认知科学导论总结-WZQ

目录

[认知科学导论总结 1](#_Toc74647456)

[第一讲 什么是认知科学 2](#_Toc74647457)

[第二讲 联想主义 2](#_Toc74647458)

[第三讲 行为主义 4](#_Toc74647459)

[第四讲 认知主义 5](#_Toc74647460)

[第五讲 脑结构 6](#_Toc74647461)

[第六讲 神经元 7](#_Toc74647462)

[第七讲 神经网络 8](#_Toc74647463)

[第八讲 认知模型范例 9](#_Toc74647464)

[第九讲 认知计算体系架构 12](#_Toc74647465)

[第十讲 认知计算表征 15](#_Toc74647466)

[第十一讲 心的数字计算理论 15](#_Toc74647467)

第一讲 什么是认知科学

1、认知：指人们获得知识或应用知识的过程，或信息加工的过程，是人的最基本的心理过程。

2、认知科学研究的对象是Mind

3、侠义的认知科学：心／脑是某种类型的计算机

4、认知科学是AI的原动力，因为

人的认知能力的提高，推动AI

 对人的认知能力模拟，实现AI

5、人工智能的出现期望通过计算机智能系统理解自然智能系统的工作原理

 认知科学的建立希望自然智能系统的工作机制引导计算机智能系统的建构。

这二者互补

第二讲 联想主义

1、定义：

联想主义从经验论（empiricism）出发，将感觉（sensation）和观念(idea)作为心理元素，并以联想作为解释心理现象的基本原则，认为一切复杂的心理现象都是通过联想整合而成。

2、基本思想：

联想主义的基本思想是在经验(experience)中“处在一起”的东西也将在思想中“处在一起”

3、创始人为霍布斯，发展于洛克和休谟

4、三个基本原则：

四个联想主义规则：近邻律……

5、洛克：英国哲学家，提出白板主义，

* Ideas：

白板学说：认为人生下来是不带任何记忆和思想的

装饰白板：利用经验，在经验的基础上，加工知识，形成自己的认知

经验分为外在经验（sensation），内在经验（reflection）

* The world：

感觉提供了对quality中的观念：

Primary quality：与个人感觉无关

Secondary quality：经个人加工形成的观念

* Idea of the world:

一个外部世界，一个内部自我

通过感觉知道存在着实体的、扩展的物质；通过反思拥有思维的物质；经验使人确信物质都是存在的

* Composition of complex ideas

所有观念都可以拆解为简单观念，反之构成复杂观念

6、补充：

洛克的学说有缺陷，即人的加工过程是推理过程，有天赋的概念在其中。康德的学说补充了这一点，即人的感性获取知识，人的理性分析知识，间接的将理性主义和经验主义合二为一。

7、詹姆斯，美国教育家、心理学家

从认知科学的角度，人的心理生活（mental life）或思维(thinking)的概念具有6个主要特征：

（1）Consciousness意识

（2）Introspective observation内省观察

（3）Privacy私有

（4）Flowing Like a stream溪流不断

（5）About something关于某事

（6）Evolution 进化

8、James学说

大脑过程(Brain Process，BP)是联想的基本载体；

思想的顺序取决于联想原则，并列举了很多法则，包括相似律和邻近律。

詹姆斯不仅仅满足各种联想方式的描述上，而且还强调如何进一步在神经水平上得到解释，即生物神经总是传到给最近的神经

James脑的两个加工原理，传递兴奋，激活总和

两种思维方式：自发性思维和自愿性思维。

第三讲 行为主义

1、定义：

行为就是一种可以外部观察的有机体的反应，其本质是人和动物对外界环境的适应，刺激—反应是有机体所有行为的共同要素。行为主义排斥意识和虚无，否认心理过程，只承认能够直接观察的事物是科学的事实，即实证主义（Positivism）哲学观点。行为主义心理学是研究行为的科学，受到当时社会的吹捧。

2、巴甫洛夫

条件反射理论，是行为主义的基础理论，具有获得、消退、恢复、泛化四个特征

巴甫洛夫生物信号系统学说，第一信号是人和动物共有的，第二信号包含语言文字等，是人类特有的

3、沃森（美国心理学家，行为主义心理学的创始人）

观点：

•心理学是一门客观的自然科学，运用的也是客观的自然科学方法；

•心理学的目的是预测和控制行为；

•人类与动物的心理并没有分界线；

•内省和意识在心理学中不起作用。

SR理论：

•刺激来自两方面：身体内部的刺激和体外环境的刺激；反应总是随着刺激而呈现

4、桑代克（联结主义创始人）

学习定律：准备律、联系律、效果律

SR理论：用刺激（**Stimulus, S**）和反应（**Response**，**R**）的层次结构来描述人类学习行为。

5、斯金纳

一切行为是由反射构成的，心理实验者的任务就是给予已知的刺激，观察学习者的反应，从而探究学习规律。是一种操作性条件反射，是有意而为之，行为发生在刺激之前。

问题：斯金纳的理论把人的学习过于简单化地归结为机械的操作条件反射，否认学习的意识特点，无视智力活动和独立思考，不承认人类学习具有特别的属性，把人的学习同动物的学习等同看待。

第四讲 认知主义

1、定义：

认知过程就是一个信息加工过程，即大脑或内心对来自环境的信息进行内在的加工

2、米勒的信息再编码过程

* 再编码：一种心理加工或变换过程，将接受的信息变换成另一种信息进行记忆，然后再解码，还原到原来的信息；最简单的方法是将输入信息归类后加以命名，储存这个命名，而非输入信息本身。然后对命名解码，得到原来的信息
* Miller‘s Law：为了理解别人说的内容，你必须首先假定别人说的是对的，并设法想象是什么支持了他所说的。也就是说，当别人说的观点你不认同时，先不要着急反驳，要穿上别人的鞋，去看待这些问题。

3、人接受信息后的处理历程：

刺激（信息）→感官记忆→选择性注意→短时记忆中心理运作→复习→长时记忆，分类组织后永久储存。

4、奈塞尔

认知心理学这个学科的定义：认知心理学是对感官接受信息后，经过转换、简化及加工等心理操作，从而获取知识。储存知识及运用知识。关于人们怎样接收信息，怎样把信息重新编码并记住它，怎样做出决定，怎样转换自己内在的知识结构，以及怎样把这种结构转译成行为的输出。

第五讲 脑结构

神经科学的历史是一个试图把作为脑功能的心灵映射到大脑结构的历程

比利时医生、解剖学奠基人维萨留斯(Andreas Vesalius，1514-1564）1543年出版了《人体结构》，他认为是脑控制思维。

1、笛卡尔

* 反射学说：人拥有心灵和身体，即mind-body二元论，身体是机械的可以用条件反射解释，但自愿性行为只能用心灵解释。提出反射的第一人。
* 松果体是人的灵魂所在，因为它是唯一的，其余都是对称存在的
* 基础问题：我思故我在。Mind和body是不同的，可以怀疑body的存在，却不可以怀疑mind的存在。

2、高尔（德国解刨学家）

第一个提出皮层定位学说

* “颅相学”：mind可以分为不同mental function，每个function对应脑中不同的位置，可以通过头骨观察出来，就有头骨测量术
* 优点：高尔的脑学说启发了后来的科学家们真正认识大脑和人体功能存在着“点对点”的关系。
* 缺点：错误的想法

3、弗洛伦斯

用切除鸽子的脑子证明颅相学的错误，使用了正确的方法得到了错误的结论

4、布洛卡+维尔尼克：

定位脑中有关听觉和理解的位置

5、拉什利

提出大脑的等势原理和整体活动原理。

等势原理：学习并非发生于大脑中的特定位置，而是发生在所有与它产生交互的大脑皮层当中；

整体活动原理：学习时大脑皮质的各部位是整体活动的，学习效率和效果与大脑切除后剩余的面积成正比。

6、赫布（拉什利的学生）

提出赫布律：cells that fire together, wire together.

行为模式是随着时间建立起来的，这些行为模式是通过特定细胞之间的联结形成的细胞集合

7、潘菲尔德

1.发展了癫痫外科治疗法；

2.更好地理解了人的大脑皮层的功能组织；

3.发现了对人类大脑皮层的某些位置进行电刺激，会唤起对过去生活经验的生动回忆，即一种早期生活事件的“倒叙”。

第六讲 神经元

1、高尔吉

观察细胞，提出神经网状学说

2、卡哈尔

凭绘画的天赋，把神经网络细致的结构描绘出来；质疑高尔吉的“神经网状学说”，提出了神经元学说。卡哈尔认为神经元是独立的单位，突起互不融合，但可以通过特殊的结构传递信号。

3、谢灵顿

他证明在反射活动中，当一群肌肉兴奋时，相对的另一群肌肉就被抑制。这种交互神经支配理论被称为谢灵顿定律。

4、神经元的数学模型

1943年麦卡洛克和皮茨(McCulloch and Pitts）根据生物神经元生物电和生物化学的运行机理提出二值神经元的数学模型，发表了神经系统开关逻辑(switching logic)理论（M-P模型），该理论在1960年代和1970年代分别产生了两个不同的研究领域：（1）数字与符号计算，（2）平行与分布式处理。





问题：

•思维与神经活动模式相对应，但神经活动模式由统计定义，并非由其构成要素的布尔函数定义。

•与单个神经元和由单个神经元构成的神经网络相联系的命题，并没有能够进行思维的适当的语义特征。人们一般所想的命题都涉及人物、地点和事件很少考虑神经元激活需要满足的条件。所以，即使逻辑命题能够模拟神经元或神经元的激活，但却不能模拟在具体实例中思维如何激活。

•思想（思维）、形式网络、神经网络与被指派到这样网络的命题之间有什么关系。

•思维的产生取决于网络组织还是网络的构成材料？

1957年罗森布拉特(Rosenblatt）提出感知机Perceptron；一个简单的感知器就是一种有着可变阈值与／或连接强度的双层M-P单元网；第一个人工神经网络（Artificial Neural Network, ANN）。

 1969年明斯基和帕佩尔特(Minsky and Papert**)出书**对感知机的提出批评，因为感知机无法实现异或运算；由于Minsky在学界的影响力，ANN研究进入低潮。

第七讲 神经网络

1、Rosenblatt, 1958

他的研究涉及神经元的结构、信息流的传递以及知识的学习和存储，他的研究打开了看待问题的全新视角！模拟人的大脑结构，是联结主义的一种。

* Discussion：

感知器不能学会归纳，随着刺激的增多正确率反而减少

它缺乏一种整体性的学习能力，即除去一部分节点过后正确率会降低，这与人脑整体性原理不符。

缺少异或运算的功能

2、联结主义历史意义

* 批评数字计算脱离了生物现实；
* 神经启发式架构；
* 强调统计与逻辑方法；
* 强调模式识别和学习；
* 运用并行加工和分布式表征。

第八讲 认知模型范例

1、三个认知模型范例

自然语言处理、心理图像表征、计算视觉理论

2、ELIZA

ELIZA是由Joseph Weizenbaum在麻省理工学院人工智能实验室创建，它是最早的机器与人自然语言对话程序，是第一个真正意义上的聊天软件。它的程序策略是，运用关键词搜索原先录制好的应答，然后根据“病人”的状态再向病人提出问题。程序实际上所做的只是重新表述用户输入，通过根据英语规则改变单词的顺序，将用户的陈述变成问题。另一个技巧是对特定的单词（例如“母亲”，“丈夫”，“家庭”等）做出反应，并提示用户多说一些。这种对话程序一直被不断复制和改进，其基本原理成为了许多现代聊天机器人的必不可缺的部分。

当然，Eliza根本“不懂”任何东西，在Joseph的论文中也对该问题做出了详尽的讨论。他提出ELIZA的“理解”能力的可扩展性是有限度的。假如我们判定一个对象是否理解，我们不是测试该对象是否有继续话题的能力，而是判断他是否从被告知的内容中总结出有效的结论。也就是说不能仅仅按照语法规则转换句子，却没有任何对单词含义的内部表示。为了做到这一点，需要机器至少有能力存储选定的输入部分，或者构建一个更庞大的可以动态学习的知识库。虽然ELIZA无法做到理解，但是针对其缺点的改进无疑会推动人工智能在自然语言理解领域的进步。

3、SHRDLU

SHRDLU是由Terry Winograd于1968年至1970年在麻省理工学院开发的一种自然语言理解程序，当它首次问世时就被誉为一种革命性的AI算法。程序提供了对问题、陈述和命令的句法以及语义分析，用户可以与电脑对话，在精简的「积木世界」的内部进行移动物体、命名集合、查询等操作。然而，它仍有许多局限性。事实上，它对“积木世界”并不真正了解，而且不知道自己在讲什么。此外，程序的扩展性也不高，除了在基于规则操作的虚拟世界中可以运行，在一个更复杂或包含更多规则的世界中就会出现问题。

但毫无疑问的是，SHRDLU的贡献同样巨大。首先，它是自然语言处理领域向前迈出的重要一步，因为它试图在语言理解过程中结合人类语言模型和推理方法，它是第一个实际上“理解”某些东西的AI程序之一；其次是其突破性的组织架构，它可以将组件有效地进行组织和交互，处理问题时能将目标和问题求解结合起来，使之类似于人的认知过程；此外，尽管它运行起来看上去有些幼稚，但是基于“积木世界”和基于规则的人工智能仍然是当今讨论的重点话题之一；再有是其所包含的认知-推理系统，展示了如何将语法和语义分析与一般知识相结合；最后，SHRDLU启发了专家系统的研究，例如1974年由斯坦福大学Edward shortliife提出的MYCIN专家系统，故它还为知识工程领域做出了重大的贡献。

4、心理图像表征

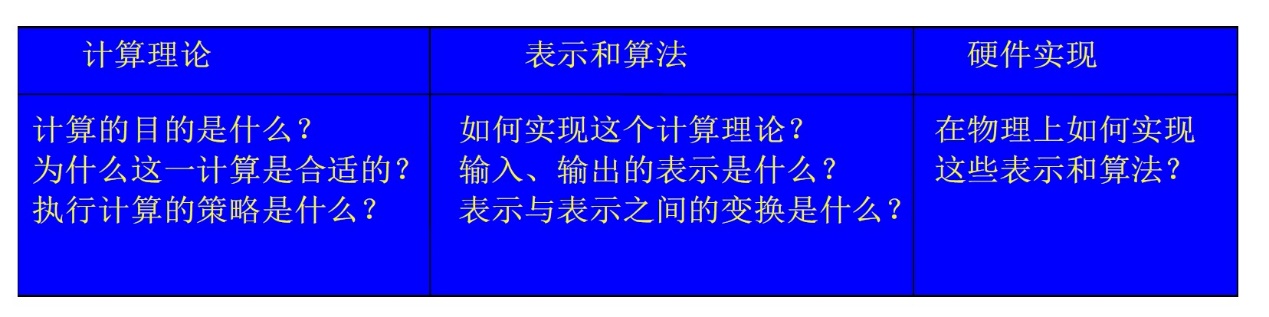
图像（**image**）和语言（**Language**）平行和联系的两个认知系统。语言系统加工离散的语言信息；图像系统对客体或事物的信息进行编码、存储、转换和提取。两类系统可分别由有关刺激激活，图像信息和语言信息可以相互转换。

表征这一心理现象是客观存在的，是可以用科学的实验方法证明的。并非像行为主义者那样，信息进入大脑就如同进了黑箱子中，无法知晓它是如何工作的。

图像表征是物体的抽象的类似物的再现。表征是真实物体的类似物，它是以观念的形式存在于头脑中的，它具有直观性。

* 视觉的正确性和整体性
* 多通道感官影响 McGurk效应，Ba/Da

5、计算视觉理论

* 信息处理的三个层次（David Marr提出）：
* 视觉表示框架（3阶段）

第一阶段**(**早期阶段**)**是将输入的原始图像进行处理，抽取图像中诸如角点、边缘、纹理、线条、边界等基本特征，这些特征的集合称为基元草图**(primitive sketch)**；

第二阶段**(**中期阶段**)**是指在以观测者为中心的坐标系中，由输入图像和基元图恢复场景可见部分的深度、法线方向、轮廓等，这些信息的包含了深度信息，但不是真正的物体三维表示，因此，称为二维半草图**(2.5 dimensional sketch)**；

第三阶段**(**后期阶段**)**是在以物体为中心的坐标系中，由输入图像、基元图、二维半图来恢复、表示和识别三维物体。

* 大脑信息加工机制是由串行和并行组成的，需要串行和并行计算理论来支撑。Marr观点中最突出的特点是分层计算。
* 只通过研究神经元来研究知觉是行不通的，Marr是基于下述逻辑提出他观点的：信息加工设备是用来解决问题的，它们凭借它们的结构来解决问题；当知道这个结构是用来解决何种问题时，我们就能解释这个设备的结构。

6、计算视觉面临的困难

（1）图像多义性：三维场景被投影为二维图像，深度和不可见部分的信息被丢失，因而会出现不同形状的三维物体投影在图像平面上产生相同图像的问题．另外，在不同角度获取同一物体的图像会有很大的差异．

（2）环境因素影响：场景中的诸多因素，包括照明、物体形状、表面颜色、摄像机以及空间关系变化都会对成像有影响，

（3）知识导引：同样的图像在不同的知识导引下，将会产生不同的识别结果．．

（4）大量数据：灰度图像，彩色图像，深度图像的信息量十分巨大，巨大的数据量需要很大的存贮空间，同时不易实现快速处理．

第九讲 认知计算体系架构

1、定义：讨论信息如何存储以及什么决定着信息在系统中流动的方式。

2、算法是一系列确保得到结果的步骤；程序是用来编码一个算法的。程序有错算法不一定有错，可能是编程语言有问题。

3、算法强等价：不仅同输入同输出，且计算的每一个中间步骤相同

算法弱等价：同输入、同输出

4、图灵机体系架构

任何一步通过五元组的形式来表示

5、冯诺依曼体系架构

* 优点：

1.冯诺依曼机允许直接（绝对、随机）和间接（或相对）访问内存，而TM只允许间接（相对）访问；

2.冯诺依曼机程序可以作为数据存储在内存中；

3.冯诺依曼机有专门的计算机构，如算术单元（TMs没有专用机构）

4. 诺依曼体系结构允许人们充分利用“子程序”；

* 缺点：

串行化的运行方式使之成为了冯诺依曼机的限制所在。

6、产生式系统(Production Systems, PSs)

* 背景：1960s人们基本上忽视了人类认知过程的研究，特别是两大主要学派：行为主义心理学家和神经科学家。行为主义者主要关注S-R模式和应用，而神经科学家关注神经信息传递机理和神经硬件实现。产生式系统在心理学中广泛应用于模拟各种认知功能，在人工智能领域成功用于构建专家系统。纽厄尔和西蒙（司马贺）提出一种新型的信息加工理论--产生式系统PSs，用来模拟人类认知过程，它代表了介于行为和神经科学之间的一个特定解释层次….
* 定义：存储了一系列”if …… then”，这个被叫做产生式，它是一种规定的系统加工程式。模拟人的思维方式，非常自然。
* 组成：包括三个主要构件：（**1**）产生式规则集**a set of production rules,**形式是：如果**A(**条件），则**B(**动作）；（**2**）记忆工作空间**(memory workspace)**（也称“语境**”);**（**3**）规则解译器**(rule interpreter),**它将相关的规则应用于工作空间中的结果（如执行顺序，消解规则等等）。它可以进行并行计算。
* 与VNMs对比：（这部分只列一部分）

它是“内容”可寻址的，而非“地址”可寻址的；

没有外部的额控制单元；

在高度模块化方面比VNMs更有优势，每个单元只是狭义地和其他单元交互；

能自然且直观地表示知识……

* 缺点：

(1) 低效率

必须执行满足条件的每个动作，因此很难让它们执行预定的步骤序列；

(2) 不透明

很难表达或看到它们的算法结构。

原因：这两个问题在很大程度上是由于产生式的模块化和统一性,它们不能像冯诺依曼程序中的子程序那样分层调用其它产生。

7、泛魔模型

* 两种学习规程：通过特征加权，子魔选择——>属于监督学习



* 四个层次：计算机的模式识别应以特征分析为基础，它包含相继发生的

(1) 获得物体的图像；（image demons）

(2) 分析它的特征；（feature demons）

(3) 更高水平上产生对事物的认知；（cognitive demons）

(4) 作出正确的决策。（decision demon, 类似softmax?）

* 它在连接主义中发挥了重要的作用，与感知机一样通过高度互联单元的并行计算，每个单元学到东西很少，但组合起来就很强大。它学习的也是权重。

8、体系架构的分类

对于一个计算设备来说，至少包含三个部分：表征、存储器、控制。

任何计算设备必须能够在存储器中储存表征，每一个这样的设备都必须能够通过某种特定方式改变状态--控制。

第十讲 认知计算表征

1、三个基本要素：表征（符号）本身、表征（符号）所指称的事物、（表征）符号的含义。

很多时候要区分表示、表达、表征。

表征是认知系统对外界激励信息作用下所产生的一种相对稳定的内在结构，这种结构不仅会影响外界信息的进一步接收和加工，而且是外界激励信息作用的反映。

2、所有认知模型需完成两步：

* 如何对表征的内容进行（创造、删除、修改）操作？
* 表征是如何进行表征的？即implement

3、表征分类：

原子表征、复合表征、组合表征（付华）

4、高级表征形式：谓词逻辑、语义网络、框架等等

5、谓词演算（Predicate Calculus，PC）

* 谓词演算中包含七类表达（expressions）：

谓词、名称、连接词、变量、开语句、量词和句子。

* 每一类表达都有相关联的语法规则，以确保良构性（well-structured）；
* 每一类表达都有语义规则，用于指称和真值。
* 优势：

具有显义特征

表达自然

良好的数学理论支撑

* 劣势：

相关信息并非在一起储存，使得搜索变得相当困难

6、语义网络：

* 推理机制：推理通过网络的“扩展激活”来进行，激活沿着箭头方向进行扩展，使该节点继承远部节点的属性。
* 问答机制：在知识库的语义网络中进行匹配。例如问：betty住在哪？则将其转换为节点及其扩展，在已构建的语义网络中进行匹配搜索，最后回答。
* problems：

1） 语义网络自然存在着表征典型性/常态性、析取和否定等概念的问题；

2）除非受到限制，否则激活会扩展的很远；

3）与谓词演算不同，语义网络不具有显式和精确的语义表征；

4）没有像谓词演算那样，获得语义网络形式化属性的结果。

7、frames and scripts

* 想法：人们对现实世界中各种对象或事物的认知都是以一种类似于框架的结构存储在记忆中；当人们遇到新情境时，就从记忆中找出合适的框架，并根据实际情况对细节进行修改、补充，从而形成对当前对象的认知。



* 优势：

结构性、深层性、继承性、自然性

* 劣势：

没有形式化的理论支撑等等

第十一讲 心的数字计算理论

太深遂，无法总结