**深度语义模型DSM简介**

# 我们需要一个深度语义模型

近几年，主要由深度学习等技术推动，人工智能技术取得了巨大进步，语音识别、图像识别、机器翻译等领域的技术水平都有了显著提升。但在向更高级的人工智能迈进时仍然遇到了很多难题和瓶颈，其中一个重要的原因是——业界一直缺乏一个统一的深度语义模型，这是AI领域最重要的自然语言处理方面缺乏实质性进展的根本原因。是时候认真思考并构建一个统一的深度语义模型了，这将为让计算机“表达人类语言的涵义”奠定一个坚实的基础。

为什么实现一个统一的深度语义模型如此重要和迫切呢？我们从目前自然语言处理采取的技术和存在的问题进行分析。

众所周知，自然语言处理（NLP）非常难。其实先不要说全面的自然语言处理，仅仅是其中的一部分即自然语言的理解（NLU）就非常非常难。自然语言理解还仅仅只是让计算机听懂人类话语的意思，而自然语言处理则要在理解的基础上实现表述、翻译、对话、查询、运算、缩写、扩写、推理、思考、规划、决策、行动等等一系列的任务。目前，最基本的自然语言理解就没有从根本上得到解决，就更不用说其它方面的处理了。

为什么自然语言理解和处理如此困难呢？这要先问一个最基本的问题“自然语言理解的结果是什么？怎么才算正确的理解了一句话？”。对此简单的回答是：“自然语言理解就是计算机对自然语言文字进行处理，解析成出文字的内在涵义——也就是俗称的语义。”。但这个回答只是把问题变成了“语义又该如何定义呢？不同领域、不同专业、不同的人采用不同的语言在不同的场景下表达出来的各种各样的文字，进行理解后产生的具体语义结构，究竟应该是怎么一个样子？”。这时我们发现，这个最基本的问题根本并没有得到令人满意的解答。

先看一看，在缺乏这样一个基础模型的情况下， NLP相关的技术是如何处理的。

目前，对自然语言的理解都是针对局部业务场景采取针对性的简化解决方案。从技术层面来看，分词、句法分析、语义分析、语用分析、模板匹配、实体标记、情感分析、指代消解、省略消解…等技术概念和手段众多，但没有一项能彻底解决问题，甚至每一项子技术自身也无法做到完全可靠。从业务层面来看，目前的语义搜索、智能客服、智能助手、聊天机器人、机器翻译、舆情分析、内容推荐、文章缩写、文章编写…等各种业务也都针对特定应用设计专用的处理方法，采用非常简化的语义体系甚至直接在语言层进行处理，不但无法统一融合，实际结果也差强人意。

一个典型的例子是人机对话处理，目前的智能助手基本都是采用对文本形式的语言模板进行匹配来实现所谓的理解，对关键字或者短语模板匹配成功了就视为正确理解了语义，然后调用对这些模板预先配置好的回复句子（同样是自然语言的）进行回复或者转到预设的一段程序进行处理。显然这些语言形式的文字模板并不真正具有描述语义信息的能力，并没有真正实现对意思的深度理解，只是依赖配置文件或者硬编码来驱动。用户话语略有变动就会导致理解错误，更难以实现多轮对话来完成更为复杂的业务应用。

近两年的主要潮流是寄希望于深度学习技术，试图把自然语言相关的问题都归纳为“串对串、端到端”的处理方法。但结果是除了机器翻译有不错的进展外，别的方面并没有实质性突破。这是由深度学习本身的原理所决定的，它适用于解决语音识别、图像识别等方面的问题，但并不是解决自然语言处理的关键。即便是机器翻译的未来也终究要基于语义理解，不基于语义理解的机器翻译技术终究一定会遇到天花板和瓶颈。这是其丢弃“语义理解”这个根本而完全依赖概率统计原理进行拟合匹配的方法本质所决定的。未来，只有在基于深度语义理解的基础之上实现的机器翻译，才能真正达到人类翻译的水准。

显然，人类的思维模式绝不是一种简单的字符串拟合匹配。人类理解别人话语的意思、进行思考推理决策、并将自己的意思生成话语表达给别人，存在着深层次的思考，运用着一种相对统一的内在思想的表达和运算体系。从这点来看，进行语义分析的方法一定是符合人类思维模式的正确方向，只是目前语义表达的模型和语义分析的技术过于薄弱因而效果不佳。各种方法和模型包括传统的语法分析、语义分析、知识图谱等等，都存在着诸多的问题，例如：

1. 语义体系缺乏深度和完备性
2. 实现结构缺乏足够的灵活性
3. 知识概念缺乏模糊性适应性
4. 语义和语言混杂不清
5. 语义和语言无法融合
6. 缺乏自我描述的体系
7. 缺乏完整的演算体系
8. ……

因为存在这诸多问题，传统的这些技术满足不了实现深度自然语言理解的要求，更无法形成支撑起完整的自然语言处理的闭环，和理想的统一语义模型相距甚远。

而这个根本问题不彻底解决，自然语言理解和更高级的智能处理也就没有一个稳固的基础和可以围绕的核心，就像没有根的水上浮萍，处理起来非常随意、非常浅薄、技术碎片化、形不成完整体系，这就是自然语言处理的根本问题。

深入进行分析，上述各个具体问题并非不能克服。而既然人类能通过不同形式的自然语言进行高效而准确的信息沟通和表达，内在就一定蕴涵着共同遵循的语义信息表达规范和标准。因此可以判定：一个统一的语义模型不但是这个世界绝对需要的，而且理论上也是存在的。终归会被构建出来，只是时间的早晚而已。

要构建这样一个统一的语义模型，需要吸取众多已有方法的优点，摒弃上述这些问题和缺点，并在多个关键点上做出很多创新突破，才能得到理想的结果。

按照这个方向进行多年研究的结果，我们设计了一种深度语义模型，简称为DSM。有望在这个重要方向上取得突破，成为自然语言处理和相关AI业务应用可以参考的统一信息模型，推动相关领域的发展。

深度语义技术的核心是深度语义模型，这个模型非常重要，是解决自然语言处理以至于认知智能方面面临的诸多问题的关键。

深度语义的理论和模型，就是为了解决上述最根本的问题而提出并设计的，有望成为彻底解决“如何完整深度的表达自然语言的语义”；回答“自然语言理解之后究竟应该产生一个什么样的结构？”；以及解决“如何用一个模型不但实现自然语言理解，还能实现更高级的记忆、生成、运算、查询、思考、学习等完整的自然语言处理以及认知智能”的基础性模型。

# 如何构建一个统一的深度语义模型

实现深度语义模型的工作量巨大，内容非常广泛，这里只能对一些要点进行简单的阐述：

## 以语义为核心，语义模型完全脱离自然语言

在深度语义模型体系里，语义模型才是本质和核心，语言只是语义的外在表现形式，语言存在的唯一意义就是为表达语义而服务。因此在研究语义模型时，需要从千变万化的语言形式的素材和表象中，概括归纳抽取其本质，并构建出一个“纯”的语义模型，这个最终的语义模型将不依赖于任何语言，而是能独立完备地描述各种知识和信息。这一点非常的重要。

## 语义模型能和任意自然语言相互转换

这个脱离自然语言形式的语义模型又需要能无缝映射到任意自然语言，同一个语义模型可以和各种自然语言——包括不同语种和同一语种的不同说话方式（包括方言等）进行灵活的转换。也就是说，一组相同的语义知识或信息可以转换生成为任意的一种自然语言形式进行表达，反过来，这些自然语言形式又可以理解转化为正确的语义知识和信息结构。简单的说“就是用语言来进行外在表达，而用语义进行内在表达、描述、存储、运算、思考…”。

这个语义模型并不是以前曾经提出的所谓“中间语言”。“中间语言”是仍然以语言概念体系为基础，没有跳出语言形式的束缚。而深度语义是真正的本体结构，完全以语义概念为核心，彻底剔除掉语言形式的“噪音”干扰，构建起完整、严谨的语义体系。然后再和语言形式进行匹配对接，可以彻底解决语义和语言的结合关系这个基础问题。

可能会提出的疑问是：“人类语言千变万化，如何能用一种结构去统一表达呢？”。实际上，这一点的难度其实并不是那么大，要点是必须把语义和语言信息分开，不要混淆在一起，该归语义的归语义，该归语言的归语言。语义体系完全成型后，语言模型表达起来也就简单了，用通用的结构设计就可以有效表达不同语言的共性。一些语言确实存在非常特殊的个性，但这些特性毕竟有限，进行针对性处理就可以。

## 语义模型要具有足够的深度和自描述性

以前也有很多语义模型，它们存在的最普遍问题是：不能形成完整的体系，缺乏深度不能自我描述和演算。而深度语义模型的重中之重就是要解决以前这些模型的弊端，构建起具有足够深度和足够的自描述能力的语义体系。

具体方法上，语义模型的构建采取自顶向下方法，形成一个形如金字塔的多层级知识和信息结构体系。

这其实是遵循了一个非常基本的道理——“人脑的知识体系是分层级的”。越高层的知识就越基础越重要，是理解和表达更低层知识的依据，而越高层的知识的数量其实是越有限的。越低层的知识的数量就越多甚至无限，但都可以用高层的知识来解释。

最基本的概念和知识是人类思考问题的最基本的一些原则和思想，就是数量非常有限的一些哲学性概念，例如离散化、概念、空间、时间、数量、程度、关联、存在…等，深度语义模型构建的基础工作是对这些概念进行了长时间的深刻的思考和组织构建，这些最基本概念必须经得起严格的检验和推敲。

核心层，是在上述基本概念和知识基础上划分出的数百个核心概念，包括事物、事件、关系、规则、公式、集合、区间、范围、比较、序列、性质、事实、目标、条件、结果、过程、愿望、情感、动作、行为…等等。

深度语义模型对核心层的描述定义，绝不仅仅停留是把这些概念名词简单罗列出来，并且要构建起了他们之间本质的、完备的、具有说服力的内在关系，这些关系足以支持它们之间的相互表达、相互解释，才能构建起对各种知识进行描述和运算的基础。

可以将深度语义的体系看着完整的一棵树，顶层的哲学概念是树根；核心概念是树干和一些主要枝干；基础概念是更细的枝干；而外围概念是树的大量的叶子。树干很粗，但只有一根并且对于树来说是最重要的。树叶虽然数目众多，但必须依附于树枝和树干。只要树干和树枝存在，树叶可以随时飘落也能随时生长，就像人脑只要记住了高层级知识，就可以用来随时学习新的低层级知识还可能随时忘却——下次遇到了再学。

知识图谱和大数据等技术的问题就是没有构建起树干和枝干等基本的知识体系——即使有一些构建也非常简陋和不完整，在这种基础并不稳固的情况下收集的大量具体化信息（对应于深度语义模型里的外围层信息）就只能实现简单的运算，而不能实现真正的深度理解和推理思考等智能化处理。一个知识图谱动辄号称数十亿条的知识量，非常庞大，在某些方面远远超出了人类的记忆力，但并未见到就产生了足够的智能来真正实现和人类进行几轮真正有意义的对话——因为缺乏足够的知识深度和完备性。

## 语义模型要足够完备

语义模型的表达能力要足够完备，既能实现对基本的事物概念和关系进行定义和描述，又要能对事件、动作、情感、意图等进行描述，还要能对因果推理等复杂体系进行描述。能描述短语，也能描述上百字的长句子以及段落、篇章、文章，还要能描述完整对话的整体过程，甚至要能对人脑的思维过程进行描述和表达。

深度语义模型在对自然语言的处理过程中，理论上对所有有效信息都不能丢弃，每一个字符的含义和作用原则上都需要能作出合理的解释。

相比较而言，目前的知识图谱等技术是很不完整的，很难完整地表达哪怕是几十个字的短文的信息，一般只是针对不同场景设计专用的一些结构来存储和处理提取局部关注的信息，更像一个具有一定灵活表达能力的词典，而词典是不具备智能的。可以说：深度语义模型关注完整信息，知识图谱只提取片段信息。

另外，目前的知识谱图也很难实现和不同自然语言双向灵活的对接。

## 基于模糊逻辑，能容纳不确定性

一些基于规则的系统不能很好地解决实际问题的原因，一方面是知识和规则缺乏层级关系，另一方面就是知识、信息、规则的定义是二值逻辑，非黑即白，在灵活多变的环境下就总是遇到不能解决的“例外”。

深度语义模型采取模糊逻辑来解决这个问题，所有的知识、信息和规则的表达都体现为概率值，可以有效解决信息不确定性、歧义处理等难题。

在深度语义模型里，语义信息的处理就是一个不断具体化的过程，模糊逻辑同时体现在“意义性”和“正确性”两个方面。随着信息的增加，表达的语义的内涵增加而外延减少，使得描述的语义越来越准确和具体，因此“意义性”得到加分，而同时因为随着具体化的提升出现错误的概率就增大，因此“正确性”将减分，两者会达到一个平衡，最终的处理结果是综合打分最优的一个结果。

## 能描述计算、查询、推理等知识和演算体系

在实现对各种知识和信息深度描述的基础上，进一步实现对规则、计算、查询、推理等的描述，形成了对各种语义进行演算甚至模拟人脑思考的基础。

深度语义还容纳了基础数学知识的表达，对自然语言和语义中蕴涵的数学知识的分析和计算已能从根本上解决实现并且遵循相同的技术原理。

## 能灵活扩展，能表达各种行业领域的语义知识

深度语义模型具有足够的深度，最核心的基础层足够抽象，具有很强的泛化表达能力，能灵活地向下和横向扩展来描述各种行业领域的语义知识。无论是常识知识，还是衣食住行、社会关系、情感交流、文化体育、自然科学、政治财经…，事实上只要适用于用语言文字表达的概念知识和思维处理，都可以用同一套深度语义模型进行描述和运算。

# 作为自然语言处理的基础表达模型

在构建起来完整的深度语义模型后，就可以回答前述的有关自然语言理解后的结构究竟应该如何的根本问题——正确的理解结果应该按照深度语义模型所定义的语义结构作为标准。

同时，不仅是自然语言理解，完整的自然语言处理技术也都将以这个模型为依据，紧紧围绕语义结构为中心进行。语义和知识结构是严谨和明确的，作为最终的处理目标同时也是正确性的评判标准。在保证语义信息本身正确的目的下，语言可以进行做各种变形、简化、省略都不是问题，甚至出现一些噪音、语病、错误都能得到解决。

这样，自然语言的信息表达、理解和处理就具有了统一的根基和核心，向彻底解决相关技术难题迈出坚实的一步。

下边解释下深度语义模型体系下的语言和语义处理算法原理：

## 理解

在深度语义模型体系下，对自然语言进行深度语义理解后，将产生语义模型定义的完整语义结构，具体来说是一个语义网络结构，结构中的每一个节点都是和具体语言无关的语义概念，语义概念之间通过各种关联连接成为一个整体。

自然语言的传统算法，包括分词、句法分析、语义分析等分析，往往组织成一个串行的流水线作业步骤，后一个步骤的正确性完全依赖前一个步骤计算结果的正确性。如果在分词环节出现了5%的错误率，那么在后续的运算过程中间，其实无法去纠正这些错误还会逐级放大错误。

深度语义的理解处理，则是一体化的运算处理过程，对于上述的分词、句法分析、语义语言转换，语义分析等进行一体化的综合处理，并没有严格的先后次序，最终选择综合权衡打分最高的结果作为最终结果，这也是符合人类大脑解决问题的本质，最终能得到最理想的理解结果。

## 生成

在深度语义模型体系下，自然语言的生成则是理解的反向过程，也就是将要表达输出的语义结构生成一种自然语言的形式。

深度语义模型里的所有知识，甚至包括规则、推理等，都可以采用自然语言的方式进行表达和输出。

并且，计算机真正知道自己说出的话语是什么含义。不会象主流技术那样，只是从预先存储的自然语言 “知识库”中选择来直接输出，这种方式下，计算机看似也说出了话语，但对自己说出的话语的含义其实都茫然不知。

## 记忆

在深度语义模型体系下，记忆就是把深度语义模型的语义和语言信息结构进行存储，在需要时再进行加载调阅。

深度语义模型的信息在内存里的形式和在磁盘等永久存储的形式是一致的，存储和加载都非常简单。并且因为在纵向上具有层级划分和在横向上有领域划分，还可以根据按需加载。

原则上，深度语义存储的是语义结构，而不存储语言形式，在使用时根据需要再动态地将语义结构生成为具体的语言形式。

## 查询

在深度语义模型体系下，查询算法或者说搜索算法，也是围绕深度语义模型来进行的。是基于语义的模式匹配运算，而不是常规的基于文本关键字的搜索运算。匹配的结果首先完全是语义结构，根据需要再动态地将语义结构生成为具体的语言形式。

## 推理、计算、思考

如前所述，深度语义模型描述了各种运算的规则和知识，直接支持各种计算、推理、思考等认知处理。

这些高级演算处理仍然围绕同一个语义模型为中心，运算主要都在语义模型上完成，只是输入输出采用不同自然语言的形式。这种方法显然符合人类大脑思维的方式。

例如：首先用中文告诉计算机一些知识“这一年，王菲嫁给摇滚歌手窦唯…”，然后用英语询问计算机“Who is Douwei's wife?”,计算机可以计算并回答“Wangfei is Douwei's wife.”。

在深度语义模型体系支撑下，这些演算的能力可以由浅到深不断进化，最终将实现完全的自动复杂演算，逐渐摆脱用硬编码来解释和驱动各种智能业务的状态。

## 学习

在深度语义模型里，学习就是围绕同一个深度语义模型增加或者更新知识的过程，增加或者更新的知识都是相同的语义结构。

学习分多种方式，最核心的知识概念需要预先手工构建，而基础知识和外围知识就可以通过手工、半自动、全自动的方式来学习和构建，甚至实时学习、随用随学，支持以自然语言的形式和素材来直接学习。

深度学习和大数据等技术在语义信息表达上缺乏深度，所以不能独立解决自然语言处理的难题。但在构建好深度语义模型的基础上，则可以发挥深度学习和大数据等技术的特点，围绕深度语义模型进行大量外围知识的自动学习和收集。

## 其它处理

自然语言处理还有其它很多处理技术，例如情感分析、领域分析、信息提取等。在深度语义模型的体系里，它们都只是特定的一些语义信息维度，将在同一个体系中和别的运算结合在一起一体化处理。

实际上，人类大脑处理问题就是将所有这些理解、生成、记忆、计算、查询、推理、情感等各种各样的处理融合在一起综合处理的，这些信息和处理本身也会相互影响、相互描述、无法分开。试图把各个维度的信息或者处理分离出来单独处理，最终结果肯定是缘木求鱼，只有将它们融合在一起形成一个完整体系进行一体化处理，才能得到理想的结果。而要这样做，就必须围绕同一个深度语义模型为核心才能成为可能。

# 为多种AI业务应用提供统一的表达信息表达模式

在作为自然语言处理技术的统一信息表达基础的同时，深度语义模型也就成为了和自然语言处理相关的众多业务应用实现的统一基础。

在深度语义模型体系下，各种业务应用都围绕深度语义模型为核心，而具体运算处理都遵循三个环节：输入环节将输入的自然语言形式的信息或者请求理解成为语义结构；运算环节则是根据输入的语义结构结合语义知识库进行查询、变换、计算、推理等各种运算处理，形成新的语义结构；输出环节则将运算处理形成的新语义结构生成为自然语言形式进行输出。对于不同业务应用来说，输入和输出环节的处理基本相同，区别主要体现在中间的运算环节也就是对语义结构的变换处理有所不同。

对于机器翻译，中间的语义结构基本不做变换，只是输入和输出对应的自然语言语种不同。

对于问题查询、语义搜索，主要是对用户用自然语言表达的问题请求理解为语义结构后，到语义知识库中进行查询匹配，并对匹配结果进行整理后再生成自然语言输出。其中用户问题的语言、语义知识的原始语言、结果输入的语言三者可以是完全不同的语言。

对于人机对话的业务，中间的语义结构变换处理环节更为复杂，需要同时对查询、计算、推理甚至学习等同时应用，最终也是对中间处理环节最终产生的语义结构生成自然语言回复给用户。

而其它业务处理也都类似，这样，机器翻译、语义搜索、智能助手、智能客服、情报搜索、信息摘要、文章编写、专家系统…等各种自然语言相关的AI业务应用，都可以围绕同一个深度语义模型和相关的算法体系来进行。

这些不同的业务应用之间可以轻而易举地通过同一个语义模型和标准实现充分的信息互通和共享，还能将不同的业务应用融合形成全新形态的一体化业务应用。

更重要的是，既然这些不同业务应用的知识和信息采用相同的模型和标准，就可以汇总在一切形成统一的语义化知识和信息库，在更广的范围内解决各种知识和信息统一表达和共享。

# 技术要点

## 整体模型由语义和语言两个模型构成，语义和语言彻底分离

**深度语义模型是核心！**

语义模型是本质，语言模型是语音模型的外在表现形式，语言存在的唯一意义就是为表达语义而服务。因此在研究语义模型时，一定要排除语言信息的干扰，摒弃掉千变万化的语言形式，抽取并构建出一个“纯”的语义模型，这一点非常非常的重要！最终的这个语义模型将不依赖于任何语言，能独立完备地描述各种知识和信息。而在需要时又可以转换表达为任意语言，或者将任意语言理解为相同的语义结构。

## 语义模型采用一个统一的DSM结构，来表达各种知识和信息

语义模型采用一个统一的DSM结构，表达能力远强于目前类似知识图谱的三元组结构。任意的知识和信息——无论是名词动词、数量程度、事物事件、关联关系，还是衣食住行、社会关系、情感交流、自然科学、数学运算、推理思考，只要能用语言文字表达的概念知识和思维处理，都可以用这同一个结构来进行表达和描述。

## 语义模型构建采取自顶向下方法，形成金字塔形的知识和信息结构

这其实是遵循了一个非常基本的一个道理——“人类的知识体系是分层级的”。越高层的知识就越基础越重要，是作为理解和表达更低层知识的依据，而越高层的知识的数量其实是越有限的。越低层的知识的数量就越多甚至无限，但都可以用高层的知识来解释。

## 语义模型的可解释性和可计算性主要由最顶层的核心层来体现

最顶层的概念是核心层，人类思考问题的最基本的一些原则和思想，就是数量非常有限的一些基础概念，例如【离散化】【概念】【空间】【时间】【数量】【程度】【关联】【存在】等基础概念，并扩展出数百个核心概念，包括事物、事件、关系、规则、公式、集合、区间、范围、比较、序列、性质、事实、目标、条件、结果、过程、愿望、情感、动作、行为…等等。

深度语义对核心层的描述定义，绝不仅仅是把这些概念名词罗列出来，而是构建起他们之间本质的、完备的、具有说服力的内在关系，这种关系，足以支持它们之间的相互表达、相互解释，以及进行相互转换运算。构建起了深度语义描述一切知识和进行一切运算的理论基础。

## 下层的大量的知识和信息由顶层知识解释，可以无限扩充

在核心层构建起来后，就可以自由地描述和扩充更下层无限的知识和信息了。这些知识和信息还可以概括地分为两层：

第三层是基础层，例如：很、不、质量、年月日、商品、价格、这、个人、男孩、城市、国家、手机、水果、飞行、结婚、夫妻、预订、辞职、购买、出售、选举、表达、宣称、询问、回答、命令、喜欢、满足、高兴、希望、货币、西甲、英超、演员、责任、品质、存款、汇率…等等。

基础概念是深度语义的骨架，这一层的概念知识构建起来后就能对常规的句子和文章进行顺畅处理。而且，基础概念数量少但频次高，实际使用的概念中95%以上都属于这些基础概念，后述的外围概念虽然为数众多但实际在每个场景中出现的占比其实很低。

第四层是外围层，就是大量非常具体的概念，这一层的数据量非常大，但处于整个深度语义知识体系的外围，没办法也不要求完全穷举。深度语义的强大之处在于：在核心层和基础层的概念和知识完备起来后，系统实际上具有动态学习新的外围概念的能力，可以随遇随学，实时解决——和人类的方法其实是一样的。

## 和其它知识和信息组织表达技术的比较分析

如果将深度语义的这个体系和【知识图谱】【大数据】等技术要做一个比较的话，可以用一棵树来进行比喻。

深度语义的体系是完整的一棵树，顶层的哲学概念是树根；核心概念是树干和一些主要枝干；基础概念是更细的枝干；而外围概念是树的大量的叶子。树干很粗，但只有一根，但其实对树来说是最重要的。树叶虽然数目众多，但必须依附于树枝和树干。

【知识图谱】和【大数据】等技术的问题就是没有构建起树干和枝干等基本的知识体系——即使有一些构建也非常简陋和不完整，然后就去收集大量信息（树叶）放入到库中——这些大量的信息其实就对应于深度语义模型里的外围层信息。这些大量的信息只能解决粗浅的查询匹配而不能自动解决需要深度思考的问题，一个知识图谱动辄号称数十亿条的知识量，非常庞大，但并未见到就产生了足够的智能来真正实现和人类进行几轮真正有意义的对话。因为缺乏知识深度，目前的一个有10亿条知识（三元组）的知识图谱，和一个传统的有10亿条数据的关系数据库相比并没有什么本质的区别，知识图谱采用的三元组提供了比关系数据库更灵活的信息表达手段，但目前的知识图谱都并没有真正充分利用起这种灵活表达的优势，只在很细的树枝层面构建了一些针对于具体领域的局部化的描述体系，而不能构建起全局和完备的语义描述体系。

数据量的“大”当然不是错误，这其实就是计算机比人的能力更强的地方。但光有大是不够的，现在需要的是“深”，也就是说人脑智能的精髓所在。假如解决了“深”的问题然后再结合数据量的“大”，那么所有的这些大量的数据久都能利用深度语义的高层次的知识来进行解释和运算，数据的价值将成几何级数的放大，最终构造出能力甚至超过人的机器智能，就像AlphaGo那样。

## 统一的语义模型可以和各种自然语言进行转换

在严格遵循语义模型和语言模型分离的基础上，同一个语义模型可以和各种自然语言——包括不同语种和同一语种的不同说话方式进行灵活的转换。也就是说，一组相同的语义知识或信息可以转换生成为任意的一种自然语言形式进行表达，反过来，这些自然语言形式又可以理解转化为相同的语义知识和信息结构。简单的说“就是用语言来进行外在表达，而用语义进行内在表达、描述、运算、思考、存储”。

以前人们曾经设计过所谓的“中间语言”来尝试对各种语言进行汇总，尤其希望解决机器翻译的问题。这种思路的缺陷是：这种中间结构是以语言的概念为基础，跳不出语言概念的束缚，所以最终都失败了。而深度语义的这种中间结构，完全以语义概念为核心，彻底剔除掉语言形式的“噪音”干扰，构建起完整、严谨的语义体系。然后再和语言形式进行匹配对接。终于彻底解决了这个问题。

迄今为止，我们可以看到机器翻译、人机对话、搜索引擎、信息提取、甚至包括知识图谱在内，都没有实现同语义多语种的统一融合处理。有一个问题值得思考：机器翻译现在号称已经达到了非常高的准确率，但是为什么不能结合这个技术来非常简单地实现一个语音助手、或者搜索引擎的多语种版本实现呢？这其实就是针对不同的问题采用专属化的解决方法以及追求所谓极简主义的“端到端”的理念方法弊端的集中表现。

## 知识和信息表达一体化，运算和记忆一体化

前边反复提到了知识和信息，看起来是两种概念，但是对于深度语义体系来说，知识和信息实际上并不是截然分开的两种数据。如果一定要界定的话，可以粗略地把深度语义的高层的核心层和基础层里的数据看着为知识，而外围层的海量数据则可以看到信息。知识和信息的表达方式其实并没有任何的差异，都是用前述的统一的DSM的方式来表达，这种高度一致的表达方式，可以保证非常强的灵活性。

同时，深度语义在运算时内存里的表达结构和进行记忆存储的结构是完全一致的，内存运算出的结构存储起来就是知识和信息，从磁盘上加载的集合信息，直接就可以参与任意的语义运算。并且基于多层级的金字塔结构使得深度语义运算可以选择性的加载结构，因为低层的信息依赖于高层的知识，而高层知识不依赖于低层信息，因此，在进行特定领域的运算时，核心层和基础层的知识必须预先加载，而外围层的信息，就可以根据需要而选择加载。

在具体存储时，理论上深度语义倾向于存储语义结构信息，而丢弃掉原始的自然语言字符串信息，语义结构信息在任何时候都可以任意选择任意语言形式再生成给人类看的自然语言自负串。只是由于深度语义结构的信息比较复杂，需要占用的存储空间比起原始的自然语言字符串信息要大得多，为了优化，可以根据实际情况将一些不常用的信息采用自然语言的方式存储，在需要计算的时候再加载并重新理解形成语义结构信息。

## 知识和信息表达采用模糊逻辑

传统的基于规则的系统不能很好地解决实际问题的原因，一方面是知识和规则缺乏层级关系，另一方面就是知识、信息、规则的定义是二值逻辑，非黑即白，在灵活多变的环境下就总是遇到不能解决的“例外”。

深度语义模型采取是模糊逻辑来解决这个问题，所有的知识、信息和规则的表达都体现为0～1之间的一个小数取值，可以有效解决信息不确定性、歧义处理等难题。

## 运算结果采用综合化评分处理

自然语言的传统算法，包括分词、语法分析、语义分析等分析，往往组织成一个串行的流水线作业步骤，后一个步骤的正确性，完全依赖前一个步骤计算结果的正确性。如果在分词环节出现了5%的错误率，那么在后续的运算过程中间，将没有办法去纠正这些错误，甚至会逐级放大这个错误，最终造成了任务的失败。

深度语义的处理，是一个一体化的运算处理过程，对于上述的分词、语法分析、语义语言转换，语义分析、语用分析、推理运算等进行一体化的综合处理印子，并没有严格的先后次序，最终选择综合权衡打分最高的结果作为最终结果。

这种方法其实和深度学习比较接近。每一个环节、每一个节点，都不保证绝对的正确性，而最终结果是各环节各节点综合运算和决策打分的相对最有结果。而这其实也符合人类大脑解决问题的根本办法！

## 对于自然语言理解的回答——深度语义网络

现在我们可以回答前边提出的那个问题“对于自然语言理解，应该产生一个什么样的结构才算是真正的理解。”。这个答案就是——用这一套深度语义理论和模型构建的深度语义结构，具体来说，就是由很多前述的DSM结构构成的深度语义网络，我们有信心相信，这样的一个解释和这样的一个模型结构，在不远的未来，应该能为大家所公认接受。

这里稍微描述一下细节，对于任何自然语言表达的任何句子、段落、文章进行理解以后，都将形成上述的深度语义网络结构。而这种理解结构的深度和详细程度，是可以根据需要扩展和细化的。通常来说，对于一个大约20个字的自然语言句子，进行初级的深度语义理解以后，可能会形成数十条DSM结构形成的网络，而在进行高级的深度语义理解以后，可能会扩展为上百条DSM结构形成的网络——不但描述了句子的字面意思，而且描述了所关联的前因、后果、情感等附带的语义信息。这绝不是为了复杂而复杂，实际上是因为人类语言蕴藏的语义知识和信息本来就是如此复杂。而我们不能回避这种复杂，必须把它探究并表达完整，才能够搞清楚事情的本质，最终才能让一切的事情都变得简单。

## 形成自然语言处理的完整体系

在回答了关于“自然语言理解”的问题基础上，我们可以探索更高级的“自然语言处理”的问题了。围绕深度语义理论和模型，不仅仅是实现理解，而是可以实现包括理解、生成、存储、查询、计算、推理、思考、决策、行为、学习等各种业务在内的完整自然语言处理的完整体系，形成一个闭环。因为有了统一的深度语义模型，就有了让所有这一切运算处理得以衔接起来的一个基础。所以，无论如何强调它的重要性都不为过！

## 各种业务应用一体化

现在来看一看，由于没有统一的这样的深度语义模型和理论，目前解决不同业务问题大多采用分散性的方法处理，针对一个业务应用，采取一种只能适合（也就是大致适合）这个场景的特定方法，搜集特定的标记语料，然后然后进行针对性的训练来解决问题——其实未必能真正解决问题。

对于机器翻译需要找很多的双语语料进行训练；对于信息摘要分析要找很多人工标签的语料进行训练；情感分析运算又要找其它打了情感信息标签的语料进行训练；对于闲聊机器人又要寻找很多聊天对话的语料进行训练；对于不同语言的聊天机器人同样也要找不同语言的聊天对话语料来进行训练。假如要实现一个双语对话聊天的功能呢？这种方法显然又要去寻找一个人用中文一个人用英文聊天的语料来进行训练——这可有些困难…。根本的原因就是没有解决根子上的问题，而试图收集大量的已有语料然后以一种具有一定灵活性和模糊性的匹配算法去匹配并映射以前类似的文字，而根本不想深入研究内在原理。按照Judea Pearl的说法“仅仅就是一种曲线拟合，并不能从根源上解决问题”。人类根本不是采用这个办法！人类从来不会单独去学习如何分词、如何学习聊天、如何学习情感分析，其实这些知识都是交织在一起同时学习，也是同时应用的。实际上所有的这些技术也都应该并且也都最终能融合在一起的，所有的问题都能从根本上统一解决并且得到巨大的提升。

一旦完整的深度语义的理论、模型和算法体系构建起来，就有望可以统一的方式去实现或者提升相关的这些业务，例如：机器翻译、语义搜索、智能助手、智能客服、智能机器人、情报搜索、文章缩写、文章编写、专家系统、语义发布、语义网、人类知识库。当然，需要做的工作内容非常多，但重要的深度语义模型指明了正确的方向，构建好了稳固的基础。

# 深入分析

## 深度学习解决感知智能， 深度语义解决认知智能

人工智能的重要特点就是要研究和模拟人类大脑的工作方式。人类大脑非常复杂，具有强大的能力，能处理各种各样的问题，其中重要的原因之一是大脑对信息处理采用了多层级的方式。初略地，可以划分为明显的两层：

* **初级信息处理层级：**

人类通过眼睛、耳朵、鼻子、皮肤等感觉器官，收集客观世界的原始信息——包括图像、声音、气味等，然后传递给大脑。大脑对这些原始信息进行加工处理，例如：从视觉图像信息里提取出道路、树木、动物、人脸等概念；从声音信息里提取出自然界声音、人类话语等概念。这个过程可以称为【感知智能】或者说是【形象思维能力】，是大脑的初级信息处理能力。说其是初级能力，一方面是因为其处于信息处理的外围，是对原始信息进行初加工；另一方面是其原理相对简单，所有的动物都具备并且天生就具备这个能力，证明实现这种能力的算法可以直接内置在初生大脑里。

* **高级信息处理层级：**

是人脑对初级处理后的低级概念进行深度处理，以建立更高级的概念，形成知识网络并实现思考、推理、策划等思维的能力。这可以称为【认知智能】或者【概念化思维能力】。这是只有人类大脑才具备的高级能力，其它任何动物终生都无法学会，即使是人类自己也是需要进行长期的学习才能逐渐掌握。但这个过程可以反复地迭代、抽象，以致于可以形成数学、物理学、哲学等高度抽象的概念和知识体系。这其实就是奠定人类在地球上的统治地位的关键能力。

现阶段的深度学习技术，比较适合解决上述初级信息处理的问题，即“感知智能”，并不适用于解决高级信息处理的问题，难以形成“认知智能”。要真正实现“认知智能”，就需要深度语义深度语义，也就是“语言-语义-知识“一体化技术和方法，具体分析如下：

现阶段的深度学习技术主要解决的问题可归纳为：输入大量初级信息，然后进行信息简化、抽取和分类。其特点是：输入的信息量较大，但比较初级、比较扁平，且信息之间局部相关而不具备全局相关性；而经过算法处理后输出的信息量则较小。总的来说只是一种【分类】方法而不涉及推理等，很难输出信息量较多的复杂结果。可以看到深度学习取得成功的领域都是输出简单结果，比如：图像识别可以从一幅数十万字节的图像中识别出“猫”这样几个字节的结果概念；语音识别从几十K字节的声音信息里边分析出十几个文字的串概念；以及AlphaGo根据数百字节的当前棋局信息分析出下一步的选点（x，y这两个字节）等。在需要输出复杂结果的领域的效果都远远谈不上成功。

反观“认知智能”面向的则是一种高度概念化的信息结构，和图像、声音等初级信息有非常大的不同，似简实繁。尤其是分析后需要得出的结果非常复杂！绝不仅是信息抽取和分类方法就能实现。

简单分析下一个普通的自然语言句子：

“昨天上飞机前，张燕在上海机场免税店花了500元钱买了一件曼妮芬的浅蓝色丝质睡衣。”

表面上看，这个句子只有30多个字，计算机存储也只需要几十个字节，其结构就是简单的一维字符串。但人对这个句子进行理解以后，实际上将在大脑里边形成一个非常复杂的信息结构，这个结构至少包含着：

1、【昨天】、【上飞机】、【张燕】、【上海】、【机场】、【免税店】等20多个显式概念和【价格】、【人民币】、【品牌】、【颜色】、【材质】等比显式概念更多的隐藏概念；

2、上述这些概念各自的类别、性质、行为；

3、上述这些概念相互之间的关系，结合起来形成的整个句子完整意思；

4、产生的联想，比如【张燕的财产减少了500元】，【张燕拥有了两件睡衣】等；

上述信息量非常多且需要有机融合形成一个具有全局关联特征的复杂概念网络结构，这个结构可以看着就是对自然语言句子“理解”后的输出结果。如果计算机能模拟人脑一样较准确、完备地形成这个结构，就在很大程度上回答并解决了“怎么定义计算机真正理解了一句话？”这个业界一直争论不决的问题。而计算机不能形成这个结构或者形成结构有缺失和错误的话，就不能说真的理解了句子，从而也绝不能做出像人一样的正确反应。

从上述的分析可以总结，高级的“认知智能”面对的问题是——显式输入的信息少而简单，而输出的信息比输入的信息大很多而且是复杂的网状结构。仅仅只是理解就非常复杂，若在理解的基础上加上推理、思考、对话、执行等任务，难度将更为巨大。现有的以分类为理论基础设计的深度学习算法无论如何不能从本质上满足这种处理要求。现在深度学习依赖的深度神经网络自身是多层级网状结构，但处理的信息对象则是如上所述的图像、声音等简单数据，和“认知智能”需要解决的复杂网状概念和知识结构不可同日而言。

还应该看到：人类的“认知智能”或者说“概念化思维能力”和“语言”实际上密不可分、紧紧地纠缠在一起。要解决自然语言的所有问题，就是要解决人脑的高级思维的问题。而解决了这方面的问题，自然语言问题也就顺利得到解决。从表象上看，自然语言的信息量少而很简单，但对看似这么简单的数据进行处理却是人工智能领域最具挑战性的问题，因为很少的字符串只是表象，就好像冰山浮在水面上的部分；本质的信息其实是隐藏在后面的整个概念网络体系，就好像冰山在水下的部分。以一副图像和一段文字来做比较：100万字节大小的一幅图像蕴涵的信息量，显然远远小于一段100万字节的文字所蕴涵的信息量；人类大脑可以用几秒钟时间分析清楚这幅图像并识别出所有的概念实体，而阅读并理解这段100万字节的文字，则至少要花上数天的时间…

在深度语义出现以前，因为无法真正攻克深度语义涉及的一系列难关，所以当前主流方法是回避深度语义处理(因此业界都不提“深度语义”这个词)，将自然语言处理和认知思维问题分开，将自然语言处理强行简化为串对串的算法问题。无论对于机器翻译、搜索引擎还是人机对话，都只抓住输入串和输出串这最终两端，而完全忽略中间的深度语义概念表达、理解、推理、思考，试图让现有的多层神经网络以黑盒的方式去自动学习就能（像简单的图像识别和语音识别一样成功地）模拟出认知智能，实际上就像希望“一场龙卷风刮过，就将垃圾场上的垃圾拼凑成一架飞机“一样，不可能得到满意的效果。例如：

1、机器翻译领域，现在主流方法就是不去考虑理解意思，把输入和输出都看着串，试图利用大量双语语料库对照和概率理论来进行串对串的匹配，整句正确率极低。

2、人机对话领域，因为涉及听说双方的思考过程，比起机器翻译来说中间增加了更多的层级，所以难度更大。现在的产品分两种应用模式：一种是只能按照规定好的“句型模板”去导向发短信、搜索、购物、查天气等数十种固定应用——微软、谷歌、苹果等巨头都推出了这方面的产品。能很好实现这些应用当然也很好，但受限于没有完整的深度语义模型，这些应用的实际效果和智能化程度离理想状态还差得非常非常远…；另一种干脆就定位于闲聊，用大量人类聊天记录进行匹配拼凑，将错误百出造成的笑料作为卖点，不具有人类沟通的连续性、逻辑性和实际意义，短期能吸引眼球，但不具备长久生命力。

3、舆情分析等文本处理应用将目标简化为对文本进行简单分类（比如“褒义、贬义”“好评、差评”“涉恐、安全”等）和提取关键字等现有大数据方法有能力实现的方法，所以能满足一定的应用效果。但这样简化和阉割后的信息分析的价值也就很小了，实际价值比起能真正深度、完整地进行语义分析的潜在价值都不到百分之一。而一旦要实现更有价值的应用比如“文章缩写”等，就遇到同样绕不过去的深度语义难题，无法实现。

归根结底，要真正解决“自然语言处理”、“认知能力”以及 “概念化思维能力”这种人类大脑的高级能力，无法回避建立完整的概念知识体系和演算问题！也就是必须要构建深度的语言-语义-知识一体化理论和技术体系。这是现阶段人工智能领域的最核心问题之一。

而好的一点是：这样做虽然难度大，但是直面问题的本质，是一劳永逸的方法。一旦建立起这个体系，那么就不只是解决了对自然语言的理解，对自然语言的各种处理以及类似人脑的计算、推理、思考等能力都同时得到实现！那将是人工智能领域的一个重大突破，将形成整个人工智能领域里最重要和最关键的基础设施之一。

而深度语义技术就是解决这些根本问题的正确方向。

## 概念化思维在人工智能领域的核心地位

前边描述了：大脑的信息处理分不同层级，其中高层级的智能主要就是概念化思维的能力，它在人工智能体系中处于最核心的地位。

这里用一个草图来表达：

**2概念化思维**

知识归纳

语言处理

逻辑推理

行为决策

……

**1感知能力**

声音感知

图像感知

触觉感知

嗅觉感知

味觉感知

体感感知

……

**3运动能力**

行走

操纵

说话

……

……

**传入**

**指挥**

深度学习技术

DSM深度语义技术

**感官**

眼睛

耳朵

皮肤

耳蜗

…

**行为器官**

四肢

口

……

传感器技术

机器人运动技术

将整个人体看着一个智能体系，可以看到有3部分功能，分别就对应着【输入】-【处理】-【输出】。人类的行为也可以看着三个过程：1、获得输入：通过各感官获得图像、声音、触觉、味觉、体感等信息并进行初步处理——主要是进行关键信息的抽取和归类，然后传递给概念化思维层——大脑核心。2、进行处理：大脑核心根据获取到的输入信息并结合概念性记忆，进行分析、思考、推理、决策，然后实施行动。3、输出行动：大脑核心形成的概念化命令，传递到肢体具体化为各种行为，如行走、操纵汽车、说话等。

模拟人体的智能体系，需要深度学习、传感器、机器人等不同技术来实现不同部分的任务。其中第2部分也就是概念化思维能力是协调和操控其它部分的中心枢纽，位于整个智能体系的最高层和最核心。可以这样界定这部分的范围：这部分的智能思维活动，就是人们清晰地意识到自己的大脑里边在“想”的内容——如果需要的话这些内容能清晰地表达为文字展现出来。例如对一个人看到一个商贩在卖桔子，上前去问“桔子多少钱一斤？”，这个过程的要点分析如下：

首先对外界信息进行初级信息处理，比如对眼睛看到的图像进行分析判断出“那是一个商贩”、“那是桔子”等概念，这都可以归于感知层的工作。但最终归纳得到【我看到了一个商贩在卖桔子】的概念化信息，传递到核心的概念化思维层，后者就开始进行一系列概念化思维活动：【我看到了一个商贩在卖桔子】->【我口渴了】->【吃桔子可以解渴】->【我希望得到桔子】->【可以通过购买桔子得到桔子】->【我要购买桔子】->【购买桔子要有钱】->【我有钱】->【购买桔子要先知道价格】->【我要去询问价格】->【执行走上前去询问价格的任务】…

这些过程步骤很多，而且可能存在很多不同分支和路径。但不管如何，每一个步骤都可以清晰地概念化，并且可以用清晰的文字进行表达出来！这也就是人类智能里最高级的概念化思维能力。而要实现这种能力就需要深度语义这种能模拟人脑概念化思维的技术…

## 深度语义实现语言、语义、知识等概念化思维的一体化处理

深度语义是一种模拟人类概念化思维的方法。

深度语义的基础是实现了语言、语义、知识的一体化处理。语义和知识是深度语义体系的核心，语义是就是自然语言蕴含的含义，各种各样的语义汇总在一起就是知识。深度语义的知识体系是用概念化的方式来描述世界的各种事物、关系、事件、规律等本质，因而脱离具体语言而自成体系，同时又能表达为任何一种自然语言。

深度语义的语义和知识体系比传统的语义网络、知识图谱等技术，在知识深度、完备性、自描述性、灵活性和可计算性方面都要完善得多。是真正能实现对人类各种概念化知识（能用文字来表达的知识）进行表达、演算的完整体系。实际上，这个体系完全超越了表达语言内在含义的任务，成为了模拟实现整个概念化思维的基础载体。

深度语义的语义和知识可以用任意语言的形式表达出来（事实上任何一种人类语言存在的唯一意义都是为了表示语义知识）。比如前述的例句中隐含地关联着一条知识【上海是一座城市】，需要对这条知识输出表达时就可以生成并输出为各种语言形式比如“上海是一座城市。”、“ShangHai is a City.“。反过来也能用自然语言的形式来输入和学习知识，比如，输入”广州是一座城市，它的英文名称是‘GuangZhou’。”这样的自然语言句子，深度语义系统就能理解成类似【广州是一座城市】、【广州的中文名称是广州】、【广州的英文名称是GuangZhou】这样的语义知识并且存储起来。新学到的知识和系统里已有知识的结构是完全相同的！在今后需要时就可以进行统一运用，比如遇到“广州的天气怎么样？”或者“What's the weather like in guangzhou?”等句子，系统都能运用已经学会的【广州是一座城市】等知识来达到对句子的深度而精确的理解！并在理解的基础上实现查询、搜索、推理计算等智能处理。

需要指出的是，深度语义对于将自然语言“理解”成为语义结构和将语义结构“生成”为自然语言这两个相反的算法采用同一套规则和数据。实际上， 深度语义首先描述了语义概念“生成”自然语言的规则，而“理解”时是对该规则进行反方向的运用。

推理和思考是深度语义更为强大的能力，因为深度语义语义和知识的体系打造得很深、很完备，在此基础上实现推理和思考能力非常自然。简单来说，深度语义的推理和思考实际就是对各种概念化语义和知识进行各种变换的过程，更具体的内容在本文后边有例子表述。

深度语义体系可以简单归纳为：**以语义为内在存储和思考，用自然语言对外表达和沟通** 。

### 深度语义知识库的结构定义

下边对深度语义知识库结构进行一个简单的解释。先来看一个根本问题：如果让计算机来模拟人脑对自然语言句子进行理解。那理解后的正确结构究竟应该是怎么样的？

可以想象，对于人脑来说，听懂一句话后人脑内部肯定将产生新的一些神经元连接。而对于计算机来说，就是要产生一个新的数据结构，和大脑产生的新的连接形成一种尽量等价的效果。那么这个结构应该是怎么定义呢？这其实就是解决自然语言处理和认知智能所要应对的基本问题！

在深度语义之前就存在着很多这类结构定义，但都存在种种问题。深度语义的语义和知识结构定义模型，经过了长期的精细设计和完善，对这些问题进行了较好的解决。

深度语义理解自然语言句子后，将产生一个多层级概念网状结构，结构中的每一个节点都是和具体语言无关的语义概念，语义概念之间通过各种基本关联连接成为一个整体。

列举一个深度语义理解的实际句子：“2016年6月的一个周末的下午，在中国北京长安街南面的一家如家酒店里，25岁的职业经理人王伟一边啃着一个苹果，一边在自己新买的新款大屏苹果上浏览一则新闻：‘由于华为、中兴的快速成长，苹果的股票大幅度下跌超过20%’。”这个句子，有107个字符约200个字节，深度语义对这个句子进行理解以后，会产生一个200多个语义概念节点连接而成的复杂语义概念网络，其中每个节点都和系统内部知识库的成千上万的概念知识存在紧密联系。达到这样的理解程度，我们可以说“深度语义较为深度、完整地理解了这个句子的含义”，从而可以开展后续的各种处理。这是目前其它的技术方法都没有实现的，却是深度语义模型的重要特点。

深度语义里的语义概念是能用语言化表达的一切事物，每个概念在人类社会中具有一致的定义和标准，这是人和人进行信息沟通的基础。具体地，在深度语义体系中的每一个概念都用数字来进行标识，比如对于【水果苹果】这个概念，系统用一个数字如【1357】来表示，而对于【ipone手机】，系统用一个数字如【2468】来表示。

每一个语义概念都对应着多种语言的表现形式，甚至对同一个语种也可能有多个不同的形式表达（比如别称、简称等）。比如上述【水果苹果】的唯一数字标识是【1357】，而对应中文的语言表示是 “苹果”、“平安果”，对应英文的表示是“apple”，对应的日语是“リンゴ”，对应维语是“alma”等等。而上述【iphone手机】的唯一数字标识是【2468】，其对应的中文表达是“苹果手机”、“苹果”，对应的英文则是“iphone”。

多个不同的语义概念有可能具有相同的语言形式，比如【苹果水果】和【iphone手机】在中文里边都有可能表达为“苹果”，这样，在理解的时候遇到字符串“苹果”就不能简单确定要说话方要表达的是【苹果水果】还是【iphone手机】，这就是自然语言处理里最为纠结的歧义问题。深度语义解决这个问题方法是参考人类解决问题的方法，也就是利用语义相关性来排除歧义，能真正彻底解决这个难题，具体在后边有较为具体的阐述。

需要重点说明的是，深度语义体系里语义概念的范畴很广，并不仅限于表现水果、手机这样具有形态和质量的事物，实际上，能用语言文字表达的一切信息（例如：数、量、时间、空间、程度、存在、关系、事件、数学公式…）在深度语义里都表达为实体化的语义概念。因此，可以也必须能理解和处理任意句子里的每一个字符、每一个词汇、每一个短语以及整个句子的完整意思，而不像其它方法那样只抽取能识别的关键词并丢弃不能识别的词汇。也就是说， 深度语义形成了完备且能自我描述的概念化知识表达和演算体系，能实现真正的深度理解和思考——这也是深度语义比起旧有的基于规则系统更为强大，并能真正解决问题的关键能力之一。

### 深度语义如何解决大量知识的构建问题

深度语义是以知识为中心的体系，大量知识的构建和完善是非常重要的基础工作。深度语义对知识的积累和人类一样，采取了层级结构的方法，比起盲目求“大”的一些大数据方法更为有效。

与深度学习联系在一起的一个概念是“大数据”，这也是用得很广泛的一个词语。不可否认，数据是包括AI在内的计算机科学的基础，没有一定量的数据就不能研究出具有普适性的规律和知识。但，数据真的是越多越好？在追求数据量大的同时，是否还应该要考虑到数据的有效性和正交化特征，尤其是层级特征呢？

有深度学习的专家说，只要数据量足够大那么就能解决所有的问题。但这是一个伪命题，就好像是阿基米德说“给我一个支点，我能撬起整个地球”一样，理论上成立而实际上没有意义。一方面且不说数据远远无法穷举；另一方面来说，数据量越多时，虽然有效信息越多但无效信息也会增多，如果无法识别有效无效那么数据再多也无用。而信息真正有效的话，只需要一条就够了——例如爱因斯坦的e=mc2就概括了世界上一切能量形式变换的本质。而人类大脑就是先从学习基础知识，然后再来处理外界的海量信息，而不是简单地以数据量取胜。人类终生会学习到很多知识，但学习每一条知识并不需要非常大的数据量，很多学习都从很小的样本中就能分析出关键特征。比如，现在的图像领域的深度学习，可能要用几十万张图片去训练才能很好地识别一只“猫”。而对于人类大脑，往往只需要一张或者几张图片，就足以学会，其中原因是人脑能将概念化知识的“层次性”“共性”“差异性”几点很好地结合在一起。例如在学习“猫”的特征前，其实学会了“动物”、“哺乳动物”等的相关特征，在看到“猫”的时候，实际上会充分和已有的知识对比，用共性去识别出大的类别，然后用发现的少量个性差异来区分小的类别。

看清楚这个问题后，我们可以认为，即使是图像分类识别这样的初级信息处理和各种深度学习方法，如果能充分运用概念和知识的层级组织，也一定可以在现有的基础上大大提高效果和效率。而深度语义研究的本来就是高度概念化的知识，对知识的层级组织要求就更加严格。根据这样的原理，深度语义的知识体系的构建，并不需要无限量的大数据，而是把数据和知识划分不同层级，对不同层级采取不同的处理手段。只有采用立体层级化的信息组织方法，才能像人脑一样，利用有限的存储，去实时解决无限的新知识问题——一个旅行者并不需要背下全世界所有的地名、风俗才敢出去环游世界，他只要具有运用已有知识，实时地解决和处理每天遇到的新问题即可。

知识是无穷尽的，而对已有的知识建立成基本的体系，就能够顺利地运转，随时可以用已有的知识体系来解释未知的知识；同时，提取基础、核心知识，加入到基础知识体系中去。因此说，盲目追求大数据量并无意义，几十万块顽石的价值加起来也不如一块钻石。大数据，如果只是求大，而不能精准地对每一条信息分析、加工；只是做概略地统计、归类等工作的话，那么这些工作确实是计算机所擅长的工作且已经在很多领域发挥了巨大价值，但这并不是人类大脑处理问题的方法，并不能体现出人脑灵活、全面、智能的优势能力，并不适合归于人工智能范畴。

深度语义里边的知识有着不同的层级。越基础的知识越重要。

例如，分析下列三个句子，也可以看着三条知识：

1. 【人都会死，人的寿命不会超过150岁。】
2. 【他父亲死了，享年83岁。】
3. 【张君宝在90岁时去世。】

显然第一条知识更基础，而且作为后边两条知识的解释依据。

具体地，深度语义 的知识可以分为金字塔型的三个层级：核心层，基础层，外围层。

核心层的知识数量非常有限，但极为重要；基础层的知识次之；外围层的知识数量很大，但每一条的价值和重要性都相对较小。深度语义整个知识体系的构建是由内而外，先建立核心知识，然后完善基础层知识。基础层知识完善后，深度语义系统就具有了思考能力，并能做到自动学习、积累知识，实时解决外围知识了。

### 核心知识：数百个核心概念和知识

核心知识库描述最核心的概念和知识，比如，数、量、时间、空间、存在、方向、代词、事物、事件、规则、数学公式、程度、范围、关系、集合、比较、序列、对应、性质、愿望、事实、情绪、目标、条件、结果、过程、…等等。

核心概念和知识是深度语义体系中最核心的部分，是描述其它概念和知识的基础，并且和算法和程序紧密相关，其它知识的处理计算都是依赖核心概念和知识来解释。有了核心知识库和对应的核心处理算法，就能理解整句、整段自然语言的完整意义，实现理解、表达、沟通、思考并解决问题。

核心概念和知识数量较少，其构建的技术水平要求最高。但核心概念和对应的核心算法一旦构建和调试好，以后基本就不会发生改变。

### 基础知识：数万个基础概念和知识

基础知识库描述的是常用的基本概念和知识，比如：很、不、质量、年月日、商品、价格、这、他、个人、城市、国家、手机、水果、飞行、结婚、预订、辞职、购买、出售、选举、表达、宣称、询问、喜欢、满足、高兴、抑郁、货币、西甲、英超、演员、责任、品质、存款、汇率…等等。

基础知识库是深度语义知识库的骨架，这一层的概念和知识构建起来后，结合前述的核心知识库和算法，对于整句整段的理解处理就已经能够顺畅完成。各种各样的自然语言句子中，95%以上的词汇都这属于这数万的基础概念，而剩下的外围词汇实际上可以根据句子中已知的基础词汇进行推算出来。

基础概念和知识的信息量大约有数万个，主要也是研发团队内部人员来配置完成，相对核心概念和知识来说，工作量大但技术难度小。而且一旦配置好以后，也不会发生大的变化。

### 外围知识：无数的外围概念和知识

外围知识库就是大量非常具体的概念和知识对象，它们由上述核心概念和基础概念来进行解释。比如北京、上海、伦敦、苹果、莲雾、美元、麦当劳、民主党、奥巴马、圣诞节、鱼香肉丝、立秋、中国银行、西甲、男款、商务舱…等等。

外围概念和知识的数量非常大，但处于整个深度语义知识体系的外围。这一部分知识库的积累不需要团队内部人员来手工完成，而是可以通过批量导入、自动搜索、用户自主录入等方式来积累，深度语义强大的能力之是：在述核心层和基础层的概念和知识完备起来后，系统实际上具有动态识别和学习外围的新概念、新词汇的能力。和外围词汇本身的不断增加完善结合起来，理解的准确度就将会越来越高。

外围知识库的变化会比较频繁，信息会随时增加也可能会随时修改和删除，但基于层级结构的设计保证，这些变化不会影响到整个体系。

### 专业知识：专业领域的概念和知识

专业领域概念知识是对概念和知识的横向切分，也就是对于各种专业学科或者行业（比如医学、物理学、法律、金融、电信、国家安全…）领域的概念和知识（比如“法人”、“通道”、“交换”、“量子”、“恐袭”等等）的构建。这项工作显然需要专业学科的人员参与进行。

专业领域的概念和知识分为两层，对应着深度语义的基础层和外围层（核心层的知识直接进行共享即可，深度语义的核心概念和知识层高度抽象，不需要做修改就能完全满足不同专业学科需要的核心表达），因此构建过程就是让各专业学科领域的人员逐步按照从内到外的方式建立知识体系——首先添加并建立该专业学科领域的基础概念和知识，然再扩展建立外围概念和知识。

各个专业领域的构建采取开放接口，让多个专业学科领域的人员可以并行进行。一个学科的知识概念在有限的时间内就能初步构建起来，今后将成为该专业领域的共享知识库。

重要的是，深度语义里的知识是真正归纳的有效知识，而不是一些大数据方法收集的大量初级数据。所以，每加一条有效知识整个系统的能力就增强一分。在各个专业领域上建立专业知识会需要付出一定量的人力和时间成本，但都将非常值得的——每个专业领域将会同步地不断向智能化方向发展。

## 深度语义和结构化大数据的结合

前边阐述了，深度语义的知识体系结构的层次性决定了它的构建并不需要像其它方法那样需要大量的初级数据，它的重点是精确地构建核心层和基础层的数据量较小的关键知识，使之学会解决问题的基本的知识。

但在应用层面的距离领域，确实有很多大数据量的结构化数据需要处理，比如维护全球数千万的地名、航班信息、商品信息等。这些确实是“大数据”，但数据都比较明确、比较扁平、缺乏深度，是传统计算机算法非常擅长做的事情。

深度语义解决这些问题的手段就是结合传统算法去处理。简单的说就是把处理工作分解成两个体系，深度语义自己的核心服务器存储核心层和基础层的知识和数据，并实现推理、思考、计算等复杂的处理，相当于核心的“大脑”。而对于众多外围结构化信息库则根据领域和类型不同分解到不同的专业化的服务器集群去维护和处理，分析处理的结果再交给大脑来整合完成，这些专业服务器变成了配合大脑工作的不同专业助手。现有的基于文本的搜索引擎服务也将成为这样的深度语义外围助手，发挥其对大量信息高速运算处理的能力，而搜索得到的目标结果信息的精细加工处理再交由深度语义核心大脑来完成。

将人脑进行深度概念化思维的能力和计算机对大量数据进行高效存储快速计算的能力这两者融合到一起，充分发挥人脑和电脑两者的优势。未来，这样的混合型系统在一些方面的智能最终将超越人脑。

## 深度语义处理不同语言的形式差异

深度语义能解决不同语言的形式和所谓文化差异时，紧紧抓住的一个原则是认清楚本质：语义永远是核心，语言是为了准确表达语义的工具。语义和知识是严谨和明确的，作为最终的评判标准。在保证语义信息本身正确的目的下，语言可以进行做各种变形、简化、省略。如果有了语义这个最终的标准，不同的语言都向语义标准去靠拢，任何语言的个性化问题就都能得到解决。而如果没有这个核心标准试图直接在两种语言的各种形式之间建立联系的话，语言层的各种错误、噪音就会放大，最终结果就会失败。这也就是现有的机器翻译等语义处理技术虽然有进展但始终无法真正突破关键台阶的原因。

深度语义理论归纳出的语言共性特点是：语言最早是以声音的形式出现，由此成熟的人类语言都是一维符号串，都要用一维符号串来表达多维的语义概念网状结构。一维字符串的组织形式决定了自然语言表达语义信息能使用的手段有限，归根结底就是两个方法：

1. 用一个字符串来表达语义信息。
2. 用两个字符串之间的相对位置来表达语义信息，体现为主语、谓语，宾语等。

具体不同语言的个性差异主要体现在：

1、语序习惯：也就是主语，谓语，宾语等的放置次序不同，不同语言的习惯都只是选择理论上可能的语序中的几种并设定不同的优先级别而已。深度语义可以支持任意的位置放置，所有语言的语序变化都超不出这个体系的范畴，因此可以描述所有的语言的语序习惯。

2、信息聚合：就是多个概念知识糅合成一个单一串，深度语义的解决手段是在理解时拆解成单一的概念知识，在生成时再拼接成整体串。

深度语义采取树形结构来管理各种语种，可以更有效地解决方言等问题。比如【汉语】是一个基本语种，而其后边派生出【普通话】【广东话】【上海话】【四川话】等子语种，所有的【子语种】共享基本语种的大部分规则和知识，并增加少量子语种独有的个性化知识。比如对于【你】这个语义概念的表达，所有的【汉语】语种都共享“你”的表达方式，而对于【上海话】则在此基础上增加了“侬”的表达方式而已。

所谓文化的差别，其实是知识的差别，这一点在所有的人类个体之间都存在，并不是语种的问题。A语种中存在而B语种中欠缺的知识，就是要在B语种补全。

口语中的音调等信息，如果语音技术能力够，可以把这些音调信息形式化为字符串符号，进行叠加处理。

## 深度语义对分词问题的解决

深度语义是以语义和知识为中心驱动的体系，不象传统的系统那样将正确的分词作为语法分析的前提，那样的问题在于：分词一旦有误，后边的词法分析和语法分析就会逐级放大这个错误。在深度语义中，分词实际上是进行深度语义分析产生的结果。具体来说，在分词阶段，系统将理论上可能存在的所有词汇全部都会分出来，并不进行计算筛选，而是在进行语义分析时再根据上下文语义来选择更为合理的词汇，具体的语义分析算法是采用独有的“竞争生长”算法，这里就不做详述。

## 深度语义对歧义问题的解决

现今主流的以字符串为研究对象的处理方法解决歧义问题，通常都是在字符串层面利用相关性统计概率理论来处理，而把实际上最为关键的中间层的语义和知识层完全忽略掉。深度语义的相关性分析是在语义和知识层面来进行，和人类自身解决此问题的思维方式是相同的，这是从根子上解决本质问题的方法，因此最终结果将更为精准和有效。

在上述的例子中，深度语义对句子里边的三个“苹果”进行了精确的理解，根据对句子的整体语义理解，将精确分析出第一个苹果是【水果】，第二个苹果是【手机】，而第三个苹果是指【苹果公司】。

现有的翻译软件产品大多都并不去区分究竟是哪个苹果，因为现有的基于字符统计翻译的理论认为翻译只是字符串游戏，不需要弄明白意思，不管是哪一种“苹果”最终翻译结果都是“apple”，干吗不取巧呢？但这就限定了目前的翻译就只是翻译，而不能做别的任何事情。深度语义的原则是自省人类看到这个句子的时候会不会分清究竟是哪个苹果？既然答案是“是的，人类看到句子是一定要理解究竟是哪一种苹果，否则就没有真正理解句子的意思”，那么深度语义就一定也要区分出来。这不但对翻译有直接的好处，而且真正理解了句子的含义，将解决翻译、语义搜索、人机对话等领域通用性的所有问题！仅对于翻译来说区分出具体语义旧有非常重要意义：看起来不区分三种不同苹果似乎都能得到“正确”的结果“apple”，但这其实是撞大运的少部分情况，更多的情况下不同的语义对应的目标语言表达根本就不同，这时如果没有分析出正确的语义对象就会出现完全错误的结果。而就算在本例中，因为深度语义做到了对语义的精确理解，在生成目标语言的时候，对于第二个“苹果”，会去选择“iphone”这个英语语境中更准确的表达，而不会生搬硬套地选用“apple”这个表达，结果实际上也更为精确。

## 深度语义实现思考

“语言即思想”,人类的概念化知识表达能力和思维能力，是和语言紧密结合在一起的，整个人类社会的进步就和语言、文字是分不开。人类有了语言，才能管理更多的知识，一方面这些知识可以更好地积累下来，另一方面文字成了人类思考处理复杂问题的一种工具。

深度语义对概念化思维的定义，就是各种概念知识的演化关系。深度语义语义结构本身就可以实现对概念化思考过程的描述。

比如，先给深度语义教一些知识，然后询问问题：

人：1996年，王菲嫁给摇滚歌手窦唯，成为当年娱乐圈最轰动的事件。

深度语义：理解。

人：ZhangMing lent an English book to WangWei.

深度语义：理解。

人：窦唯的妻子是谁？

深度语义：王菲。

人：Who is Wangfei's husband?

深度语义：Douwei。

人：王伟向谁借了英语书？

深度语义：张明。

在这几个简短的例句中，其实涉及到非常复杂的理解-存储-推理-查询-转换-生成等自然语言和语义处理和思考的过程。首先来说，在第一步学习知识的时候，深度语义理解了句子，从中得到了新的知识，并且从“王菲嫁给摇滚歌手窦唯“这种显式知识推理和思考了”窦唯和王菲是夫妻关系“等隐藏知识，将这些知识都以语义结构的方式存储了起来。然后在接受到询问问题的时候，深度语义会进行思考和查询，回答出“窦唯的妻子是谁？”和“Who is Wangfei's husband? ”这样的问题——且知识、问题、答案几者的具体语言形式（中文、英文…）可以任意变换。

这里列举的例子比较简单，但复杂的东西往往可以由简单的基础结构组合人而成，只要对简单问题的解决方法具有本质性和共性并具有自描述的特征，就可以组合来解决所有的复杂问题。深度语义的技术解决了概念化思维和思考的根本问题，知识足够丰富后，通过反复的嵌套和迭代——这是计算机擅长的事情，将能解决非常复杂的推理和思考问题，不仅能去实现精准和智能的语义搜索，还能彻底地利用知识驱动的方式实现多轮对话交互完成为客户订购机票等智能化业务——最终完全到达人类客服人员的水准！

## 深度语义实现计算

深度语义的基本体系完全容纳了基础数学知识的表达，实际上，在深度语义体系中数学知识的表达、解析和计算和普通的自然语言是相同的。从理解问题到处理成数学式，以及对数学式求解的过程完全一体化。所以，对语言中蕴涵的数学信息的分析和计算也可以依托深度语义模型从根本上进行实现。

比如问题：

问：北京距离上海1200公里，飞机以800公里每小时从北京飞到上海需要多少时间？

深度语义回答：1.5小时。

深度语义解决这个例子的方法，并不是现在的语音助手等采用的固定句式模板并对应硬编码一段程序来实现，而是模拟人类解决这些应用题的方法，也就是根据各种知识（如：距离=速度╳时间、乘法<=>除法、单位换算）等知识来自动推导运算的！具有广泛的适应性和灵活性，实际上，深度语义完全成熟后，将轻而易举地完成各种应用题的运算，也显然地可以结合到实际场景，比如购物、订票等等场景中。

# 对比

## 和符号主义对比

众所周知，人工智能领域里最主要的划分是两个流派：符号学派和连接学派。符号学派由来已久，但多年来技术进展不显著，近年来被连接学派方面的深度学习夺去了光芒。

用概括性的符号来表达知识（起码是概括性的知识）这个基本理念并没有任何错, DSM本身也归于符号主义的范畴！问题在需要有很多具体的工作，尤其需要构建起完整的基础概念体系！对于这个工作，以前哲学家、语言学家、人工智能学者们都只能浅尝即止，只在局部提出一些片段假说，做不到深入透彻地研究形成完整的体系。从这个角度来看：符号主义只是有一个名字，而并没有真正构建起来！

可以归纳传统符号学说的根本问题在于：

1，符号的体系不完整；

2，知识缺少合理的层级关系；

3，缺乏概率体系；

4，具体实现结构和方法不够灵活；

5、重实体而轻连接；

DSM但做出了比较重要的创新和突破就是：构建起了完整的基本概念和运算体系，解决了传统符号主义存在的上述诸多难题。

DSM也对连接学派里真正精华的部分比如概率理论进行了吸取和融合，并探索出全新的一些理论和方法，实际上已经成为一种全新的体系，将有力推动AI技术的进展。

## 和知识图谱对比

知识图谱就是符号主义的一种实现，现在也是自然语言处理的主要方法之一。这也充分证明符号方法本质上并不是错的，但目前的知识图谱也就没有解决上述符号主义存在的根本问题，例如：

1，不够立体，不能形成完整的知识体系，不能形成包含运算和推理思考的智能知识库。

2，比如分别提出很多实体图谱、知识图谱、社交图谱、事理图谱、因果图、图-图网络，却不能整合成为一体。

3，实际还是割裂的专题数据库，为特定运算提供数据素材而已。

4，因此，现有的知识图谱的宣传还是突出其“大”，而没有表现出“深”。光是数据量大不能说明什么，仅仅是构造一个10亿个人信息的库，就能轻而易举能凑够百亿条的知识，这和以前基于传统结构化数据库存储的海量数据并没有体现出本质区别。构建起完整的多层级知识并能充分灵活应用到运算中，能用来解释新的知识并支撑真正的推理思考，才是更重要的。

5，不能和自然语言无缝地结合。

DSM比现有的知识图谱强大得多，不光是表达实体、属性和关系，对于事件、事理、因果、推理、场景、对话等，都进行了完整的表达。而且所有的信息——包括语义和语言的信息甚至演算规则，都融合在一个结构中，远超越了现有知识谱图的范畴。

如果一定要用知识图谱来理解DSM结构的话，可以认为是一个超级知识图谱。DSM研究的成果，是业界进行知识图谱研究应用的公司和团队所梦寐以求的。

## 和深度学习对比

如前所述，深度学习是典型的连接学派，近年来因为技术上有很大提升并解决了一些实际问题而成为目前最为主流的方法。但显然，目前包括深度学派自己的几大创始人也分别表达了深度学习只适用于解决部分问题，而并不是解决强人工智能的终极手段。

深度学习试图用一种数学方法描述一切智能问题，然后“把其它问题都丢给计算机，让它自己去学习知识，避免了人工来归纳处理知识”这种捷径方法，是大家非常喜欢的。问题是这种方法其实只能学到一些简单的数理统计规律的浅层知识，人脑那样真正概括性的知识用这种方法是不可能学会的！

深度学习试图用大量的【输入-输出】信息素材来让计算机自己找输入-输出的黑盒的匹配关系，这只使用于输入和输出间确实存在一定函数关系的部分领域（例如语音识别和机器翻译）。

DSM的核心是手工构建高层最具概括性的知识为基础，在这个基础构建起来后，再借鉴深度学习等方法实现在大量信息里边查找初级知识，进行提炼后再纳入到DSM的范畴。那将会在自然语言语义处理和强人工智能的目标方向上取得真正的进展。

DSM的理念不是从黑盒视角看待输入和输出，而要介入【输入】到【输出】整个过程的中间环节，进行完整的表达、解释和演算。在DSM的理论中，深度学习的几大问题实际上都能得到很好的解决：

1，可解释性

2、因果关系

3，小样本学习

4、通用性