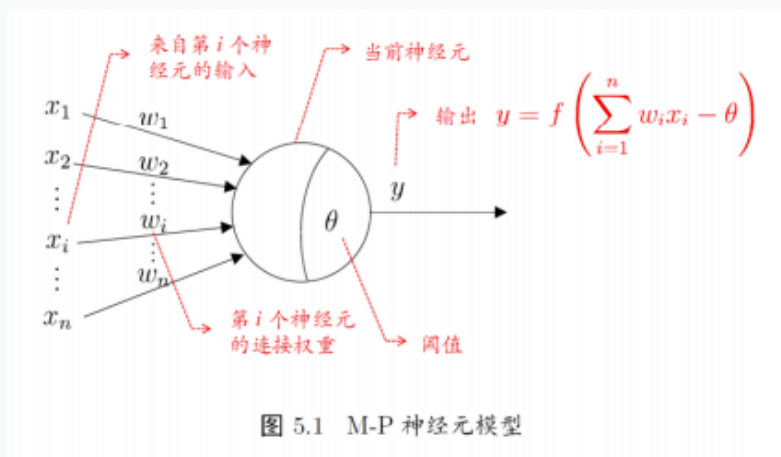


脑与认知总结（上）

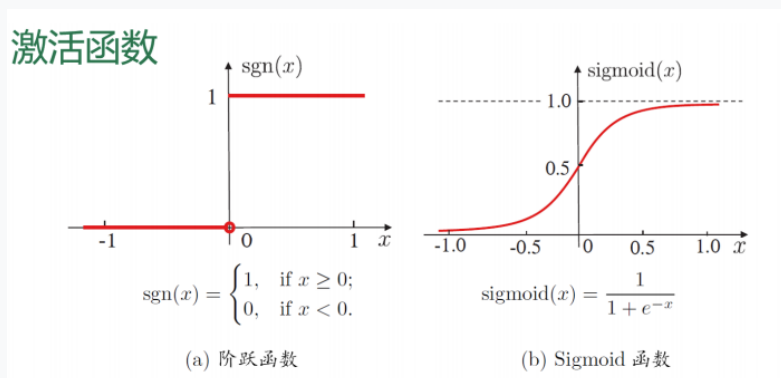
一、绪论及脑科学的发展历史

- 1. 通用人工智能的目标是制造出真正能推理和解决问题的智能机器。
- 2. 研究类脑的智能计算并非复制人的大脑，而是模拟人类大脑的功能，仅研究人的思维活动或记录脑中所有神经元不可能研制出真正的智能机器。从人脑中借鉴原理，受到启发。
- 3. 特定的大脑区域负责某项独立的功能，但这些区域组成的网络以及它们之间的连接才是人类表现出整体行为的原因。
- 4. 人工智能的联结主义：通过模拟人脑中神经元网络来处理信号，信号通过类似于神经元间的连接方式从一个节点传到另一个节点。
- 5. 神经元模型：

- M-P神经元模型：



- 激活函数



- 6. 感知机模型：

- 单层感知机：1 单层感知机无法解决非线性可分的问题 2 当时电脑算力不足

感知机学习规则

对训练样例 (\mathbf{x}, y) 若当前感知机的输出为 \hat{y} , 则感知机权重调整规则为:

$$w_i \leftarrow w_i + \Delta w_i$$

$$\Delta w_i = \eta(y - \hat{y})x_i$$

其中 $\eta \in (0, 1)$ 称为学习率

若感知机对训练样例 (\mathbf{x}, y) 预测正确, 则感知机不发生变化; 否则根据错误程度进行权重的调整.

• 多层感知机: 包含隐层的网络

- 神经网络的泛逼近定理: 只需要一个包含足够多神经元的隐层, 多层前馈神经网络就能以任意精度逼近任意复杂度的连续函数

- 7. 深度神经网络的不足: 1 需要大量的训练数据和时间 2 会犯人类不会犯的常识性错误 3 缺乏完善的理论基础

• 8. 人工神经网络的训练:

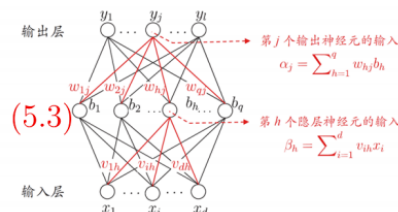
- 对于样例 $(\mathbf{x}_k, \mathbf{y}_k)$, 假设网络的实际输出为 $\hat{\mathbf{y}}_k$

前向计算

step1: $b_h = f(\beta_h - \gamma_h), \beta_h = \sum_{i=1}^d v_{ih}x_i$

step2: $\hat{y}_j^k = f(\alpha_j - \theta_j), \alpha_j = \sum_{h=1}^q w_{hj}b_h$

step3: $E_k = \frac{1}{2} \sum_{j=1}^l (\hat{y}_j^k - y_j^k)^2$



参数数目

权重: v_{ih}, w_{hj} 阈值: θ_j, γ_h ($i = 1, \dots, d, h = 1, \dots, q, j = 1, \dots, l$)

因此网络中需要 $(d + l + 1)q + l$ 个参数需要优化

参数优化

BP是一个迭代学习算法, 在迭代的每一轮中采用广义的感知机学习规则对参数进行更新估计, 任意的参数 v 的更新估计式为

$$v \leftarrow v + \Delta v.$$

《机器学习》, 周志华著

• BP 学习算法

- BP算法基于梯度下降策略, 以目标的负梯度方向对参数进行调整. 对误差 E_k , 给定学习率 η

$$\Delta w_{hj} = -\eta \frac{\partial E_k}{\partial w_{hj}}$$

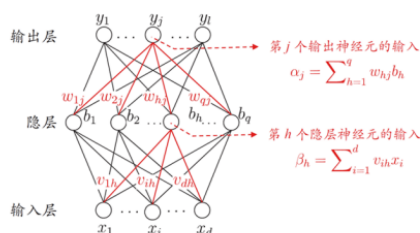
$$\frac{\partial E_k}{\partial w_{hj}} = \frac{\partial E_k}{\partial \hat{y}_j^k} \cdot \frac{\partial \hat{y}_j^k}{\partial \beta_j} \cdot \frac{\partial \beta_j}{\partial w_{hj}}$$

$$g_j = -\frac{\partial E_k}{\partial \hat{y}_j^k} \cdot \frac{\partial \hat{y}_j^k}{\partial \beta_j}$$

$$= -(\hat{y}_j^k - y_j^k) f'(\beta_j - \theta_j)$$

$$= \hat{y}_j^k (1 - \hat{y}_j^k) (y_j^k - \hat{y}_j^k) \quad (5.10)$$

$$\Delta w_{hj} = \eta b_{hj} g_h \quad (5.11)$$



《机器学习》, 周志华著

• BP 学习算法

- 类似的可以推导出：

$$\Delta \theta_j = -\eta g_j, \quad (5.12)$$

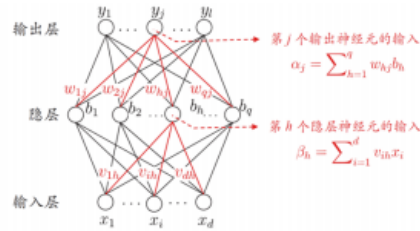
$$\Delta v_{ih} = \eta e_h x_i, \quad (5.13)$$

$$\Delta \gamma_h = -\eta e_h, \quad (5.14)$$

其中

$$\begin{aligned} e_h &= -\frac{\partial E_k}{\partial b_h} \cdot \frac{\partial b_h}{\partial \alpha_h} &= \sum_{j=1}^h w_{hj} g_j f'(\alpha_h - \gamma_h) \\ &= -\sum_{j=1}^l \frac{\partial E_k}{\partial \beta_j} \frac{\partial \beta_j}{\partial b_h} \frac{\partial b_h}{\partial \alpha_h} f'(\alpha_h - \gamma_h) &= b_h(1 - b_h) \sum_{j=1}^h w_{hj} g_j. \end{aligned} \quad (5.15)$$

- 学习率 $\eta \in (0, 1)$ 控制着算法每一轮迭代中的更新步长，若太长则让容易震荡，太小则收敛速度又会过慢。



《机器学习》，周志华著

• 二、脑科学基础：神经系统的细胞机制

- 1. 神经系统细胞分为神经元和神经胶质细胞。神经元进行信息传递和加工；神经胶质细胞对神经元提供结构支持，保证神经元间的信息传递更为有效。
- 2. 神经元包含细胞体，树突（突触后），轴突（突触前）。
- 3. 棘：树突表面的球状突起，突触通常位于这些棘上，有时也存在于细胞体等神经元的其他部位。
- 4. 突触负责神经元之间的信息传递，在树突上的空间分布具有特异性。（感知位置重叠，突触位置近）
- 5. 根据树突和轴突的形态，将神经元分为4类：
 - 1. 单极神经元：只有一个远离细胞体的突起，分支形成树突和轴突，常见于无脊椎动物。
 - 2. 双极神经元：通过树突接收来自一端信息，并通过轴突传至另一端。参与感觉信息传递和加工。
 - 3. 假单极神经元：树突和轴突融合的双极。将躯体感觉信息传递到中枢神经系统。
 - 4. 多极神经元：存在与神经系统的多个区域，参与运动和感觉信息的传递与加工，如人脑中的神经元。形态学相似的神经元倾向于集中在神经系统的某一特有区域，并具有相似的功能。
- 6. 胶质细胞：数量远多于神经元，占脑容量一半以上。包括：星形胶质细胞；小胶质细胞；少突胶质细胞和许旺氏细胞，以同心缠绕的方式包绕在轴突周围形成髓鞘。
- 7. 髓鞘损伤会导致信号传递减慢或阻断。刻意练习可使髓鞘增厚。神经元的信息传递
- 8. 神经元是神经系统传递信息的基本功能单位，通过对信息的接收和加工，并传递给其他神经元，构成了局部或长程的神经环路。
- 9. 主要过程：1. 信息接收：化学信号（神经递质或气味），物理信号（触视觉信号）2. 动作电位产生3. 动作电位传递4. 神经递质释放

- 10. 磷脂双分子层：离子、蛋白质和其他水溶性分子无法直接通过。
- 11. 离子通道：由跨膜蛋白质构成，允许带电离子通过。被动型离子通道：状态不变，只对某些离子开放；主动型离子通道：在电、化学、物理刺激下开放或关闭，参与产生动作电位。
- 12. 静息电位：胞内K高，胞外Na高，离子通道渗透性 $K \gg Na$ ，形成细胞膜外正内负的电荷梯度，阻止离子进一步沿浓度梯度流动。离子在浓度梯度和电荷梯度的共同作用下达到动态平衡，产生静息电位（-40 ~ -90mV）
- 13. 接收信号后，电位两种变化：
 - 1. 去极化：兴奋性突触后电位，使电位差变小，更容易产生动作电位。
 - 2. 超极化：抑制性突触后电位，电位差变大，不容易产生动作电位。
- 14. 去极化程度未达到启动阈值，则不会引起动作电位。达到或超过阈值，均引起相同振幅的动作电位，具有全或无的特点。
- 15. 动作电位两阶段：
 - 1. 快速去极化：Na离子通道开放，进入神经元，进一步加强膜去极化，使Na离子快速流入，动作电位快速上升形成尖峰。该正反馈：霍奇金-赫胥黎循环。
 - 2. 复极化阶段：Na离子通道快速减弱，直至完全关闭。K通道缓慢开放，流出神经元，动作电位下降。复极化的最后阶段完全由K通道控制，此时动作电位处于静息电位之下，进入短暂的超极化状态，随后恢复。
- 16. 不应期：复极化过程中暂不能产生新的动作电位。1. 绝对不应期：Na通道失活，无论如何刺激均不产生。2. 相对不应期：在1之后，因超极化现象导致更难达到启动阈值，只有高于正常强度的刺激才能重新产生动作电位。
- 17. 高度绝缘性的髓鞘使动作电位沿轴突的传递速度显著增加。相邻髓鞘之间的郎飞氏结上通过跳跃式传导进行快速传递，120m/s。
- 18. 神经元通过动作电位的发放率对信息进行编码。受限于动作电位的不应期，发放率存在上限，例如人类听觉神经元发放率<1kHz。
- 19. 还有相关性编码，信息通过两个或多个邻近神经元的共同激活进行编码。发放率编码可以完成快速的信息编码，更具时效性；相关性编码可以利用更少的动作电位来传递信息。
- 20. 使用ReLU代替sigmoid作为激活函数，模仿神经元发放率编码方式的稀疏性特点，只激活部分神经元，增加稀疏性。降低层数较多时反向传播训练误差。
- 21. 突触类型：
 - 1. 化学突触：动作电位达到轴突末梢，引起末梢去极化，钙离子流入末梢，信息编码为不同类型/数量的神经递质分子，释放至突触间隙，扩散至突触后膜，引起突触后神经元去极化或超极化。
 - 2. 电突触：缝隙连接形成穿膜的孔道，直接传递电信号，快速传导信息。但不能传递抑制性信息，也不能放大信号。
- 22. 兴奋性神经元引起突触后神经元更容易产生动作电位，抑制性神经元引起更难产生动作电位。

- 23. 神经回路：神经元之间相互连接，共同实现某一特定功能的神经元集合。最简单的神经回路仅包含一个感觉神经元和一个运动神经元，如膝跳反射的神经回路。
(伸肌收缩，屈肌松弛)
- 24. 神经回路的基本模式：收敛式兴奋，发散式兴奋，前馈兴奋，反馈兴奋，复发性兴奋（侧兴奋），前馈抑制（对兴奋时间、强度进行控制），反馈抑制，复发性抑制，侧抑制，去抑制。
- 25. 突触输入的空间分布：
 - 1. 兴奋性：树突棘
 - 2. 抑制性：树突棘、树突轴、胞体和轴突起始段
 - 3. 调节性输入：树突、胞体、突触前末梢。释放调节性神经递质，可影响神经元的兴奋性和突触传递效率。空间分布对信息编码有重要意义。细胞体上输入的信号，电压最大，上升和衰减最快。距离细胞体越远，电压越小，上升和衰减速度越慢。
- 26. 多个突触输入信号整合：空间整合，不同的树突上同时收到两个输入；时间整合，同一突触先后收到两个输入。
- 27. 神经元静息电位的计算：
 - 神经元的电路模型：
 - 利用RC电路模型对、膜电压 V_m 进行计算
 - 无离子通道部分的细胞膜具有电容特性，可用一个理想电容元件等效(C_m)
 - 细胞膜上存在 K^+ 离子通道、 Na^+ 离子通道以及 Cl^- 等其它离子通道(统称leak通道),可用理想电压源+等效电阻进行等效，其中电压源电压 E 为细胞膜只通透该离子时的静息电位，等效电阻的参数为 $1/g$

列写各支路的电路方程（电压参考方向：胞外+/胞内-，电流参考方向：胞外→胞内）

- $V_m = \frac{I_{Na}}{g_{Na}} + E_{Na}$

- $V_m = \frac{I_k}{g_k} + E_k$

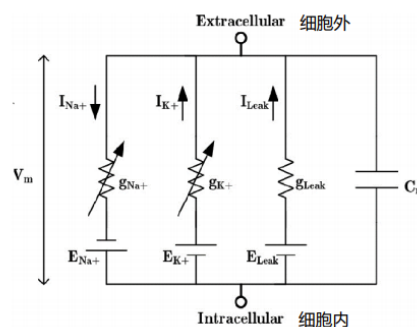
- $V_m = \frac{I_l}{g_l} + E_l$

- $I_c = C_m \frac{dV_m}{dt}$

神经元处于静息状态时，有 $I_c = 0$ 和细胞膜电流 $I_m = 0$

- 根据KCL有 $I_{Na} + I_k + I_l = 0$

- 得到实际静息电位的计算公式：
$$V_m = \frac{E_{Na}g_{Na} + E_k g_k + E_l g_l}{g_{Na} + g_k + g_l}$$

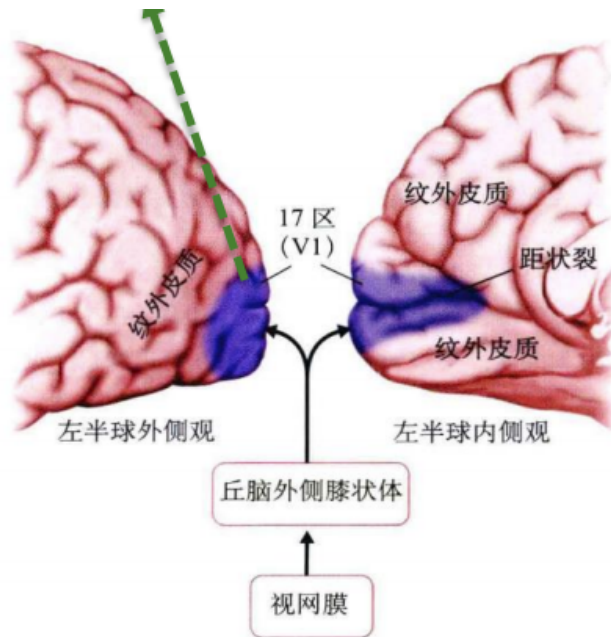


神经元动作电位的计算

- 由 $I_m = I_c + I_{Na} + I_k + I_l$ 得到
 - $I_m = C_m \frac{dV_m}{dt} + g_{Na}(V_m - E_{Na}) + g_k(V_m - E_k) + g_l(V_m - E_l)$
- 非静息状态下离子通道的电导 g 与细胞膜电压和时间均相关，因此可令 $g_i = F(V_m, t) * \bar{g}_i$ ，其中 \bar{g}_i 为该类离子的最大电导值
- 如何确定离子通道函数 $F(V_m, t)$?

• 三、人脑的基本结构和功能

- 1. 对大脑神经结构研究的层面：微观层面，精细解剖学，神经元甚至亚细胞结构之间的组织关联；宏观层面，大体解剖学，可以用肉眼区分的整体结构及关联。
- 2. 神经系统构成：中枢神经系统，脑和脊髓，进行命令和控制；周围神经系统，负责传递信息。
- 3. 额叶，顶叶，颞叶，枕叶。Brodmann分区，按细胞形态和结构划分，52个分区，与脑功能分区不完全重合。
- 4. 胼胝体：由发源于皮质神经元的轴突构成，是神经系统中最大的白质联合，是连接左右大脑半球的神经束。
- 5. 大脑皮质上存在大量沟回，增加了皮质的总面积，便于神经元之间形成紧密的三维联系。
- 6. 皮质由多层细胞构成，平均厚度3mm，包含神经元的细胞体、树突和部分轴突，被称为灰质。皮质下是由神经元轴突构成的神经束，成为白质。
- 7. 大脑皮质中包含300亿个神经元，每个神经元产生约1万个突触。大脑皮质的微观结构模式相对统一，由多层神经元构成，可能采用某种通用的计算模式进行信息处理。
- 8. 分层模式：
 - 1. 新皮质：占90%，6层细胞，神经元组织方式具有高度特异性。包括感觉皮质，运动皮质和联合皮质。
 - 2. 中间皮质：6层神经元。包括扣带回、海马旁回等边缘系统中的皮质。
 - 3. 异质皮质：1-4层，包括海马、初级嗅皮质。
- 大脑皮质的功能分区
 - 1. 额叶中的运动皮质，在运动的执行方面起重要作用。
 - 2. 顶叶中的躯体感觉皮质，接受来自丘脑的躯体感觉输入，包括触觉、痛觉、温度和本体感觉。
 - 3. 枕叶中的视觉加工区。初级视觉皮质接收丘脑外膝状体传来的视觉信息，6层细胞对信息精细编码和加工，将视觉信息传递给高级视觉皮质进一步处理。



位于枕叶的视皮质。BA17区，也称为初级视皮质（V1），位于枕极（occipital pole），并且一直延伸至半球的内侧面，大部分埋藏于距状裂中。

- 4. 颞叶中的听觉加工区，来自耳蜗的听觉信号通过丘脑的内侧膝状体最终到达颞叶上部的听觉皮质，经信息加工后形成对声音的感觉。
- 5. 对于躯体感觉和运动信息的加工，大脑皮质（感觉、运动）和身体之间存在空间拓扑关系。
- 6. 新皮质中不能单纯划分为感觉或运动的皮质称为联合皮质，接收多个皮质区域的输入信息，细胞被不止一个感觉通道的刺激激活。哺乳动物进化过程中，脑特别是新皮质面积增大，其中联合皮质占比最大。
- 7. 前额叶皮质属于联合皮质，占额叶皮质的一半，包含大量的神经网络，与几乎所有的大脑皮质都存在直接或间接联结。前额叶皮质在人对实现某种目标所采取行动的计划、执行、记忆等认知过程中发挥重要作用。
- 8. 边缘系统，参与情绪、学习和记忆的加工：
 - 1. 海马 位于颞叶的腹内侧，属于异质皮质，仅有3-4层神经元构成。在记忆和学习方面有重要作用。
 - 2. 丘脑皮质的关口，除了嗅觉外其他感觉信息都要经过。
 - 1. 视网膜神经元，外侧膝状体，初级视皮质
 - 2. 内耳听觉神经元，内侧膝状体，初级听皮质
 - 3. 躯体感觉神经元，腹后侧核团，初级躯体感觉皮质还接收来自相同皮质区域的输入信息。
 - 3. 下丘脑。神经系统，内分泌系统。接收边缘系统等其他脑区的输入信息以调节生理周期的节律，输出信息到前额叶皮质、垂体，向血液中释放激素。
 - 4. 扣带回、海马旁回、杏仁核
 - 9. 基底神经节，皮质下多个神经组织的集合，在运动控制中起重要作用。参与监控运动活动的进程，不直接控制。

- 10. 脑干位于脑和脊髓之间，包括中脑、脑桥和延髓，控制呼吸、睡眠等意识，损伤相对于皮质大多致命。
- 11. 小脑覆盖于脑干上部，处于脑桥水平位置，有约110亿神经元，与中枢神经系统其余部分相当。参与运动和感觉信息加工，整合身体和运动信息并调整运动，协调流畅。
- 12. 脊髓：后角的感觉神经元接收感觉信息并传导到大脑皮质，前角的运动神经元将大脑产生的运动控制信息传递到相应的肌肉组织
- 13. 输出相关的神经系统：
 - 1 运动系统，控制骨骼肌实现人体运动。
 - 2 自主神经系统，控制平滑肌心肌收缩，调控心脏功能。
 - 3 神经内分泌系统，分泌激素调节人的生理和行为。丘脑是后两者的共同调节中心。
- 14. 神经系统的运动控制具有分布式、跨越不同层级的特点：
 - 1. 底层脊髓提供神经系统和肌肉的连接，进行简单的反射运动。接收神经元输入的感觉信息，将运动控制信息转变成具体动作。
 - 2. 高层的大脑皮质处理当前感觉输入信息，对运动目标进行执行和计划（联合皮质），在小脑和基底神经节的辅助下，运动皮质和脑干将动作指令转化成动作。不关心运动的细节

• 四、大脑半球特异化

- 1. 左右大脑半球在整体结构和功能上有很高的相似性，但在少数特定功能上有明显的差异。
- 2. 胼胝体包含超过2亿条神经纤维，是连接两大脑半球相互交流信息的重要途径，多数连接等位区域，少数连接异位区域。
- 3. 裂脑人可以用语言确认从右眼输入的视觉信息，不能确认左眼。但可以通过左手动作确认从左眼输入的视觉信息。因为左侧大脑半球具有语言处理能力，右侧大脑半球具有物体辨识能力。
- 4. 左侧：语言信息处理能力（裂脑人、双耳分听实验）、逻辑推理能力。右侧：空间信息处理、陌生人脸识别、语言中的情绪信息处理。
- 5. 空间能力：右侧擅长定量空间坐标关系，左侧擅长抽象的空间范畴关系。
- 6. 左：语言、逻辑、计算、分析，逻辑规则和概念知识的因果推理。右：空间、音乐、直觉，物体时空关系的因果感知。
- 7. 大脑半球相互激活会有信息处理和输出控制一致性的问题。输入信息独立，信息传递延迟大。将输出功能集中于大脑的某一侧半球（左侧），实现输出一致性。
- 8. 生成装配器（GAD）假说：大脑的某个单侧半球进化出一个从单词、手势等少量基本元素中产生复杂连续表征的装置，从而实现以生成式认知为主的行为和思考模式。（生成模型）
- 9. 大脑半球的视觉信息处理特异化：
 - 1. 视觉信息在大脑中被层级化表征，左脑擅长表征局部信息，右脑擅长表征整体信息

- 2. 空间频率假说：人脑的视觉信息处理系统由空间滤波器构成，右脑低通，左脑高通。
- 3. 两侧信息进一步通过胼胝体整合后形成对物体的全面感知。
- 4. 可解释右脑擅长性别判断，左脑擅长熟悉人脸识别。
- 10. 记忆表征特异化：左脑擅长抽象、范畴化的原型记忆，右脑擅长具体、特例化的样例记忆。
- 11. 情感信息处理特异化：左脑前额叶皮质擅长积极情感信息处理，右脑前额叶皮质擅长消极情感信息处理。
- 12. 概率决策特异化：左脑解释器构建理论，解释所观察的现象（解释器假说），右脑更擅长简单现象的决策。
- 13. 胼胝体功能假说：抑制性，两侧大脑竞争信息处理的控制权；激活性，整合输入信息，促进信息处理。

以上内容整理于 [幕布文档](#)