

1 简介

各位老师好，很荣幸能在这里参加各位老师的面试，我叫黄君扬，出生于上海市杨浦区，是 2015 级（2019 届）上海理工大学光电信息与计算机工程学院计算机科学与技术系的一名本科生。我的特点是：在拥有良好的学科理论基础的同时也有优秀的代码能力（同时包含算法能力和工程能力）。下面我将从四个方面来介绍自己，分别是：一学习情况 / 二学科竞赛 / 三项目经历 / 四读研动机。

学习情况 在本科阶段，我一直以优异的成绩位列年级第一，曾获上海市奖学金，连续五学期获得校学习优秀奖学金一等奖，多次获得专项奖学金，优秀学生等荣誉。

学科竞赛 在课余之时，我积极参加各类学科竞赛，曾获“ACM-ICPC 亚洲区域赛银奖”、“美国大学生数学建模竞赛二等奖”、“中国计算机程序设计天梯赛上海市特等/全国三等奖”、“蓝桥杯 Java 软件开发上海市一等奖/全国二等奖”等多种洲际级、国家级奖项。

项目经历 参与多个项目的开发，在过去的一年中参与了如下几个项目：1) 参与企业项目“iSmart 创伤中心”并独立开发项目中的规则引擎。2) 软件著作权“一种便携的集成身份认证方法”。3) 基于 SpringCloud 的分布式“机票销售系统——USST Travelling”。4) 在项目“文档协同编辑系统”中提出了一种全新的算法以保证数据的一致性。

读研动机 在过去的三年时间里，我发现自己对计算机科学已经产生了浓厚的兴趣，我越发地学习，却越来越发现自己懂得知识很少，我也体会到了确实要做到“术业有专攻”，因此我迫切的希望能找到一位导师，为我指明一个领域中值得研究的方向或问题来让我进行研究工作。希望通过自己的努力在研究生阶段能在科研上有所建树。

Good afternoon. Dear teachers! It's my great honor to be here for this interview. My name is Huang Junyang. I was born in Yangpu District, Shanghai. Currently, I'm a student from University of Shanghai for Science and Technology majoring Computer Science and Technology. My characteristic is: I have both good theory basis of computer science and excellent coding ability (including algorithm design and engineering ability). I will introduce myself from four aspects, they are: academic records / competitions / project experiences / motivation for postgraduate study.

Academic Records. During the last three years, thanks to my teachers' serious teaching and my hard-working, I always ranked first in my grade. Each semester I won the First-Class Scholarship and I also won the Shanghai Scholarship granted by the Shanghai Municipal Commission of Education in 2017.

Competitions. In my free time, I learned 'Algorithm design and analysis' and 'mathematical modeling' through self study, and won many awards in various disciplines, such as silver medal of the ACM-ICPC Aisa Regional Contest and second prize of the American College Students Mathematics Modeling Contest and so on.

Project Experiences. Besides, I also learned several popular software development frameworks in my free time, and implemented several projects, some are implemented by myself and some are implemented with my best friends. The source code of these projects can be found on my github. The two most representative projects are a wechat APP called “a portable integrated identity authentication method” and a distributed Document collaborative editing system where I proposed a new algorithm to ensure data consistency.

Motivation for Postgraduate Study. During my undergraduate years, I found that I have developed a strong interest in computer science. The more I learnt, the more I found myself lack of knowledge. So I do want to have a tutor who could help me find the research direction and problems that are worthwhile for me to focus on. I hope that through my hardworking, I can produce high-quality papers and achieve certain scientific research outcomes in my postgraduate studies.

2 项目介绍

1. 一种便携的集成身份认证方法 A portable integrated identity authentication method

提供了一种多层次多维度且易定制的身份认证服务。曾入围中国计算机设计大赛决赛。核心技术：TOTP（基于时间的一次性密码生成算法）、人脸识别、全球定位和微信用户唯一识别码 OpenID。项目心得：

This project implement a multi-dimensional and customized identity authentication service using Wechat OpenID, Global Positioning, Time-Based One-Time Password Algorithm and Face Recognition techniques. I tried to implement a face recognition algorithm, but since the algorithm is very general, I ended up using the FACE++ API.

2. iSmart 创伤中心

这个项目是我目前在实习的公司正在参与的项目。我主要负责规则引擎的开发。规则引擎就是对定期传入的数据进行检测，若一个数据或一组数据满足某个或多个规则，就会报警（fire），它实际上就是设计模式里的观察者模式的一种应用。在做这个项目的时候我学到了公司里做项目的规范和流程，从我的带教人那里学会了很多开发时的注意事项和开发技巧。

3. 分布式机票销售系统

这个项目是一个学校大作业利用分布式框架 Spring Cloud，构建高可用、支持高并发的服务。在做这个项目的过程中，我深刻体会到了要让一个项目上线，这个项目必须做到高可用、支持高并发和注重信息安全。同时我也对服务熔断和服务降级有了更好的理解。服务熔断一般是某个服务（下游服务）故障引起，而服务降级一般是从整体负荷考虑。

4. NT 文档协同编辑系统 NT Document collaborative editing system

在这个项目中，我提出了一个全新的算法来保证数据的一致性。这个算法有点类似于数据库中的写锁。算法如下：当一个人要对某一处或某一段进行修改的时候，首先要将那一段选中，然后点击右上角弹出的锁定按钮，这时候除了锁定的人以外其他人都没法修改这一段文字。当对某一段的修改完成后，点击释放按钮，这时候原先被锁定的文字会被解除锁定，别人就可以修改了。注意每次修改都要加锁，修改完后要释放。

In this project, I proposed a new algorithm to ensure data consistency. This algorithm is somewhat similar to a write lock in the database. The steps are:

- (a) When you want to modify a certain paragraph or a certain word in the article, first select the section.
- (b) Then click the lock button that appears in the upper right of your web-page. The selected area can not be modified by others.
- (c) When the modification of your selected area is finished, click the release button. The originally locked area will be unlocked, and others can modify it, however, they should lock it.

Note that each modification must be locked, and then released after modification.

3 研究方向

我的研究兴趣是知识图谱。目前越来越多的行业或者企业积累了规模可观的大数据，但这些数据并未发挥出应有的价值，我认为这是因为及其缺乏诸如知识图谱这样的背景知识，从而限制了大数据的精准与精细的分析，从而大大降低了了大数据的潜在价值。

目前，我认为知识图谱可以用于以下领域：

My research interest lies in knowledge graph. At present, more and more industries or enterprises have accumulated large amounts of data, but these data have not played the value if deserves. I think this is because it lacks background knowledge such as knowledge graph, which limits the analysis of big data. This greatly reduces the potential value of big data.

Currently, I think the knowledge graph can be used in the following areas:

1. 搜索引擎 Search Engine

在分析用户的搜索意图时，我们希望能做到对搜索意图的精确理解，知识图谱可以帮助我们更好地理解用户的意图。例如两个用户分别输入“toy kids”和“kids toy”，这两个词都是名词，机器需要分辨哪一个核心词，哪一个是修饰词，借助知识图谱，我们就可以知道两位用户的真实意图都是搜索儿童玩具。

When analyzing the user's search intent, we hope to have a precise understanding of the search intent, and the knowledge graph can help us better understand users' intent. For example, two users input "toy kids" and "kids toy" respectively. These two words are nouns. The machine needs to distinguish which one is the core word and which one is the modifier. With knowledge graph, we can know that the true intent of both users is to search children's toys.

2. 推荐系统 Recommender System

如果用户搜索了羊肉卷、牛肉卷、菠菜，利用知识图谱我们可以知道用户很可能想要做一顿火锅，这时我们就推荐火锅底料、火锅电磁炉。又比如说冷启动下的推荐，是基于传统的统计行为的推荐方法难以有效解决的问题，但利用知识图谱，特别是关于用户与物品的知识指引冷启动阶段的匹配与推荐，是有可能让系统尽快渡过这个阶段的。

The recommendation under cold start is a problem that is difficult to solve effectively based on the traditional statistical recommendation method. However, the use of knowledge graph, especially the knowledge and guidance of users and articles to guide the cold start phase of match and recommendation, is likely to let the system pass this stage as soon as possible.

3. 可解释性 Interpretability

机器学习，特别是深度学习的黑盒特性，日益成为学习模型实际应用的主要障碍之一。知识图谱在可解释性方面有着重要的作用。比如 AI 医生所做的任何治疗方案，都必须配备解释，否则人类永远不可能为它买单。知识图谱是一种语义网络，包含大量实体和概念及其之间的语义关系，他拥有海量规模、语义丰富、结构友好、质量精良的优点，能帮助机器理解人类知识，从而解释结果。

Machine learning, especially deep learning, is a black box and is becoming one of the main obstacles to the practical application of learning models. Knowledge graph plays an important role in interpretability. For example, any treatment plan done by AI doctors must be equipped with an explanation, otherwise humans will never be able to pay for it. Knowledge map is a semantic network, which contains a large number of entities and concepts and their semantic relationship. He has the advantages of massive scale, rich semantics, friendly structure and excellent quality. It can help machines understand human knowledge and explain the results.

4 兴趣爱好

我的兴趣爱好是打羽毛球，乒乓球和编程。

My hobbies are programming, playing badminton and table tennis.

5 你联系了哪些老师

我联系了肖仰华老师，因为在未来我想从事知识图谱方向的研究工作。

I contacted Prof. Yanghua Xiao because I want to do research works in the area of knowledge graph in the future.

6 阐述一个你熟悉的排序算法

Quick Sort. Quicksort first divides a large array into two smaller sub-arrays: the low elements and the high elements. Quicksort can then recursively sort the sub-arrays. The steps are:

1. Pick an element, called a pivot, from the array.
2. Partitioning: reorder the array so that all elements with values less than the pivot come before the pivot, while all elements with values greater than the pivot come after it (equal values can go either way). After this partitioning, the pivot is in its final position. This is called the partition operation.
3. Recursively apply the above steps to the sub-array of elements with smaller values and separately to the sub-array of elements with greater values.

The base case of the recursion is arrays of size zero or one, which are in order by definition, so they never need to be sorted.

The pivot selection and partitioning steps can be done in several different ways; the choice of specific implementation schemes greatly affects the algorithm's performance.