

# 基底神经节的机器人模型：行为选择与内部处理

A ROBOT MODEL OF THE BASAL GANGLIA: BEHAVIOR AND INTRINSIC PROCESSING

Tony J. Prescott<sup>1</sup>, Fernando M. Montes Gonzalez<sup>2</sup>, Kevin Gurney<sup>1</sup>, Mark D. Humphries<sup>1</sup>, and Peter Redgrave<sup>1</sup>

<sup>1</sup>Adaptive Behavior Research Group, University of Sheffield, Sheffield, South Yorkshire, UK.

<sup>2</sup>Department de Física e Inteligencia Artificial, Universidad Veracruzana, Xalapa, Veracruz, Mexico.

(Song Jian, translate)

**摘要：**由于感觉运动系统与基底神经节之间存在多个平行回路，因此有人提出，这些连接机制可以作为一种选择机制，解决在特定环境下可用且可选择的动作之间的竞争问题。对这一假设的一个强有力的检验是，需要一个基底神经节的计算模型来在一个自治体中生成综合选择序列，因此，我们描述了一个嵌入这种模型的机器人架构，并要求它在受动物观察启发的机器人任务中控制动作选择。我们的结果证明，嵌入式模型在多感官和动机促进条件下，会有有效的行动选择。当面对多个高度显著的替代方案时，机器人还表现出行为选择崩溃的情况，这与动物在冲突情况下的行为相似。该模型还揭示了最近关于行为转换和测序的神经生物学发现。

**关键词：**基底神经节；行为选择；行为切换；嵌入计算神经科学；机器人；老鼠

## 一、引言

基底神经节是一组高度互联的中枢大脑结构，对运动和认知有重要影响。对这些结构的兴趣部分源于它们对包括帕金森氏病、亨廷顿氏病、图雷特综合征、精神分裂症和注意力缺陷多动障碍在内的一系列大脑疾病的重要影响，而且对这些结构的研究也推动了一个多世纪的神经生物学研究。这种广泛的研究工作已经产生了大量的相关数据信息，因此迫切需要更好地理解这些结构的功能。因此，基底神经节是脑功能计算建模最令人期待的前景之一，也是当下用于广泛建模研究的重点（Houk、Davis 和 Beiser, 1995；Wickens, 1997；Gillies 和 Arbuthnott, 2000；Prescott、Gurney 和 Redgrave, 2002；Gurney、Prescott、Wickens 和 Redgrave, 2004）。

基底神经节相关文献中反复出现的一个主题是，这些结构的作用是解除所需动作的抑制，同时保持或增加对不希望动作的抑制（Denny Brown 和 Yanagisawa, 1976；Cools, 1980；Robbins 和 Brown, 1990；Hikosaka, 1994；Mink, 1996；Wickens, 1997）。在我们自己的理论工作（Prescott, Redgrave 和 Gurney, 1999；Redgrave, Prescott 和 Gurney, 1999a）中，我们提出了这样一种观点：基底神经节作为一种动作选择机制，解决了大脑内物理分离但在行为表达上存在竞争的功能单元之间的冲突。我们已经展示了这个观点如何与已知的解剖学和生理学相关理论建立联系，并且满足了有效动作选择装置的几个高级计算要求。根据这一假设，我们还设计了一个程序，在几个抽象层次上模拟基底神经节和相关结构的电路。重点是研究受已知功能解剖学约束的基底神经节系统级模型，在功能解剖学中神经群由简单的漏积分子单元表示（Gurney、Prescott 和 Redgrave, 2001a、2001b；Humphries 和 Gurney, 2002；Gurney、Humphries、Wood、Prescott 和 Redgrave, 2004）。在较低水平的神经生物学细节中，我们研究了动作电位序列（峰值）编码的信号模式（Humphries 和 Gurney, 2001；Humphries, 2002），并研究了纹状体神经元膜动力学的生物物理模型（Wood、Gurney 和 Wilson, 2004）。所有这些层次的研究都产生了互补的结果，证实了基底神经节的生物学结构可以作为一种有效的选择机制发挥作用。然而，在我们看来，这种计算神经科学的方法，即在更广泛的背景下对特定的大脑系统进行建模，仍然留下许多重要的问题没有解答。首先，

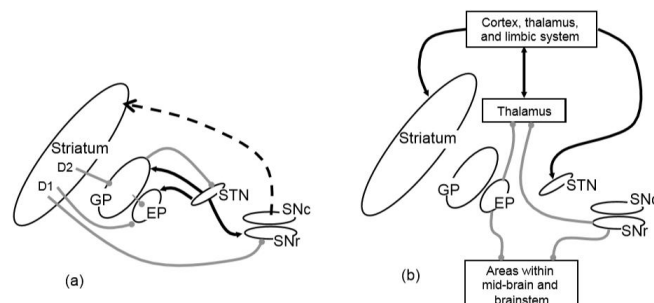
我们想知道如何最好地解释模型的输入和输出，我们可以选择将输入视为，例如，感觉信号，或输出视为运动信号，但这种分配基本上是毫无根据的。第二，如果不在更广泛的背景中建立模型，我们就无法判断它是否能够在更为全面指定的控制架构中实现其假设的功能角色。第三，如果基底神经节与感觉和运动系统之间没有任何联系，我们可能会质疑一个模型是否能够处理噪音或模糊的感觉数据，或者作为一个系统的一部分，对协调实际效应器系统的运动提出质疑。最后，如果没有多个需求的背景，例如需要保持物理完整性、避免威胁、探索和利用资源等，我们将无法确定模型是否能够满足适应性行为的一些基本要求。因此，在本文中，我们描述了一种将基底神经节和相关的丘脑皮质连接的系统级模型嵌入到一个小型移动机器人的控制结构中，该小型移动机器人执行模拟觅食任务，要求机器人在不断变化的感觉和动机条件下选择适当的动作，从而生成集成行为序列。我们把现在正在应用的描述这一套方法，作为具体的计算神经科学（embodied computational neuroscience）。机器人模型的初步结果已在（Montes Gonzalez, Prescott, Gurney, Humphries 和 Redgrave, 2000）中进行了描述，并且该模型的一个版本已显示出比标准“赢者通吃（winner-takes-all）”在机器人生存任务中的选择机制具有更好的选择属性（Girard, Cuzin, Gullot, Gurney 和 Prescott, 2003）。然而，本文首先全面介绍了机器人嵌入基底神经节模型，并与相关的神经行为研究进行了比较，对模型行为进行了广泛的评估。我们还展示了机器人模型在面对多个高显著性选择时的行为，并与冲突情况下的动物行为研究进行了比较。

文章的其余部分按如下方式展开论述：第 2 节回顾了基底节的动作选择假设和相关的建模工作。机器人基础神经节模型的**动机**、机器人实施的全部细节以及动作选择指标的总结，这些都是在第 3 节（以及随附的附录）中进行了描述。接着第 4 节描述了三个实验的结果：

- 实验一：使用扩展基底节模型的无实体版本系统搜索特征空间（Humphries 和 Gurney (2002) 对该模型的早期分析进行了扩展）；
  - 实验二：我们主要研究机器人基底节模型的动作选择；
  - 实验三：高特征性选择背景下的机器人行为研究。
- 第 5 节提供了我们对实验结果的讨论，重点是与生物数据的比较。

## 二、背景介绍：基底神经节被视为一种动作选择装置

已经有很多研究者对基底神经节的功能解剖进行了许多精辟的总结（例如 Gerfen 和 Wilson, 1996; Mink, 1996; Smith、Bevan、Shink 和 Bolam, 1998），因此以下重点介绍，与理解下面将要讨论的模型最相关的方面。



抑制 (Inhibition) → 兴奋 (Excitation) → 多巴胺 (Dopamine)

图 1：大鼠基底神经解剖：(a) 内部通路；(b) 外部通路。没有显示所有连接。缩写：STN-下丘脑核、EP-脚内核、GP-苍白球、SN-黑质致密部、SNr-黑质网状部、D1、D2-纹状体神经元优先表达多巴胺受体亚型 D1 和 D2。

啮齿动物基底神经节（图 1a）的主要结构是纹状体（由尾状体（caudate）、壳核（putamen）

和腹侧纹状体（ventral striatum）组成）、下丘脑核（STN, subthalamic nucleus）、苍白球（GP, globus pallidus）、黑质（SN, substantia nigra（由网状部 SNr, pars reticulata 和紧密部 SNc, pars compacta 组成））和脚内核（EP, entopeduncular nucleus）（与灵长类中的苍白球内段或 GPi 类似）。这些结构相互连接，在整个大脑结构中形成了功能子系统（图 1b）。

基底神经节的输入核为纹状体和 STN。传入连接到这两种结构的信号实际上来源于整个大脑，包括大脑皮层、脑干的许多部分（通过丘脑）和边缘系统。这些连接提供相位兴奋输入。主要输出核为黑质网状部分（SNr）和脚内核（EP）。这些结构提供广泛的分支输出到丘脑（丘脑再反过来投射回大脑皮层），以及中脑和脑干的运动前区。大多数输出投射是兴奋和抑制的。

为了了解基底节的内在连通性，其中最为重要的是通过传出连通性和神经化学，认识到纹状体（中等多棘细胞，medium spiny cells）的主要投射神经元形成两个广泛分布的神经群之间的区别。其中一个神经群含有神经肽物质 P 和强啡肽，优先表达多巴胺受体的 D1 亚型，主要投射到输出核（SNr 和 EP）。在基底神经节的普遍非正式模型中（Albin, Young 和 Penney, 1989），这种“D1 条纹”投射构成了到输出核的所谓的直接途径。这些神经元的传出活动抑制输出结构中的高频抑制放电，输出结构转而解除抑制丘脑和脑干中的对象。第二组纹状体投射神经群含有脑啡肽，优先表达 D2 亚型多巴胺受体。该组神经群主要针对苍白球（GP），其高频抑制输出既指向输出核（SNr 和 EP），也指向 STN。这些“D2 纹状体”神经元的抑制投射构成了通向输出核的间接通路的第一段。由于此途径有两个抑制环节（纹状体 GP, GP-STN），其次是兴奋环节（STN-EP/SNr），其主要有效作用是激活输出核，从而增加对丘脑和脑干的抑制控制。

纹状体多巴胺神经支配的主要来源是黑质致密部（SNc）。基底神经节的多巴胺调节通常被认为是在两个时间尺度上起作用（Grace, 1991; Walters, Ruskin, Allers 和 Bergstrom, 2000）。其中一种尺度是短延迟相位响应（100ms 突发），这与生物显著刺激的发生有关，而且似乎对某些形式的激励学习至关重要（Schultz、Dayan 和 Montague, 1997; Redgrave、Prescott 和 Gurney, 1999b），另一种尺度是受各种脑病理改变的活动强度水平（1-8Hz）。如帕金森病，在正常大脑中可能受到额叶皮质等结构的调节。有趣的是，D1 和 D2 纹状体群体对多巴胺能传递的变化反应不同。虽然已经有研究者报道了一系列影响，但一个解释了大量现有发现的简化假设是，多巴胺通过 D1 受体（Akkal、Burbaud、Audin 和 Bioulac, 1996 年）作用时增强了其他突触输入的有效性，同时降低了在 D2 受体作用时的有效性（Gerfen 等人, 1996 年），1990; Harsing 和 Zigmond, 1997 年）。这种安排似乎为直接（抑制）和间接（有效兴奋）基底神经节通路所支持的“推/拉”机制提供多巴胺能控制。我们的另一篇文章的主题研究了这种兴奋性多巴胺水平的变化对我们的机器人模型的影响；在当前的工作中，我们报告了模拟多巴胺水平固定在中间水平的结果。同样，本篇文章也没有讨论学习的问题（以及多巴胺在其中的作用），而是逻辑上不同的问题，即基底神经节-基底神经节被适当配置以支持在嵌套体中的作用选择。

我们的基础神经节模型的一个关键假设是，大脑同时处理大量的感觉（sensory）和认知流（cognitive streams）或通道（channels），每个通道都可能存在一个需要采取行动的请求信息。为了做出有效的行为，必须抑制其中的大多数请求，以便只允许有限数量的表达式（可能只有一个）表达出来。这种基于通道的方案与基底神经节通过一系列局部组织的并行处理流输入明显一致（Alexander 和 Crutcher, 1990）。基底神经节的行为选择假说进一步表明，在纹状体和 STN 中细胞群的活动编码候选行为的显著性（salience）或选择倾向（propensity for selection）。同时，基底神经节输出结构，SNr 和 EP，被认为是通过降低抑

制输出，从而门控胜出通道的候选动作。当孤立于更广泛的大脑结构时，就“信号选择”的无太大相关性的问题而言，最好强调这一行动选择理论；换言之，建议在纹状体处选择大信号输入，在 EP/SNr 处选择 STN 作为低信号输出。

从信号选择的角度来看，基底神经节和相关电路内的多种机制似乎都经过适当配置，以解决竞争渠道之间的冲突，并在获胜者之间提供所需的平稳和快速的切换。我们的基底神经节初始系统水平模型（Gurney 等人 2001a, 2001b）重点介绍了以下候选选择机制。

首先，在细胞水平上，较为关注的研究目标集中在纹状体投射神经元的内在特性上，这样，在任何给定时刻，大多数细胞都处于非活动的兴奋性“下降状态”，并且只能通过大量的同时输入（Wilson 和 Kawaguchi, 1996 年）触发到活动的“上升状态”（在那里它们可以激发动作电位）。这种双稳态行为可以充当高通滤波器（high-pass filter）来排除低频率的“请求”。

其次，计算理论表明，前馈（feedforward）、偏心（off-centre）、环绕网络（on-surround network）是一种适当的增强信号选择的机制。在基底神经节中，这种类型的选择电路似乎是通过抑制对输出核（偏心）的中心纹状体进行抑制和对同一输出核（环绕网络）的分散的 STN 进行激发（Parent 和 Hazrati, 1995）的组合模式来实现选择机制。然而，仔细观察，似乎在基底节固有电路中有两个这样的前馈网络（见图 2a, b），这两个网络是由纹状体神经元的 D1 型和 D2 型亚群的投射目标区分的。一个实例（图 2a）利用 EP/SNr 作为其输出层；由于这与我们对基底神经节的信号选择假设明显一致，因此我们将此电路指定为选择路径（selection pathway）。然而，还有第二个前馈体系结构的实现，其目标是 GP（图 2b）。由于 GP 的传出连接仅限于其他基底神经节，因此尚不清楚第二个实现电路在什么意义上有助于总体选择任务。这个问题可以通过假设第二个子系统形成一个控制路径来解决，该控制路径（control pathway）的功能是调节主选择机制的特性。当两个子系统结合在一起时，给出如图 2c 所示的总体功能结构，从 GP 发出的控制信号是明显的。

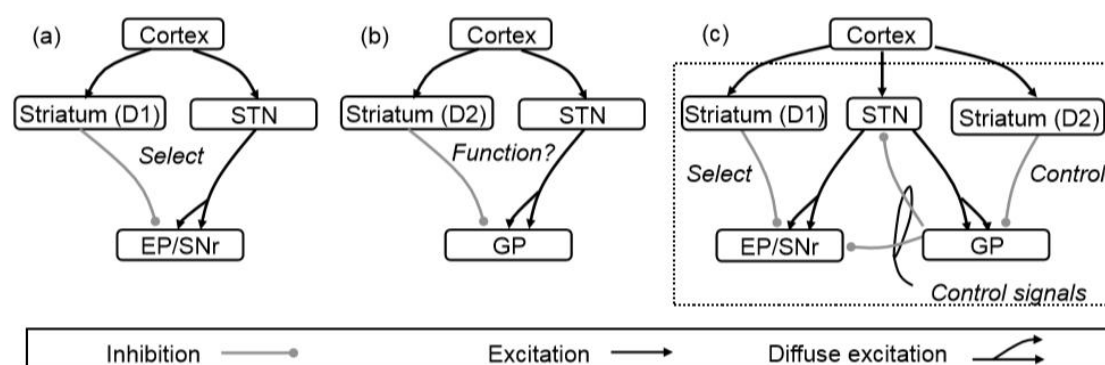


图 2：基底神经节被视为一种行为选择机制。缩写如图 1 所示。我们对基底神经节内在连通性的分析（Gurney 等人，2001ab）表明存在两个偏心、环绕、前馈网络。一个实例（a）利用 EP/SNr 作为输出层，指定选择路径，第二个实例（b）针对 GP，指定控制路径。当两个子系统结合在一起时，从 GP 发出的控制信号是明显的，以给出如图 2c 所示的总体功能结构。

在我们最初的系统级模型中，我们将上述电路（图 2c）作为多通道系统进行操作，其中，对于每个基底神经节核，编码每个通道的神经群由一个适当配置的漏积分（leaky integrator）单元进行模拟。利用该模型进行的分析和模拟研究（Gurney 等人，2001a, 2001b）表明，该模型能够支持多个竞争对手之间的有效切换。在仿真中，模型的两个或多个通道以不同幅度的手动信号的形式提供输入。结果表明，最大的信号输入总是产生最小的信号输出（从而显示信号选择），系统迅速从当前选择的通道切换到突然具有较大输入的竞争通道。我们还能够在我们的基底神经节模型的组成电路中生成信号特征，该模型的时间模式与 GP（Ryan 和 Clark, 1991）和 SNr（Schultz, 1986）中神经放电的各单位记录相似。

Humphries 和 Gurney 扩展了固有基底神经节处理的原始模型，包括基底神经节丘脑皮质环路（Humphries 和 Gurney，2002 年）。这项研究提出了丘脑与腹侧（VL）丘脑和丘脑网状核（TRN）的复合作用，以提供额外的选择相关功能。具体来说，如图 3 所示，这些电路可以理解为两个重要角色。首先，通过 EP/SNr 对 VL 丘脑靶点的抑制，从而使获得基底神经节通道的正反馈回路（positive feedback loop）能够增加其自身皮质输入的激活。其次，TRN 和 VL 丘脑之间的通道内和通道间连接似乎实现了末端横向抑制网络，以牺牲相邻通道为代价，增加最强烈神经支配通道的活动。在模拟中，再次利用手动信号，发现这些额外的基底神经节机制的额外选择性功能可促进几个理想的选择特征，包括在高度匹配的通道之间进行更平稳的切换，以及忽略瞬时显著性中断的能力。最近，我们还发现，该模型可以容纳纹状体苍白样投射，以及苍白球和黑质内局部抑制连接的新数据（Humphries、Prescott 和 Gurney，2003；Gurney、Humphries 等人，2004）。这两个扩展似乎都增强了系统的选择性，并且在增加更多的生物学现实性的同时，进一步支持了基底神经节功能的选择假设。

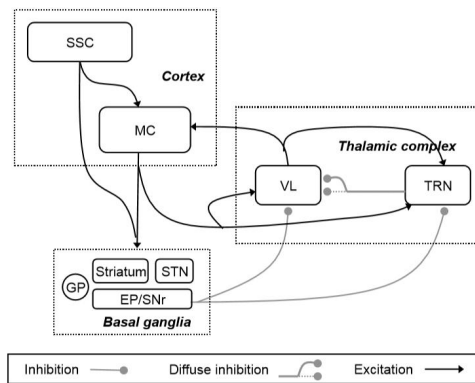


图 3: Humphries 和 Gurney 的扩展基底节模型 (2002 年)。缩写: SSC-体感皮层、MC-运动皮层、VL-腹侧丘脑、TRN-丘脑网状核, 其他如图 1 所示。模型的基底神经节组成部分内的连通性如图 2c 所示。基底神经节丘脑皮质环路可以理解为提供了有助于有效作用选择的额外机制。首先, 从 VL 中去除基底节抑制, 完成一个到运动皮层的正反馈回路。然后, 从 TRN 到 VL 的扩散抑制连接, 在通道之间比在通道内更强 (如图中的平面和点状抑制连接所示), 以及从 VL 到 TRN 的通道内激励, 在通道之间产生一种相互抑制的形式。更多解释请参见文本和 Humphries 和 Gurney (2002)。

一个有效的行为选择机制应该对显著性权重的变化较为敏感, 这种变化会改变特定环境下竞争行为的相对紧迫性。然而, 不太明显的是, 选择机制应该如何应对显著性权重的变化, 而这些变化会使相对显著性保持不变, 同时也会提升选择竞争的整体水平。以广义使用的“赢者通吃 (winnertakes-all)”选择机制为例, 整体的显著性水平是不相关的 (具有最高显著性的竞争对手总是首选)。我们之前已经证明了内在 (Gurney 等人, 2001b) 和扩展 (Humphries 和 Gurney, 2002) 基底神经节模型的选择特性不符合这一假设, 而是根据选择竞争的总体强度而变化。我们通过展示获胜子系统的滞后程度或持久性, 可能会由于整体显著性水平的变化而发生变化, 将这项工作拓展到下面的内容。我们之前的研究指出, 当模型具有多个高显著性的备选方案时, 会产生多个渠道选择的有趣模式。因此, 我们研究机器人模型在这些情况下的行为, 并考虑可能与行为冲突的行为学研究得出的观察结果平行。

### 三、基底节动作选择机器人模型的建立

上述建模工作旨在说明基底神经节的信号选择, 而不是动作选择本身。为了令人信服地证明基底神经节模型能够作为一种有效的动作选择装置运行, 我们认为它需要嵌入到与物理世界的实时感觉运动交互中。因此, 一个重要的目标是构建一个嵌入式基底神经节模型, 在该模型中, 在移动机器人的多个物理实现行为之间进行选择。由于机器人技术在计算神经科

学中的应用相对较新，因此，我们在对该模型的描述之前，简要说明了如何将计算神经科学模型嵌入生成可观察行为的机器人架构中。

### 3.1 一种具体的计算神经科学方法

任何计算神经科学模型，无论是机器人还是其他，都是由仿生的组件组成的，也就是说，它们的目的是直接模拟神经生物学过程（在适当的水平上），而那些组件仅仅“被设计”，以便提供一个接口，使模型能够被询问和评估。在机器人模型的情况下，对工程部件的需求尤为明显，在这种情况下，神经电路的模拟必须在某一点上与（通常）完全非神经机器人硬件连接。此外，在尝试模拟完整行为能力的模型中，由于任务的规模，或由于缺乏必要的神经生物学数据，在给定的细节水平上模拟目标能力的神经基质的所有成分通常是不现实的。在这种情况下，工程部件也需要替代一些神经回路的功能，这些神经回路是已知的或未知的，它们与动物的某些能力的产生有关。在目前的模型中，由于感兴趣的生物底物是基底神经节，因此提供机器人硬件（和低级控制器）与基底神经节和相关核的模型被构建为一组工程化的子系统，我们统称为“嵌入体系结构（embedding architecture）”。虽然从生物学角度来看，我们会强调，这种嵌入体系结构并不打算直接模拟任何特定的神经基质。

Latimer（1995）已经确定存在两种不同类型的模型成分，作为认知科学研究的普遍特征。在他的阐述中，被设计成仿生的组件被认为是“理论相关”，而那些仅仅使模型可用的组件则是“理论无关”。事实上，Latimer这样做只是为了立即解构它！毕竟，“理论相关性”在很大程度上是非专业的人认为的，而对于一个研究者来说，是一个“理论无关”的假设，使模拟运行，而对一位评论家而言，模型的结果至关重要是十分不合理的。例如，这些问题为心理过程中的连接主义模型的反对者提供了一条重要的攻击线索（见 Massaro, 1988; Pinker 和 Prince, 1988）。此外，模型的某个特定部分是否被认为是“理论相关”，可以在很大程度上取决于正在测试的假设以及模型本身的性质。因此，虽然我们在这里讨论如何实现（机器人）运动行为的具体细节与基底神经节功能的动作选择假设不相关，但如果将相同的模型作为一个完整的运动回路的理论进行评估，这些细节将变得相关。

由于在认知科学中不可避免地要使用工程组件，但它们的“理论无关性”不能被认为是理所当然的，因此重要的是：

- (i) 模型的这些组件被足够详细地描述，以允许它们的评估和复制；
- (ii) 与仿生成分的输入和输出的接口符合重要的生物限制条件；
- (iii) 当工程部件最终具有显著的行为或功能结果（即潜在的理论相关）时，将对这些进行探索，并在可能的情况下与相关神经系统的功能特性相关。

在本文中，我们尝试了三个任务。首先，在本节的其余部分以及附录中，我们对嵌入体系结构进行了全面的描述，并证明在功能层面上，将该体系结构分解为行为相关（行动）子系统与行为学证据是一致的。第二，在第2节（上文）中，我们回顾了证明我们的一般假设的证据，即基底神经节输入的输出编码动作显著性，其输出作为门控信号抑制不想要的运动行为（我们也在第5节重新讨论这些问题）。第三，在第5.2节中，我们将考虑嵌入体系结构的特定组件是否存在可能的神经关联，这些关联具有显著的功能结果。

### 3.2 模型任务

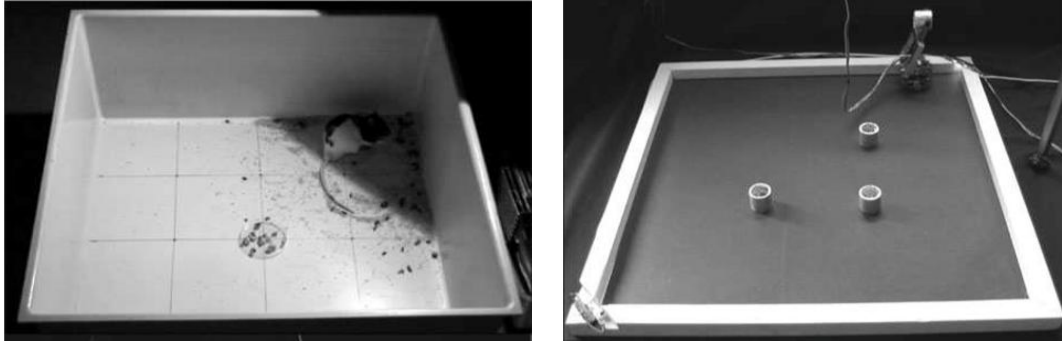


图 4: 成年大鼠在阴凉巢区 (右上) 和中央食物资源区 (中间) 的方形实验平台中的行为为基底神经节 (右图) 动作选择机器人模型中的任务设置提供了启示。

为了评估包含的基底神经节模型的动作选择特性, 需要一个足够复杂的环境/任务设置 (task-setting), 从而在可选择的行为系统之间呈现有趣的竞争, 但也要足够简单, 以便建立基线性能水平, 并允许对行为转换的结果模式进行详细分析。在制定一项满足这些需求的任务时, 我们的灵感来自于观察被放置在一个不熟悉的矩形实验平台中的缺乏食物的老鼠的行为转换, 这个实验平台包含一盘半圆形的食物颗粒 (见图 4, 左)。这种动物的最初行为通常是探索性的、防御性的, 其特点是避免开放空间。被放置在实验平台中心的动物会迅速移动到外围, 然后倾向于靠近实验平台墙壁, 表现出对实验平台角落的偏好, 对食物消费几乎没有或没有明显的兴趣。随着动物越来越习惯新的环境, 与饥饿相关的行为变得更加明显。一种常见的觅食行为是定位食物放置位置, 再收集食物颗粒, 然后将其带回实验平台的一个巢穴角落 (通过已放置的用于休息的垫层材料来识别) 供食用。运动、进食和休息之间的平衡关系对动物的饥饿程度及其对竞技场的熟悉程度很敏感。

我们致力于创造一个环境来测试嵌入的基底神经节模型, 重点是为一个小型移动机器人创造一个类似的相对简单的任务设置。机器人被放置在一个有围墙的方形实验平台 (55×55cm) 中, 实验平台中有许多小的圆柱形物体 (见图 4, 右)。圆筒代替了食物颗粒, 因此食物的收集和消费是通过收集圆筒并将它们存放在实验平台的一个巢穴角落来模拟的。巢穴是通过局部光源 (8V 白炽灯泡) 的存在来识别的, 其中有两个光源, 放置在对角相对的角落。模拟动机被用来调节机器人的倾向, 以避免空旷的空间 (模拟“恐惧”), 并通过每次实验的时间过程收集圆柱体 (模拟“饥饿”)。

### 3.3 机器人控制体系结构

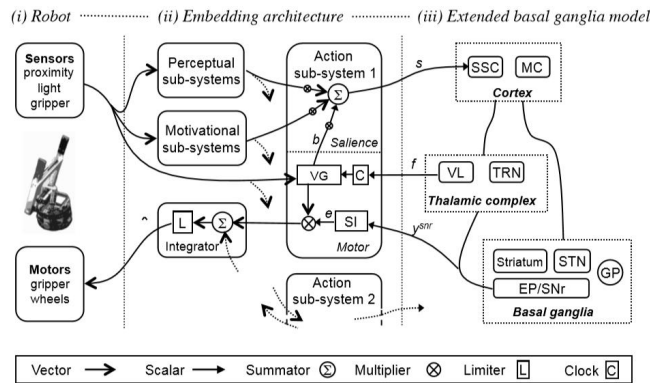


图 5: 基底神经节嵌合体模型。该模型由三部分组成: (i) 机器人及其感官和电机动力装置; (ii) 嵌入体系结构——一系列感知、动机、动作 (行为) 子系统及其接口; (iii) 仿生扩展基底节模型 (其完全连通性如图 3 和 2c 所示)。显示了五个作用子系统中第一个的连接 (与其他作用子系统之间的投射用虚线表示)。缩写: VG- (电机) 矢量发生器, SI-分流抑制,  $b$ -忙信号,

完整的机器人控制体系结构如图 5 所示,下面的一节连同附录一起对该体系结构的各个组件进行了完整的描述。图 5 区分了:

- (i) 机器人及其可用的原始感觉和运动系统;
- (ii) 提供一系列动作(行为)子系统的嵌入体系结构,计算其相对显著性,并将其输出与基底神经节的选通相结合;
- (iii) 扩展的基底神经节模型为解决行动选择冲突提供了基础。

如前所述,只有模型的第三个要素有志于模拟脊椎动物大脑功能的特定方面。然而,为满足工作控制系统的要求而包含的架构的其他组成部分提供了有用的假设,关于基底神经节在更广泛的大脑架构中的嵌入,我们在讨论中再来说明这个问题。

机器人控制体系结构有大量的自由参数,例如,这些参数规定了行为模式(第 3.3.2 节)子元素的时间和动作显著性计算的权重(第 3.3.3 节)。我们一直选择使用手动调整的参数,因为嵌入体系结构中的自由参数不需要是最优的,而只需简单地足以生成所需的行为。这一实践与许多现有的关于经常使用手动编码系统的行为选择的研究是一致的(Maes, 1995)。从机器人控制的角度来看,从经验中学习的能力显然会使我们的体系结构更具适应性,然而,当前研究的目标仅限于研究一个具体的基底神经节模型的选择能力,因此对嵌入体系结构的优化不是一个关键的要求。

### 3.3.1 机器人的感觉和运动系统

Khepera-I 是一个直径 60 毫米的小型圆柱形机器人,有两个驱动轮和一个可拆卸的夹持臂。机器人通过八个外围传感器阵列感知环境,这些传感器可以利用主动模式,利用红外距离传感器进行操作,和被动模式,利用环境光传感器进行操作。在主动模式下,传感器阵列的范围非常有限,能够可靠地检测距离不超过 25 毫米的附近垂直表面。在这方面,它有点类似于生物触觉系统,如大鼠的触须。此外,机器人在夹爪臂上有一个位置传感器,用于确定其当前是升高还是降低,以及一个二值光学传感器,用于检测夹爪爪之间是否存在物体。机器人感官系统的更多细节见附录 1。两轮电机可以独立向前或向后驱动,夹持器转塔由两个电机驱动,一个用于提升/降低夹持臂,另一个用于打开/关闭夹持臂。电机系统的更多细节见附录 4。一个由 Webots2.0 机器人接口软件控制的串行链路,用于向承载嵌入体系结构和基底神经节模型的计算机发送传感器读数,并接收来自该计算机的电机命令。该接口在一系列离散时间步骤上运行,以大约 7Hz 的速率提供更新。为了便于记法,在下面的嵌入体系结构描述中假定根据当前“机器人时间步”进行索引(但是,注意,嵌入的基底神经节模型具有不同的固有时间步,如下面 3.3.4 中所述)。

### 3.3.2 嵌入体系结构动作子系统

机器人的控制结构包括五种行为,即动作子系统,可以在任意时间进行切换。这些步骤包括:柱面搜索(cylinder-*seek*, Cs),圆柱拾取(cylinder-*pickup*, Cp),壁面搜索(wall-*seek*, Ws),壁面跟踪(wall-*follow*, Wf),并将圆柱放置(cylinder-*deposit*, Cd)。每个动作子系统独立运行,以便计算指向机器人电机系统的输出信号流。因此,举例来说,柱面搜索使用红外线距离传感器来检测附近的表面,并将可能是圆柱体的物体与其他轮廓(如墙壁)区分开来,并生成电机输出,以根据情况指定朝向或远离目标物体的运动。

我们将机器人活动分解为这五个子系统的灵感来自行为的行为学分类。每个动作子系统由一组条件动作映射和五个动作子系统中的一个组成:柱面搜索、壁面搜索和壁面跟踪从外



围传感器阵列输入的映射模式，以使机器人朝向或远离特定类型的刺激（如物体轮廓）的运动。这些行为可以被视为属于定向反应（orienting responses）或排序（taxes）的行为学范畴（参见，Hinde, 1966）。剩下的两个子系统——圆柱拾取和圆柱放置——根据固定动作模式（FAP）的行为学概念，生成了定型的、仔细计时的运动模式。正如 Lorenz（1935）最初定义的那样，FAP 是对特定刺激模式的特定物种、本能反应。尽管 FAP 可以描述复杂的时空运动模式，但一个独特的特征是，一旦引出，模式的整体形式（尽管不是特定运动元素的参数）不会受到进一步外部线索的影响（Colgan, 1989）。也许 FAP 最著名的例子是鸡蛋回收模式，许多地面筑巢鸟展示了这种模式，这种模式被描述为具有三个顺序要素（Tinbergen, 1951; Hinde, 1966）：

- (i) 站起来；
- (ii) 把喙放在蛋的外面；
- (iii) 把鸡蛋滚回巢穴（把喙从一边移到另一边以防止蛋打滑）。

相比之下，机器人模型中的圆柱拾取固定动作模式由五个顺序要素组成：

- (i) 缓慢接近圆柱体（确保正确识别和正确对准取圆柱体）；
- (ii) 打开抓手时后退（留出空间降低臂）；
- (iii) 将臂降到地面；
- (iv) 关闭抓手（希望在圆柱体周围）；
- (v) 将臂垂直放回原位。

固定动作模式这一术语在行为学中受到了批评，因为它过分强调由此产生的行为模式的刻板和本能（Barlow, 1977; Colgan, 1989），因此有时它更可取。我们在这里使用了最初的术语，这是因为我们的 FAPS 机器人的实现与 Lorenz 的定义是一致的，并且因为神经行为学的研究人员继续发现这是一个有用的概念（Hoyle, 1984; Ewert, 1987; McFarland 和 Bossler, 1993; Casseday 和 Covey, 1996; Toates, 1998）。

虽然一般来说，动作模式可以利用感觉或本体感受数据来形成持续的运动输出，但一些 FAP 在本质上被认为是自然成型的（Hinde, 1966; Ewert, 1987），这表明了内在模式生成机制的参与。在机器人模型中，适当的感官数据相对缺乏，导致我们研究了使用这种固有模式来调节 FAP 中的序列。具体地说，模式的子元素的时间是相对于内部时钟的状态（图 5 中的 C）来确定的，并且完整的行为是作为一组从经过的时间（由该时钟记录）到指定的电机输出模式的映射来实现的。考虑到机器人 FAP 的时空组织完全由这个固有的时间信号来调节，一个关键问题是子系统时钟是如何被控制的。我们的结构通过使时钟的状态依赖于丘脑皮质的反馈，为基底神经节的输出分配了一个重要的角色。具体来说，只有当相关基底神经节通道中的 VL 丘脑发出非零反馈信号（如图 5 中的符号  $f$  所示）时，子系统时钟才启用（非零）。由于嵌入体系结构的这一方面具有重要的功能作用（即，可能与行为选择假设相关），我们将在讨论中考虑基底神经节在行为时间安排中所起作用的证据（第 5.2.2 节）。

我们将  $\mathbf{v}$  指定为任何给定动作子系统产生的矢量，该子系统对其电流电机输出进行编码。为了有效地被基底神经节门控，我们要求  $\mathbf{v}$  的所有元素都是正值并且位于  $[0,1]$  区间内，为了与 Khepera 机器人接合，我们把  $\mathbf{v}$  构造一个包含九个元素的矢量，其中包含目标左右轮速度（各两个元素）、臂位置（三个元素）和夹持器位置（两个元素）。该电机编码方案的更多细节见附录 2，附录 2 还提供了每个动作子系统实现的条件动作映射的完整描述。机器人行为指令表的要素如图 6 所示。

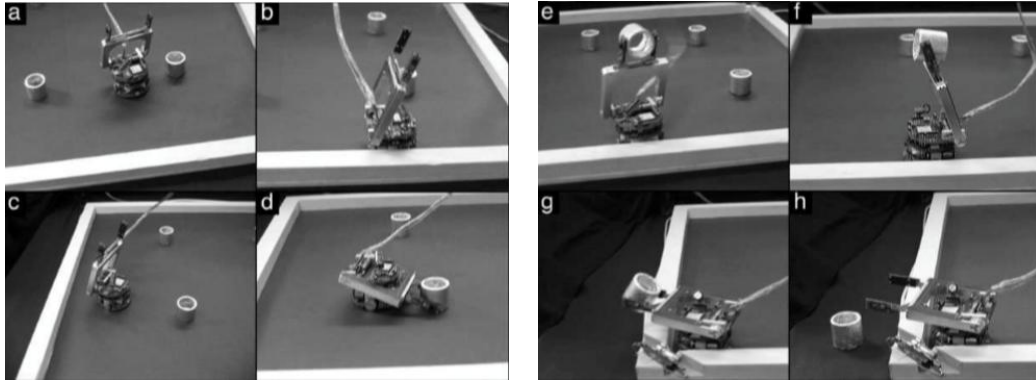


图 6: 模拟觅食任务中机器人行为的要素: (a) 壁面搜索, (b) 壁面跟踪, (c) 柱面搜索, (d) 圆柱拾取, (e) 壁面搜索 (携带圆柱体), (f) 壁面跟踪 (再次携带圆柱体), (g, h) 圆柱放置。

### 3.3.3 嵌入体系结构——决定显著性

中央集中控制的行动选择系统需要能够获取相关知觉、动机和语境信号的机制,以某种广为接受的形式确定每一种竞争行为的相对显著性或紧迫性 (McFarland, 1989; Redgrave 等人, 1999a)。在我们的模型的嵌入架构中,在每个时间步骤中,每个动作子系统的显著值被计算为相关知觉 (perceptual) 和动机变量 (motivational variables) 的加权和,并且也可能依赖于动作子系统本身的当前活动状态。下面简要介绍了这些对显著性计算的贡献。

正如行为学家所指出的,引起行为选择的知觉刺激,通常被称为符号刺激 (sign stimuli),通常与用于控制所选行为执行的刺激截然不同 (Ewert, 1987; Colgan, 1989)。符号刺激通常表示在不同行为的相关供给的直接环境中存在 (或不存在)。在机器人模型中,这些值由感知子系统 (perceptual sub-systems) 根据机器人可用的原始感官数据计算得出,并生成四个双极信号,指示附近墙壁 ( $p_{wall}$ )、嵌套 ( $p_{nest}$ )、圆柱体 ( $p_{cyl}$ ) 或机器人抓取器 ( $p_{grip}$ ) 中对象的存在 (+1) 或不存在 (-1)。

行动选择机制还需要有关内在状态的信息,例如指示当前的能源储备水平 (McFarland, 1989; McFarland 和 Bossler, 1993)。在目前的模型中,两个简单的内在驱动力,与饥饿和恐惧大致相似,是由两个激励子系统 (motivational sub-systems) 计算出来的。恐惧 ( $m_{fear}$ ) 是暴露在环境中的一个函数,随着在环境中的时间的推移而减少,而饥饿 ( $m_{hung}$ ) 随着时间的推移而逐渐增加,当圆柱被放置在实验平台的巢穴角落时,饥饿感也会减少。关于知觉和动机变量计算的更多细节,见附录 3。

除了感知和激励输入,我们的模型还允许动作子系统通过生成一个信号来促进其自身的显著性计算,该信号指示完成正在进行的任务的紧迫性或重要性。我们采用术语忙碌信号 (busy signal) 来描述对加权显著性计算的贡献,加权显著性计算编码了动作子系统活动状态的这一方面 (如图 5 中的符号 b 所示)。在机器人模型中,这些信号为三个子系统的显著性计算提供了贡献。在圆柱拾取中,忙碌信号 ( $b_{pick}$ ) 会提高子系统的显著性,同时机器人会后退以抓取目标圆柱体 (在该操作过程中,机器人通常无法检测到圆柱体),然后在圆柱体被抓取后继续忙碌信号,直到机器人臂返回到安全的垂直位置。信号的作用是补偿感知变量  $p_{cyl}$  (圆柱检测) 和  $p_{grip}$  (夹持器状态) 开关在任务期间发生的显著变化。以类似的方式,圆柱沉积物使用一个忙碌信号 ( $b_{dep}$ ),在圆柱被释放后,直到机器人手臂回到垂直位置,来增强其显著性。最后,如果六个距离传感器中只有一个检测到附近的表面,则“壁面跟踪”会生成信号  $b_{foll}$ 。在这种情况下,墙壁感知 ( $p_{wall}$ ) 通常有一个负值 (因为没有足够的感官

输入来可靠地识别墙壁)，因此信号鼓励保持壁面跟踪，直到恢复墙壁的明确感知或机器人完全失去墙壁的轨迹。附录 2 中给出了这些忙碌信号产生时间的全部细节。第 5.2.1 节考虑了此类信号的一些可能的神经相关因素。

第  $i$  个动作子系统的整体显著性  $S_i$  是相关知觉和动机变量以及该子系统的忙碌信号（如果需要）的线性加权和（包括阈值项）。下面给出的权重是手动选择的，以提供近距离的动作选择竞赛，然后在观察机器人的同时进行调整，直到观察到合适和取巧的动作选择。

柱面搜索（cylinder-*seek*,  $C_s$ ) :

$$s_1 = s_{seek} = -0.12p_{cyl} - 0.12p_{grip} - 0.06m_{fear} + 0.45m_{hung}$$

圆柱拾取（cylinder-*pickup*,  $C_p$ ) :

$$s_2 = s_{pick} = 0.21p_{cyl} - 0.15p_{grip} - 0.18m_{fear} + 0.18m_{hung} + 0.78b_{pick} + 0.25$$

壁面搜索（wall-*seek*,  $W_s$ ) :

$$s_3 = s_{wall} = -0.12p_{wall} + 0.14p_{grip} + 0.18m_{fear} + 0.25$$

壁面跟踪（wall-*follow*,  $W_f$ ) :

$$s_4 = s_{foll} = 0.12p_{wall} + 0.14p_{grip} + 0.21p_{fear} + 0.25b_{foll} + 0.25$$

圆柱放置（cylinder-*deposit*,  $C_d$ ) :

$$s_5 = s_{dep} = 0.33p_{nest} + 0.33p_{grip} + 0.18m_{hung} + 0.40b_{dep} + 0.13$$

这些显著的信号形成了对模型基底神经节的输入，我们接着转向基底神经节。第 3.3.5 节描述了负责将基底神经节输出与电机系统连接的嵌入架构元素。

### 3.3.4 扩展基底神经节模型一种仿生基质的作用选择

在机器人中，基底神经节模型的任务是在五个可用动作子系统之间的每个时间步内进行仲裁，并生成一个随时间变化的动作选择模式，从而产生连贯的行为序列。对于下面报告的实验，我们使用了 Gurney 等人（2001a, b）描述的基底神经节固有电路模型，扩展到包括由 Humphries 和 Gurney（2002）描述的 VL 丘脑和丘脑网状核模型。因此，模型的完整功能架构结合了图 2c 和图 3 中的元素。

在活体动物中，大脑和身体的活动都是在连续时间内展开的，因此建模的理想方案是以相同的、高时间速率模拟神经过程和行为过程。不幸的是，目前的机器人模型存在一个问题，即感知-动作循环以大约 7Hz 的相对稳定的速度在离散的步骤中运行。这就提出了一个问题，即我们如何将基底神经节模型的更新率与机器人的更新率联系起来。在这项研究中，我们选择不在两个更新周期之间执行特定的固定映射，因为这可能导致由非平衡值的基础神经节活动的间歇性采样引起干扰。相反，我们选择在每个机器人的时间步收敛时运行基底神经节模型。显然，只允许平衡基底神经节状态影响运动输出，我们就失去了以高时间速率观察基底节动力学行为后果的机会，这一课题本身就值得我们将进一步研究。

为了方便起见，下面简要总结了完整的基底节模型；读者可参考 Gurney 等人（2001a, b）和 Humphries 和 Gurney（2002）对其形式进行了详细的论证。整个模型中使用的标准漏积分器单元定义如下：设  $a$  为单位激活， $u$  是输入到单位的输入所产生的净输入（或整个突触后电位）。给定速率常数  $k$ （对应于细胞膜电容和电阻） $\dot{a} = da/dt$  由下式给出

$$\dot{a} = -k(a - u) \quad (1)$$

单位的输出  $y$  对应于平均激发率，其下界为 0，上界为 1，并由下述分段线性函数给出：

$$y = L(a, \theta) = \begin{cases} 0 & a < \theta \\ (a - \theta) & \theta \leq a \leq 1/(1 + \theta) \\ 1 & a > 1/(1 + \theta) \end{cases} \quad (2)$$

注意， $\theta$  是一个阈值，低于这个阈值，任何值都会输出零，因此负阈值表示高频强激活，而正值表示对突触输入的抵抗。

下面的公式为模型的每个组件中的第  $i$  个通道指定净输入  $u_i$  和输出  $y_i$ 。净输入  $u_i$  使用模型其他组件的输出  $y_j$  计算，除了模型体感皮层，其中它等于该通道的当前显著性输入  $s_i$ （见第 3.3.3 节）。模型中的多巴胺调制是通过在公式中引入一个乘法因子来提供的，该公式规定了纹状体 D1 通道的传入输入，其中多巴胺调制增加了突触效率。有效比重为  $(1 + \lambda)$ ，其中  $0 \leq \lambda \leq 1$ ；在 D2 通道中，有效比重为  $(1 - \lambda)$ 。除了对从 TRN 到 VL 丘脑<sup>1</sup> 的输入进行了加权，所有参数和阈值与 Humphries 和 Gurney (2002) 和 Gurney (2001a, b) 等人使用的相同。请注意，由于这种结构形成了一个连续的循环，所以运动皮层净输入的定义需要列表后面定义的 VL 丘脑 ( $y_i^{vl}$ ) 的输出。所有其他值都是连续定义的：

躯体感觉皮质：

$$u_i^{ssc} = s_i, y_i^{ssc} = L(a_i^{ssc}, 0.0) \quad (3)$$

运动皮质（层）： $u_i^{mc} = y_i^{ssc} + y_i^{vl}, y_i^{mc} = L(a_i^{mc}, 0.0)$

纹状体 D1（d1）： $u_i^{d1} = (1 + \lambda) \frac{1}{2} (y_i^{ssc} + y_i^{mc}), y_i^{d1} = L(a_i^{d1}, 0.2)$

纹状体 D2（d2）： $u_i^{d2} = (1 - \lambda) \frac{1}{2} (y_i^{ssc} + y_i^{mc}), y_i^{d2} = L(a_i^{d2}, 0.2)$

底丘脑核（stn）： $u_i^{stm} = \frac{1}{2} (y_i^{ssc} + y_i^{mc}) - y_i^{gp}, y_i^{stm} = L(a_i^{stm}, -0.25)$

内侧苍白球（gp）： $u_i^{gp} = 0.9 \sum_i y_i^{stm} - y_i^{d2}, y_i^{gp} = L(a_i^{gp}, -0.2)$

EP/SNr（snr）： $u_i^{snr} = 0.9 \sum_i y_i^{stm} - y_i^{d1} - 0.3 y_i^{gp}, y_i^{snr} = L(a_i^{snr}, -0.2)$

VL 丘脑（vl）： $u_i^{vl} = y_i^{mc} - y_i^{snr} - (0.125 y_i^{trn} + 0.4 \sum_{j \neq i} y_j^{trn}), y_i^{vl} = L(a_i^{vl}, 0.0)$

丘脑网状核（TRN）： $u_i^{trn} = y_i^{mc} + y_i^{vl} - 0.2 y_i^{snr}, y_i^{trn} = L(a_i^{trn}, 0.0)$

在机器人的实现中，利用方程（1）的欧拉解 2 来模拟基底节模型的时间过程：

$$\begin{aligned} \Delta a_i(t) &= -k(a_i(t-1) - u_i(t)) \Delta t \\ a_i(t) &= a_i(t-1) + \Delta a_i(t) \end{aligned} \quad (4)$$

因此，净输入  $u_i(t)$  是使用先前迭代步骤中其他模型组件的输出  $y_j(t-1)$  和为当前机器人时间步（因此在整个收敛过程中）固定的显著性  $s_i$  来计算的。输出  $y_i(t)$  是通过用适当的阈值代替等式（2）中的  $a_i(t)$  得到的。在这里报告的实验中使用了速率常数  $k=25$  和步长  $\Delta t=0.012$ ，并且当两个连续时间步长上最小的  $\Delta a$  小于 0.0001 时，该模型被认为已经收敛。

先前的研究（Gurney 等人，2001b, Humphries 和 Gurney，2002）已经证实，基底神经

节模型在广泛的显著性配对中显示出良好的选择特性，模拟多巴胺水平设为 $\lambda=0.2$ 。因此，该值被用于当前文章中描述的所有实验；模拟多巴胺水平变化的后果将在单独的文章中详细研究。

### 3.3.5 嵌入式结构——门控电机输出

模型基底神经节的输出通过减少或增加对相应运动路径的抑制来控制每个动作子系统产生的运动载体。这在我们的嵌入架构中被实现，使用通过分流抑制（shunting inhibition）机制（图 5 中标记为 SI）生成的门控信号  $e_i$ ，该机制定义如下，对于第  $i$  个动作子系统， $e_i$  ( $0 \leq e_i \leq 1$ ) 由 (5) 给出：

$$e_i = L(1 - y_i^{snr} / y_{ic}^{snr}, 0.0) \quad (5)$$

这里， $y_{ic}^{snr}$  是一个常数，等于当模型运行到收敛时，所有信道上的零显著性输入所得到的 EP/SNr 的高频输出。对于上面列出的基底神经节模型的参数， $y_{ic}^{snr} = 0.169$ 。我们的模型假定，基底神经节的输出水平完全抑制了靶结构，当输出低于  $y_{ic}^{snr}$  时，靶的解除抑制开始，随着输出的减少成线性增加，当输出为零时最大。由于这种嵌入结构的元素调和了基底神经节输出对运动系统的影响，因此它在我们的模型中起到了“理论相关”的作用。因此，在第 5.2.3 节中，我们将考虑到电机和电机前系统的基底节输出也可能具有类似于分流抑制的门控效应的证据。

所有动作子系统的门控电机输出在所有通道上求和，所得结果通过另一个限制器给出聚合矢量，如下：

$$\hat{v} = L(\sum_i e_i v_i, 0.0) \quad (6)$$

最后，电机装置将矢量  $\hat{v}$  映射到机器人可以理解的电机命令中，该映射的详细信息在附录 4 中给出。

请注意，聚合电机矢量表示电机状态的目标值，即目标轮速度、夹持器臂高度和夹持器钳口位置。如果所有通道的基底神经节完全受到抑制，则此聚合命令将为零值，机器人将冻结在当前位置。如果一个或多个通道被部分（但不是完全）抑制，机器人将采取行动，但其移动可能会因电机信号尺寸减小而减慢。最后，请注意，任何未被基底神经节完全抑制的丢失行为所产生的运动信号将与获胜的运动信号结合。这种机制允许在基底神经节无法有效抑制竞争对手的情况下发生扭曲（distortion）（机器人试图同时做两件事）。

### 3.3.6 机器人模型的时间历程

为了明确机器人的更新周期、嵌入结构和基础神经节模型之间的关系，可以将一个机器人的时间步内所发生的具体模型中的活动概括为：

- (i) 制定机器人电流聚合电机指令  $\hat{v}$  并获得新的传感器数据。
- (ii) 对于每个动作子系统  $i$  更新显著性  $s_i$  并生成新的运动矢量  $v_i$ 。
- (iii) 运行基底节模型直至收敛。
- (iv) 利用聚合基底节模型的输出  $y_i^{snr}$  生成门控信号  $e_i$ ，并计算出新的聚合电机指令矢量  $\hat{v}$ 。
- (v) 使用 VL 丘脑、 $y_i^v$  的输出，启用或禁用执行固定动作模式的任何子系统的子系统时钟。对于嵌入的基底神经节模型，在收敛时保留所有漏积分单元的激活值作为下一时间步的起始值。这使得机器人的时间步长可能存在滞后现象，并对行为产生潜在的重要影响。

## 3.4 行动选择指标

为了帮助展示模型结果，有助于定义一些描述行动选择竞争结果的术语。

首先，我们注意到，方程（5）中定义的选通信号  $e$  提供了一个通过嵌入的基底节模型进行选择的有用的标准化测量。因此，在评估模型性能时，我们将使用  $e_i$  作为第  $i$  个动作子系统的电机输出传输到电机资源的效率（efficiency）度量。考虑到 5% 的绝对利率，我们将第  $i$  个竞争对手的选择状态定义为：如果  $0.95 \leq e_i \leq 1$ ，则称为完全选择（fully selected）；如果  $0.05 \leq e_i < 0.95$ ，则称为部分选择（partially selected）；如果  $e_i < 0.05$ ，则称为取消选择（unselected）。

对与获胜子系统相关的具体指标是很有帮助的，因此，我们定义

$$e_w = \max_{\forall i} e_i \quad (7)$$

把（7）作为当前机器人时间步的获胜者的效率，以及

$$d_w = \frac{2(\sum_i e_i - e_w)}{\sum_i e_i} \quad (8)$$

把（8）作为影响获胜者输出的失真程度。请注意，当所有其他竞争对手的效率为零时， $d_w$  等于零，它将随着部分不受限制的输家数量的增加而增加，如果两个或多个渠道完全不受限制，则  $d_w$  等于 1.0 或更大。最后，在行为学研究的启发下（Lehner, 1996），我们描述了一系列不间断的时间步骤，这些步骤共享同一个获胜者，并且对于这些步骤， $e_w \neq 0$ ，作为一个单独的行为回合（bout）。

基底神经节选择竞争的结果，作为一个整体，可以用向量  $e$  来描述。使用刚刚为单个竞争者定义的标准，我们为完整竞争的可能结果分别标记以下标签：

绝对选择（Clean selection）：完全选择一个竞争对手，未选择所有其他竞争对手。

部分选择（Partial selection）：部分选择一个或多个竞争对手，没有完全选择的竞争对手。

扭曲选择（Distorted selection）：一个竞争者只能完全选择，至少一个竞争者只能部分选择。

多项选则（Multiple selection）：完全选择两个或两个以上的竞争对手。

没有选择（No selection）：所有竞争对手都没有选择。

## 四、实验

采用基底神经节模型进行了三个实验。实验 1（第 4.1 节）旨在帮助解释机器人模型的行为，并使用扩展基底神经节模型的无实体版本对显著空间进行了系统搜索。实验 2（第 4.2 节）是我们对使用第 3.3.3 节中描述的手动调节显著度输入实现的机器人模型的主要研究。实验 3（第 4.3 节）研究了在机器人模型中使用多个高显著性输入的行为结果。

### 4.1 实验 1——非实体模型的选择性质

#### 方法：

为了提供一个解释机器人模型行为的框架，我们使用扩展基底神经节模型的无实体版本对显著性空间进行了系统搜索。具体地说，我们模拟了一个五通道模型，其中有两个活跃通道，系统地将通道 1 中的显著性  $s_1$  从 0 变为 1，步幅为 0.01，然后对通道 1 显著性的每个值，对应的将通道 2 中的显著性  $s_2$  从 0 变为 1，步幅为 0.01。对于每个产生的显著性向量  $(s_1, s_2, 0, 0, 0)$ ，模型运行直至收敛，结果按照第 3.4 节中的方案分类。重要的是，选择竞争是按从低值到高值的顺序进行的。模型中所有漏积分器的激活水平对于  $s_1$  的每一个新值都初始化为零，但此后，当测试显著性指标的值时，该值从一个竞争者保留到下一个竞争者。换言之，我们模拟了通道 1 最初是唯一活跃通道的情况，并在保持通道 1 不变的情况下逐渐增加通道 2，目的是模拟机器人模型经验连续性的某些方面，其中基底神经节选择竞争的最近历史经历可能会通过滞后竞争影响电流。

## 结果:

非实体模型显示绝对选择的比例较高 (79%)，部分选择 (17%)。选择效率降低和扭曲 (选择) 只发生在均匀匹配、高度显著的竞争中。该模型还显示了滞后的证据，并随显著性强度的变化而变化。

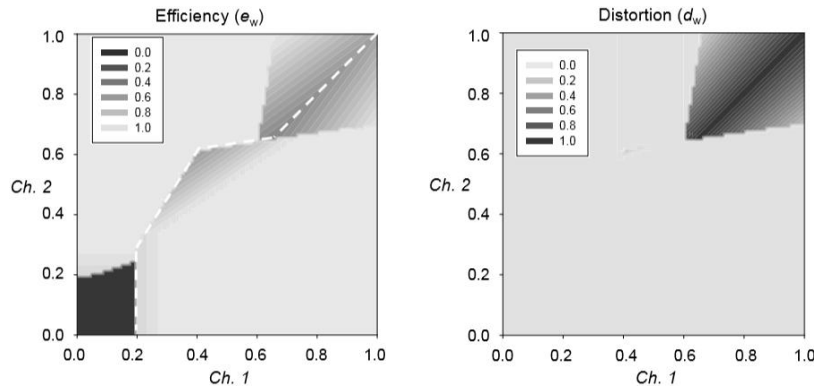


图 7: 在取胜通道中进行效率 (左) 和失真 (右) 两种有效通道的系统显著空间搜索。逐渐变浅的阴影显示出更高的效率 (顶部) 或更低的失真 (底部)。显著空间的采样分辨率为 0.01。对于通道 1 的每个显著值, 通道 2 从 0 开始逐渐增加。仅当选择了新的通道 1 值时, 才重新初始化模型, 从而允许滞后的可能性。左图中的虚线表示选择竞争的下边界, 有利于通道 1。

上述状态空间搜索产生了 10000 场 (100×100) 显著性竞争, 其中 78.6% 的竞争结果为绝对选择, 16.7% 的竞争结果为部分选择, 4.3% 的比赛结果为没有选择, 0.3% 的竞争结果为扭曲选择, 0% 的比赛结果为多项选择。该分析的一些进一步结果如图 7 所示。在上图中, 我们展示了在 ( $s_1$ ,  $s_2$ ) 平面上绘制的每个显著性竞争的获胜通道的效率  $e_w$ , 下面是获胜通道的变形等效图  $d_w$ 。逐渐变浅的阴影分别表示提高效率 (顶部) 和减少变形 (底部)。上图中的虚线还显示了选择竞争被解决的边界, 有利于通道 1 (即通道 1 的效率超过通道 2 的效率)。图 7 的几个特性值得评论。首先, 在大多数状态空间中, 效率高、失真小, 因此“绝对选择”。第二, “没有选择”只有非常低的显著配对。第三, 对于强、均匀匹配的显著值 ( $s_1$ ,  $s_2 > 0.6$ ), 选择效率降低, 显著失真, 导致状态空间这一区域的“部分选择”。最后, 我们注意到模型显示了明显的滞后 (如上图中的虚线所示)。特别是, 1 号通道的显著性值介于 0.2 和 0.6 之间, 能够抵抗 2 号通道显著性的上涨, 直到 2 号通道显著性在数值上显著增强。此外, 显著性竞争的整体强度的变化导致不同程度的滞后, 在中等显著性水平上效果最为显著。

## 4.2 实验 2——机器人基底神经节

### 方法:

机器人进行了五次试验, 每次持续 300 秒 (2000 个机器人时间步), 模型的所有参数如第 3.3 节所述。实验步骤如下。在每次试验开始时, 机器人被放置在实验平台的中心 (见图 4), 面向四面墙中的一面, 四个圆柱体从每个角落斜向放置相距 18cm。所有的运动输出最初设置为零, 基底神经节模型运行直至收敛, 所有通道上的显著性为零。每一次试验都会生成自动日志, 详细记录机器人在每个时间步的感觉、动机和基底神经节状态, 以及其行为选择的总体结构。大多数试验都使用了一个位于实验平台上约 1 米处的摄像头被记录在数字视频中, 可以详细检查机器人的行为及其与环境中物体和表面的相互作用。

### 结果:

在下文中, 我们描述了:

- (i) 机器人模型的一般选择特性;
- (ii) 机器人行为过程中模型基底神经节的内在处理;

(iii) 机器人模型的行为。

(i) 机器人在激烈的选择竞争中所占的比例很高，结果主要是绝对选拔（84%），在 10.1% 的竞争中表现出行为持续性的迟滞现象。

图 8 提供了机器人探索的显著空间的局部视图。每个单位方格显示（大约 10000 个）基底神经节竞争（每个机器人时间步 1 个）的比例，其中获胜通道（水平轴）和最显著输者（垂直轴）的显著性在给定范围内（注意，通常在其他未显示的失败通道中也存在非零显著性。

（图中）颜色较深，表示比例较大。该图显示，有很大比例的竞争激烈的显著性竞争，但高显著性竞争的地区（可以预期效率降低）是相对稀疏的抽样。我们的分析再次根据 3.4 节中规定的标准对基底神经节选择竞争的结果进行分类。在所有五个试验中，84.4% 的显著性竞争结果是绝对选择，9.7% 的部分选择，5.9% 的没有选择。没有扭曲选择或多项选择。这一结果表明，基底神经节模型产生（主要）绝对选择的操作范围足以满足相当复杂的机器人任务的操作选择要求。

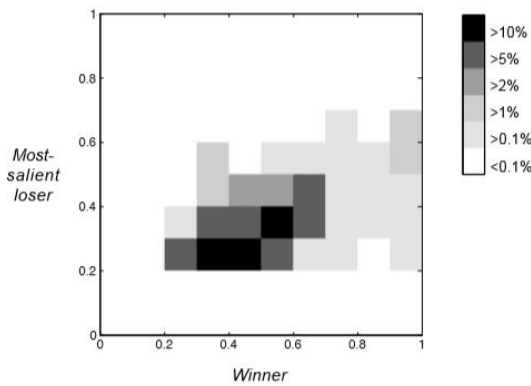


图 8: 机器人在典型试验中对突出空间采样的局部视图。轴（水平）表示获胜通道和（垂直）最显著输者的显著性。阴影表示在给定 (0.1X0.1) 箱中约 10000 对显著性对的比例（较暗=较大）。平均频道显著性为 0.235（跨所有频道和所有时间步），平均制胜显著性为 0.475，平均利润（赢者和最显著的输者之间）为 0.154。

图 8 中对角线 ( $x = y$ ) 上方的一些阴影单位方格的存在提供了嵌入基底神经节模型后的证据（即胜出通道的显著性低于其最近竞争者的竞争）。在机器人中，迟滞现象转化为行为持续性，在这种持续性中，机器人继续显示选定的行为，超过了“赢者通吃”选择器时，将切换到更高的显著性任务。在所有 5 次试验中，在 10.1% 的时间步上显示持续时间，因此对观察到的机器人行为有显著影响。

(ii) 具体模型的内在处理

图 9 说明了在典型试验的前 180 秒（约 1200 个机器人时间步）期间，嵌入式基底神经节模型中发生的一些内在处理。图 9a 显示了每个通道（动作子系统）的两个基底节输入结构中的活动：体感皮层， $y^{sc}$ （实线），其输出与净显著性成比例，以及 D1 纹状体， $y^{d1}$ （虚线）。在图 9a 中的最后一个图显示了胜出通道（实线）的皮质输出与最显著的输者（虚线）的皮质输出。

图 9b 显示了基底神经节输出核 EP/SNr， $y^{snr}$  的每个通道的模型活性。图 9b 中的最后一个图显示了四个丢失通道（实线）的平均 EP/SNr 输出与获胜通道（虚线）的平均 EP/SNr 输出。下面我们简要地将模型内在处理的一些关键特征与基底神经节模型和嵌入结构的计算特性以及机器人的观察行为联系起来。我们还提出了一些定量的措施，将嵌入式模型的选择行为与 EP/SNr 和 STN 中的内在活动联系起来。



纹状体的活动反映了通道显著性的变化，受丘脑皮质反馈的调节。

皮质和纹状体 D1 活动图 (9a) 显示，显著性随时间逐渐和急剧变化，反映了连续和一般缓慢变化、动机或离散和快速变化的知觉变量的变化。皮质和纹状体活动 (填充线和虚线) 之间的差异说明了丘脑皮质反馈增强选定通道的效果，以及纹状体阈值抑制低显著性输入的效果。在  $t=160-180s$  的时间间隔内，本试验出现一段无选择期，其中最活跃通道的皮质输出处于较低水平 ( $<0.3$ )，第二个通道的输出接近相等水平 (见 9a 中的最终图)，另外两个通道的输出显著非零。这一结果表明，在五通道机器人模型中，选择阈值通常高于图 7 所示的两通道无实体模型。这一结果与先前的发现 (Gurney 等人, 2001b) 一致，即多个活跃通道的存在使任何给定通道的选择变得更加困难，这是我们称之为选择限制的模型的一个特性。

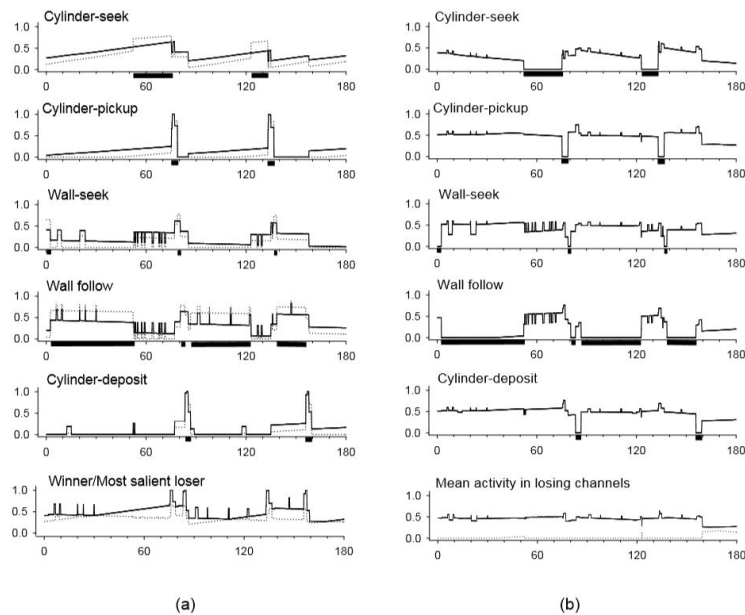


图 9: 典型试验前 180s 嵌入基底神经节模型的内在活性。(a, 左图) 前五张图显示, 对于五个基底神经节通道中的每一个, 躯体感觉皮质 (实线) 和 D1 纹状体 (虚线) 的输出与时间对应。最后的图显示了胜出通道 (实线) 的输出与最显著的输者 (虚线) 的输出。(b, 右) 前五张图显示了 EP/SNr 的每个通道输出, 而最后一张图显示了与取胜通道 (虚线) 相比, 丢失通道 (实线) 的平均输出。子系统图下方的实心条表示相应作用子系统的完全选择周期。注意, 在  $t=160-180s$  期间没有选择。

“忙碌信号”在维持行为选择方面起着重要作用。

在皮质和纹状体输出图 (9a) 中也可以看到子系统忙碌信号在维持选定行为方面的效用。例如, 在圆柱拾取过程中, 通过执行与实现抓取圆柱体的主要目标相对应的行为 ( $t=75-79s$ ), 会有一个明显的变化。此时, 忙碌信号被激活, 以保持其高于竞争者的行为的显著性, 同时完成完整的移动 (使机器人手臂回到垂直位置)。如果没有这个信号, 一旦夹住圆柱体, 行为的显著性就会大大下降, 导致无法完成完整的动作模式。换句话说, 忙碌信号允许在完成任务的一个基本的“内务 (house-keeping)”时维持行为。在执行圆柱放置 ( $t=83-87s$ ) 过程中, 忙碌信号也起着类似的作用。在壁面跟踪过程中忙碌信号的作用可以在一系列间歇的显著性峰值 ( $t=5-40s$ ) 中看到, 这些峰值补偿了墙感知的暂时中断, 因此阻止了早期返回壁面搜索行为。在这个角色中, 忙碌的信号有助于避免由于噪声或不明确的感知数据而导致不必要的行为切换。

基底神经节的输出主要表现为对胜利者的完全抑制和对失败者的增强抑制。在行为选择过程中, 平均 EP/SNr 活性增加。

基底神经节输出图 (9b) 显示了更加本质的基底神经节选择机制 (前馈, 偏离中心, 回转和 GP 控制电路) 的结果, 这导致在选定子系统完全不受抑制 (零 EP/SNr 输出) 的动作子系统之间产生明确决策边界。注意, 在没有选择行为 ( $t=160-180s$ ) 期间, 与所有赢的行为的时间相比, 丢失通道的平均活动 (见下图) 有显著差异。为了量化这一效应, 对不同选择结果的五次试验进行了平均 EP/SNr 活性计算。以未选择期间的 EP/SNr 输出为基线, 部分选择期间所有通道 (丢失通道) 的平均输出为基线的 123% (134%), 绝对选择期间的平均输出为 156% (182%)。更普遍地说, 在所有的竞争中, 获胜通道的活动与失败通道的平均活动之间存在着很强的负相关 ( $r=-0.895$ )。因此, 当一个赢的通道被解除抑制时, 对输的通道的抑制输出水平增加。

在行为选择过程中, STN 活性增加, 是增加对丢失通道抑制的主要原因。

在基底神经节模型中, 唯一为 EP/SNr 提供兴奋性输入的结构是 STN, 因此 STN 与丢失通道中的活动具有很高的相关性 ( $r=0.884$ )。STN 本身是由皮质输入 (体感和运动) 编码通道显著性和丘脑皮质反馈 (EP/SNr-VL-MC) 驱动的, 并由一个带有 GP 的环路驱动。对模型中 STN 激活的分析表明, 随着皮质活动的增加, 显著性逐渐增加 ( $r=0.696$ , SSC), 当没有选择的基线期与部分选择 (基线的 176%) 或绝对选择 (272%) 相比较时, 也显示出急剧增加。从这些数据中, 我们可以得出结论, 当选择一个赢的通道时, 通过丘脑皮质反馈, 增加了对该通道的 EP/SNr 的抑制 (通过直接纹状体神经通路), 同样的正反馈信号也通过 STN, 导致丢失通道的 EP/SNr 活性增加。

(iii) 观察到的机器人行为显示出在选定通道之间清晰而果断的切换, 在延长的回合和目标实现序列中被组织, 并在序列结构和持续时间上显示出变异性。

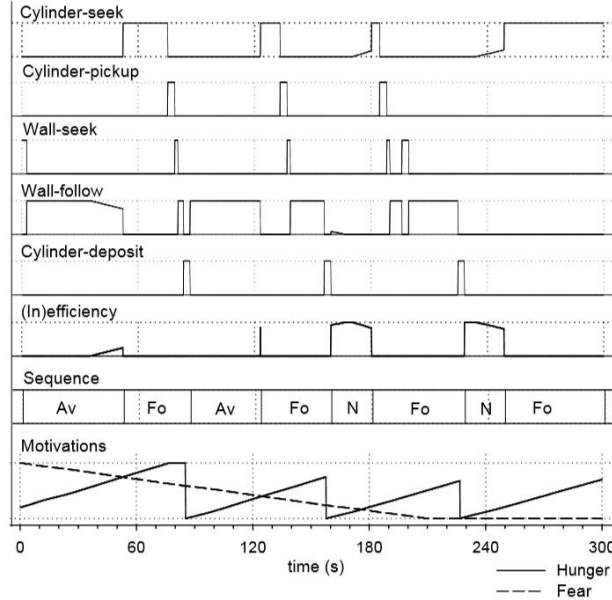


图 10: 全 300s 试验机器人模型中动作选择的回合/顺序结构。从上到下, 前五个图显示了针对时间绘制的给定行动子系统的选择效率 ( $e$ ), 第六个图显示了当前胜利者的效率 ( $1 - e_w$ ), 第七个图显示了关于序列的高阶结构 ( $Av$  = 回避,  $Fo$  = 觅食), 最后一个图显示了两个模拟动机的级别。y 轴上的所有测量值都在 0 和 1 之间。机器人开始这项试验时, 高度的模拟恐惧导致了壁面搜索比其他动作更为突出。在快速找到一个墙 ( $t=3s$ ) 后, 壁面跟踪, 变得更加突出并被选中。这两个回合形成了一个更高阶的回避行为序列。回避行为被饥饿感的增加和恐惧感的减少所打断, 从而提高了圆柱体搜寻的相对显著性。一旦壁面跟踪的显著性明显低于柱面搜索的显著性, 机器人就切换到柱面搜索的显著性 ( $t=52s$ )。当它找到一个圆柱体时, 选择了圆柱拾取 ( $t=75s$ ), 然后选择了壁面搜索 ( $t=79s$ , 这次携带一个圆柱体), 壁面跟踪 ( $t=81s$ ), 最后在检测到嵌套区域时选择了圆柱拾取 ( $t=84s$ )。这四个回合构成了一系列的食欲或觅食 ( $Fo$ ) 活动。完成觅食序列后, 模拟饥饿程度暂时降至零 ( $t=87s$ ), 机器人进入了一个新的回避期 (壁面跟踪, 因为机器人已经在竞技场外围)。饥饿感的增加导致了另外三个觅食序列 (最后一个未完成), 中间是两段时间的不活动 (当恐惧接近零时, 没有动机执行回避行为)。

根据我们先前对行为回合的定义，可以看出机器人的活动（如图 9 所示）由适当的、适当扩展的个体活动回合组成，随着时间的推移，这些个体活动被整合到适当的、高阶的目标实现行为序列中。图 10 中更容易看到这种回合/序列结构，在图 9 中我们显示了整个 300 s 试验的子系统活性，其中样本基底神经节活性如图 9 所示。从上到下，图中的前五个图显示了五个动作子系统中每个系统随时间变化的效率 E，其中一系列完整的选择显示为实体块（以动作行为图的形式）。下一（第六）张图显示了一段时间内  $(1-e_w)$  的图，从而显示了机器人在多大程度上效率低下地表达其当前动作或处于静止状态（请注意，本次试验中没有扭曲的选择，因此无需绘制扭曲的图）。下一幅图以高阶行为序列的形式显示了机器人行为的结构，而最后一幅图则显示了两个模拟的动机。每个动作子系统的选择是由相关的感知供给触发的，维持一个适当的时间，然后迅速而果断地切换到下一回合。当正在进行的活动的显著性显著低于竞争者时，或者竞争者的显著性显著高于当前选定的行为时，就会发生行为转换。在这两种情况下，正在进行的行为突然终止，新的活动开始时几乎没有延迟（通常在下一个机器人时间步中）。

最初的“壁面搜索”和“壁面跟踪”序列反映了模拟恐惧的高初始水平，可以被视为形成一个高阶的避免（*Av*）行为序列，使机器人远离开放空间。随着恐惧感的减少和饥饿感的增加，这之后又发生了第二次行为选择：寻找和收集一个圆柱体（柱面搜索和圆柱拾取），把它带到一个巢穴（壁面搜索，壁面跟踪），然后把它放在那里（圆柱放置），这可以被视为一系列的食欲或觅食（*Fo*）活动。随后，机器人进行了进一步的回避和觅食序列，其间有短暂的不活动期，此时机器人没有显示任何运动，这与两个人工激励水平都较低的时间相对应。

图 10 所示的模型行为是在所有五个试验中观察到的典型行为，然而，有许多因素导致试验内外的显著差异。这些包括机器人和圆柱体初始位置的微小变化、传感器噪声、感知混叠（例如，可能来自墙或圆柱体的模糊信号）、车轮打滑、与实验平台地板和墙壁的摩擦。这种变异性的一些影响如表 1 所示，表 1 描述了所有行为对的转换频率（垂直轴为前一行为，水平轴为后续行为）。标准觅食序列，*Cp*, *Ws*, *Wf*, *Cd* 的优势，从对角线上的高比例过渡可以清楚地看到。*Cd*-*Wf* 和 *Wf*-*Cs* 转换反映了在觅食之后或之前，作为一种回避行为而出现的墙面跟踪。较少的 *Cp*-*Cs* 转换和 *Wf*-*Ws* 反映了这样一个事实：机器人偶尔未能正确地抓住一个圆柱体（因此返回到圆柱体搜索），或在墙面跟踪过程中失去与墙的接触（因此返回到墙搜索）。变异性的第二个结果是，觅食序列的持续时间分布广泛（范围 23.7-126.4s，中位数 36.63s），部分原因是机器人使用的简单搜索策略及其传感器的短范围。

	每次试验回合	相对频率 (%)	转换频率 (%)					
			Cs	Cp	Ws	Wf	Cd	No
Cs	3.6	15.0	0.0	<b>100.0</b>	0.0	0.0	0.0	0.0
Cp	3.6	15.0	16.7	0.0	<b>83.3</b>	0.0	0.0	0.0
Ws	5.2	21.7	0.0	0.0	0.0	<b>100.0</b>	0.0	0.0
Wf	6.6	27.5	30.3	0.0	18.2	0.0	<b>45.5</b>	6.1
Cd	3.0	12.5	0.0	0.0	0.0	46.7	0.0	<b>53.3</b>
No	2.0	8.3	100.0	0.0	0.0	0.0	0.0	0.0

表 1: 对于每个动作子系统，该表显示了每次试验的平均回合数、替代行为的相对频率以及不同过渡的相对频率（垂直轴上的先前行为、水平轴上的后续行为）。否表示有一次不活动。过渡矩阵主要由标准觅食序列 *Cs*-*Cp*-*Ws*-*Wf*-*Cd*（以粗体显示）控制。

#### 4.3 实验 3——机器人模型在高显著性竞争中的行为

### 方法:

我们之前注意到（第 4.1 节），当两个竞争者的显著性水平较高时，无实体模型中可能会出现绝对选择的崩溃。为了检验这种选择模式的行为结果，对机器人模型进行了 5 次 120s（约 800 个机器人时间步）的试验，其中每个通道的显著性在每个时间步上都增加了一个恒定的量（+0.4）。实验程序的所有其他方面如实验 2（第 4.2 节）所述。

### 结果:

在一系列连续的高显著性竞争中，机器人表现出行为崩溃的模式，其特征是：

- (i) 降低了选定行为的效率和扭曲；
- (ii) 快速切换和不完成觅食行为。

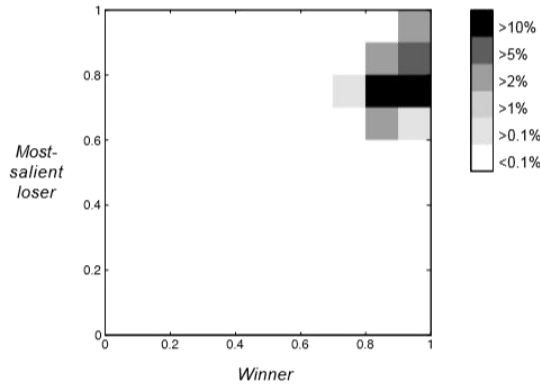


图 11: 所有通道的显著性均匀增加 (+0.4) 后的显著性空间探索。轴（水平）表示获胜通道和（垂直）最显著输者的显著性。阴影表示在给定 (0.1X0.1) 箱中约 4000 对突出显示的比例（较暗=较大）。平均频道显著性为 0.576（所有频道和所有时间步骤），平均制胜显著性为 0.935，平均利润（赢者和最显著输者之间）为 0.173。

图 11 显示了在所有试验中，显著性强度增加对赢者/最显著输者显著性空间探索的影响。该图显示，几乎所有（~4000 个）显著性竞争都出现在显著性空间区域（与图 7 相比），在该区域可以预期效率降低和选择失真。

觅食行为之间的快速转换构成了行为维护错误而导致的“行为陷阱”。

在非常突出的竞争中出现的效率降低和选择扭曲，通常与图 7 所示的分析一致。然而，在两种觅食行为之间的觅食序列振荡中观察到的崩溃，在未能抓住圆柱的情况下，需要进一步的解释。回想一下，在机器人倒车为抓爪臂腾出空间的地方，拾捡圆柱体的 FAP 利用了一个忙碌信号，因此不再能够检测圆柱体。在正常情况下，该信号将足以维持行为选择，直到收集到圆柱体，并将夹持臂提升到垂直位置。

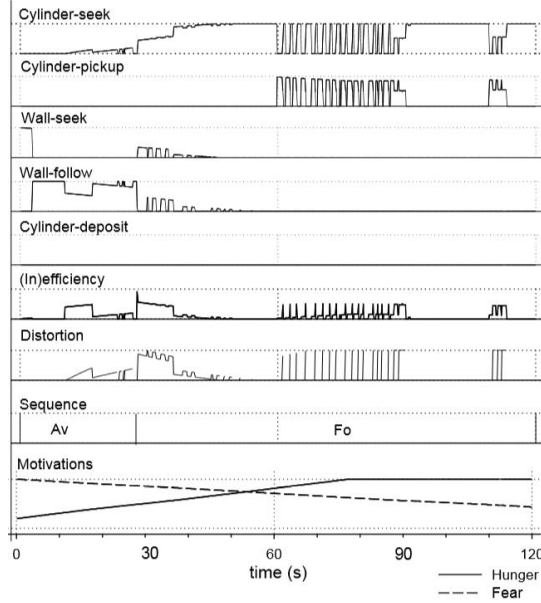


图 12: 在所有通道的显著性均匀增加 (+0.4) 后, 机器人模型试验 120s 中动作选择的回合/顺序结构。从上到下, 前五个图显示了针对时间绘制的给定动作子系统的选择效率 ( $e$ ), 第六和第七个图显示了当前胜利者的效率 ( $1-e_w$ ) 和失真 ( $d_w$ ), 第八个图显示了关于序列的高阶结构, ( $A_v$ =回避,  $F_o$ =觅食), 最后一个图显示了两个模拟的动机。y 轴上的所有测量值都在 0 和 1 之间。最初的回避顺序遵循预期模式, 壁面搜索成功一段时间壁面跟踪, 然而, 从大约  $t=15s$  开始, 柱面搜索显著性逐渐增加, 导致壁面跟踪行为的效率降低, 导致移动明显放缓。到觅食活动的转变发生在  $t=30s$  左右, 选择柱面搜索, 但在间歇部分选择回避行为 (壁面搜索、壁面跟踪) 时效率降低。当机器人找到一个圆柱体 ( $t=60s$ ) 时, 圆柱拾取行为被清晰地选择, 但随后过早中断。机器人后退准备放下夹爪臂, 但随后突然切换回柱面搜索。接着, 机器人反复接近圆柱体, 但未能抓住圆柱体, 在圆柱拾取和柱面搜索之间进行了一段时间的快速切换。随着饥饿动机的进一步增强, 机器人显示出两种行为的混合 (柱面搜索和圆柱拾取) (约  $T=90s$  和  $T=110s$ ), 但效率降低, 仍然无法正确地抓住圆柱体。

图 12 说明了机器人在典型试验中的行为。最初的回避序列遵循预期的模式, 但向觅食活动的过渡并没有完全开始, 而是表现出效率降低和间歇、部分选择 (失去) 回避行为。在观察者看来, 机器人在过渡过程中的运动似乎有些缓慢和不确定。在觅食回合中, 机器人反复接近圆柱体, 但未能抓住圆柱体, 柱面搜索和圆柱拾取之间的快速切换时间延长。最初观察到的模式 ( $t=60-85s$ ) 是机器人接近圆柱体; 倒车时好像要将其收集到夹持器中; 然后向前移动而不降低夹持器臂, 轻轻向前推动圆柱体。稍后 ( $t=85-90s, 110-115s$ ), 两种行为都显示出部分选择, 机器人会在向前移动的同时降低手臂, 但由于未正确对齐而无法抓住圆柱体。在所有的五个试验中, 机器人的选择行为同样低效和扭曲, 经常显示觅食行为的快速交替。这在表 2 中的转换矩阵中进行了说明, 表明机器人的行为主要由序列 Cs-Cp-Cs-Cp 控制, 没有任何试验导致成功的觅食序列。

	每次试验回合	相对频率 (%)	转换频率 (%)					
			Cs	Cp	Ws	Wf	Cd	No
Cs	14.6	<b>44.5</b>	0.0	<b>98.6</b>	0.0	1.4	0.0	0.0
Cp	14.4	<b>43.9</b>	<b>100.0</b>	0.0	0.0	0.0	0.0	0.0
Ws	1.8	5.5	0.0	0.0	0.0	100.0	0.0	0.0
Wf	2.0	6.1	60.0	0.0	40.0	0.0	0.0	0.0
Cd	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

表 2: 对于每个动作子系统, 该表显示了每次试验的平均回合数、替代行为的相对频率以及不同过渡的相对频率 (垂直轴上的先前行为、水平轴上的后续行为)。否表示有一次不活动。过渡矩阵主要由圆柱拾取 (Cs) 和柱面搜索 (Cs) 之间的切换 (大约和以粗体突出显示的过渡频率)。

然而，在高显著性模型中，圆柱体的消失增加了柱面搜索的显著性，从而引发一个短暂的扭曲选择周期（注意图 12 中扭曲测量图中的短暂峰值），在此期间，圆柱拾取和柱面搜索同时被选择。在扩展的基底神经节模型中，高水平的畸变有一个有趣的结果，即减少取胜通道上的丘脑皮质反馈（圆柱拾取）。这是因为丢失通道中减少的 SNr 活动会增加这些通道的 TRN 活动，从而抑制获胜者的 VL 丘脑活动（见图 3 和第 3.3.4 节）。在模型中，圆柱拾取模式的时间，以及忙碌信号的时间，依赖于持续的丘脑皮质反馈。当此反馈丢失时，子系统时钟断开，忙碌信号取消。圆柱拾取的显著性随之急剧下降，其竞争行为，柱面搜索，赢得了随后的基底节选择竞争。当机器人重新建立与圆柱体的接触时，这种模式就会重复。总而言之，快速行为转换的崩溃模式是由于行为过早中断而产生的，这种行为的完全依赖于内在的显著性增强（忙碌信号）。这可以被描述为行为维护的错误。未能成功执行动作模式（从而触发行为序列的后续元素）将机器人置于行为陷阱中，在该陷阱中，它会重复执行不完整且无效的动作序列。

## 五、讨论

### 5.1 主要研究结果总结

嵌入的基底神经节成功地在一个机器人模型中生成了集成行为序列，该模型提供了一系列替代行为和不同程度的模拟动机。机器人在连续的行为之间清晰而果断地切换，每当有明显更突出的竞争对手时，都会中断正在进行的行为。这个结果支持这样一个假设：基底神经节电路的功能特性（在我们的计算模型所捕获的范围内）使其适合于解决选择冲突的任务。因此，机器人模型支持基底神经节的有效动作选择的主张，超越了先前对这些电路良好信号选择特性的论证。虽然该模型有一个操作范围，支持对大多数突出竞争级别的绝对选择，并且足以在我们的机器人任务中提供适当的选择，但当模型具有多个突出竞争对手时，可能会降低选择效率和失败的竞争者产生部分选择。因此，当模型的某个版本被测试时，所有通道的显著性都显著增加，观察到运动输出失真和行为崩溃。

在余下的讨论中，我们考虑：

- (i) 嵌入架构的一些元素及其可能的神经关联；
- (ii) 机器人模型的结果与动物行为的神经行为学观察结果的比较；
- (iii) 当前研究与基底神经节其他计算模型的关系。

### 5.2 机器人嵌入体系结构的可能的神经关联

如前（第 3.1 节）所述，我们对基底神经节进行动作选择假设的进行机器人测试的方法是基于这样一个假设：信息模型可以通过将仿生组件（此处为扩展基底神经节模型）与足够的工程组件相结合来创建完整的工作。然而，这种策略要求我们在仿生组件和工程组件之间设计一个合适的接口，以便为嵌入的神经模型提供适当的输入信号，并为其输出分配一个生物学上合理的角色。我们还认为，有必要评估这样构建的模型的工程组成部分对正在研究的理论问题可能产生的影响。以下简要考虑了与这些需求相关的架构/模型接口的关键元素。

#### 5.2.1 基底节输入编码与持续行为的选择和维持相关的信号

假设基底神经节参与选择行为意味着基底神经节输入编码竞争行为的相对显著性（Redgrave 等人，1999a；Zink、Pagnoni、Martin、Dhamala 和 Berns，2003）。有证据支持这一观点，其中有研究显示纹状体多棘神经元在运动前活动（参见 Mink，1996 年的评论）。我们认为基底神经节是一种动作选择装置，进一步猜想基底神经节的活动不仅对选择获胜动

作很重要，而且对所选动作的适当维持和终止也很重要（Redgrave 等人，1999a）。数据表明，基底神经节参与了选择的维持，运动后纹状体投射神经元中有相当一部分发生了激活，纹状体中与运动相关的活动的时间分布在与运动开始相关的较宽延迟范围内。（Aldridge、Anderson 和 Murphy，1980；DeLong 等人，1984；Jaeger、Gilman 和 Aldridge，1995；Mink，1996）。基底神经节输出神经元（灵长类中的 EP/SNr 或 GPi/SNr）的活动抑制也被记录为发生在肢体运动（Schultz，1986；Mink 和 Thach，1991），扫视（Hikosaka 和 Wurtz，1989；Handel 和 Glimcher，1999；Basso 和 Wurtz，2002），以及行为切换（Joseph，Boussaoud 和 Biguer，1985）的执行过程中。我们在这里描述的机器人模型突出了行为选择的适当维护中的一个关键问题，即导致选择给定行为的感知和激励条件通常不会在该行为的整个执行期间持续存在。这意味着，一种机制，如“赢者通吃”，忽略了最近的选择历史，并将电机系统的控制权分配给具有最高瞬时显著性的动作子系统，这将容易出现行为维护错误（errors of behavior maintenance），如正在进行的行为的过早中断，或在两个有相似的显著性的动作之间徘徊（快速切换）。这一问题的两个逻辑上可能的解决方案是：

(i) 胜出的竞争者发起某种形式的“相互排斥锁（mutual exclusion lock）”，以防止竞争对手在完成预期的电动机行为（Motor Act）之前访问电动机（然后释放锁）；

(ii) 正在进行的竞争偏向于当前获胜的竞争者，使其能够在适当延长的时间内，在低于（外在的）显著性水平（比发起竞争所需的水平）上保持对其竞争对手的优势。

虽然实时计算机操作系统中经常采用 (i) 的解决方案（Ganssle 和 Barr，2003），但我们的模型建议，基底神经节通过在行为期间为当前赢者提供额外的显著性支持来实现维护问题的 (ii) 的解决方案。更具体地说，我们提出了两种显著的递增机制。首先，我们提出基底神经节丘脑皮质环路例示一个正反馈电路，可以为获胜的子系统提供显著的促进（Redgrave 等人，1999a）。在扩展的基底神经节模型中，这种反馈会导致明显的滞后（见图 7），从而在机器人中产生行为持续性。第二，我们发现，当一个动作子系统能够产生一个精确定时的、内在产生的对其显著性的贡献（我们称之为忙碌信号）时，对动作模式的维护和选择的终止的精确控制可能是最好的。因此，我们假设，基底神经节活动在进行中的行为可能部分反映了类似的选择维持机制的运作。

认为基底神经节丘脑皮质环路在当前选择的通道中起作用以增加显著性，这与行为学中的一项重要研究一致，表明正向反馈机制在维持行为选择中的作用（Roeder，1975；Houston 和 Sumida，1985；McFarland，1989）。进一步的观点是，持续运动活动产生的信号对维持行为选择很重要，这可能解释了为什么大脑皮层（Cowan 和 Wilson，1994；Levesque，Charara，Gagnon，Parent 和 Deschenes，1996）和皮质下来源（Chevalier 和 Deniau，1984；Krout，Loewy，Westby 和 Redgrave，2001）对基底神经节的输入很重要。通常包括从纤维延伸到脑干和脊髓运动区的侧支。目前还没有对这些数据的一致解释，因此，我们的机器人模型中使用的忙碌信号提出了一个关于这些与电机相关的输入可能的功能作用的新假设。

注意，虽然我们的模型在每个通道的每个核中调用一个单一的漏积分器，但这是为了表示每个目标结构中的神经元群。因此，在考虑纹状体时，例如，我们的模型与不同的纹状体神经元亚群在执行运动行为期间在不同时间编码当前选择显著性的不同方面的可能性是一致的。换句话说，一些纹状体神经元的亚群可能与行为的开始（从而在行为开始前发生）有关，而其他亚群则与维持当前选择（从而在行为的表达过程中发生）有关。

有趣的是，对比我们模型中实现的 (ii) 的维护机制（并假设为基底神经节），这将允许更强大的竞争对手中断当前行为，(i) 的相互排斥锁，或互斥锁，是实时操作系统中的首选。有趣的是，使用互斥（mutex）可以导致一个称为优先级反转的资源分配问题，这在 1997 年火星探路者任务（Reeves，1998）遇到的重大计算机困难中最为著名。尽管存在可

以避免反转问题的解决方法，但这些方法在计算上是非常重要的，例如，涉及从一个任务到另一个任务的优先级继承（Sha、Rajkumar 和 Lehoczky, 1990）。因此，可以想象，较弱的资源锁定形式（如本文所述的维护机制）可以在人工调度系统中应用。

### 5.2.2 丘脑皮质反馈可能在时序动作模式中起作用。

基底神经节在行为时间选择中的作用与我们的一般假设一致，即基底神经节调节维持和适当终止动作，作为整体动作选择问题解决方案的一部分。我们的机器人模型使用 VL 丘脑的输出作为一个信号来调节每个动作子系统使用的内部时钟，该子系统生成一个固有模式的行为序列（即固定动作模式）。虽然用基底神经节的输出来控制这些固有的模式系统的方式并不打算是接近仿生的，但以下证据支持丘脑皮质反馈对行为时间很重要的观点。首先，皮质基底神经节丘脑皮质环路被特别假设为间隔时间可能的神经基质（Meck, 1983; Meck 和 Benson, 2002），并且是神经生物学上可信的计算模型的基础（Matell 和 Meck, 2000）。第二，多巴胺能药物已经被发现影响内部时钟的速度（Buhusi 和 Meck, 2002），进一步暗示基底神经节是大脑功能系统的一部分，用于时间估计。最后，已知帕金森病患者的重复性、固有序列（如带节奏的手指敲击）的时间安排受损，丘脑活动减弱是一个可能的原因（Marsden 和 Obeso, 1994; Elsinger 等人, 2003）。Meck 及其同事（Gibbon, Church, 和 Meck, 1984; Meck 和 Benson, 2002）阐述了许多关于基底神经节在间隔时间中作用的模型，我们认为其中最简单的模型与我们在模型中使用的机制有着惊人的相似性。具体来说，Gibbon 等人（1984）已经提出：

- (i) 基底神经节可以作为一种开关形式，根据对时间重要信息的检测可以打开或关闭；
- (ii) 当关闭时，该开关允许起搏器脉冲流到目标系统；
- (iii) 当时间重要信息结束时，该开关打开停止脉冲流。

同样地，在我们的模型中，目标 FAP 子系统的内部时钟被基底神经节解除抑制（关闭开关）启用（即开始测量从零开始的经过时间），并且当基底神经节通道的输出返回到阈值以上（打开开关）时被禁用（重置为零）。虽然机器人模型将受益于对间隔时间进行更具生物现实性的模拟，但这可能需要对振荡神经元群进行建模（Matell 和 Meck, 2000 年），而不是目前使用的更简单的速率编码单元。

如上所述，我们的机器人嵌入架构允许获胜的子系统在关键的行为通道中增强自己的显著性。如果这种额外的显著性输入是由内部时钟触发的，则可以合理地假设丘脑皮质反馈的丢失以及随后子统计时的中断也应中断模型中实现的忙碌信号。

### 5.2.3 输出到脑干的基底神经节可作为运动门控信号

计算研究表明，轴突末端在目标神经元树突树上的位置有助于确定抑制输入的非线性效应。更具体地说，细胞体上或靠近细胞体的终端被认为具有非线性、乘性分流效应，可缩放传入的兴奋信号（Blomfield, 1974; Koch、Poggio 和 Torre, 1983）。我们的机器人模型通过门控信号  $e$ （方程式 5）举例说明了基底输出核对电机模式发生器影响的分流抑制形式。GABA-氨基丁酸能的终端被视为在神经系统其他部位实施分流抑制（Ulrich, 2003），这种对 GABA-氨基丁酸能的基底神经节输出到脑干运动系统的作用的解释似乎与现有的（如果有限）电子显微镜发现的证据一致。例如，Tsumori 和 Yasui（1997）在大鼠上丘神经元的体细胞和近端树突上发现了 SNr 轴突末端，Shink、Sidibe 和 Smith（1997）发现松鼠猴足脑桥核的 GPi 输出主要与近端树突形成对称接触。由于动作选择假设认为基底神经节作用于进入最终运动路径，因此我们期望在未来对运动和前运动系统的 EP/SNr 输出的研究中找到进一步的分流抑制证据。



### 5.3 动物行为的行为学和神经行为学研究比较

#### 行为选择和行为排序

在我们的机器人任务中，基底神经节的嵌入模型展示了生成适当和目标导向行为的扩展序列的能力，在两个时间尺度上组织（柱面搜索、圆柱拾取等）和行为序列（避免、觅食）。这一结果与多种研究一致，在这些研究中，脊椎动物基底神经节在产生顺序行为中发挥了作用。例如，Kermadi 及其同事（Kermadi 和 Boussaoud, 1995; Kermadi 和 Joseph, 1995）发现猴子的尾状核活动与记忆的步态和手臂运动序列有关。Berridge 及其同事（Cromwell 和 Berridge, 1996; Aldridge 和 Berridge, 1998; Meyer Luehmann、Thompson、Berridge 和 Aldridge, 2002）证明纹状体对于啮齿类动物的特殊的梳理行为序列的表达是必要的，并记录了背外侧纹状体和腹内侧纹状体以及黑质的相关活动。行为动物的电生理学研究还发现了大鼠腹侧纹状体（Shibata、Mulder、Trullier 和 Wiener, 2001; Mulder、Tabuchi 和 Wiener, 2004）和背侧纹状体（Schmitzer Torbert 和 Redish, 2004）中编码迷宫穿越行为连续阶段的活动。最后，非哺乳动物脊椎动物中的行为序列，例如由 Ewert 和同事研究的蟾蜍捕食序列（Ewert 1987; Ewert, Buxbaum-Conradi, Dreisvogt 等人, 2001），也可能被基底神经节环路（与哺乳动物中发现的基本神经节环大体相同）所取代（Marin, Gonzalez, 和 Smeets, 1997; Redgrave 等人, 1999a）。

我们对机器人模型的实验提出了一些有趣的问题，这些关于基底神经节序列的神经行为数据，最好参考 Berridge 和同事对啮齿动物梳理的研究（Cromwell 和 Berridge, 1996; Aldridge 和 Berridge, 1998）。啮齿类动物的梳理模式通常出现在 Aldridge 和 Berridge（1998）描述为“四阶段句法链（four-phase syntactic chain）”的一个定型序列中。虽然梳理模式本身被认为是在基底神经节外产生的（见下文），但其行为表现已被证明是严重依赖于前背外侧纹状体的一小部分区域的完整性。在行为动物的电生理单细胞记录中（Aldridge 和 Berridge, 1998），背外侧纹状体神经元的活动在梳理序列的一个或两个阶段显示出明显的增加。此外，当在序列外进行类似的梳理运动时，大多数神经元没有反应（表明它们的活动与序列相关而不是运动相关）。最后，腹内侧纹状体中的一些神经元损伤不影响句法修饰链的产生，在修饰序列中也表现出增加的活动，但是这些增加比在背外侧纹状体中看到的要小。如图 9a 中模型“D1”单位的痕迹所示，机器人基底神经节模型的内在活动表明，纹状体神经元的活动可以同时发生在多个通道中，而那些显著性基于重叠特征集的通道中会发生相关变化。然而，它通常只是最活跃的通道，其电机输出为行为表达门控（参见图 9b 中的模拟 EP/SNr 输出）。这表明，腹内侧纹状体神经元（较弱）的活动可能是由句法修饰序列中的背景显著性线索部分启动的丢失行为编码。一个相关的假设也可以被表述为关于那些背外侧神经元的活动，这些神经元在梳理规则的多个阶段中激活。具体来说，一个神经元在梳理的一个阶段中被调到最大限度地激活，由于传入输入（与共享的背景线索相关）与另一个阶段编码的神经元部分重叠，它也可能在早期或后期表现出活动，这似乎是合理的。同样地，这种活动不会通过行为表达而产生，因为在基底神经节其他部位的竞争解决过程中，较弱通道的活动会消失。图 9a 中可以看到这方面的几个例子，例如，壁面搜索通道在前一个圆柱拾取过程中（以及壁面跟踪过程中的一些活动）显示出强烈的活跃（ $t=77-79s$ ），而在前一个壁面跟踪过程中（ $t=80-83s$ ），在所有情况下，壁面搜索通道显示出显著的活跃（ $t=80-83s$ ）。当 EP/SNr 输出值降低很小时，这些通道不足以扭曲获胜者的行为表达。注意，背外侧纹状体神经元的多相活动是由于重叠的特征集的假设不同于 Aldridge 和 Berridge（1998）提出的建议，即在多相期间激发的神经元编码为“序列作为高阶属性在链的持续时间内分布”（第 2784 页）。

虽然我们的机器人模型的活动与哺乳动物的实验研究显示出惊人的相似性，但在纯粹的

行为层面上,最明显的相似性可能是与蟾蜍的预备捕获序列的行为相似,这一序列已由 Ewert (1987) 仔细描述和分析。蟾蜍捕食是由一系列动作模式组成的,这些动作模式指向猎物(o)、接近(a)、固定(f)和捕捉(s),可通过涉及腹侧纹状体的抑制回路在蟾蜍大脑中实现(Ewert 等人, 2001)。总结了一系列关于捕捉猎物行为的广泛实验, Ewert 得出的结论是,决定后续反应的并不是先前的行为,而是正在进行的刺激情况(Ewert, 1987年,第340页)。我们的机器人模型也可以这么说,在这里,决定在任何给定时刻选择哪种行为的主要是感知/激励环境。此外,虽然标准的蟾蜍捕食行为由动作序列 o、a、f、s 描述,但观察到的行为往往以一种既显示灵活性又显示机会主义的方式偏离该模板。因此,如果猎物与蟾蜍之间的距离较短,则捕捉到的猎物由 o、f、s 组成;如果猎物突然出现在动物附近,则只会引出 o、s 或 f、s 或只是 s;如果猎物逃离蟾蜍,则会根据猎物的行为学变量(如 o、o、o、a、o、a、a、o、a、f、o、f、f、f、s)发生食欲追求反应(Ewert, 1987,第340页)。机器人产生的行为序列同样依赖于背景信息。因此,例如,如果机器人在第一次尝试时未能正确地抓住圆柱体,或者如果壁面搜索行为意外地将机器人放置在巢穴区域,那么标准的觅食序列 Cs、Cp、Ws、Wf、Cd 可能将显示为 Cs、Cp、Cs、Cp、Ws、Cd。

虽然机器人行为的两级结构暗示着分层组织,但从嵌入体系结构的设计来看,观察到的行为序列并不是控制的任何显式分层分解的结果。更确切地说,机器人的活动是由机器人感知和运动与其环境的相互作用、其内部(动机)状态以及嵌入的基底神经节模型的选择/切换特性所构成的一系列瞬间动作选择决策来组织的。根据 Ewert (1987) 的研究,我们得出结论,控制的等级组织对于顺序活动的出现并不重要。这并不是说我们会排除脊椎动物行为选择中的等级组织的可能性。事实上,我们认为(Prescott 等人, 1999; Redgrave 等人, 1999a)不同的基底神经节区域(边缘、联想和运动)之间的部分闭环、部分开环相互关系(Joel 和 Weiner, 1994)强烈建议控制的某种层次分解形式,人工制剂的研究证实了其益处(Prescott 等人, 1999; Bryson, 2000)。相反,机器人模型清楚地表明,任何有关动物行为的描述,其目的是显示层次分解,都必须证明,这里描述的紧急排序,不是一个可行的替代解释。

相关问题涉及基底神经节提供的行为选择的间隔尺寸。例如,一些研究者已经提出了基底神经节在更细间隔运动序列中的作用,而不是在相互竞争的行为选择中进行选择(参见 Mink, 1996 年的评论)。事实上,这一级别的动作选择相当于我们固定动作模式子系统(如圆柱拾取)中当前执行的动作顺序类型。然而,关于基底神经节参与运动序列的细节的建议可以与基底神经节作为动作选择装置的观点相一致,因为这样的任务可以在更短的时间尺度上被视为动作选择问题。同样,这与多个基底神经节域的证据和在基底神经节的不同区域采用类似开关电路来解决不同功能整合水平下的选择问题的一般假设是一致的(Redgrave 等人 1999)。然而,似乎有可能,在先天或实践良好的运动模式的情况下,运动的短时间尺度控制一般发生在基底神经节外。上述对大鼠梳理行为规则的研究旨在证明这一点,Cromwell 和 Berridge (1996 年)提出,基本梳理行为的组分运动排序的作用是由脑干中的模式发生器来满足的,而基底神经节的作用对于“序列顺序模式<...>的生成和在正常行为流中实现该模式并不是那么重要。”(第 3455 页)。对纹状体学习的研究表明,一个更有趣的可能性是与后天行为有关。Carelli、Wolske 和 West (1997 年)已经证明,当老鼠学习压杆任务时,纹状体神经元会在这种行为得到很好的练习后停止放电。这些研究人员从这一发现中得出的结论是,一旦动作变得自动化,学习特定运动反应所需的纹状体活动可能就不需要它的表现了。这一结果对许多解释是开放的,但是,一种可能性是,基底神经节的动作选择可能参与构建新的运动序列,实践之后,这些序列将作为更大的行为“模块”被选择(Graybiel, 1998)。

在行为过程中黑质网状部神经元的活动

最近对行为大鼠的一些研究发现，SNr 活性与运动行为的发作之间存在相关性（Gulley, Kuwajima, Mayhill 和 Rebec, 1999; Gulley, Kosobud 和 Rebec, 2002; Meyer-Luehmann 等人, 2002）。鉴于目前普遍认为基底神经节通过消除 EP/SNr 抑制选择性地阻断运动系统，一个特别有趣的发现是，大鼠的 SNr 细胞表现出与行为相关的放电增加，而那些表现出放电率降低的细胞则普遍不存在。例如，Gulley 等人（1999）比较了运动期间和安静休息期间相同神经元的 SNr 细胞电生理记录。在显示与运动有总体相关性的细胞中，79%的细胞显示出放电增加，而 21%的细胞显示出放电减少。在激活次数增加的细胞中，比率比基线休息期高出 38%。在进一步的研究中（Gulley 等人 2002），将条件强化任务期间的 SNr 活性与暴露于奖励相关装置之前的基线期进行比较，225 个 SNr 细胞中的 110 个（48%）的活化率增加了 200%或更多，而只有 17 个（8%）的活化率下降了 25%或更大。

上述发现与我们的 EP/SNr 模型在不同机器人活动模式下的活动水平一致。具体来说，我们发现 EP/SNr 输出的波动与通道显著性的变化相关，与无选择期（不活跃期）相比，部分或完全选择期内丢失通道的平均输出显著增加（34-82%）。似乎可以合理地预期，在行动中，由老鼠基底神经节输者调和的选择竞争将超过赢者（就像在机器人中，通常有 1 个赢者和 4 个输者一样）。因此，我们的模型与 Gulley 等人（1999, 2002）的数据一致。显示只有少数细胞在行为过程中减少了他们的活动（这里解释为赢的通道），而大多数细胞则增加了活动（输的通道）。我们对这些变化背后的机制的分析表明，显示放电增强的 SNr 神经元对 STN 的相关增加作出反应，这反过来是由于皮质 STN 传入的更大活性，特别是在编码来自获胜通道的丘脑皮质反馈的途径中。与不活动相关的活动期间（如 Gulley 等人, 1999），或奖励行为的供给突然增加时（如 Gulley 等人, 2002），预计这种增加最为明显。

#### 多个高显著性竞争对手的行为崩溃

当出现连续的高显著性选择竞赛序列时，机器人表现出两种可识别的行为崩溃模式。首先，它显示出胜利子系统的效率降低，以及部分失败子系统的活动。这导致最活跃行为的执行速度变慢和扭曲。第二，由两种行为之间的反复转换形成了一种行为陷阱。后一种模式是通过固定动作模式（圆柱拾取）的完全选择（但过早中断）产生的，并且取决于嵌入模型的两个特征：（i）固有的显著性信号用于保持持续选择；（ii）这些信号的定时和维护依赖于来自选择机制的反馈信号（在强、均匀匹配的显著性情况下被破坏）。上一节讨论了这些机制的可能神经相关因素（第 5.2 节）

虽然从纯行为选择的观点来看，模型在这些情况下的表现明显是次优的，但它与大量研究冲突情况下动物行为的研究结果显示有趣的相似性（Hinde, 1953, 1966; Fentres, 1973; Roeder, 1975）。例如，Hinde（1966）描述了对强烈行为冲突的行为学研究中观察到的一些可能的结果：

- (i) 除了一种反应之外的所有反应的抑制；
- (ii) 行为的不完全表达（通常是行为的准备阶段）；
- (iii) 行为之间的交替（或徘徊）；
- (iv) 双重行为（运动反应的混合物）；
- (v) 妥协行为（类似于矛盾心理，只是运动活动的模式与两种行为倾向兼容）；
- (vi) 自主反应（例如排便或排尿）；
- (vii) 无关行为（与当前动机环境无关的行为表现，如在战斗或逃跑冲突情况下的仪容仪表）。

在这些结果中，有几个与机器人在高显著性条件下的行为表现出明显的相似性。具体地说，在试验的早期阶段观察到的扭曲可以理解作为一种双重行为（iv），而随后的重复性行为转换

具有行为不完全表达 (ii) 和行为交替 (iii) 的元素。

更一般地说, 包含的基底神经节模型的行为是一致的, 心理学和行为学的广泛研究表明行为过程在激活的中间水平最有效 (Malmo, 1959; Berlyne, 1960; Bindra, 1969; Fentres, 1973), 这些发现也可以被视为表达 Yerkes-Dodson 法则 (Yerkes 和 Dodson, 1908), 该定律预测了觉醒和行为之间的倒 U 型关系。我们的模型与这一规律一致, 即当只存在低显著性输入时, 机器人很少或没有表现出行为表现, 对一系列中等显著性输入 (以及对于没有高显著性竞争对手的高显著性输入) 显示出有效的动作选择, 并且在多个高显著性系统之间的冲突的电路中显示出崩溃行为。因此, 机器人模型表明, 基底神经节是神经基质的一个重要组成部分, 调和觉醒对行为有效性的影响。

#### 5.4 基底神经的其他模型研究比较

关于基底神经节计算建模的文献已在其他地方进行了广泛的研究 (Houk, Davis 等人, 1995; Gillies 和 Arbuthnott, 2000; Prescott 等人, 2002; Gurney, Prescott 等人, 2004)。虽然行为选择是本文献中一个强有力的新兴主题, 但我们嵌入在机器人架构中的 Gurney 等人 (2001a, b) 模型在解释基底神经节固有电路时具有独特性, 它包含协同选择和控制路径。目前的研究表明, 当这些机制与基底神经节丘脑皮质环路相结合时, 在广泛的竞争性显著输入中提供有效的机器人动作选择。大量的模型还研究了基底神经节作为涉及运动控制或行为排序的更广泛回路的一部分的作用 (例如, Dominey 和 Arbib, 1992; Contrerasvidal 和 Stelmach, 1995; Houk 和 Wise, 1995; Dominey 和 Boussaoud, 1997; Beiser 和 Houk, 1998; Fukai, 1999; Taylor 和 Taylor, 2000; Frank, Loughry 和 O'Reilly, 2001; Bargad, Morris 和 Bergman, 2003; Brown, Bullock 和 Grossberg, 2004), 但是没有研究充分体现 (机器人) 的实现方法。在本篇文章中, 我们采用了一种不同的策略, 强调实施方式作为验证基底神经节功能假设的试验台, 同时作为一个“直觉泵”, 产生对神经生物学数据处理的新见解。例如, 要求提供完整的机器人行为序列, 以实现真正的目标, 集中我们的注意力, 在面对不同的动机和感官输入时, 保持持续的行为选择问题。解决机器人模型的这些问题, 然后促使我们重新考虑在持续行为期间纹状体和丘脑皮质活动的证据, 作为选择维护功能的可能神经基质。虽然一些作者已经将维持工作记忆模式中的基底神经节丘脑皮质环路 (例如 Beiser 和 Houk, 1998; Frank 等人, 2001) 指定为研究对象, 但我们建议更一般的假设, 即这些回路的运作是维持行为或工作记忆区域中的持续选择。此外, Frank 等人 (2001) 提出了基底神经节间歇放电与额叶皮质持续放电之间的分离, 前者起到门控作用, 后者起到维持作用; 我们已经注意到基底神经节在运动过程中的持续活动, 这些活动与新选择的开始无关。我们认为, 对这种活动的一个狭隘的解释, 与基底神经节选择的更广泛假设相一致, 是它有助于维持选择, 并随当前任务完成的紧迫性 (维护信号) 和竞争对手显著值的变化而变化 (由于基底神经节通道间相互作用的动力学)。

基底神经节与目标导向或激励性学习密切相关 (Kimura, 1995; Hollerman, Tremblay, 和 Schultz, 2000; Dayan 和 Balleine, 2002), 在这方面的一个关键发现是中脑基底神经节核中的多巴胺能神经元似乎与奖励性事件同时发生, 或在预期奖励性事件之前发生。

(Schultz, Apicella 和 Ljungberg, 1993; Schultz 等人, 1997)。Montague 及其同事 (Montague, Dayan 和 Sejnowski, 1996; Schultz 等人, 1997) 提出, 这些结构对纹状体神经元的传入可能提供类似于人工强化学习方法中使用的时差误差的训练信号, 而 Houk, Adams 和 Barto (1995) 是第一批提出类似于演员-评论 (actor-critic) 学习系统 (Sutton 和 Barto, 1998) 可能在基底神经节中运行的作者。这些提出观点中有各种各样的计算公式 (见 Schultz, 1997; Doya, Dayan, 和 Hasselmo, 2002; Montague, Hyman, 和 Cohen, 2004; Worgotter 和 Porr,

2005 评论), 包括描述的基于多巴胺的演员-评论模型的机器人演示 (Sporns 和 Alexander, 2002)。尽管如此, 在系统层面上, 对于这种学习如何在基底神经节的回路中实现, 还没有一个公认的理论 (Joel, Niv 和 Ruppin, 2002; Worgotter 和 Porr, 2005), 而这仍然是一个非常活跃的研究领域。我们的文章讨论了基底神经节的具体模型是否能够进行适当的动作选择, 而不管具体动作的显著性相关参数是如何确定的; 虽然在使用手动调节参数的地方, 该模型原则上可以扩展, 以便使用上述模型学习系统从经验中学习。

基底神经节的建模有一个普遍的趋势, 即使用更符合生物学实际但计算密集的基于“漏积分和发放 (LIF)”的神经回路模型或单个神经元的间隔模型 (Gurney, Prescott 等人, 2004)。并行计算集群的可用性开始使大规模电路的模拟成为可能, 包括这些更详细的单神经元模型, 速度允许实时控制机器人的行为。因此, 我们的机器人模型的未来版本将寻求纳入更大的生物学细节, 例如, 通过使用尖峰编码 (而不是速率编码) 神经元作为我们的基本模型元素 (Humphries, 2002), 通过合并额外的路径 (Gurney, Humphries 等人, 2004), 以及利用单个神经元的生物物理建模的想法 (例如 (Wood 等人, 2004 年)。一个特别有希望的途径可能是使用所谓的“简化”模型 (例如 (Rinzel 和 Ermentrout, 1998; Izhikevich, 2003), 它可以显示真实神经元中的许多兴奋性模式, 而无需跨多个离子电流实例化的 Hodgkin-Huxley 动力学的完整装置 (如生物物理模型所需)。另一个目标将是结合目标传感器输入和运动输出系统的现实模型, 例如, 通过对基底神经节在视觉控制等感觉运动任务中的作用的具体建模, 以便更直接地处理广泛的神经行为数据。

## 六、结论

我们描述了一个基底神经节和相关核的高级模型的机器人嵌入, 前提是这些神经回路在动作选择中起着关键作用。该模型面临的任务是在不同的动机和感官输入的情况下, 在五个可选的行为子系统之间进行选择, 并需要生成机器人行为的一致序列。结果表明, 模型基底神经节根据其相对显著性的动态变化, 在竞争的子系统之间进行有效的切换。因此, 该结构提供了一个存在性证明, 证明基底神经节在嵌入物理设备时能够作为有效的动作选择机制发挥作用。此外, 通过生成一个行为可以直接观察到的模型 (而不仅仅是可以解释为无实体模型), 我们能够与动物行为实验的结果进行一些有趣的比较, 尤其是关于:

- (i) 基底神经节在行为排序中的作用;
- (ii) 持续行为过程中基底神经节输入 (纹状体和 STN) 和输出 (SNr) 核神经元的活动;
- (iii) 动物在行为冲突情况下的行为。

## 感谢

作者希望感谢 Jonathan Chambers、Ric Wood、Michael Port、Andy Ham 和 Olivier Michel 的帮助, 以及英国工程和物理科学研究委员会 (EPSRC) 的支持, 批准号: GR/R95722。

## Appendices

### Appendix 1. Sensory systems

The first six peripheral sensors (1–6) are arranged in a semi-circle at the front of the robot, sensor 1 is furthest left, sensor 6 furthest right, with sensors 3 and 4 covering a narrow field of view directly ahead of the robot. Sensors 7 and 8 are directed towards the rear of the robot and are not used in the current model. The  $i^{\text{th}}$  peripheral sensor generates both an infra-red proximity reading,  $ir(i)$ , which is integer valued in the range 0–1023 with higher values indicating greater proximity to a nearby surface, and an ambient-light reading,  $amb(i)$ , in the range 0–450 with lower values indicate greater luminance. The optical gripper sensor,  $opt()$ , provides a binary signal, 1 if there is an object in the gripper, 0 otherwise. The arm position sensor,  $arm()$ , returns a value in the range 255 (lowered in front) to 152 (raised overhead). The following variables are computed from the current infra-red and ambient light readings for use in determining motor vector values and perceptual and motivational variables:

$$ir_{tot} = \sum_{i=1}^6 ir(i), \quad ir_{left} = \sum_{i=1}^3 ir(i), \quad ir_{right} = \sum_{i=4}^6 ir(i),$$

$$ir_{diff} = |ir_{left} - ir_{right}|, \quad side = \begin{cases} \text{left} : ir_{left} \geq ir_{right} \\ \text{right} : \text{otherwise} \end{cases}$$

$$detect(i) = \begin{cases} 1 : ir(i) > 30 \\ 0 : \text{otherwise} \end{cases}, \quad n_{touch} = \sum_{i=1}^6 detect(i),$$

$$lit(i) = \begin{cases} 1 : amb(i) < 100 \\ 0 : \text{otherwise} \end{cases}, \quad n_{lit} = \sum_{i=1}^6 lit(i).$$

### Appendix 2. Action sub-systems

Each action sub-system generates a motor vector

$$\mathbf{v} = [v_{lws-}, v_{lws+}, v_{rws-}, v_{rws+}, v_{vert}, v_{horiz}, v_{floor}, v_{open}, v_{close}]$$

where  $0 \leq v_j \leq 1 \forall v_j \in \mathbf{v}$ . The first four elements of  $\mathbf{v}$  correspond to the backward and forward components of the desired left ( $v_{lws-}, v_{lws+}$ ) and right ( $v_{rws-}, v_{rws+}$ ) wheel speeds, the next three to alternate positions for the arm ( $v_{vert}, v_{horiz}, v_{floor}$ ) and the last two to instructions to open or close the gripper ( $v_{open}, v_{close}$ ).

The following variable wheel speed values are computed based on current sensory input, for use by action sub-systems:

$$nws = \begin{cases} 0.07 : ir_{diff} < 30 \\ ir_{diff}/450 : ir_{diff} < 450 \\ 1.0 : \text{otherwise} \end{cases},$$

$$sws = 0.4 - 5.0 \times 10^{-4} |ir_{tot} - 1200|,$$

$$fws = 0.4 - 3.5 \times 10^{-4} |ir_{tot} - 1200|.$$

The condition-action mapping employed by each action sub-system to generate a motor vector and a busy signal value (where needed) at each time-step, are given in Table 3 in pseudo code. Note that for the two ‘fixed action patterns’—*cylinder-pickup* and *cylinder-deposit*—the condition element of the mapping indicates dependency on elapsed time (in seconds) since the start of the behavior as recorded by the relevant sub-system clock ( $t_{pick}$  or  $t_{dep}$ ), see section 3.3.2 for details.

### Appendix 3. Perceptual and motivational sub-systems

*Detecting a wall:* A wall is detected if the sum of infrared readings across all forward-facing sensors indicates a nearby surface *and either* the left-most(1) *or* right-most(6) sensor input suggest a nearby surface on that side *or* three or more of the forward sensors detect a surface at any distance (input>30). These conditions are required to allow a wall to be detected when it is approached at any angle, or when the robot is moving parallel to a wall. A wall can only be detected when the arm is raised above horizontal ( $arm() \leq 227$ ) since, otherwise, the gripper arm will be detected as a wall (note that the gripper arm may still be responded to as a ‘surface’ by action sub-systems).

$$P_{wall} = \begin{cases} +1 : (ir_{tot} > 800) \\ \quad \wedge \left( \begin{array}{l} ir(1) > 800 \vee ir(6) > 800 \\ \vee \sum_{i=1}^6 detect(i) \geq 3 \end{array} \right) \\ \quad \wedge (arm() \leq 227) \\ -1 : \text{otherwise} \end{cases}$$

*Detecting a nest:* A nest is detected if the ambient light reading on at least two of the forward peripheral sensors is below a threshold, hence:

$$P_{nest} = \begin{cases} +1 : n_{lit} \geq 2 \\ -1 : \text{otherwise} \end{cases}$$

*Detecting a cylinder:* A cylinder is detected when the two front-most sensors (3 and 4) detect a surface at very close proximity, and the two sensors either side (2 and 5) of the front detect no surface. A cylinder cannot be detected in the nest (to prevent perceptual aliasing).

$$P_{cyl} = \begin{cases} +1 : \left( \begin{array}{l} ir(2) < 10 \wedge ir(3) > 1000 \wedge ir(4) \\ > 1000 \wedge ir(5) < 10 \end{array} \right) \\ \quad \wedge (P_{nest} \neq +1) \\ -1 : \text{otherwise} \end{cases}$$

*Gripper status:* The gripper is considered to contain a cylinder when the optical sensor is triggered.

$$P_{grip} = \begin{cases} +1 : grip() = 1 \\ -1 : \text{otherwise} \end{cases}$$

The simulated motivation  $m_{fear}$  is initialized to 1.0 and decays toward a minimum value of 0.0 at a rate of  $-0.0007$  per step; the motivation  $m_{hung}$  is initialized to 0.2 and increases at a constant rate of  $+0.0015$  per step toward a maximum of 1.0, except on any time-step where a cylinder is deposited in a ‘nest’ area at which point it falls immediately to 0.0.

#### Appendix 4. The motor plant

The two wheel motors can be powered forwards and backwards and are controlled by integer-valued motor commands ranging from  $-20$  (maximum reverse) to  $+20$  (maximum advance). The robot gripper turret is powered by two motors, one to lift/lower the arm, the other to open/close the gripper. For the current model the useful range of operation for the arm motor varies from touching the floor (255), to overhead/vertical (152). The gripper motor is controlled by a binary command signal of 1 to close, 0 to open. To operate within these constraints the elements of the aggregate motor vector

$$\hat{v} = [\hat{v}_{lws+}, \hat{v}_{lws-}, \hat{v}_{rws+}, \hat{v}_{rws-}, \hat{v}_{up}, \hat{v}_{middle}, \hat{v}_{down}, \hat{v}_{open}, \hat{v}_{closed}]$$

are converted into instructions to the four robot motors as follows:

$$\text{wheels: } lws = 15(\hat{v}_{lws+} - \hat{v}_{lws-}), \quad rws = 15(\hat{v}_{rws+} - \hat{v}_{rws-}).$$

$$\text{arm: } \text{unless } \hat{v}_{vert} + \hat{v}_{horiz} + \hat{v}_{floor} = 0.0$$

$$\text{arm\_position} = \frac{152 \times \hat{v}_{vert} + 227 \times \hat{v}_{horiz} + 255 \times \hat{v}_{floor}}{\hat{v}_{vert} + \hat{v}_{horiz} + \hat{v}_{floor}}.$$

$$\text{gripper: } \text{unless } \hat{v}_{open} + \hat{v}_{closed} = 0.0$$

$$\text{gripper\_position} = \begin{cases} 1 \text{ (closed)} : \hat{v}_{closed} - \hat{v}_{open} > 0.0 \\ 0 \text{ (open)} : \text{otherwise} \end{cases}$$

where all fractional values are rounded to the nearest integer.

```

cylinder-seek:
ir_tot≤500          v_seek= (0,1.00, 0,1.00, 0,0,0, 0,0) // no nearby objects
// fast ahead
ir_tot>500 & n_lit ≥2 // strong light (nest)
sd= left:          v_seek= (0.27,0, 1.00,0, 0,0,0, 0,0) // backup, rotating right
sd= right:         v_seek= (1.00,0, 0.27,0, 0,0,0, 0,0) // backup, rotating left
500<ir_tot≤1025 & n_lit <2 v_seek= (0,1.00, 0,1.00, 0,0,0, 0,0) // nearby object
// fast ahead
1025<ir_tot≤2000 & n_lit <2 // possible cylinder
side= left:        v_seek= (0.20,0, 0,0.15, 0,0,0, 0,0) // rotate toward object
side= right:       v_seek= (0,0.15, 0.20,0, 0,0,0, 0,0) // rotate toward object
ir_tot>2000 & n_lit <2 // probable wall
side= left:        v_seek= (0,nws, nws,0, 0,0,0, 0,0) // rotate away (right)
side= right:       v_seek= (nws,0, 0,nws, 0,0,0, 0,0) // rotate away (left)

wall-seek:
ir_tot≤10          v_wall= (0,1.0, 0,1.0, 0,0,0, 0,0) // in 'free space',
// fast ahead
10<ir_tot≤500     v_wall= (0,0.50, 0,0.50, 0,0,0, 0,0) // some contact
// slow ahead
ir_tot>500        // near an obstacle
side= left:       v_wall= (0,nws, nws,0, 0,0,0, 0,0) // rotate right
side= right:      v_wall= (nws,0, 0,nws, 0,0,0, 0,0) // rotate left

wall-follow:
ir_tot≤600        // well away from wall
side= left:       v_foll= (0,0.20, 0,0.27, 0,0,0, 0,0) // veer-in sharp left
side= right:      v_foll= (0,0.27, 0,0.20, 0,0,0, 0,0) // veer-in sharp right
if 600<ir_tot<1200 // away from wall
side= left:       v_foll= (0,sws, 0,fws, 0,0,0, 0,0) // veer-in left
side= right:      v_foll= (0,fws, 0,sws, 0,0,0, 0,0) // veer-in right
1200≤ir_tot<2000 // quite near wall
side= left:       v_foll= (0,fws, 0,sws, 0,0,0, 0,0) // veer-out gently right
side= right:      v_foll= (0,sws, 0,fws, 0,0,0, 0,0) // veer-out gently left
ir_tot>2000      // very close to wall
side= left:       v_foll= (0,0.15, 0.15,0, 0,0,0, 0,0) // rotate right
side= right:      v_foll= (0.15,0, 0,0.15, 0,0,0, 0,0) // rotate left
n_touch≠1        b_foll=0
n_touch=1        b_foll=1

cylinder-pickup:
0<t_pick<0.3     b_pick= 0, v_pick= (0,0.10, 0,0.10, 0,0,0, 0,0) // slow approach
0.3≤t_pick<1.4   b_pick= 1, v_pick= (0.20,0, 0.20,0, 0,0,0, 1.0,0) // backup, open gripper
1.4≤t_pick<1.8   b_pick= 1, v_pick= (0,0, 0,0, 0,0,1.0, 0,0) // lower arm (floor)
1.8≤t_pick<2.8   b_pick= 1, v_pick= (0,0, 0,0, 0,0,0, 0,1.0) // close gripper
2.8≤t<3.5       b_pick= 1, v_pick= (0,0, 0,0, 1.0,0,0, 0,0) // raise arm (vertical)
3.6≤t_pick       b_pick= 0, v_pick= (0,0, 0,0, 0,0,0, 0,0), t= 0.0 // idle

cylinder-deposit:
0<t_dep<0.8     b_dep= 1, v_dep= (0,0, 0,0, 0,1.0,0, 0,0) // lower arm (horizontal)
0.8≤t_dep<1.6   b_dep= 1, v_dep= (0,0, 0,0, 0,0,0, 1.0,0) // release cylinder
1.6≤t_dep<2.4   b_dep= 1, v_dep= (0,0, 0,0, 1.0,0,0, 0,0) // raise arm (vertical)
2.4≤t_dep       b_dep= 0, v_dep= (0,0, 0,0, 0,0,0, 0,0), t= 0.0 // idle

```

**Table 3.** The condition-action mapping employed by each action sub-system to generate a motor vector and a busy signal value (where needed) at each time-step.



## References

- Akkal, D., Burbaud, P., Audin, J., & Bioulac, B. (1996). Responses of substantia nigra pars reticulata neurons to intrastriatal D1 and D2 dopaminergic agonist injections in the rat. *Neuroscience Letters*, 213(1), 66-70.
- Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences*, 12(10), 366-375.
- Aldridge, J. W., Anderson, R. J., & Murphy, J. T. (1980). The role of the basal ganglia in controlling a movement initiated by a visually presented cue. *Brain Research*, 192(1), 3-16.
- Aldridge, J. W., & Berridge, K. C. (1998). Coding of serial order by neostriatal neurons: a 'natural action' approach to movement sequence. *Journal of Neuroscience*, 18(7), 2777-2787.
- Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neurosciences*, 13(7), 266-271.
- Bar-Gad, J., Morris, G., & Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Progress in Neurobiology*, 71, 439-473.
- Barlow, G. W. (1977). Modal action patterns. In T. A. Sebeok (Ed.), *How Animals Communicate* (pp. 98-134). Bloomington, Indiana: Indiana University Press.
- Basso, M. A., & Wurtz, R. H. (2002). Neuronal activity in substantia nigra pars reticulata during target selection. *Journal of Neuroscience*, 22(5), 1883-1894.
- Beiser, D. G., & Houk, J. C. (1998). Model of cortical-basal ganglionic processing: Encoding the serial order of sensory events. *Journal of Neurophysiology*, 79(6), 3168-3188.
- Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill.
- Bindra, D. (1969). An interpretation of the "displacement" phenomenon. *British Journal of Psychology*, 50, 263-268.
- Blomfield, S. (1974). Arithmetical operations performed by nerve cells. *Brain Research*, 69(1), 115-124.
- Brown, J. W., Bullock, D., & Grossberg, S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks*, 17, 471-510.
- Bryson, J. (2000). Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental and Theoretical Artificial Intelligence*, 12(2), 165-189.
- Buhusi, C. V., & Meck, W. H. (2002). Differential effects of methamphetamine and haloperidol on the control of an internal clock. *Behavioral Neuroscience*, 116(2), 291-297.
- Carelli, R. M., Wolske, M., & West, M. O. (1997). Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. *Journal Of Neuroscience*, 17(5), 1804-1814.
- Casseday, J. H., & Covey, E. (1996). A neuroethological theory of the operation of the inferior colliculus. *Brain, Behavior, and Evolution*, 47(6), 311-336.

- Chevalier, G., & Deniau, J. M. (1984). Spatio-temporal organization of a branched tecto-spinal/tecto-diencephalic neuronal system. *Neuroscience*, *12*(2), 427-439.
- Colgan, P. (1989). *Animal Motivation*. London: Chapman & Hall.
- Contrerasvidal, J. L., & Stelmach, G. E. (1995). A neural model of basal ganglia-thalamocortical relations in normal and parkinsonian movement. *Biological Cybernetics*, *73*(5), 467-476.
- Cools, A. R. (1980). Role of the neostriatal dopaminergic activity in sequencing and selecting behavioural strategies: facilitation of processes involved in selecting the best strategy in a stressful situation. *Behavioural Brain Research*, *1*, 361-378.
- Cowan, R. L., & Wilson, C. J. (1994). Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *Journal of Neurophysiology*, *71*(1), 17-32.
- Cromwell, H. C., & Berridge, K. C. (1996). Implementation of action sequences by a neostriatal site - a lesion mapping study of grooming syntax. *Journal Of Neuroscience*, *16*(10), 3444-3458.
- Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*(2), 285-298.
- DeLong, M. R., Alexander, G. E., Georgopoulos, A. P., Crutcher, M. D., Mitchell, S. J., & Richardson, R. T. (1984). Role of basal ganglia in limb movements. *Human Neurobiology*, *2*(4), 235-244.
- Denny-Brown, D., & Yanagisawa, N. (1976). The role of the basal ganglia in the initiation of movement. In M. D. Yahr (Ed.), *The basal ganglia* (pp. 115-149). New York: Raven Press.
- Dominey, P. F., & Arbib, M. A. (1992). A Cortico-subcortical model for generation of spatially accurate sequential saccades. *Cerebral Cortex*, *2*, 153-175.
- Dominey, P. F., & Boussaoud, D. (1997). Encoding behavioral context in recurrent networks of the fronto-striatal system: a simulation study. *Cognitive Brain Research*, *6*, 53-65.
- Doya, K., Dayan, P., & Hasselmo, M. E. (2002). Special Issue: Computational Models of Neuromodulation. *Neural Networks*, *15*, 475-774.
- Elsinger, C. L., Rao, S. M., Zimelman, J. L., Reynolds, N. C., Blindauer, K. A., & Hoffmann, R. G. (2003). Neural basis for impaired time reproduction in Parkinson's disease: an fMRI study. *Journal of the International Neuropsychological Society*, *9*(7), 1088-1098.
- Ewert, J.-P. (1987). Neuroethology of releasing mechanisms: Prey-catching in toads. *Behavioral and Brain Sciences*, *10*, 337-389.
- Ewert, J. P., Buxbaum-Conradi, H., Dreisvogt, F., Glasgow, M., Merkel-Harff, C., Rottgen, A., et al. (2001). Neural modulation of visuomotor functions underlying prey-catching behaviour in anurans: perception, attention, motor performance, learning. *Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology*, *128*(3), 417-461.
- Fentress, J. C. (1973). Specific and nonspecific factors in the causation of behavior. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in Ethology*. New York: Plenum Press.
- Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, and Behavioral Neuroscience*, *1*(2), 137-160.
- Fukai, T. (1999). Sequence generation in arbitrary temporal patterns from theta- nested gamma oscillations: a model of the basal ganglia- thalamo-cortical loops. *Neural Networks*, *12*(7-8), 975-987.
- Ganssle, J., & Barr, M. (2003). *Embedded Systems Dictionary*. Lawrence, KS: CMP Books.
- Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., et al. (1990). D1 and D2 Dopamine Receptor Regulated Gene-Expression of Striatonigral and Striatopallidal Neurons. *Science*, *250*(4986), 1429-1432.
- Gerfen, C. R., & Wilson, C. J. (1996). The basal ganglia. In L. W. Swanson, A. Bjorklund & T. Hokfelt (Eds.), *Handbook of chemical neuroanatomy, Vol 12: Integrated systems of the CNS, Part III*. (pp. 371-468). Amsterdam: Elsevier.
- Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, *423*, 52-77.
- Gillies, A., & Arbuthnott, G. (2000). Computational models of the basal ganglia. *Movement Disorders*, *15*(5), 762-770.
- Girard, B., Cuzin, V., Guillot, A., Gurney, K. N., & Prescott, T. J. (2003). A basal ganglia inspired model of action selection evaluated in a robotic survival task. *Journal of Integrative Neuroscience*, *2*(2), 179-200.
- Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity - a hypothesis for the etiology of schizophrenia. *Neuroscience*, *41*(1), 1-24.
- Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*, *70*, 119-136.
- Gulley, J. M., Kosobud, A. E., & Rebec, G. V. (2002). Behavior-related modulation of substantia nigra pars reticulata neurons in rats performing a conditioned reinforcement task. *Neuroscience*, *111*(2), 337-349.

- Gulley, J. M., Kuwajima, M., Mayhill, E., & Rebec, G. V. (1999). Behavior-related changes in the activity of substantia nigra pars reticulata neurons in freely moving rats. *Brain Research*, 845(1), 68-76.
- Gurney, K., Prescott, T. J., & Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*, 84(6), 401-410.
- Gurney, K., Prescott, T. J., & Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics*, 84(6), 411-423.
- Gurney, K., Prescott, T. J., Wickens, J., & Redgrave, P. (2004). Computational models of the basal ganglia: from membranes to robots. *Trends in Neurosciences*, 27, 453-459.
- Gurney, K. N., Humphries, M., Wood, R., Prescott, T. J., & Redgrave, P. (2004). Testing computational hypotheses of brain systems function: a case study with the basal ganglia. *Network*, 15(4), 263-290.
- Handel, A., & Glimcher, P. W. (1999). Quantitative analysis of substantia nigra pars reticulata activity during a visually guided saccade task. *Journal of Neurophysiology*, 82(6), 3458-3475.
- Harsing, L. G., & Zigmond, M. J. (1997). Influence of dopamine on GABA release in striatum: Evidence for D-1-D-2 interactions and non-synaptic influences. *Neuroscience*, 77(2), 419-429.
- Hikosaka, O. (1994). Role of basal ganglia in control of innate movements, learned behavior and cognition—a hypothesis. In G. Percheron, J. S. McKenzie & J. Feger (Eds.), *The Basal Ganglia IV*. New York: Plenum.
- Hikosaka, O., & Wurtz, R. H. (1989). The basal ganglia. In E. H. Wurtz & M. E. Goldberg (Eds.), *The Neurobiology of Saccadic Eye Movements* (pp. 257-281). Amsterdam: Elsevier.
- Hinde, R. A. (1953). The Conflict between Drives in the Courtship and Copulation of the Chaffinch. *Behaviour*, 5, 1-31.
- Hinde, R. A. (1966). *Animal Behaviour: a Synthesis of Ethology and Comparative Psychology*. London: McGraw-Hill.
- Hollerman, J. R., Tremblay, L., & Schultz, W. (2000). Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior. *Progress in Brain Research*, 126, 193-215.
- Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis & D. G. Beiser (Eds.), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press.
- Houk, J. C., Davis, J. L., & Beiser, D. G. (1995). *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press.
- Houk, J. C., & Wise, S. P. (1995). Distributed Modular Architectures Linking Basal Ganglia, Cerebellum, and Cerebral-Cortex - Their Role in Planning and Controlling Action. *Cerebral Cortex*, 5(2), 95-110.
- Houston, A., & Sumida, B. (1985). A positive feedback model for switching between 2 activities. *Animal Behaviour*, 33(FEB), 315-325.
- Hoyle, G. (1984). The scope of neuroethology. *The Behavioural and Brain Sciences*, 7(3), 367-412.
- Humphries, M. D. (2002). *The basal ganglia and action: a computational study at multiple levels of abstraction*. Unpublished PhD, University of Sheffield, Sheffield.
- Humphries, M. D., & Gurney, K. N. (2001). A pulsed neural network model of bursting in the basal ganglia. *Neural Networks*, 14, 845-863.
- Humphries, M. D., & Gurney, K. N. (2002). The role of intra-thalamic and thalamocortical circuits in action selection. *Network: Computation in Neural Systems*, 13, 131-156.
- Humphries, M. D., Prescott, T. J., & Gurney, K. N. (2003). The interaction of recurrent axon collateral networks in the basal ganglia. In E. A. O. Kaynak, E. Oja & L. Xu (Ed.), *Artificial Neural Networks and Neural Information Processing - ICANN/ICONIP 2003 (Lecture Notes in Computer Science 2714)* (pp. 797-804). Berlin: Springer-Verlag.
- Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6), 1569-1572.
- Jaeger, D., Gilman, S., & Aldridge, J. W. (1995). Neuronal activity in the striatum and pallidum of primates related to the execution of externally cued reaching movements. *Brain Research*, 694, 111-127.
- Joel, D., Niv, Y., & Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Networks*, 15(4-6), 535-547.
- Joel, D., & Weiner, I. (1994). The organization of the basal ganglia-thalamocortical circuits - open interconnected rather than closed segregated. *Neuroscience*, 63(2), 363-379.
- Joseph, J. P., Boussaoud, D., & Biguer, B. (1985). Activity of neurons in the cat substantia nigra pars reticulata during drinking. *Experimental Brain Research*, 60(2), 375-379.
- Kermadi, I., & Boussaoud, D. (1995). Role of the primate striatum in attention and sensorimotor processes:

- comparison with premotor cortex. *Neuroreport*, 6(8), 1177-1181.
- Kermadi, I., & Joseph, J. P. (1995). Activity in the caudate nucleus of monkey during spatial sequencing. *Journal of Neurophysiology*, 74(3), 911-933.
- Kimura, M. (1995). Role of basal ganglia in behavioral learning. *Neurosci Res*, 22(4), 353-358.
- Koch, C., Poggio, T., & Torre, V. (1983). Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proceedings of the National Academy of Science USA*, 80(9), 2799-2802.
- Krout, K. E., Loewy, A. D., Westby, G. W., & Redgrave, P. (2001). Superior colliculus projections to midline and intralaminar thalamic nuclei of the rat. *Journal of Comparative Neurology*, 431(2), 198-216.
- Latimer, C. (1995). Computer modelling of cognitive processes. *Noetica: Open Forum*, 1(1), <http://psy.uq.edu.au/CogPsych/Noetica/>.
- Lehner, P. N. (1996). *Handbook of Ethological Methods* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Levesque, M., Charara, A., Gagnon, S., Parent, A., & Deschenes, M. (1996). Corticostriatal projections from layer V cells in rat are collaterals of long-range corticofugal axons. *Brain Research*, 709(2), 311-315.
- Lorenz, K. (1935). Der Kumpan in der Umwelt des Vogels. *Journal of Ornithology*, 83, 137-213, 289-413.
- Maes, P. (1995). Modelling adaptive autonomous agents. In C. G. Langton (Ed.), *Artificial Life: An Overview*. Cambridge, MA: MIT Press.
- Malmo, R. (1959). Activation: a neuropsychological dimension. *Psychological Review*, 66, 367-386.
- Marin, O., Gonzalez, A., & Smeets, W. (1997). Anatomical substrate of amphibian basal ganglia involvement in visuomotor behaviour. *European Journal of Neuroscience*, 9(10), 2100-2109.
- Marsden, C. D., & Obeso, J. A. (1994). The function of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. *Brain*, 117, 877-897.
- Massaro, D. W. (1988). Some criticisms of connectionist models of human performance. *Journal of Memory and Language*, 27, 213-234.
- Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays*, 22(1), 94-103.
- McFarland, D. (1989). *Problems of Animal Behaviour*. Harlow, UK: Longman.
- McFarland, D., & Bosser, T. (1993). *Intelligent Behaviour in Animals and Robots*. Cambridge, MA: MIT Press.
- Meck, W. H. (1983). Selective adjustment of the speed of internal clock and memory processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(2), 171-201.
- Meck, W. H., & Benson, A. M. (2002). Dissecting the brain's internal clock: how frontal-striatal circuitry keeps time and shifts attention. *Brain and Cognition*, 48(1), 195-211.
- Meyer-Luehmann, M., Thompson, J. F., Berridge, K. C., & Aldridge, J. W. (2002). Substantia nigra pars reticulata neurons code initiation of a serial pattern: implications for natural action sequences and sequential disorders. *European Journal of Neuroscience*, 16(8), 1599-1608.
- Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress In Neurobiology*, 50(4), 381-425.
- Mink, J. W., & Thach, W. T. (1991). Basal ganglia motor control. II. Late pallidal timing relative to movement onset and inconsistent pallidal coding of movement parameters. *Journal of Neurophysiology*, 65(2), 301-329.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal Of Neuroscience*, 16(5), 1936-1947.
- Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, 431(7010), 760-767.
- Montes Gonzalez, F., Prescott, T. J., Gurney, K., Humphries, M., & Redgrave, P. (2000). An embodied model of action selection mechanisms in the vertebrate brain. In J. A. Meyer (Ed.), *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior* (pp. 157-166). Cambridge, MA: MIT Press.
- Mulder, A. B., Tabuchi, E., & Wiener, S. I. (2004). Neurons in hippocampal afferent zones of rat striatum parse routes into multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7), 1923-1932.
- Parent, A., & Hazrati, L. N. (1995). Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Research Reviews*, 20(1), 128-154.
- Pinker, S., & Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.
- Prescott, T. J., Gurney, K., & Redgrave, P. (2002). Basal Ganglia. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (2nd ed., pp. 147-151). Cambridge, MA: MIT Press.

- Prescott, T. J., Redgrave, P., & Gurney, K. N. (1999). Layered control architectures in robots and vertebrates. *Adaptive Behavior*, 7(1), 99-127.
- Redgrave, P., Prescott, T., & Gurney, K. N. (1999a). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, 89, 1009-1023.
- Redgrave, P., Prescott, T. J., & Gurney, K. (1999b). Is the short latency dopamine burst too short to signal reward error? *Trends in Neurosciences*, 22, 146-150.
- Reeves, G. (1998). *Re: What really happened on Mars*. Retrieved 54, 19, from <http://catless.ncl.ac.uk/Risks/19.54.html>
- Rinzel, J., & Ermentrout, E. (1998). Analysis of neural excitability and oscillations. In C. Koch & I. Segev (Eds.), *Methods in Neuronal Modelling*. Cambridge, MA: MIT Press.
- Robbins, T. W., & Brown, V. J. (1990). The role of the striatum in the mental chronometry of action: a theoretical review. *Reviews in the Neurosciences*, 2, 181-213.
- Roeder, K. D. (1975). Feedback, spontaneous activity, and behaviour. In G. Baerends, C. Beer & A. Manning (Eds.), *Function and Evolution in Behaviour*. Oxford: Clarendon Press.
- Ryan, L. J., & Clark, K. B. (1991). The role of the subthalamic nucleus in the response of globus- pallidus neurons to stimulation of the prelimbic and agranular frontal cortices in rats. *Experimental Brain Research*, 86(3), 641-651.
- Schmitzer-Torbert, N., & Redish, A. D. (2004). Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. *Journal of Neurophysiology*, 91(5), 2259-2272.
- Schultz, W. (1986). Activity of pars reticulata neurons of monkey substantia nigra in relation to motor, sensory, and complex events. *J Neurophysiol*, 55(4), 660-677.
- Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. *Current Opinion in Neurobiology*, 7, 191-197.
- Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal Of Neuroscience*, 13(3), 900-913.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate for prediction and reward. *Science*, 275, 1593-1599.
- Sha, L., Rajkumar, R., & Lehoczky, J. P. (1990). Priority inheritance protocols: An approach to real-time synchronization. *IEEE Transactions on Computers*, 39, 1175-1185.
- Shibata, R., Mulder, A. B., Trullier, O., & Wiener, S. I. (2001). Position sensitivity in phasically discharging nucleus accumbens neurons of rats alternating between tasks requiring complementary types of spatial cues. *Neuroscience*, 108(3), 391-411.
- Shink, E., Sidibe, M., & Smith, Y. (1997). Efferent connections of the internal globus pallidus in the squirrel monkey: II. Topography and synaptic organization of pallidal efferents to the pedunculopontine nucleus. *Journal of Comparative Neurology*, 382(3), 348-363.
- Smith, Y., Bevan, M. D., Shink, E., & Bolam, J. P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience*, 86, 353-387.
- Sporns, O., & Alexander, W. H. (2002). Neuromodulation and plasticity in an autonomous robot. *Neural Networks*, 15(4-6), 761-774.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.
- Taylor, J. G., & Taylor, N. R. (2000). Analysis of recurrent cortico-basal ganglia-thalamic loops for working memory. *Biological Cybernetics*, 82(5), 415-432.
- Tinbergen, N. (1951). *The Study of Instinct*. Oxford: Clarendon Press.
- Toates, F. (1998). The interaction of cognitive and stimulus-response processes in the control of behaviour. *Neuroscience and Biobehavioral Reviews*, 22(1), 59-83.
- Tsumori, T., & Yasui, Y. (1997). Organization of the nigro-tecto-bulbar pathway to the parvocellular reticular formation: a light- and electron-microscopic study in the rat. *Experimental Brain Research*, 116(2), 341-350.
- Ulrich, D. (2003). Differential arithmetic of shunting inhibition for voltage and spike rate in neocortical pyramidal cells. *European Journal of Neuroscience*, 18(8), 2159-2165.
- Walters, J. R., Ruskin, D. N., Allers, K. A., & Bergstrom, D. A. (2000). Pre- and postsynaptic aspects of dopamine-mediated transmission. *Trends in Neurosciences*, 23(10), S41-S47.
- Wickens, J. (1997). Basal ganglia: structure and computations. *Network-Computation in Neural Systems*, 8(4), R77-R109.
- Wilson, C. J., & Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *Journal of Neuroscience*, 16(7), 2397-2410.
- Wood, R., Gurney, K. N., & Wilson, C. J. (2004). A novel parameter optimisation technique for compartmental models applied to a model of a striatal

medium spiny neuron. *Neurocomputing*, 58-60, 1109-1116.

Worgotter, F., & Porr, B. (2005). Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Computation*, 17(2), 245-319.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.

Zink, C. F., Pagnoni, G., Martin, M. E., Dhamala, M., & Berns, G. S. (2003). Human striatal response to salient nonrewarding stimuli. *Journal of Neuroscience*, 23(22), 8092-8097.