

《大脑与意识》特刊简介

Introduction to the special issue on 'Brain and Consciousness'

J.G. Taylor¹, W. Freeman², and A. Cleeremans³

¹Department of Mathematics, King's College, Strand, London, UK

*²Department of Molecular & Cell Biology, University of California at Berkeley,
Berkeley, California, USA*

*³Cognitive Science Research Unit, Universite Libre de Bruxelles,
Brussels, Belgium*

(translated by zang jie)

自从我们两个人（WF 和 JGT）编辑了本期刊关于“意识的神经网络模型”（Freeman & Taylor, 1997）的特刊以来，现在已经十年了。在此研究领域的十年间发生了很多事情，原因是与任何大脑和心理研究相关的实验结果的大量增加，以及声称能够解释意识活动的更复杂模型的创建和进一步发展。意识是通过适当的大脑活动产生的。因此，要问的一个问题是：我们在意识研究中真正向前推进了多少？很明显，随着时间的推移，我们越来越充分地认识到意识如何呈现是一个非常深刻和困难的问题。早先在（Freeman & Taylor, 1997）的论文中提出的观点已被更谨慎的方法所取代。从本期特刊的论文中可以清楚地看出这一点。

我们以更具理论性的论文开始此次特刊。第一篇论文（Aleksander & Morton, 2007）着手制定一些深层原则，以在信息处理系统中创建意识。它进一步发展了作者已经在别处发表的公理意识理论（ACT），以构建一种方法，该方法稍后在论文中通过模拟实现。ACT 的基本公理包括现实存在、想象力、注意力、意志和情感五个组成部分。这些在标志性的神经状态机中实现，该状态机以外部刺激的方式激活。进一步的组件被认为是模拟与眼球运动相关的感觉发展的一部分。

我们继续通过（Sanz, López, Rodríguez, & Hernández, 2007）的控制理论来研究意识。本文将认知系统定义为一种利用其他系统内部模型的系统。这导致系统意识，定义为其他系统内部模型的不断更新。然后，通过更新系统为自身创建的内部模型，进而在给定系统中产生意识。这导致了系统属性的五个公理，使其有望拥有意识，这些公理与上一篇文章中提出的公理非常接近，即 ACT 公理。

本期特刊的第三篇论文（Coward & Sun, 2007）是一种有价值且有益的意识研究方法。它考虑了科学理论的一般性质，特别强调了我們目前最有效的科学理论的层次性质，即物理世界的层次性质。然后作者认为大脑本身在其处理过程中具有层次特征，因此应该能够以类似的方式进行处理。在讨论了以这种方式了解意识的可能性之后，对这种方式的评价是對几种当前的意识方式进行的。

然后有人提交了一篇论文（Baars & Franklin, 2007），其中一位作者（BB）之前提出的有影响力的“全局工作空间”（GW）想法被用于模拟商业价值。这是 IDA 系统，一个用于测试 GW 假设的有趣工作区。IDA 解决了评估进入海军的水手以确定他们最适合哪种工作的问题。它被用作各种机制的平台，这些机制已被提议作为 GW 方法的组成部分。因此，IDA 系统是一个有价值的基础神经网络模拟器，可以通过详细的模拟测试关于意识和更一般的大脑处理的进一步想法。

本期特刊中的第五篇论文（Rolls, 2007）为我们带来了一种在全局层面上的更直接的基于大脑的意识方法。该论文首先仔细讨论了颞叶皮层编码的性质（这表明不一定涉及神经元活动的同步），并且颞叶层神经元中对给定刺激的意识阈值高于意识之外的活动的阈值（被视为可能的噪音）。然后将这些结果和其他结

果用于支持意识的高阶语法思想 (HOST) 方法。当需要纠正多步骤计划时, 这种高阶方法被视为很自然; 感受质的出现被视为次要于这种高阶校正系统。

下一篇论文 (**Taylor, 2007a**) 也使用了基于大脑的分层方法, 并且还采用了与注意力有关的工程控制思想 (与 Sanz 等人有关)。该模型的构建 (称为 CODAM) 使用了大脑成像和单细胞注意力实验的结果。建立了一个详细的注意力模型, 扩展到运动反应的注意力控制模型以及建模者使用的弹道注意力模型。CODAM 中的额外功能是注意力运动的必然释放。CODAM 允许探索动力学以了解意识如何从注意力运动中产生, 并探索这种意识的结果属性。特别是解释了重要的“通过错误识别第一人称代词而免于错误”; 由“所有权”信号 (必然放电) 解释, 创造了“我”的感觉。

第七篇论文 (**Taylor & Fragopanagos, 2007**) 解决了注意力和意识之间的关系问题, 该问题目前有些争议。对批判范式 (例如注意眨眼和变化盲) 的结果进行了各种成功的模拟, 并讨论了其他一些据称表明注意和意识的独立性质的模拟。这些模拟是在前一篇文章的 CODAM 注意力控制框架内创建的, 因此暗示所声称的注意力和意识分离无需充分依据。从这项研究中可以看出, 注意力仍然是通向意识的门户。

下面的论文 (**LaBerge & Kasevich, 2007**) 继续使用直接基于大脑的方法, 但现在更仔细地研究大脑活动中使用的基质, 即皮层中的神经元。该论文详细分析了皮层神经元的结构和动力学, 并提出意识是在皮层神经元顶端树突的神经活动中找到其关键动力学的。垂直于皮质表面 (在丘脑皮质环中) 的活动流被建议在时间上适当持续, 以支持意识的创造。与此活动流相关联的电磁场进一步辅助了这种创造, 从而为意识提供了场论基础 (另请参阅以下论文中提出的想法)。

在 (**Freeman, 2007**) 的论文中继续采用一种更直接的基于大脑的方法, 该方法基于他使用多个表面电极记录对兔子、猫和人的皮层活动进行的开创性测量。他考虑了一般动力系统如何处理大脑活动, 尤其是与使用同步神经活动向吸引子运动相关的大脑活动, 从而使大脑中的反射弧 (在所有反应的基础上) 详细接地。这种动态系统方法导致意识作为场现象的图景, 甚至建议将其扩展到亚原子和社会上的个人之间的相互作用。

论文 10 (**Cleeremans, Timmermans & Pasquali, 2007**) 从意识的高阶思想方法开始, 并考虑了可以实现它的可能的计算机制。模拟由两个交互网络执行: 一个, 一个一阶网络被训练执行一个简单的分类任务, 输入到一个被训练为编码器的二阶网络, 从而“观察”一阶网络的状态和在其输出单元上再现这些状态。这被提议作为解释心理态度的计算机制的开始, 即认知系统对其一阶知识的持有方式的

理解。在这种方法中，意识是一个人自己内部表征的地理知识，随着时间的推移，它会根据其组成部分对嵌入在物理世界中的代理的相对重要性来学习。

下面的论文 (**Grossberg, 2007**) 通过自适应共振理论 (ART) 和“所有意识状态都是共振状态”的猜想返回到基于大脑的方法。这篇文章回顾了这种方法的理论考虑以及一些支持它的快速增长的行为和大脑数据。本文总结了预测层状丘脑皮质回路中已识别细胞的功能作用的 ART 模型。这些预测包括解释在没有意识的情况下知觉学习是如何缓慢发生的，以及为什么新皮层下层的振荡频率有时是较慢的 β 振荡，而不是更频繁地发生在表层皮层中的高频伽马振荡。ART 将这些特性追溯到皮层内反馈回路的存在，以及各种详细的重置机制。

可以这么说，最后一篇论文 (**Herzog, Esfeld, & Gerstner, 2007**) “将猫置于鸽子之中”。这表明目前市场上的大多数意识神经模型都遭受“小网络”悖论的命运：它们都暗示意识应该存在于小网络中（由不超过十或二十个神经元组成）。这样的结果可以理解为所有这些模型都缺少某些东西（因此需要大量神经元来实现它们）或者这些小型网络系统确实是有意识的。我们参考 (**Taylor**) 在本期的评论，以进一步讨论这一结果的影响。

鉴于这组论文，尤其是 (**Herzog 等人**) 的论文，我们需要评估意识作为一种基于大脑的现象的研究的发展方向。当然，这里展示的论文集只是该学科整体实力的一小部分。例如，我们将任何感兴趣的读者引导至意识科学研究协会年度会议的论文集，以及“意识研究杂志”和“意识与认知”等期刊，以获得对当今的更广泛的看法意识研究。然而，我们声称，目前的论文为意识研究中可能处于“前沿”的基于神经网络的模型提供了一个很好的视角。它们涵盖了广泛的现存模型，并表明建模者在面对计算神经科学家的意识挑战方面取得的进展。那么他们在多大程度上表明正在取得真正的进展？是否有任何真正的突破？

由于其令人惊讶和重要的结果，我们分别对 (**Herzog et al., 2007**) 的论文进行了评论；正如评论 (**Taylor, 2007**) 所暗示的那样，它显然需要谨慎对待。因此，让我们转向解决刚刚提出的问题：自 1997 年特刊以来，在意识建模方面是否有任何突破？我们不得不承认没有，但本期特刊中的论文表明，在过去十年中，该主题取得了稳步进展。实验方面更清楚地说明了伴随不同输入产生或以其他方式产生意识的各种大脑活动。正是这些必须用于测试任何提出的意识神经模型。这种基于实验的建模方法在本期 (**Freeman, 2007; LaBerge, 2007; Rolls, 2007; Taylor & Fragopanagos, 2007**) 的论文中得到了很好的体现。正如 (**Coward & Sun, 2007; Aleksander & Morton, 2007; Herzog et al., 2007**) 的论文所证明的那样，对意识模型和意识本身的理论分析也已经发展到更深层次。

结束本期特刊有两条评论。第一个（Cooper, 2007）考虑了通过在物理世界中受过训练的人的眼睛观察时意识所呈现的问题。他雄辩地指出，首先需要探索可能（但非常微妙）的机制，这些机制可以被合适的神经元集合用来创造有意识的体验。只有当这种为意识提供神经基础的尝试失败时，我们才应该考虑第二步，即假设物质有一个独立的意识“成分”，与现在已知的构成它的夸克和其他成分不同。最后，他呼吁发挥想象力，使第二步变得多余。

第二条评论（Taylor, 2007b）更深入地讨论了（Herzog 等人）在本期中的小网络论证的含义。这些影响从不同的角度进行了讨论，尤其是从进化的角度。还需要对输入刺激进行较低级别的编码，从而增加任何意识系统所需的神经元数量。然而，如果这种编码和相关的意识网络是稀疏的，那么预计可能确实存在意识体验，但具有非常“薄”的特征。

我们从这里去哪里？更多相同的还是我们认为需要一个全新的方向？这涵盖了一种应用于模型的新型标准，例如（Herzog et al., 2007）的“小网络”论证给出的。它还将涵盖一种从大脑活动中创造意识的新机制。虽然我们没有任何此类全新的贡献，但我们特刊中包含的理论分析为任何此类新发展需要采取的形式提供了有用的限制。

我们似乎正处在更长的途中，逐渐爬升到意识山的顶峰。我们还没有做到这一点，但在这个问题上的贡献增加了一种普遍的感觉，即真正的进步正在发生，我们正在通过我们的努力朝着顶峰越来越高。

参考文献

- [1] Aleksander, I., & Morton, H. (2007). Phenomenology and digital neural architectures. *Neural Networks*, 20(9), 932–937.
- [2] Baars, B. J., & Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA. *Neural Networks*, 20(9), 955–961.
- [3] Cleeremans, A., Timmermans, B., & Pasquali, A. (2007). Consciousness and metarepresentation: A computational sketch. *Neural Networks*, 20(9), 1032–1039.
- [4] Cooper, L. N. (2007). On the problem of consciousness. *Neural Networks*, 20(9), 1057–1058.
- [5] Coward, L. A., & Sun, R. (2007). Hierarchical approaches to understanding consciousness. *Neural Networks*, 20(9), 947–954.
- [6] Freeman, W. J. (2007). Indirect biological measures of consciousness from field studies of brains as dynamical systems. *Neural Networks*, 20(9), 1021–1031.
- [7] Grossberg, S. (2007). Consciousness CLEARs the mind. *Neural Networks*, 20(9), 1040–1053.
- [8] Herzog, M. H., Esfeld, M., & Gerstner, W. (2007). Consciousness & the small network argument. *Neural Networks*, 20(9), 1054–1056.
- [9] LaBerge, D., & Kasevich, R. (2007). The apical dendrite theory of consciousness. *Neural Networks*, 20(9), 1004–1020.
- [10] Rolls, E. T. (2007). A computational neuroscience approach to consciousness. *Neural Networks*, 20(9), 962–982.
- [11] Sanz, R., López, I., Rodríguez, M., & Hernández, C. (2007). Principles for consciousness in integrated cognitive control. *Neural Networks*, 20(9), 938–946.
- [12] Taylor, J. G. (2007a). CODAM: A neural network model of consciousness. *Neural Networks*, 20(9), 983–992.
- [13] Taylor, J. G. (2007b). Commentary on the ‘small network’ argument. *Neural Networks*, 20(9), 1059–1060.
- [14] Taylor, J. G., & Fragopanagos, N. (2007). Resolving some confusions over attention and consciousness. *Neural Networks*, 20(9), 993–1003.