



第4章 思维运作



第11讲 意识整合

意识活动指的是与感知、认知和记忆等有意
识心理活动相伴随的一种脑活动，

心智活动的其他能力，包括感知、理解、推断、学习、
预想、创造、情感、行为等都伴随着意识活动。

通过这样的意识活动，我们不仅因此能够自觉
到统一心理活动的具体表现过程，而且能够具有
创建世界模型并用以模拟未来的能力。



1、机器意识研究概况

早期有关机器意识的研究比较初步，研究工作较少得到学术界的认同，甚至前些年提到“机器意识”还有不合时宜的顾虑。但随着最近十年发展，**机器意识**现在已经成为广泛关注的合法研究领域。在过去十年，人们再次高度关注意识问题并有众多的哲学家、心理学家和神经科学家在此领域开展工作。

人们也开始使用**计算模型**来检验**意识理论**，并已有一些推测认为这最终可能导致实际**拥有意识**这种奇妙状态的、更为智能的机器。这类研究逐渐被称为“**机器意识**”的研究。



1、机器意识研究概况

从研究策略来看，机器意识的研究主要分为**算法构造策略**（A）与**仿脑构造策略**（B）两种途径。

所谓**算法构造策略**，就是不考虑人类脑机制的借鉴，**纯粹采用机器算法策略**来进行机器意识的研究（Algorithm）；

所谓**仿脑构造策略**，就是充分**借鉴人脑意识**的发生机制（**Brain-Inspiration**），并利用一切可利用的**生物物理机制**来进行机器意识的研究。

从具体的实现方法上，则主要又可以分为如下**三种具体方法**。



1、机器意识研究概况

(1) 规则计算方法 (R)：规则计算方法与认知科学中的符号逻辑范式相对应，其实质性观点简单说就是：心智内部具有对世界的“表征”，并可以根据“规则”来操作或操纵这些表征。这一范式的基本方法是探究智能或认知是怎样经由理性符号表征的操作而成为可能的，并认为这就是认知的本质。

(2) 人工神经网络的方法 (N)：人工神经网络是由具有各种相互联系的神经元组成的集合，每个单元具有极为简化的人脑神经元的特性。人工神经网络方法与认知科学中的联结主义范式相对应，其核心概念是“并行分布处理”的概念，即认知或智能是从大量单一处理单元的相互作用中产生。



1、机器意识研究概况

(3) 量子计算方法 (Q)：机器意识实现的另一种途径是采用量子物理学方法来进行意识的建模研究。在这方面，Amoveso等人首先开启了深入研究的先河。到了21世纪，Woolf给出了用于视觉意识的量子方法。作为这种机器意识途径的进一步发展，Gao提出了意识的一种量子理论，研究了量子塌缩与意识之间的关系，假定量子塌缩是一种客观的动态过程，可以描述意识的产生过程。

在上述三种方法中，值得注意的是最后的量子计算方法。



1、机器意识研究概况

量子理论描述意识产生机制的有效性并不是说物质的量子活动可以直接产生意识，而是强调意识产生机制与量子机制具有跨越尺度的自相似性。

正是因为这样，近年来，意识的量子模型有快速发展的动向：比如：

从类量子计算的角度提出意识涌现理论、
给出的一种同步振荡神经网络，
以及提出量子力学的形式化代数性质来描述现象意识等，

都为意识的量子计算建模建立了新的基础，值得我们特别关注。



1、机器意识研究概况

从研究的**目标**来看，迄今为止的研究大致可以分为：

面向**感知**意识实现的（MC-P, Perception）、
面向**具体特性**意识实现的（MC-C, Characteristics）、
面向**机制**意识实现的（MC-A, Architecture）、
面向**自我**意识实现的（MC-S, Self）、
以及面向**感受**意识实现的（MC-Q, Qualia）

等五个方面。



1、机器意识研究概况

(1) 具有意识表现的机器感知模型（MC-P）：

早期的机器意识研究主要是在机器感知处理中引入意识机制，因此是属于MC-P类型的机器意识研究。

比如Aleksander有关视觉觉知模型，及其Taylor提出的三段阶意识的神经网络模型，都是这方面的研究工作。不过，总体上讲，由于此类研究工作的特点是帮助机器更加有效地进行感知，采取的策略往往是某种更高阶的监督或注意机制，因此，还谈不上真正意义上的伴随性意识的实现。



1、机器意识研究概况

(2) 具有认知特性及其行为表现的机器意识 (MC-C) :

迄今为止，已经研制了许多介乎于意识与认知能力之间，诸如语言、想象、情感等的、基于神经连接主义方法实现的系统。

应该指出，MC-C类型的机器意识研究是一种**强势模式**，从简单的算法程序到基于模拟神经元的系统等都有研究，这种类型的**实现方法比较广泛**，所**强调**的是**有意识的认知能力**，包括语言、想象、情感等**内部模型**以及适应环境的**外部表现**行为等。



1、机器意识研究概况

(3) 具有人类意识相对应机制的意识机器 (MC-A) :

对意识产生机制及其机器模拟实现的研究，无疑对于机器意识研究而言是最具有吸引的研究方面，因此也吸引了众多研究人员开展有关这方面的研究工作。

此类研究的一个关键前提就是首先必须对人类的意识产生机制有全面的了解，才能基于这样的了解基础上，来实现机器的意识机制。因此，很多这类研究通常采用仿脑构造策略来进行。



1、机器意识研究概况

（4）拥有自我反思的机器意识（MC-S）：

在机器意识实现研究中，还有一个重要方面就是有关机器自我意识的研究。这一方面，比较系统的研究是Samsonovich所开展的研究工作，主要是基于图式建立了一个自我意识模型并进行系统实现。

在该机器系统中，其认知体系结构是围绕着对应于系统自我“意识”的图式（schemas）展开的，据称这些图式能够处理的数据项不但包括语义知识、行动要素，而且还包括感受体验（qualia）。



1、机器意识研究概况

(5) 拥有奇妙意识体验状态的机器（MC-Q）：

前四种类型的机器意识研究都是相对无争议的，由于这些方法都是针对与意识相关联的现象进行建模而并没有声称具有真正的意识体验状态。

但机器意识这第五个方面的研究却涉及更多的哲学问题，因为其涉及到拥有真正意识体验的机器——这种意义上的机器已经不仅仅是意识研究的工具，而是实际具有意识体验的机器。



1、机器意识研究概况

当然，不管采用什么实现策略、使用什么计算方法、以及面向什么内容方面，机器意识的研究目的都是要围绕着让机器拥有某种程度的意识能力展开的。

从目前机器意识已经开展研究工作的实现目标来看，机器意识的研究主要围绕着功能意识、自我意识和现象意识三个方面展开的。

不过由于涉及到主观体验的不可还原性，真正具有实际意义的主要还是在功能意识的机器建模方面。



2、全局工作空间理论

在机器意识研究中，最具有代表性的、也最具有影响力的研究工作，就是以**意识的全局工作空间理论**为指导所开展的一系列研究工作。

全局工作空间理论（Global workspace theory）是由美国加利福尼亚大学圣地亚哥分校神经科学研究所研究员**巴尔斯**（Bernard J. Baars）在1988提出的有影响的意识解释理论。在该理论的指导下，由巴尔斯等人（Baars、Stan Franklin和Uma Ramamurthy）组成的研究团队开展长达20多年的机器意识研究工作，最终开发完成了**LIDA**认知系统。



2、全局工作空间理论

LIDA(Learning Intelligent Distribution Agent)，是在该研究团队等人早期开发的IDA(Intelligent Distribution Agent)基础上发展起来的。

主要依据**巴尔斯全局工作空间理论**，采用**神经网络与符号规则**混合计算方法，通过**为每个软件主体建立内部认知模型来实现**诸多方面的**意识认知能力**，如注意、情感与想象等。

该系统可以区分**有无意识状态**、有效**运用有意识状态**，以及具备一定的**内省反思能力**等，并得到一些应用和扩展。



2、全局工作空间理论

首先，在巴尔斯提出的全局工作空间理论中，整个理论架构是由三个部分构成的，即：

背景（contexts）、
全局工作空间（global work-space）
和专门处理器（specialized processors）。



2、全局工作空间理论

背景
幕后导演 注意控制 局域背景

如图所示。图中：

全局工作空间是核心，作为中枢信息交换的中心，也是每时每刻意识内容的处所。

在各自专长领域有效进行信息处理，众多无意识的专门处理器，都与全局工作空间建立联系，并通过相互之间的协作与竞争来占据全局工作空间，从而可以一时成为意识的内容。

全局工作空间理论架构

专门处理器（观众与演职人员）

外部
感知

内部
感知

思维
内容

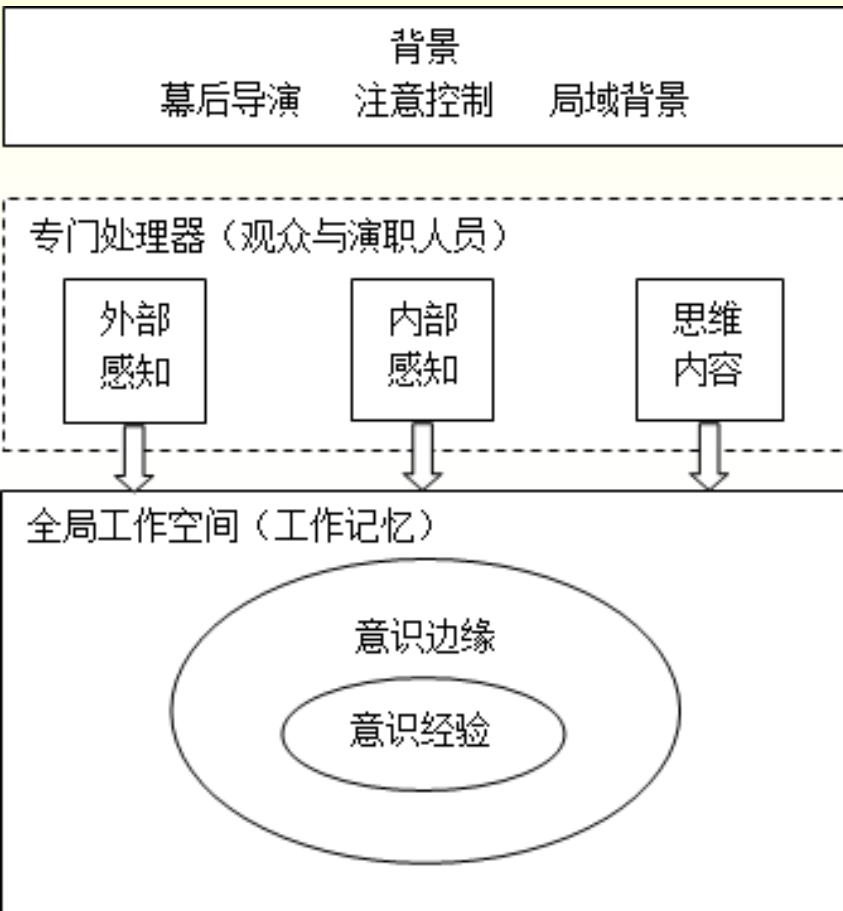
全局工作空间（工作记忆）

意识边缘

意识经验



2、全局工作空间理论

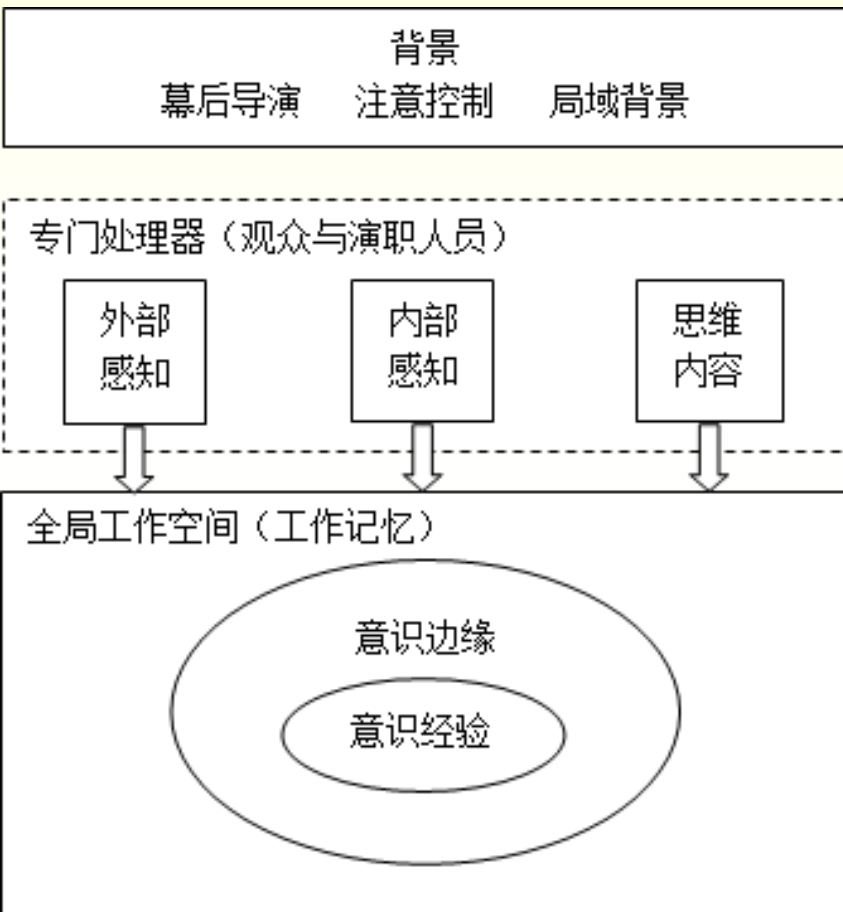


在全局工作空间理论模型中，所谓背景主要模仿的是人类无意识心理活动，看作是一种可以引起意识体验的无意识心像网络。

比如像知觉表象背景、思维概念背景、内在动机背景、甚至文化观念背景，等等都是不同的局域背景。由于所有这些背景都是潜在地可以相互影响、相互作用和相互转化，所以可以看作是神经系统已经建立好联系的一张无意识的网络，随时都可以潜在指导或注意控制地激发或抑制有意识的心理活动。



2、全局工作空间理论



专门处理器是指各种专门的信息加工过程，

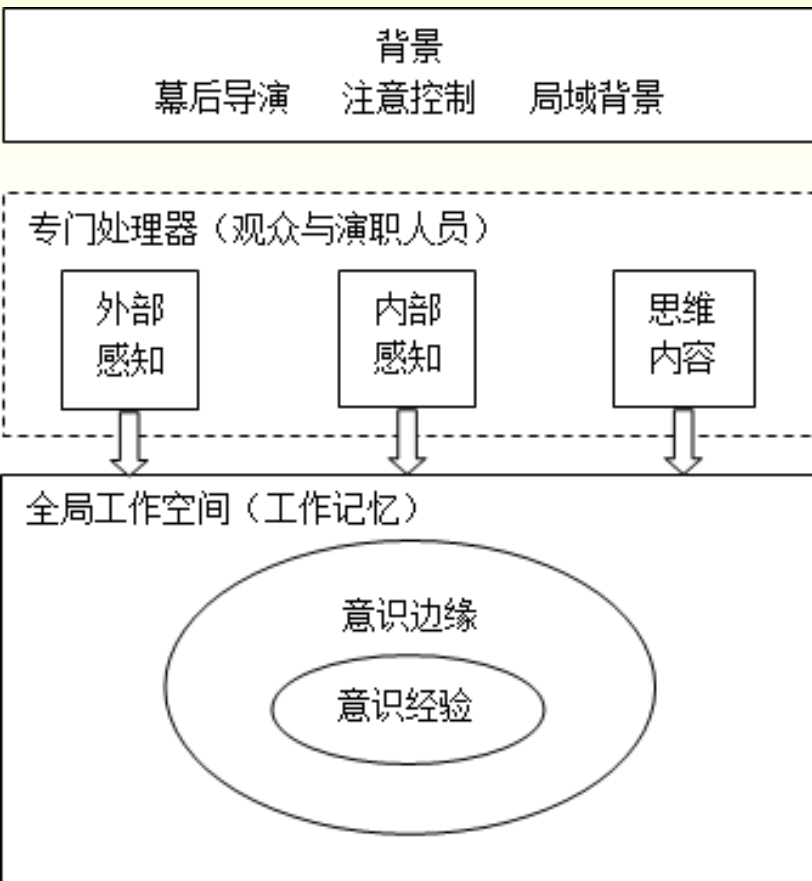
1、包括视觉、听觉、嗅觉、味觉、触觉、平衡觉等各种感官获取外部环境的**外部感知处理器**，

2、也包括视觉想象、内部语言、想象感受和梦幻等各种**内部感觉信息处理的内部感知处理器**。

3、以及像思想抽象、心像内容、想象思维、语言思维、直觉思维等**思维内容的信息处理器**。



2、全局工作空间理论



如果说专门处理器是意识加工信息的来源，背景是提供了意识信息加工的温床，那么全局工作空间就是意识内容产生的处所。

因此，全局工作空间就是一个意识加工形成的中心舞台。在全局工作空间这个舞台中，不仅可以协调集成来自不同专门处理器的信息输入，而且可以潜在受到无意识背景的信息激发或牵制，从而能够形成每时每刻的意识内容，并掌控着整体信息加工的过程。



2、全局工作空间理论

为了更好地解释全局工作空间理论对意识活动过程的刻画，巴尔斯又提出了意识剧院类比模型，进一步来阐述意识、无意识、注意和工作记忆等的相互联系和区分。

在剧院模型中，巴尔斯把工作记忆类比到剧院的舞台，把注意作用类比为一只聚光灯，而把聚光灯照亮的舞台部分类比为意识呈现的内容，等等。



2、全局工作空间理论

从上述剧院类比模型中可知，**某一时刻的意识内容**是**不同专门处理器在幕后背景作用下相互竞争与协调的结果**。就这一点而言，巴尔斯的剧院模型，不仅澄清了意识研究中一些关键概念，而且也为机器意识系统的构建提供了重要的原型模型。我们期待能够从中能够部分解决意识的计算建模问题，特别是功能意识的机器实现，并具体应用到各类智能系统之中。

End 2



3、机器意识困难所在

实际上，现象意识完全是一种主观体验，目前尚没有测量与检测这种体验状态的科学实证方法，因此要让机器拥有内在的意识体验与前面提到的其他四种方法性质完全不同。

也就是说，即使外部观测一个机器系统完全具备人类所有意识行为，我们也无法断定该机器系统内部一定拥有这种体验状态。

著名的蛇神（**zombie**）思想实验充分说明了这一点：因为一个蛇神机器人能够再现人类意识行为而无须体验状态。这样，就必然带来了更多的哲学争论。



3、机器意识困难所在

确实，由于涉及到心灵的一些本质问题，机器意识研究一开始就引起了哲学上的广泛关注，有专门讨论机器意识研究的哲学基础，也有讨论机器意识所会面临的困难，包括像心灵（mind）、感受质（qualia）和自我觉知（self-awareness）这些回避不了的、显而易见的困难问题，以及一些与意识相关的认知加工，如感知、想象、动机和内部言语等上的技术挑战。



3、机器意识困难所在

但更多则是延续早期对人工智能的哲学反思，对机器意识的可能性提出质疑。

涉及到强弱人工智能之争、通用人工智能问题、意识的困难问题、中文之屋的新应用、人工算法在实现意义能力方面的局限性、蛇神机器人不可能拥有主观性、人工奇妙状态体验学等众多方面的争论。

只有人工合成现象学（**Synthetic phenomenology**）才是机器意识哲学讨论中涌现出的一个新术语，专指现象意识的人工合成。



3、机器意识困难所在

那么机器能够拥有这种奇妙的意识状态吗？

Kiverstein分析指出了怀疑机器意识的可能性主要有两个方面的论点：

一方面认为我们神经生物系统惟有共有的就是主观体验；

另一方面认为，机器仅仅是一个蛇神（**zombie**）而已，除了机器还是机器，不可能具有任何主观体验的东西。

显然，这样的分析过于简略，难以对机器意识问题的澄清有所实质性的帮助。



3、机器意识困难所在

我们的结论是，对于机器意识研究与开发，应该搁置有争论的主观体验方面（身心感受）的实现研究，围绕**意向性意识能力**（环境感知、认知推理、语言交流、想象思维、情感发生、行为控制），采用**仿脑**与**量子计算**思想相结合的策略与仿脑，来开发具有一定意向能力的机器人，并应用到社会服务领域。