laura Antanas, Plinio Moreno, and Luc De Raedt.基于关系核的数字特征抓取。在《国际归纳逻辑编程会议上》, 第1-14页。Springer, 2015.

Thomas Anthony, Zheng Tian, and David Barber.用深度学习和树状搜索来思考快与慢。在《神经信息处理系统的进展中》, 第5360-5370页。 2017.

Iro Armeni, Zhi-Yang He, JunYoung Gwak, Amir R Zamir, Martin Fischer, Jitendra Malik, and Silvio Savarese.3D场景图。一个用于统一语义、三维空间和相机的结构。在《IEEE/CVF计算机视觉国际会议论文集中》, 第5664-5673页。 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.深潜空间中的经典规划。弥合亚符号-符号边界。在《AAAI人工智能会议 (AAAI) 的论文集中》。 2018.

Nikolay Atanasov, Menglong Zhu, Kostas Daniilidis, and George J Pappas.通过矩阵永久进行语义定位。In *Robotics: 科学与系统》* , 第1卷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. Xin-jilefu.不跌倒, 不重置。DARPA机器人挑战赛中可靠的仿人行为。在2015年IEEE-RAS第15届仿人机器人国际会议 (Humanoids) 上, 第623-630页。 2015.

Bernard J Baars.*A cognitive theory of consciousness*.剑桥大学出版社。 1993.

Bernard J Baars.在意识的剧场里。全球工作空间理论, 一个严格的科学的意识理论。《意识研究杂志》, 4 (4) : 292-309。 1997.

Akhil Bagaria 和 George Konidaris.使用深度技能链的期权发现。In *International Conference on Learning Representations*, 2019.

Dzmitry Bahdanau, Shikhar Murty, Michael Noukhovitch, Thien Huu Nguyen, Harm de

Vries, and Aaron Courville.系统泛化：需要什么，能学会吗？*ArXiv:1811.12889*, *ICLR'2019*, 2018.

Ruzena Bajcsy. 信号到符号的转换, 反之亦然。From fundamental processes to representation. *ACM 计算调查 (CSUR)*, 27 (3) : 310-313, 1995.

丹-巴恩斯, 威尔-马德恩, 和英格玛-波斯纳。找到你自己的路。用于城市自治的弱监督分割路径建议。在 *IEEE 国家间机器人和自动化会议 (ICRA)* 的论文集中, 新加坡, 6月 2017.

Peter W Battaglia, Jessica B Hamrick, and Joshua B Tenenbaum. 仿真作为物理场景理解的引擎。 *美国国家科学院院刊*, 110 (45) : 18327-18332. 2013.

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. 基于流网络的生成模型, 用于非迭代的多样化候选人生成。 *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. 令人惊讶的强大的手部操纵。一项经验研究。In *Proceedings of Robotics: 科学与系统*. 2021.

Mathew Botvinick, Sam Ritter, Jane X Wang, Zeb Kurth-Nelson, Charles Blundell, and Demis Hassabis. 强化学习, 快与慢。 *认知科学的趋势*, 2019.

Ivan Bratko. 用于自主机器人发现的机器学习的比较。在 *机器学习的进展I* 中, 第441-456页。Springer, 2010.

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

Rodney A. Brooks 和 Maja J. Mataric. 真正的机器人, 真正的学习问题。在 Jonathan H. Connell 和 Sridhar Mahadevan, 编辑的 *《机器人学习》* 中, 第193-213页。Springer US, Boston, MA, 1993.

Wolfram Burgard, Abhinav Valada, Noha Radwan, Tayyab Naseer, Jingwei Zhang, Johan Vertens, Oier Mees, Andreas Eitel, and Gabriel Oliveira. 关于深度多模型机器人学习的观点。在 *机器人研究* 中, 第17-24页。斯普林格。 2020.

Christopher P Burgess, Loic Matthey, Nicholas Watters, Rishabh Kabra, Irina Higgins, Matt Botvinick, and Alexander Lerchner. MONet: *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. *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。用丰富的地图进行三维跟踪和预测。在*IEEE/CVF 计算机视觉和模式识别会议*上, 第8748-8757页。2019.

Javier Civera, Dorian Alvez-López, Luis Riazuelo, Juan D. Luis, and JMM Montiel.使用单眼相机实现语义上的抨击。在2011年*IEEE/RSJ 智能机器人和系统国际会议*上, 第1277-1284页。IEEE, 2011.

Silvia Coradeschi, Amy Loutfi, and Britta Wrede.机器人和智能系统中符号基础的简短回顾。 *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.第一届亚马逊采摘挑战赛的分析和观察。 *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. 计算中的元认知。一个精选的研究回顾。 *Artificial intelligence*, 169(2):104-141, 2005.

Sven Cremer, Lawrence Mastromoro, and Dan O Popa.论巴克斯特研究机器人的性能。在2016年*IEEE 装配与制造国际研讨会 (ISAM)* 上, 第106-111页。IEEE。2016.

Andrew Cropper, Sebastijan Dumančić, Richard Evans, and Stephen H Muggleton. 诱导性逻辑编程在30.*arXiv 预印本 arXiv:2102.10556*。2021.

Marco F Cusumano-Towner, Feras A Saad, Alexander K Lew, and Vikash K Mansinghka. Gen : 一个具有可编程推理功能的通用概率编程系统。在第40届*ACM SIGPLAN 编程语言设计与实现会议*上, 第221-236页。2019.

Shreyansh Daftry, Sam Zeng, J Andrew Bagnell, and Martial Hebert. 自省的感知。学习预测视觉系统的故障。在*IEEE/2016 RSJ 国际智能机器人和系统会议 (IROS)* 上, 第1743-1750页。IEEE, 2016.

Christoph Dann, Tor Lattimore, and Emma Brunskill. 统一PAC和遗憾。偶发强化学习的统一PAC界限。In *Advances in Neural Information Processing Systems 30* , 2017.

朱莉-德奎尔, 池海棠, 温斯顿-丘吉尔和英格玛-波斯纳。Offthe beaten track:预测视觉教学和重复中的定位性能。在2016年IEEE国际机器人和自动化会议 (ICRA) 上, 第795-800页。IEEE, 2016.

Ophelia Deroy, Charles Spence, and Uta Noppeney.多感官知觉中的元认知。《认知科学的趋势》, 20 (10) : 736-747。 2016.

Thomas G Dietterich.用MAXQ值函数分解进行分层强化学习。*Journal of Artificial Intelligence Research*, 13:227-303, 2000.

董宏华, 毛嘉远, 田林, 王崇, 李丽红, 和周丹妮。神经逻辑机。In *International Conference on Learning Representations*, 2019.

Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun.Carla : 一个开放的城市驾驶模拟器。在《机器人学习会议上》, 第1-16页。PMLR。 2017.

Martin Engelcke, Adam R Kosiorek, Oiwi Parker Jones, and Ingmar Posner.GENESIS : 以物体为中心的潜在表征的生成性场景推理和取样。《国际学习表征会议》, 2020.

Clemens Eppner, Sebastian Höfer, Rico Jonschkowski, Roberto Martín-Martín, Arne Sieverling, Vincent Wall, and Oliver Brock.构建机器人系统的四个方面 : 亚马逊采摘挑战赛的教训。《自主2015.机器人》, 42 (7) : 1459-1475。 2018.

S.M. Ali Eslami, Nicolas Heess, Theophane Weber, Yuval Tassa, David Szepesvari, and Geoffrey E Hinton.注意、推断、重复 : 用生成模型快速理解场景。《神经信息处理系统的进展》, 2016.

Richard Evans and Edward Grefenstette.从嘈杂的数据中学习解释规则。

《人工智能研究》杂志, 61:1-64。 2018.

Benjamin Eysenbach, Abhishek Gupta, Julian Ibarz, and Sergey Levine.Diversity is all you need:arXiv preprint arXiv:1802.06070, 2018.

Richard E Fikes, Peter E Hart, and Nils J Nilsson.学习和执行广义的机器人计划。《人工智能》, 3 : 251-288。 1972.

Chelsea Finn, Pieter Abbeel, and Sergey Levine.用于快速适应深度网络的模型诊断元学习。在《第34届国际机器学习70会议论文集-卷》, 第1126-1135页。JMLR. org, 2017.

Jerry A Fodor and Zenon W Pylyshyn.连接主义和认知结构。A critical analysis.《认知》, 28 (1-2) : 3-71。 1988.

Luca Franceschi, Mathias Niepert, Massimiliano Pontil, and Xiao He.学习图神经网络的

离散结构. *arXiv 预印本* *arXiv:1903.11960*, 2019.

Yarin Gal and Zoubin Ghahramani. 作为贝叶斯近似的辍学。代表深度学习中的模型不确定性。在*国际机器学习会议上*, 第1050-1059页。PMLR, 2016.

Charles R Gallistel. *学习的组织*. The MIT Press, 1990.



Caelan Reed Garrett, Tomas Lozano-Perez, and Leslie Pack Kaelbling. FFRob : 利用符号规划来实现精确的任务和运动规划。 *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. 综合任务和运动规划。 *控制、机器人学和自主系统* 4, 的年度评论。 2021.

Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. 视觉与机器人技术的结合。 KITTI数据集。 *国际机器人研究杂志 (IJRR)* 。 2013.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *神经信息处理系统* 27, 的进展, 2014.

Anirudh Goyal 和 Yoshua Bengio. *arXiv preprint arXiv:2011.15091*, 2020.

Anirudh Goyal, Alex Lamb, Jordan Hoffmann, Shagun Sodhani, Sergey Levine, Yoshua Bengio, and Bernhard Schölkopf. 循环性独立机制。 *ArXiv:1909.10893, ICLR'2021*, 2019.

Anirudh Goyal, Aniket Didolkar, Nan Rosemary Ke, Charles Blundell, Philippe Beaudoin, Nicolas Heess, Michael Mozer, and Yoshua Bengio. 神经生产系统。 *arXiv 预印本 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. *arXiv 预印本 arXiv:2103.01197*, 2021b.

Klaus Greff, Antti Rasmus, Mathias Berglund, Tele Hao, Harri Valpola, and Jürgen Schmidhuber. Tagger: 深度无监督的知觉分组。 *神经信息处理系统的进展*。 2016.

Klaus Greff, Raphaël Lopez Kaufmann, Rishab Kabra, Nick Watters, Chris Burgess, Daniel Zoran, Loic Matthey, Matthew Botvinick, and Alexander Lerchner. 迭代变异推理的多目标再现学习。 *International Conference on Machine Learning*, 2019.

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

Hugo Grimmer, Rudolph Triebel, Rohan Paul, and Ingmar Posner. 机器人感知的内视分类。 *The International Journal of Robotics Research*, 35(7):743-762, 2016.

Oliver Groth, Fabian B Fuchs, Ingmar Posner, and Andrea Vedaldi. Shapestacks: 学习基于

视觉的物理直觉的通用物体堆叠。 *欧洲计算机视觉会议 (ECCV) 论文集*, 第702-717页。  
。 2018.

- Charlie Guan, William Vega-Brown, and Nicholas Roy.利用环境接触进行近乎最优的顺应性操纵的智能规划。在2018年IEEE国际机器人和自动化会议 (ICRA) 上, 第215-222页。IEEE, 2018.
- Corina Gurău, Chi Hay Tong, and Ingmar Posner. 适合的目的? 基于过去的经验来预测感知性能。在实验机器人学国际研讨会上, 第454-464页。Springer, 2016.
- Corina Gurau, Alex Bewley, and Ingmar Posner.Dropout distillation for efficiently estimating model confidence. *arXiv preprint arXiv:1809.10562*, 2018.
- David Ha和Jürgen Schmidhuber.World models. *arXiv preprint arXiv:1803.10122*,2018 .Stevan Harnad.符号接地问题。 *Physica D*, 43:335-346, 1990.
- Irina Higgins, David Amos, David Pfau, Sebastien Racaniere, Loic Matthey, Danilo Rezende, and Alexander Lerchner. 邁向不相干表征的定義。 *arXiv 預印本 arXiv:1812.02230*。 2018.
- Sebastian Hofer, Kostas Bekris, Ankur Handa, Juan Camilo Gamboa, Florian Golemo, and Melissa Mozifian.第二届缩小现实差距的机器人模拟转移研讨会, 2020. URL [2https://simreal.github.io/](https://simreal.github.io/).
- N.Hogan.阻抗控制。一种操纵方法。第1部分-理论。 *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.从材料到机器的体现性智能的发展。 *Nature Machine Intelligence*, 1:12-19, 2019.
- Daniel Joseph Hsu.*Algorithms for active learning*.博士论文, 加州大学圣地亚哥分校, 2010年。Dieuwke Hupkes, Verna Dankers, Mathijs Mul, and Elia Bruni. 构成性的符号主义和连接主义的整合.....*arXiv preprint arXiv:1908.08351*, 2019.
- Moksh Jain, Salem Lahlou, Hadi Nekoei, Victor Butoi, Paul Bertin, Jarrod Rector-Brooks, Maksym Korablyov, and Yoshua Bengio.Deup:*arXiv 预印本 arXiv:2102.08501*。 2021.
- Michael Janner, Sergey Levine, William T. Freeman, Joshua B. Tenenbaum, Chelsea Finn, and Jiajun Wu.用面向对象的预测和规划来推理物理交互。In *International Conference on Learning Representations*, 2019.
- Jindong Jiang, Sepehr Janghorbani, Gerard De Melo, and Sungjin Ahn.SCALOR : 具有可扩展对象表示的生成性世界模型。在*国际学习表征会议上*。 2020.

Yifeng Jiang, Tingnan Zhang, Daniel Ho, Yunfei Bai, C. Karen Liu, Sergey Levine, and Jie Tan. Simgan : 通过对抗性强化学习进行领域适应的混合模拟器识别。 *CoRR*, abs/2101.06005, 2021.

Leslie Pack Kaelbling. 学习的基础。《科学》, 369(6506):915916,- 2020.

丹尼尔-卡尼曼《思考的快与慢》。Farrar, Straus and Giroux, 2011.

Daniel Kappler, Jeannette Bohg, and Stefan Schaal. 利用大数据进行抓取计划。在2015年IEEE国际机器人和自动化会议 (ICRA) 上, 第4304-4311页。IEEE, 2015.

Peter Karkus, Xiao Ma, David Hsu, Leslie Pack Kaelbling, Wee Sun Lee, and Tomas Lozano-Perez. 用于可组合机器人学习的可区分算法网络。在《机器人学》。Science and Systems (RSS), 2019.

Maximilian Karl, Maximilian Soelch, Justin Bayer, and Patrick Van der Smagt. Deep variational Bayes filters: arXiv preprint arXiv:1605.06432, 2016.

Diederik P Kingma and Max Welling. 自动编码的变异贝叶斯。ArXiv:1312.6114, ICLR'2014, 2013.

Marius Kloetzer 和 Calin Belta. 一个完全自动化的框架, 用于控制来自时间逻辑规格的线性系统。IEEE Transactions on Automatic Control, 53(1):287297,- 2008.

Daniel E. Koditschek. 什么是机器人学? 我们为什么需要它, 我们如何才能得到它? 控制、机器人学和自主系统年度评论, 4(1):1-33, 5月 2021.

George Konidaris, Leslie Kaelbling, and Tomas Lozano-Perez. 构建高层规划的符号代表。在第二十八届AAAI人工智能会议上。2014.

George Konidaris, Leslie Pack Kaelbling, and Tomas Lozano-Perez. 从技能到符号。为抽象的高层规划学习符号表征。人工智能研究杂志, 61:215-289。2018.

Adam Kosior, Hyunjik Kim, Yee Whye Teh, and Ingmar Posner. 顺序出席, 推断, 重复: 移动物体的生成模型。神经信息处理系统的进展, 2018.

Ioannis Kostavelis, Konstantinos Charalampous, Antonios Gasteratos, and John K Tsotsos. 通过空间和时间连贯语义图的机器人导航。人工智能的工程应用, 48:173-187, 2016.

Hadas Kress-Gazit, Georgios E Fainekos, and George J Pappas. 基于时间逻辑的再主动任务和运动规划。IEEE 机器人交易, 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. 视觉基因组。使用众包的密集图像注释连接语言和视觉。国际计算机视觉杂

志, 123 (1) : 32-73。 2017.

Robert Krug, Achim J Lilienthal, Danica Kragic, and Yasemin Bekiroglu.带触觉反馈的分析性抓取成功预测。在2016年IEEE国际机器人和自动化会议 (ICRA) 上, 第165-171页。IEEE, 2016.

Norbert K̈ger, Christopher Geib, Justus Piater, Ronald Petrick, Mark Steedman, Florentin Ẅgitter, Aleř Ude, Tamim Asfour, Dirk Kraft, Damir Omṙen, Alejandro Agostini, and R̈diger Dillmann.物体-行动复合体。感觉-运动过程的基础抽象。机器人学与自主系统, 59(10): 740-757。2011.

Tejas D Kulkarni, Pushmeet Kohli, Joshua B Tenenbaum, and Vikash Mansinghka.图片。一种用于场景感知的概率编程语言。在计算机视觉和模式识别中, 第4390-4399页。2015.

Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman.建立像人一样学习和思考的机器。行为和脑科学40,。2017.

G Lakoff和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. 层次模仿和强化学习。在国际机器学习会议上, 第2917-2926页。PMLR。2018.

Gregor Leban, Jure Źabkar, and Ivan Bratko. 用ilp进行机器人发现的实验。在归纳逻辑编程国际会议, 第77-90页。Springer, 2008.

Adam Lerer, Sam Gross, and Rob Fergus.通过实例学习块状塔的物理直觉。arXiv预印本arXiv:1603.01312。2016.

Sergey Levine, Peter Pastor, Alex Krizhevsky, and Deirdre Quillen.用大规模数据收集学习机器人抓取的手眼协同排序。在Dana Kulić, Yoshihiko Nakamura, Oussama Khatib, and Gentiane Venture, editors, 2016 International Symposium on Experimental Robotics, pages 173-184.斯普林格国际出版公司。2017.

Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu.Offline reinforcement learning:arXiv预印本arXiv:2005.01643, 2020.

Andrew Levy, George Konidaris, Robert Platt, and Kate Saenko.arXiv preprint arXiv:1712.00948, 2017.

Lihong Li, Michael L Littman, Thomas J Walsh, and Alexander L Strehl.知道它知道什么

: 一个自我意识学习的框架。 *机器学习*, 82 (3) : 399-443。 2011.

李锐, 罗伯特-普拉特, 袁文振, 安德烈亚斯-特帕斯, 内森-罗斯库普, 曼达亚姆-斯里尼瓦桑, 以及爱德华-阿德尔森。使用GelSight触觉感应定位和操纵小部件。在2014年 *IEEE/RSJ 智能机器人和系统国际会议*上, 第3988-3993页。IEEE, 2014.



- 李文斌, Aleš Leonardis, 和 Mario Fritz. 机器人操纵的视觉稳定性预测。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. 通过模拟稳健的抓取序列为机器人拣选垃圾学习深度策略。在 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, 11月13-15日 2017.
- Perla Maiolino, Marco Maggiali, Giorgio Cannata, Giorgio Metta, and Lorenzo Natale. 一个用于机器人的灵活而坚固的大规模电容式触觉系统。 *IEEE 传感器杂志*, 13 (10) : 3910-3917. 2013.
- Cynthia Matuszek, Michael Witbrock, Robert C Kahlert, John Cabral, Dave Schneider, Purvush Shah, and Doug Lenat. 搜索常识。从网络中填充CYC。 *UMBC 计算机科学和电子工程系收藏*. 2005.
- Cynthia Matuszek, Evan Herbst, Luke Zettlemoyer, and Dieter Fox. 学习解析机器人控制系统的自然语言命令。在 *实验性机器人学* 中, 第403-415页。斯普林格。 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. 永无止境的学习。在 *第二十九届 AAAI 人工智能会议 (AAAI-15) 的论文集中*. 2015.
- Tom M Mitchell. *学习概括中的偏见需求*. 计算机科学系, 计算机科学研究实验室。 1980.
- Douglas Morrison, Peter Corke, and Jürgen Leitner. 学习稳健、实时、反应式的机器人抓取。 *国际机器人研究杂志*, 39 (2-3) : 183-201. 2020.
- Jonathan Mugan 和 Benjamin Kuipers. 连续环境中高层次状态和行动的自主学习。 *IEEE Transactions on Autonomous Mental Development*, 4 (1) : 70-86. 2011.
- Vincent C Müller 和 Matej Hoffmann. 什么是形态学计算? 关于身体如何有助于认知和控制的问题。 *人工生命*, 23 (1) : 1-24. 2017.
- Ofir Nachum, Honglak Lee, Shixiang Gu, and Sergey Levine. 数据-efficient 科学的分层强化学习

。在*神经信息处理系统的进展中*，第3303-3313页。2018.

Sanmit Narvekar, Jivko Sinapov, and Peter Stone. 强化学习中自主任务排序的用户化课程设计。在*IJCAI*上，第2536-2542页。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: 用马尔科夫逻辑网络进行非结构化信息处理的集合学习。在 *IEEE 机器人和自动化国际会议 (ICRA)* 上, 中国香港, 5月31日-6月7月。 2014.

Alberto Olivares-Arcos, Daniel Beßler, Alaa Khamis, Paulo G. Alves, Maki Habib, J. Bermejo, Marcos Barreto, Mohammed Diab, Jan Rosell, João Quintas, Joanna Ol-szewska, Hirenkumar Nakawala, Edison Pignaton de Freitas, Amelie Gyrard, Stefano Borgo, Guillem Alenyà, Michael Beetz, and Howard Li. 基于本体的机器人自主性方法的回顾和比较。 *The Knowledge Engineering Review*, 34:12, 2019.

Peter Pastor, Mrinal Kalakrishnan, Sachin Chitta, Evangelos Theodorou, and Stefan Schaal. 操纵的技能学习和任务结果预测。在 *2011年IEEE国际机器人和自动化会议上*, 第3828-3834页。 IEEE, 2011.

Hanna M Pasula, Luke S Zettlemoyer, and Leslie Pack Kaelbling. 学习随机领域的符号模型。 *Journal of Artificial Intelligence Research*, 29:309-352, 2007.

Siddharth Patki, Andrea F Daniele, Matthew R Walter, and Thomas M Howard. 推断紧凑表征以实现机器人指令的精确自然语言理解。在 *2019年国际机器人和自动化会议 (ICRA)* 上, 第6926-6933页。 IEEE, 2019.

Rohan Paul, Jacob Arkin, Nicholas Roy, and Thomas M Howard. 为与机器人操纵器的自然语言交互提供抽象空间概念的科学依据。在 *机器人学。 Science and Systems Foundation*, 2016.

Ali Payani. *Differentiable Neural Logic Networks And Their Application Onto Inductive Logic Programming*. 博士论文, 乔治理工学院。 2020.

Anastasia Pentina and Christoph Lampert. 终身学习的PAC-Bayesian约束。在 *国际机器学习会议上*, 第991-999页。 2014.

Rolf Pfeifer 和 Josh Bongard. *身体如何塑造我们的思维方式: 关于智力的新观点*. 麻省理工学院出版社。 2006.

Lerrel Pinto 和 Abhinav Gupta. 超大的自我监督。从50k次尝试和700个机器人小时中学习

抓取。在2016年IEEE国际机器人和自动化会议 (ICRA) 上, 第3406-3413页。IEEE, 2016.

Horia Porav, Will Maddern, and Paul Newman.恶劣环境下的对抗性训练。使用外观转移的鲁棒性度量定位。在2018年IEEE机器人和自动化国际会议 (ICRA) 上, 第1011-1018页。IEEE, 2018.

- Meng Qu, Yoshua Bengio, and Jian Tang. GMNN : 图马尔科夫神经网络。在 *ICML会议记录*, 2019年。
- Vasumathi Raman, Alexandre Donzé, Dorsa Sadigh, Richard M Murray, and Sanjit A Seshia. 从信号时序逻辑规格的反应式合成。在 *第18届混合系统国际会议论文集中*。计算和控制, 第239-248页。 2015.
- Fabio Ramos, Rafael Carvalhaes Possas, and Dieter Fox. BayesSim : 通过机器人模拟器的概率推理实现自适应做主随机化。 *arXiv预印本arXiv:1906.01728*. 2019.
- Anurag Ranjan, Joel Janai, Andreas Geiger, and Michael J. Black. 攻击光流。在 *国际计算机视觉会议 (ICCV)* 上。 2019.
- Dushyant Rao, Francesco Visin, Andrei Rusu, Razvan Pascanu, Yee Whye Teh, and Raia Hadsell. 持续的无监督表示学习。In *Advances in Neural Information Processing Systems* 32, 2019.
- Lynne M Reder和Frank E Ritter. 是什么决定了最初的认知感觉? 对问题术语的熟悉程度, 而不是对答案的熟悉程度。 *实验心理学杂志. 学习、记忆和认知*, 18 (3) : 435. 1992.
- Charles Richter和Nicholas Roy. 通过深度学习和新颖性检测的安全视觉导航。在 *机器人学。科学与系统基金会*。 2017.
- Ludovic Righetti, Mrinal Kalakrishnan, Peter Pastor, Jonathan Binney, Jonathan Kelly, Randolph C. Voorhies, Gaurav S. Sukhatme, and Stefan Schaal. 一个基于力控制和优化的自主机械手系统。 *自主机器人*, 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. 机器人工程师。在 *ILP(Late Breaking Papers)* 中, 第101-106页。 2015.
- Adam N Sanborn, Vikash K Mansinghka, and Thomas L Griffiths. 调和碰撞物体的直觉物理学和牛顿力学。 *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: Embodied AI研究的平台。在 *IEEE/CVF国际计算机视觉会议 (ICCV)* 的论文集中。 2019.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, et al. 掌握Atari, Go, chess and shogi是通过学习模型的计划。《自然》, 588(7839): 604-609, 2020.

- Amanda Seed和Richard Byrne.动物的工具使用。《当前生物学》, 20 (23) : R1032-R1039。2010。
- Burr Settles.主动学习文献调查.技术报告, 威斯康星大学麦迪逊分校计算机科学系。2009。
- 拉里-夏皮罗体现认知的研究计划。《哲学指南针》, 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, 大型现实场景中互动任务的模拟环境。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 and G van den Driessche, Thore Graepel, and Demis Hassabis. 在没有人类知识的情况下掌握围棋的游戏。《自然》, 550(7676):354359,- 2017。
- H.A. Simon.理性选择和环境的结构。《Psychological Review》, 63(2):129-138, 1956。
- 赫伯特-A-西蒙。《The sciences of the artificial》.麻省理工学院出版社, 版本3。1996。
- Martin Simonovsky and Nikos Komodakis.GraphVAE : 使用变异自动编码器生成小图的方向。在《国际人工神经网络工程会议上》, 第412-422页。Springer, 2018。
- Aman Sinha, Matthew O'Kelly, Russ Tedrake, and John C. Duchi.用于评估安全关键型自主系统的神经桥采样。在《NeurIPS》中。2020。
- 伊丽莎白-S-斯佩尔克和凯瑟琳-D-金兹勒。核心知识。《发展科学》, 10 (1) : 89-96。2007。
- 玛丽莲-斯特拉瑟恩。改善评级”。英国大学系统中的审计。《欧洲评论》, 5 (3) : 305-321。1997。
- 孙培, Henrik Kretschmar, 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.自动驾驶感知的可扩展性。Waymo开放数据集。2019。
- Richard S Sutton, Doina Precup, and Satinder Singh.在MDPs和半MDPs之间 : 强化学习中的时间抽象的框架。《Artificial intelligence》, 112 (1-2):181-211, 1999。

Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus.神经网络的诱人特性。在国际学习代表会议上。 2014.



Tadahiro Taniguchi, Emre Ugur, Matej Hoffmann, Lorenzo Jamone, Takayuki Nagai, Ben-jamin Rosman, Toshihiko Matsuka, Naoto Iwahashi, Erhan Oztog, Justus Piater, and Florentin Wötter. 认知发展系统中的符号涌现：一项调查。 *IEEE 认知和发展系统交易*, 11 (4) : 494-516. 2018.

Stefanie Tellex, Thomas Kollar, Steven Dickerson, Matthew R Walter, Ashis Gopal Banerjee, Seth Teller, and Nicholas Roy. 用概率图形模型处理符号接地问题。 *AI 杂志*, 32 (4) : 64-76. 2011.

Jesse Thomason, Shiqi Zhang, Raymond J Mooney, and Peter Stone. 通过人与机器人的对话学习解释自然语言命令。在 *第二十四届国际人工智能联合会议* 上。 Citeseer, 2015.

Sebastian Thrun 和 Tom M Mitchell. 终身的机器人学习。 *机器人学和自主系统*, 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. *arXiv preprint arXiv:1903.07438*, 2019.

Marc Toussaint. 逻辑-几何编程。基于优化的方法，将任务和运动规划结合起来。在 *IJCAI 上*, 第1930-1936页。 2015.

Marc Toussaint, Kelsey R Allen, Kevin A Smith, and Josh B Tenenbaum. 用于工具使用和操纵规划的可分辨物理学和稳定模式。In *Proc. of Robotics: 科学与系统 (RSS 2018)*。 2018.

James Tu, Mengye Ren, Siva Manivasagam, Ming Liang, Bin Yang, Richard Du, Frank Cheng, and Raquel Urtasun. 用于激光雷达物体检测的物理上可实现的对抗性例子。 2020.

Vasileios Vasilopoulos, William Vega-Brown, Omur Arslan, Nicholas Roy, and Daniel Koditschek. 部分已知环境中基于传感器的反应性符号规划。在 *国际机器人和自动化会议 (ICRA)* 上, 澳大利亚布里斯班, IEEE2018.。

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N 戈麦斯, Łukasz Kaiser, 和 Illia Polosukhin. 注意力是你所需要的一切。在 *进展中 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. 基于视觉模型的强化学习中的实体抽象。在 *机器人学习会议* 上, 第1439-1456页。 PMLR, 2020.

Julius von Kgelgen, Ivan Ustyuzhaninov, Peter Gehler, Matthias Bethge, and Bernhard

Elkpf. 通过专家的竞争建立因果生成的场景模型。 *ICLR 决策中的因果学习研讨会*。  
2020.

L.S. Vygotsky. *社会中的心智：高级心理过程的发展*。哈佛大学出版社，剑桥，马。 1978.

Po-Wei Wang, Priya Donti, Bryan Wilder, and Zico Kolter. SATnet: 使用一个可满足性求解器来连接深度学习和逻辑推理。在*国际机器学习会议上*, 第6545-6554页。PMLR, 2019.

William Yang Wang, Kathryn Mazaitis, and William W Cohen. 基于结构化稀疏性的软版谓词发明。在*第二十四届国际人工智能联合会议上*。 2015.

Manuel Watter, Jost Springenberg, Joshka Boedecker, and Martin Riedmiller. 嵌入控制。一个用于控制原始图像的局部线性潜在动力学模型。在*神经信息处理系统的进展中*, 第2746-2754页。 2015.

罗伯-韦斯顿, 奥伊维-帕克-琼斯, 以及英格玛-波斯纳。在那里又回来了。学习模拟现实世界应用的雷达数据。在*IEEE国际机器人和自动化会议 (ICRA)* 上。 2021.

希蒙-怀特森。TreeQN和ATreeC: 用于深度强化学习的可选择的树计划。在*国际学习表征会议的论文集中*。 2018.

David H Wolpert. 监督学习的无自由午餐定理。在*软计算和工业*, 第25-42页。Springer, 2002.

David H Wolpert and William G Macready. 没有免费午餐的优化定理。*IEEE transactions on evolutionary computation*, 1(1):67-82, 1997.

吴家俊, Ilker Yildirim, Joseph J Lim, Bill Freeman, 和Josh Tenenbaum. 伽利略。通过整合物理引擎和深度学习来感知物理对象的属性。在*神经信息处理系统的进展中*, 第127-135页。 2015.

吴佳俊, Erika Lu, Pushmeet Kohli, Bill Freeman, 和Josh Tenenbaum. 通过视觉去动画化学习看物理学。在*神经信息处理系统的进展中*, 第153-164页。 2017.

杨振培, 柴玉宁, Dragomir Anguelov, 周寅, 孙沛, Dumitru Erhan, Sean Rafferty, 和Henrik Kretzschmar. Surfegan: 为自动驾驶合成现实的传感器数据。在*IEEE/CVF计算机视觉和模式识别会议上*, 第1118-1127页。 2020.

于天河, Deirdre Quillen, 何展鹏, Ryan Julian, Karol Hausman, Chelsea Finn, 和Sergey Levine. Meta-world: 多任务和元强化学习的基准和评估。在*机器人学习会议上*, 第1094-1100页。PMLR, 2020.

Andy Zeng, Shuran Song, Johnny Lee, Alberto Rodriguez, and Thomas Funkhouser. Toss-

ingbot:学习用残余物理学来投掷任意的物体。 *IEEE Transactions on Robotics*, 36 (4) : 1307-1319。 2020.

Wenshuai Zhao, Jorge Ba Queralta, and Tomi Westerlund. 机器人深度强化学习中的模拟到现实的转移：一项调查。在2020年IEEE商业智能系列研讨会 (SSCI) 上, 第737-744页。 2020.

Gabriel Ziller, Vincent Wall, and Oliver Brock. 软性气动执行器的主动声学接触感应。在 2020 年 IEEE 国际机器人和自动化会议 (ICRA) 上。IEEE。2020.