

# 《图神经网络导论》

## 课程报告

报告人           ：       练炳诚      

报告批阅教师：                                

计算机科学与技术学院

2020 年 12 月 6 日

---

# 华中科技大学课程实验报告

---

## 目 录

|                                                              |    |
|--------------------------------------------------------------|----|
| 1 摘要.....                                                    | 1  |
| 1.1 论文相关概述.....                                              | 1  |
| 1.2 报告组织结构.....                                              | 1  |
| 2 阅读心得.....                                                  | 2  |
| 2.1 Abstract.....                                            | 2  |
| 2.2 Introduction.....                                        | 2  |
| 2.3 Relational inductive biases.....                         | 4  |
| 2.4 Graph networks.....                                      | 4  |
| 2.4.1 Background.....                                        | 4  |
| 2.4.2 Graph network (GN) block.....                          | 5  |
| 2.4.3 Computational steps within a GN block.....             | 5  |
| 2.4.4 Relational inductive biases in graph networks.....     | 6  |
| 2.5 Design principles for graph network architectures.....   | 6  |
| 2.6 Discussion.....                                          | 8  |
| 2.6.1 Combinatorial generalization in graph networks.....    | 9  |
| 2.6.2 Limitations of graph network.....                      | 9  |
| 2.6.3 Open questions.....                                    | 9  |
| 2.6.4 Integrative approaches for learning and structure..... | 10 |
| 2.6.5 Conclusion.....                                        | 10 |
| 3 课程体会.....                                                  | 11 |

## 1 摘要

### 1.1 论文相关概述

**论文题目:** Relational inductive biases, deep learning, and graph networks

**作者:** Peter W. Battaglia, Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, Caglar Gulcehre, Francis Song, Andrew Ballard, Justin Gilmer, George Dahl, Ashish Vaswani, Kelsey Allen, Charles Nash, Victoria Langston, Chris Dyer, Nicolas Heess, Daan Wierstra, Pushmeet Kohli, Matt Botvinick, Oriol Vinyals, Yujia Li, Razvan Pascanu

**论文链接:** <https://arxiv.org/abs/1806.01261>

**选择该篇论文的原因:** 这篇论文是一篇在 GNN 方向的综述性论文，被引用数很高，就这方面上看应该是一篇不错的论文。从论文的题目上看，论文主要介绍了关系归纳偏差，深度学习和图网络，比较适合入门去读。

### 1.2 报告组织结构

报告后续组织结构分为阅读心得和课程体会，阅读心得中主要撰写了本人在阅读论文部分章节中了解到的知识与个人见解，由于本人才疏学浅和时间不充足的原因，并没有完整地读完整篇论文，而是有选择地阅读了部分章节。课程体会中主要撰写了本人对这学期图神经网络课程的一些体会。

## 2 阅读心得

### 2.1 Abstract

摘要中表达的核心思想是反对单纯使用人工特征或者单纯使用端到端系统的方法。AI 中首要的问题是要解决组合泛化的问题，而解决这个问题的方法是要解决结构化表征方法和结构化计算的问题。

这里的结构化表征方法应该针对的是如今的单纯数值类型的词向量输入，虽然词向量有很多维度，但是知识之间没有组织结构。

这里的结构化计算应该针对的是如今单纯的数值运算法，虽然存在着 pooling 等等的非连续性计算单元，但是缺乏逻辑性的本质还是没有改变，现在使用的神经网络虽说是利用数值运算去模拟逻辑运算，但是由于数值运算的连续性和逻辑运算的离散性的差异，使得两者之间存在着一定的距离。

论文的研究问题是如何利用关系归纳偏置帮助深度模型去学习实体，关系以及组织他们的方法。作者用 GNN 的模型建立了支持关系推理和组合泛化的推理系统，系统是复杂的，可解释的以及灵活的。关于可解释性，应该是针对的网络模型的缺点，灵活性针对的是纯逻辑推理系统，至于复杂性，无论是逻辑推理系统还是网络模型也好的，都是十分复杂的。

### 2.2 Introduction

第一句就是 Humboldt 说的“A key signature of human intelligence is the ability to make “infinite use of finite means”，之后便举例说明了人类如何利用组合泛化（Combinatorial Generalization）进行认知推理。人类可以利用层级结构去高度抽象细粒度的差异，并且可以找到表征和行为的一般性行为，比如“泄漏”一词，可能会想到“泄漏情报”和“液体泄漏”两个用法，这同一个词汇在两个意思下描述的是不同

# 华中科技大学课程实验报告

---

的事情,但是两种存在着一种共性.这种共性就是认知属性,个人理解“泄漏”的认知属性就是:实体  $E$  在被动的情况下从空间  $S_1$  以某种形式移动到了空间  $S_2$ 。那在得知泄漏这个词汇可以和情报一起使用后,可以对泄漏进行泛化,假设已经知道了泄漏拥有一个认知属性,但是不知道其相应的搭配,通过实例之后,我们不仅知道了泄漏可以和情报搭配使用,我们还知道了这种认知属性可以和那种类型属性一起搭配使用,因此我们就可以将“泄漏”搭配上其他同属于“信息”类型的词汇,例如“泄漏数据”。这种泛化其实是基于一个很强的假设:实例类的可组合性是抽象类可组合下的充分必要条件。

接着论文引用了一段文字,用来解释“解释的本质”。是学术界对人去理解解释世界的模型。

简单来说,上面我个人的理解是将抽象和类型完全进行了分离,一个类型中的一些原型词汇的使用方法,决定了同一个类型内的其他词汇和概念的使用方法。这就使得有限的结构化的知识在接受有限的实例后可以拥有了无限表示的能力,因此我们在学习的时候,可以做两件事情:

将新的知识放入已有的结构化知识框架;

调整框架去适应新的知识。

现阶段的实现组合泛化的理论如 logic, grammars, classic planning, graphical models 等其下面的子领域都明确地以实体和关系为中心进行学习。本文强调了结构化方法的重要性,即关系归纳偏置,其对应的神经网络模型,深度网络模型中的弱归纳偏置对应的是数据的低学习效率,也就意味着需要大量的数据,但数据是非常昂贵的,因此通过增强归纳偏置来提供一个结构化的方法是非常重要的。在早期的连接主义提出了一些解决方法如 analogy-making, 到近期出现一些使用深度模型去利用分布式的连续的向量表示的文本,图,几何,逻辑的方法,表示结构化和灵活性不是对立的,这些方法中有一个很大的特点是: performing computation over discrete entities and

# 华中科技大学课程实验报告

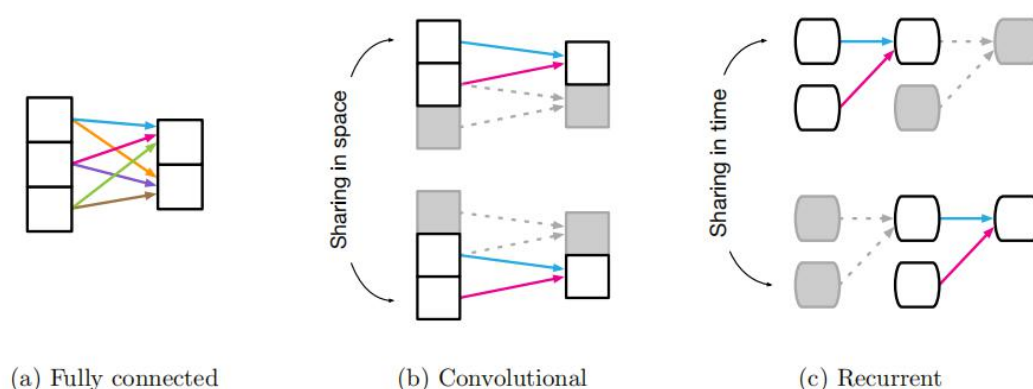
---

the relations between them, 于是作者对这一特点进行了分析, 并设计出了一种通用框架去描述基于实体和基于关系的推理, 即 graph network。

后续章节先介绍了目前一些强归纳偏置模型, 然后简介过往有关 graph neural network 任务.其次,提出自己的 graph network framework。

## 2.3 Relational inductive biases

这部分内容主要介绍了深度学习中的强归纳偏置模型, 并不是我本次阅读论文关注的重点, 论文中介绍了全连接层、卷积层、循环层的各类特点见下图。



总之就是最后总结了具有处理结构性知识的模型, 一句话总结就是顺序的不变性, 应该是需要通过深度模型的结构或者组件进行规范和反映的, 这样的模型结构就具有了一定程度的 relational reasoning 的能力。

## 2.4 Graph networks

本章首先总结了一些过往的 GNN 的任务, 然后提出了自己的 graph network framework。

### 2.4.1 Background

第一段先介绍了很多利用网络模型做的任务, 第二段介绍了一些利用网络模型去解决需要对离散的实体和关系进行推理的任务, 例如组合优化问题, SAT 问题等, 总

# 华中科技大学课程实验报告

之就是介绍了很多方向。

## 2.4.2 Graph network (GN) block

这一节介绍了 GN 框架的定义：GN 框架基于图结构表征(表示)，定义了一组用来进行关系推理的函数。功能：supports constructing complex architectures from simple building blocks，就是用一些基本的 building blocks 去构建复杂的结构。组成：GN 框架的基本单元是 GN block，输入是 graph，输出是 graph，执行的是基于结构的计算，也就是最最开始说的 structural computation，graph 的 node 是实体(entities)，edges 是关系(relations)。之后就是用图和公式进行了规范定义，包括状态更新函数等。

总之这一节的内容这和课程里面对图网络的介绍差不多，是一些基本的定义介绍。

## 2.4.3 Computational steps within a GN block

这一节主要介绍了 GN block 的计算步骤，用算法来描述如下：

---

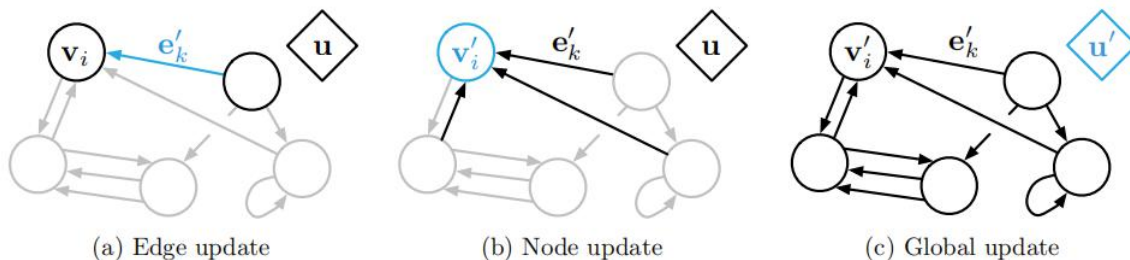
**Algorithm 1** Steps of computation in a full GN block.

---

```
function GRAPHNETWORK( $E, V, \mathbf{u}$ )
  for  $k \in \{1 \dots N^e\}$  do
     $\mathbf{e}'_k \leftarrow \phi^e(\mathbf{e}_k, \mathbf{v}_{r_k}, \mathbf{v}_{s_k}, \mathbf{u})$            ▷ 1. Compute updated edge attributes
  end for
  for  $i \in \{1 \dots N^n\}$  do
    let  $E'_i = \{(\mathbf{e}'_k, r_k, s_k)\}_{r_k=i, k=1:N^e}$ 
     $\bar{\mathbf{e}}'_i \leftarrow \rho^{e \rightarrow v}(E'_i)$            ▷ 2. Aggregate edge attributes per node
     $\mathbf{v}'_i \leftarrow \phi^v(\bar{\mathbf{e}}'_i, \mathbf{v}_i, \mathbf{u})$            ▷ 3. Compute updated node attributes
  end for
  let  $V' = \{\mathbf{v}'_i\}_{i=1:N^n}$ 
  let  $E' = \{(\mathbf{e}'_k, r_k, s_k)\}_{k=1:N^e}$ 
   $\bar{\mathbf{e}}' \leftarrow \rho^{e \rightarrow u}(E')$            ▷ 4. Aggregate edge attributes globally
   $\bar{\mathbf{v}}' \leftarrow \rho^{v \rightarrow u}(V')$            ▷ 5. Aggregate node attributes globally
   $\mathbf{u}' \leftarrow \phi^u(\bar{\mathbf{e}}', \bar{\mathbf{v}}', \mathbf{u})$            ▷ 6. Compute updated global attribute
  return  $(E', V', \mathbf{u}')$ 
end function
```

---

其实就是对 2.4.2 节中的定义用伪代码进行了算法描述，概况来说包括以下 3 个步骤：



在学习了图网络的知识后还是能看懂的。

## 2.4.4 Relational inductive biases in graph networks

在前面的叙述中反复强调了归纳偏置的重要性，这一节终于介绍这个架构如何去实现强归纳偏置了。

graph 本身可以根据实体和关系去设计。在之前的网络结构中，我们通常是对 node 没有进行关于输入的分别，而是一视同仁，即任何词汇都可以 instantiated 到每个 node，但是，原文中说到：

graphs can express arbitrary relationships among entities, which means the GN's input determines how representations interact and are isolated, rather than those choices being determined by the fixed architecture.

个人理解是 graph 的 node 之间的关系是可以根据对 node 的输入来定的，这样的话，网络架构就会变成一个动态的架构。

图关于 entities 和他们之间的 relations 是的表征是基于 set 的。也就是说，对于参数位置的交换是没有反应的。GN block 对于参数的顺序不敏感，这个就支持了对于很多对顺序不敏感的模型的支持。

每个 node 和每个 edge 都是可以被复用的。这就支持了组合泛化，是相当重要的，在决定了 GN block 后可以根据实例改变 size 以及 shape。

## 2.5 Design principles for graph network architectures

这一章主要介绍了对于深度模型的 graph network 设计规则。

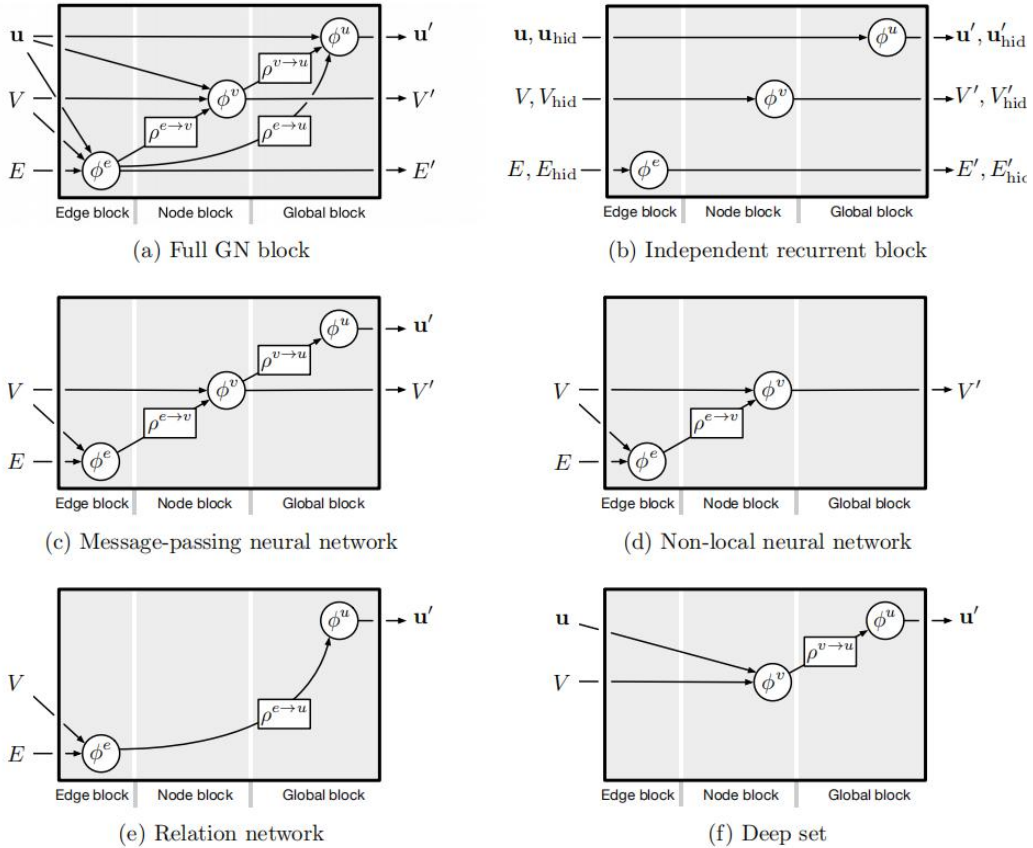


# 华中科技大学课程实验报告

首先介绍了图网络的灵活性体现在两点，一是属性值表征的灵活性，二是图结构本身的灵活性。对于属性，其实就是 `edges,nodes,global` 的值，在深度模型中，一般是采用 `real value vector or tensor` 的形式。输出的单元根据任务的不同也可分为三种类型：`edge-focused GN,node-focused GN,graph-focused GN`。

然后介绍了图结构，文章中提到图的结构有两个极端，一个是实体显式地决定其之间的关系结构，一个是关系结构必须被推断或假定，也可以采取两者的混合。第一个的例子是知识图谱以及句法树以及社会网络问题，这样的问题中，`edge` 是根据其相连实体来确定的。第二个的例子是 `visual scenes`，自然语料库等等，这样的初始的情况下我们不知道实体之间的任何关系。

接下来就介绍了很多图模型的配置，如 `Full GN block`，`Independent recurrent block`，`Message-passing neural network` 等等，具体见下图。



# 华中科技大学课程实验报告

---

总之这一节就是主要介绍各类模型，这并不是我关注论文的重点，因此就只是大概的看了下，其中一些模型在课上讲过的还能理解，一些还未听说过的模型就没去深入的看了。

接下来的一节就是讲了由于一个 GN block 的输入和输出都是 graph，那么一个 Block 的输出就可以当成另外一个 Block 的输入，即：

$$G' = \text{GN}_2(\text{GN}_1(G)).$$

关于组合的方式，论文中提到了以下三种：

Composition of GN blocks

Encode-process-decode

Recurrent GN architecture

第一种方式的具体例子是文章中提到的另一篇文献的物理动力学系统，这个系统中每个 GN block 可以视为对一个 dt 后系统中 object 的预测，组合 M 个这样的 Block 就可以视为是预测 M \* dt 后 object 的状态；第二种方式就是 encoder、decoder 分别用一个 GN block，还有个处理中间状态的 GN block，输入到输出经过 encoder、process、decoder；第三种方式文章中没说太清楚，不过通过参考文献中的介绍可以用来预测时序的图结构数据，例如物理动力学系统中随着时间而不停变化的图的状态。

最后一节是给出了实现代码的例子，包括 github 仓库地址，没有具体去看。

最后做个总结就是，本章从三个角度讲了 GNN 的设计原则。在最底层，从 GN 的属性设置和 graph 边的设计方法阐释了 GNN 的灵活性，从 GN 的多种变种阐述了 GNN 的灵活性，最后从 GN 的组合来阐述灵活性。

## 2.6 Discussion

之前的内容讲了 GN 的基本结构和各类模型等等，这一章主要讲了如何使用 GN 去设计复杂的结构去包含强归纳偏执到模型中。

# 华中科技大学课程实验报告

---

## 2.6.1 Combinatorial generalization in graph networks

这一节介绍了证明 graph network 具有组合泛化能力的论文，以下特性：

被设计用来预测 One-step 的模型可以用来预测上千步；

在 object 的数量发生变化时也可以保持高正确率；

GN-based 决策模型，在 agents 数发生变化时也能很好的应对；

还可以生成一些之前没有的 useful node embedding。

这些特性都证明了 graph network 具有“从有限中生成无限”的能力，即组合泛化能力。

## 2.6.2 Limitations of graph network

这一节介绍了 graph network 不能解决的一些问题，并列出了相关的论文列表，典型问题如下：

非同构图的判别的问题；

一些无法用图表示的结构的问题：recursion，control flow，and conditional iteration。

这些“computer-like”processing”是图无法处理的结构，但是这些对于人的认知判断是极为重要的，也就是说，graph network 虽然很强大，但是想相比于人的认知能力还是差很多必要因素。

## 2.6.3 Open questions

这一章的中心点在于，认为在相比于某一个框架下进行调整（即在某一个模型上添添补补），更加重要的是认识到 graph network 的全部潜力。这里的问题就是围绕这“潜力”来的。

question 1 : graph network 操作的 graph 是怎么得来的？

# 华中科技大学课程实验报告

---

这个其实是一个最根本的问题。也就是如何把原始感官数据转化为有结构的数据，如今的 wordnet, dependency parsing 其实都是在做这个工作。一个方法就是使用全连接图，但是实际的结构应该是更为稀疏的。有一些研究是做推导出稀疏结构的，文章列出来相关的论文列表以供参考。

question 2：如何在计算的过程中动态的改变 graph 的结构？

例如，在一个物体分成几个部分之后，这个 entity 的 node 也应该分为几个 node。算法应该具有在计算过程中对 edge 进行增加和取消的能力，文章列出了两个针对这个问题的研究文献。

总之，这一节提出了两个 Open Question，并列举出了相关研究的文献。

## 2.6.4 Integrative approaches for learning and structure

强调本文主要在 graph network 计算部分，而在设计思路上只给出了以下的几个方向：

linguistic trees;

partial tree traversals in a state-action graph;

hierarchical action policies;

capsules;

programs。

原文中每个方向都列举了相关研究的文献，这里就不一一列举出来了。

## 2.6.5 Conclusion

这一小节强调了组合优化的重要性。作者在肯定 graph network 的同时，也承认了他的不足，并且提供了几个作者认为有意思的有前途的但是不是很热门的方向，具体都在原文中。

## 3 课程体会

在这门导论课中，通过张老师的介绍和自己课外在 web 的了解，我对图神经网络的概念和基本原理有了一些理解，不过感觉要深入下去，还是得先具备扎实的数学推导基础和机器学习的相关知识。这一点在上手实验的时候深有体会，虽然参考手册完成了代码，但对每个 layer 的定义及激活函数的选择上实在是懂原理，正如课件三说的，现在我只是一个不求甚解的炼丹师（甚至算不上）。

此外，在查找文献的过程中，发现在(图)神经网络方向的论文近两年有非常多，各种研究方向的都有，五花八门，真是万物皆可 AI。这里要讲一件趣事，前段时间我在实验室做课程实验的时候，一个博后跟我工位旁边的博士正在闲聊吹水，无意间发现我看的東西，我们实验室主要是做 AI for Storage 及 AI for Data Management 的，刚好这两位师兄的研究方向就是图神经网络在存储系统上的应用，并且也发了几篇 paper，于是博后师兄非常兴奋，告诉我我们实验室搞这个的很少，希望以后能多跟我交流交流。当时我十分汗颜，立即表达了我不是在做这个方向，而是在做课程实验而已。

总之，对这个今年的热门方向，相关的研究仍在不断进展，想要在这一方向有所建树，还是得要抓紧前沿问题，这门课算是为我揭开了图神经网络的面纱，也许今后就用到了呢？