|  |
| --- |
| **生涯人物访谈记录**  **一、信息采集**  **考虑到当事人的隐私问题，这部分删去** |

**二、背景介绍**

胡津铭学长是我校计算机学院17级优秀研究生，研究生在读期间在顶级会议中发表了若干篇论文，拥有丰富的科研经历，曾多次担任学校高难度课程《数据挖掘导论》的助教和图灵班课程《机器学习》的课程设计者和主讲老师，并在秋招和社招等多次招聘中收获了数十个国内互联网大厂，AI独角兽企业和知名外资企业的高级别Offer

日常生活中胡学长也勤于学习，乐于助人，在知乎和CC98等网络社区中多次发表学习国外高质量公开课的经验文章，并组织若干交流群帮助学习课程的同学答疑，共同进步，广受好评。

因此我认为胡学长是本次职业生涯规划讲座布置的生涯人物访谈中非常合适的访谈对象，无论是在学习，科研，实习，招聘还是工作中，胡学长都取得了骄人的成绩，当然这离不开胡学长个人的努力和勤奋。

本次访谈就将围绕这些主题展开，我通过QQ电话连线的方式和胡学长进行了交流，并讲访谈过程中的对话整理成了如下访谈记录。

**三、访谈记录**

以下内容根据访谈录音整理而成，大致可以分为工作、学习、职业规划三个方面。

**1.我了解到您现在是DolphinDB担任研发工程师，请问您现在主要的工作内容是什么？**

DolphinDB是一家分布式时序数据库公司。传统的关系型数据库在时序场景下没有做一些针对性的优化，时序数据库是专门设计出来针对时间序列数据进行存储和查询，主要的应用场景就是金融和物联网。那么什么是分布式数据库呢？就比如很多场景下，数据量会比较大，用单机的数据库是不可能存储的，这时候就需要多台机器支持，把分布式和时序数据相结合就是DolphinDB在做的事情。

我在DolphinDB工作的主要分为存储引擎和计算引擎两大块，大家都知道数据库要支持增删查改，那其实增删改都是和存储引擎挂钩的，而查和存储引擎计算引擎都挂钩，所以数据库就可以分成这样两个大的部分，存储引擎就是解决怎么样高效地将数据存储到硬盘上，并且能让后续的查询尽可能高效。计算引擎就是说如何高效地处理输入的一个SQL语句，另外还有一块叫做logging and recovery，用来保证数据库的一致性和事务的原子性。

我目前主要负责了存储引擎的一部分和计算引擎的工作。

插曲：中间我聊到了学校里一门课程目前让我们做一个分布式关系型数据库，不过因为作业要求只有几页PPT，没有提供任何的框架代码和测试用例，这时候胡神给我安利了MIT的6.824课程。

**2.那么您现在平时工作使用的技术栈是怎么样的？是主要用的C++吗？**

对，我们就是都用C++来写的，因为C++兼具高性能和灵活性，所以我们的主力编程语言就是C++，我们会用到比较多的C++11的特性，相对还是比较现代的，但你要具体到C++的话，我们使用的C++语言特性可能没这么多，主要还是利用C++实现一些高性能的东西。我们内部有一套自己实现的类型系统，我们一般基于这套类型系统进行一些开发。

**3.可以简单描述一下您工作日的一天吗？大概是怎么度过的。**

OK，我们的工作时间是965，中午有一小时半的休息时间。通常我九点到公司，等我坐到位置上的时候会先check一下我的邮件，看看有没有重要的事情，然后会刷一下内部的一个软件，它主要是用来跟踪我们的数据库有没有一些bug，如果有的话我就会判断bug的重要性，然后决定自己今天要解决哪些bug，以及今天要做哪些新功能的开发，**我会写一个简短的今日规划**，列一下今天打算做的事情，也许就只有三四条，然后就会开始写代码了，一直写到中午休息左右，当然我会时不时起来休息一下或者走动一下。我们这边的一个特点是你的代码经常需要跟别人合作，所以我们也会做比较多的沟通，有些时候也需要沟通来梳理一下代码的设计思路。

中午的话我会吃个午饭然后回家睡个午觉然后回到公司开始下午的工作，下午基本也还是和上午一样的工作内容。

偶尔我会思考接下去的一个时间段，我可能回去做怎么样的事情。比如说如果这天是周一的话，我会去想这一周我打算做哪些事情。对未来更长时间的规划也是会有的，但是频率没有这么高，大概就是这样子。

我的一天差不多就是这样，看代码写代码调bug，跟人沟通以及开会，有的时候我们还有一些内部的会议和分享，或者是做一些讨论，进行一些系统设计的交流，有的时候也会去面试一些候选人。下班之后应该不要说吧(笑)

【这里我说您觉得方便的话也可以分享，然后胡学长继续分享下班后的生活】

ok那我下班后一般会吃个晚饭消化一下，大概就是七点左右，然后这个时候我会**自己去学一些东西，有可能是看一些公开课的课程，也可能是做一些我比较感兴趣的research和coding**，**或者是看看博客和最新的论文**，到八点半左右我会去健身房锻炼一下，锻炼到十点左右就会回家洗个澡再学一波，到十一点多，然后睡觉，大概就是这样。

【我】感觉这样的一天还是非常work life balance的。

对，我们公司的工作还是比较work life balance的，我个人感觉这个还是非常的重要，因为**人还是要保持持续性的成长，虽然公司也会给你带来一些成长，但是更多的时候可能公司不一定能给你带来很多的成长，这个时候就需要你自己去成长。**

或者说公司带来的成长可能还不足够，可能比较偏科只是在某个方向上的成长，这个时候就需要你自己，你可能需要在某个方向上由你自己的成长，这个时候你可能就需要自己去学习

【我】我也非常羡慕这样work life balance的工作生活

（笑）我相信你们也有能力，如果你们希望的话你们也可以获得一份work life balance的工作，这个没有这么难。

不过麻烦的是，现在国内整个互联网的就业环境是比较糟糕一些，大部分互联网公司可能都是996这样的模式，你会很难做到wlb，我也经历过这样的模式。这不仅是时间的问题，还有一部分是**你的精力或者说心情都会在日复一日的加班中磨损掉，你下班之后很难再会有心情去学一些新东西**，你会觉得我加班已经这么累了不想再学了，我想休息，这样就很难再成长下去。

所以说这个东西就是一个trade-off吧，看你觉得什么更重要，对我来说wlb更重要，所以我会选择wlb的工作，但是对于很多人来说这个不一定，这个也是个人选择问题。

**4.您刚才提到您会学一些国外的公开课的课程，那您现在一般都在学什么课程呢？**

我现阶段在学的是CMU的数据库课程15-445，这个的话我在来我们公司之前是没有什么数据库背景的，所以我也希望自己能把数据库学的扎实一点，对，这个是一点。这门课还有一个Advanced版本是15-721，不过这门课是更偏向于做research和读paper的，对于目标是工作的学生来说15-445就足够了。我也有在看，不过比较忙吧，所以没有怎么推进。

**5.您在学习外国名校公开课的时候更倾向于用什么样的方式学习呢？想我的话会更喜欢直接看提供的notes和slides，然后尝试着做一下提供的lab，请问您平时一般是用什么样的方式进行的呢？**

我其实还蛮喜欢他们讲课的视频的，我一般挑课程的时候倾向于去挑有视频的课程，所以我看的课程一般都是有视频的，因为我觉得slides和notes制作的时候可能没有这么self-contained，你直接看slides可能很多东西是缺失的。

第二个很重要的原因是，除了授课老师会讲一些slides上没有的东西以外，还有一个就是**视频里会有学生在提问，很多斯坦福，MIT等学校的学生他们本身的水平也是相当高的，他们可能会提出一些很有价值的问题，不管是问题本身还是教授的回答都会很有价值**。有的时候我会去思考他们提问题的角度，我会去想为什么我没有提出这样的问题，我的思维是缺了哪一关。

还有一个原因就是看视频可能能用一种比较放松的态度去看，因为我白天工作也比较累了，所以也更喜欢轻松的方式。

不过我认为看不看视频都可以，算是自己的一种选择，不过他们提供的lab一定要去做。

【我提到学校里很多课程的作业可能质量比较糟糕】

对，这个是因为国内对本科基础教育可能没有这么重视，没有投入很多精力在作业的设计上，这点国内可能还没有这个条件，也没有这么重视。

我可以讲讲之前的一些经历。之前那门《机器学习》课程的lab实际上是我设计的，我之前出这些lab可能前前后后花了将近一个月的时间，每天八到十个小时的时间，才出好的。这种时间我觉得可能没有几个TA愿意去花的。但是像国外他们的TA团队比较庞大，而且他们的TA都是全世界同龄人中最顶尖的一批人，水平非常高，另一方面他们本身也特别重视教学质量，还有一个原因他们TA的工资非常高。所以因为种种原因，你在国内课程很难碰到高质量的作业，除非TA愿意用爱发电，这也是个很大的问题。

刚刚我也说了，你如果要把一门课掌握清楚的话，视频可以不看但是lab必须要做，我认为这是非常重要的一部分，尤其是计算机，你学了之后不去写代码你很难真正把它学懂。但是有了lab，你去写代码然后你去跑出一个结果，它相当于是**提供了一个监督信息，来检测你的学习情况**，经过这个过程你会对知识掌握的更加扎实。学习这个东西还是没有捷径可以走的。

所以我觉得学公开课一定要做lab，另一个的话我们都是程序员，coding是我们的根本竞争能力。我觉得我们作为程序员一定不要轻视工程能力，这是我们的立身之本。

**6.我看到您大力提倡组队学习、组队刷课这样的结对学习的模式，您也做出了很多具体的实践，不过我感觉还是难以找到一起学习的“队友”，您认为组队学习应该怎么解决找不到“队友”的问题呢？**

组队学习这个对我自己倒是没有多少帮助，因为我基本一个人就可以搞定公开课程的学习，就算我一个人不行，身边也有不少小伙伴在学习可以互相交流。

但是大家自己进行公开课课程学习的时候是没有professor和TA的，只是利用他们的资源，如果我们碰到作业里的坑，可能掉进去很久。我们不像这些学校的在校学生可以直接去问TA，会掉进坑很久，很多时候浪费掉大量时间，这是不值得的，这个时候我就想，我们组个队大家互相讨论就可以避免很多坑。

第二点的话，可能很多人学习的时候会产生孤军奋战的感觉，如果认识一群人一起学习一门课程，那可能就会感觉很快乐，所以我就想着建个群大家一起交流，一起答疑。我和我的朋友们有时候会在群里进行一些义务答疑或者讨论之类的工作，那我觉得这样对大家而言都很意义和帮助。

回到你的问题来说，找不到队友怎么办呢？其实找到周围刚好有人在学这门课是一件很凑巧的事情，找不到队友才是很正常的情况。这怎么办呢，有几种做法：

* 安利同学一起学，这个事情我做的很多，比如说之前有做斯坦福的计算机网络和MIT的6.S081我都能抓到朋友一起学，你要告诉他这些课有多重要
* 第二个做法就是不局限于身边的人，要不就去98知乎上找，有时候你能找到2-3个人一起学可能效果就会不一样。

**7.我了解到您从校招到社招收获了十几个高级别的offer，算法和研发的都有，但我看到您两次择业都选择了研发岗，我想知道您是怎么考虑的？您为什么每次都会放弃算法而选择研发岗？您刚才已经提到work life balance，但我觉得这可能不是您每次都选择开发的最重要理由，因为开发岗的工作可能会更累，那么您为什么会坚定选择研发岗呢？**

这个其实是涉及到一个个人喜好的问题。其实我很喜欢机器学习这个领域，我研究这方面已经有四五年了，做过很多项目，也在一些比赛中拿到过TOP名次，也有顶会论文的产出，我读研的几年其实主要都是在搞算法岗的工作。

我做过比较多的机器学习方面的事情，也有过比较多的积累，所以这个就更加有意思了，我在算法岗的招聘市场上是比较有竞争力的，如果要排个名的话，我觉得我在算法岗中的排名肯定是要比研发岗中的排名要靠前的，那为什么我最后要选研发的工作呢？那这个有很多原因：

* 首先work life balance是一个比较重要的原因，你在国内的互联网公司的算法岗要做到wlb是比较困难的
* 还有一个是个人喜好的原因，我其实很喜欢机器学习，包括它的教学工作，我之前也做过机器学习的授课工作，但是我不太喜欢机器学习尤其是深度学习的工作，我也不是很喜欢深度学习的research。因为现在的深度学习的岗位可能很多都是清洗数据，你很多时候都在清洗数据或者观察数据，然后呢你可能要训练一个模型，然后你要进入一个调参的阶段，然后要把它部署下去，但这个过程中的环节太玄学了，**我会特别希望我对这个领域我能做到in-and-out，我能把细节和逻辑理清楚，我希望知道为什么这个方法效果好，那个方法为什么效果不好，碰到效果不好的时候呢我希望能有比较系统化的方式能诊断哪里出了问题**。我会更希望我的工作流程是一个scientific的过程，但我觉得我现在还没有到能scientific做深度学习任务的水平
* 还有一个原因是我觉得深度学习要写的代码比较少，可能大部分都是些没有什么技术含量和工程上的难度，我作为工程师还是希望自己的工程能力非常强，那研发很能锻炼我的工程能力。
* 还有最后一点是我纯粹的个人的看法，如果有冒犯的话我要说声抱歉。我会觉得，现在的深度学习领域，我做了两三年，我并不觉得2020年的我比2019年的我更了解深度学习多少，也不觉得我比2018年的我强很多。如果说2018年的我和2020年的我同时去做一个深度学习项目，完全有可能2018年的我做好了，而现在的我没做好。然而，我希望我的工作能给我带来实实在在的积累，也就是一年后的我一定要比一年前的我强，这样我等到十年二十年以后，我会成为这方面真正的专家，新进来的后浪可能就没法把我淘汰掉，但深度学习没有给我这种感觉，我要是干到35岁，25岁毕业的学生可能就能把我淘汰掉。

**8.那您研究生阶段主要是做算法的，但是现在工作做得是数据库研发的岗位，您觉得您在研究生期间的学习对于现在而言，给您带来了什么收获和帮助？**

我觉得收获最大的应该是一种系统解决问题的方式，当你拿到这个问题的时候，你先要知道这个问题的定义是什么，像机器学习问题你就应该知道input和output是什么。

举个例子就比如人脸识别。你的问题可能是判断一个在你面前的人和数据库里的照片是不是同一个人，那你的输出结果就是“是与不是”，接下来你就要对这个问题进行抽象。那怎么判断呢？首先我可以获得面前的人的人脸的照片，那现在的问题就变成了判断两张照片是不是同一个人，那这又可以抽象成一个0-1的二分类问题，对吧，这就是核心的任务。

那接下来要对这个问题进行思考要用什么方式解决它，如果你不熟悉就需要调研别人的解决方法，你要去分析别人的方法是否适用于你的场景。在你定好了使用什么方法以后，你就需要去写一些模型的代码，当然也可以直接git clone之后做一些修改，接下来你要对模型进行训练。等模型训练出来之后你要把这个模型集成到一个人脸识别的系统中，所以你要写一些后端的代码来调用你的模型。

写好之后你可能又会发现我这个系统可能跑在服务器上，有很多客户端要同时访问，那你可能就会遇到性能瓶颈，那你就要考虑怎么优化系统提高并发量，那优化完了之后你还要测试系统本身的准确度和稳定性，要做很多很多测试，等他们都测试完了，OK恭喜你有了一个产品，那接下来你还要跟人沟通把你的产品发布出去。

我讲了这么多其实就是想说，整个过程不仅仅是训练一个模型这么简单，你要考虑的东西还有很多，所谓的解决问题的能力就是说，给你一个问题之后你要系统地分析抽象，知道自己需要做哪些事情把它做出来，把一个问题做好，你要考虑很多很多事情，而怎么样把一个问题做好，就是我在研究生阶段学到的最重要的东西。因为不管我是做机器学习，还是金融交易系统还是数据库，他们都是一个解决问题的过程。所以大家平时在学校里，或者实习做项目，也要注意培养自己解决问题的能力。

**9.您刚才举的人脸识别的例子里不仅有训练模型，也有很多像后台开发这样和研发比较相关的工作，那您在工作中是否考虑过做一些算法和开发结合的方向呢？就比如R&D岗位和机器学习系统的开发，您是否对这些方向产生过兴趣呢？**

实际上我们DolphinDB这边就有类似的机器学习系统开发的工作，我们叫in-database learning，就是在数据库中完成机器学习这件事，这个工作我们是有在做，并且还有功能的缺失，可能未来会由我和同事一起来完成，所以我是有这方面的兴趣的，并且也在做这方面的事情。

**10.您对学计算机的学弟学妹的职业规划有什么建议或者经验吗？**

我之前写了一篇文章提到了一个乔布斯的演讲，当然真的能像乔布斯这样能追寻理想是一件很luxury的事情，乔布斯当时能说这些话是因为，他面对的是斯坦福的毕业生，他们不仅优秀，而且家里有钱，所以他可以说这个话，但并不是我们所有人都适用这种情况。

我在这边想提的是，他在文章中提到了一种connecting the dots这样的感觉，就是说你在人生中可能会做很多选择，当然你后面可能会觉得之前的某个选择是错的，我觉得这时候就需要抱有乔布斯的这种心态，就是你这些选择，在未来某些时候可能会起到某些作用。现在我自己感觉，虽然我去年的半年工作经历可能不怎么愉快，但我觉得当时这个选择是一个非常合适的选择，因为我能在刚毕业的这个时间第一次就发现高频交易这个行业不适合我，我还是适合做科技行业，我很快就排除了一个错误选项。

我觉得学弟学妹们在做选择的时候不要特别焦虑，而且在做了选择之后也不要反复去想或者后悔怎么样，你可能可以报着这样一种信念，你不管做了什么选择，你在过程中磨练的能力都会在未来帮上你，这样可能就会使得你没有这么焦虑，这是我想说的第一点就是说以什么样的心态来看待人生中做出的选择，不要太后悔过去所做出的选择。

第二点我想说，做选择很多时候真的还是看运气居多，一个人尤其是刚毕业的学生，你看不出一个部门或者说一个岗位的未来的发展，很多时候你做这个而选择真的要看运气，不管你找过多少人交流过。我当年也找过很多人聊，但是我做出的选择还是不适合我，这个事情也很有可能发生在学弟学妹身上，你很难说我做出的就一定是正确的选择。所以我想说的是，对我们个人来说，**我们不能把握住的事情就随他去吧，但是我们要把握住我们能把握住的事情，这就是持续性的成长**，花时间提升自己的能力。

第三点我想说，职业生涯的选择这件事还是**建立在个人实力的基础上**，你实力越强你的选择就越多，及时止损的可能性也就越大，而且你碰到不好的选择的几率也就越小，所以我还是建议学弟学妹们先把基础打扎实。

还有一个是工作之后的心态，比如有个人冲着wlb可能毕业去了微软，结果看到他去了阿里的同学，一年升p6，两年升p7，三年升p8，可能也会心态失衡。我想说还是要**控制好自己的心态，有所失不一定有所得，但是有所得必有所失**。所有的offer都会有自己的优缺点，我建议大家选完之后一定要保持心态平衡。

最后一点是说不同的公司，不同的部门之间可能千差万别，可能这个组996那个组就是965，你可能很难分析清楚这个部门这个公司未来的发展，但你能够把握的，可能是你认识一个很靠谱的学长学姐，找他们去打听哪个组哪个leader比较靠谱。比起选部门，不如选leader，而且靠谱的直系学长学姐说好的leader，一般就可以考虑跟着走。因为如果你自己去选，你选的部门可能很random，你也不知道里面的leader如何；但是如果你跟着靠谱的leader，你自己也能提升比较快，这是我的Google实习的时候leader传达给我的。不同的leader差别非常大，我也经历过这些，跟着一个好的leader你可能会过得非常开心还能学到很多知识，碰到障碍也能帮你提点，碰到不好的leader可能你工作心情也会非常糟糕。

**11.您刚才提到建议学弟学妹们在学校里夯实基础，您觉得这些基础学到什么样的程度可以称得上是扎实？**

我觉得有一个很简单的评价指标，你选一门国外的课程，你把它的lab都做了就说明，你基础扎实了，这是一个非常简单的评价标准(笑)

我觉得大多数人对自己基础扎不扎实是心里有数的，当然我也碰到过自我感觉良好的人谜之自信的。

还有一个可以考虑的地方是，如果你能把一个知识给人讲清楚，那就说明你也搞清楚了。

公司招聘的考察面试可能都是八股文或者怎么样，我觉得你真的把基础学扎实之后这些东西看看面经就可以应该没有什么难度。

【我提到学校里的操作系统和计算机网络等课程中感觉只学了一些理论，而缺少实践】

所以我建议你们，看一看更大的世界，看看外面的课，选用更加合适的资源来学习，如果只是按照国内这一套去学的话，大家心里也都有数，应该是学得不扎实的。

从事程序员行业，多学一门好课可能就值几百万，多学一门好课在你求职的时候可能就能让你的package高五万左右，考虑到你工作二三十年还有升职加薪，可能有几百万了，我不知道这种说法能不能吸引更多人去学外国的公开课(笑)。

【提到大家平时都疲于奔命各种学校里的琐事，包括绩点什么的，可能没有很多时间去学国外名校的一些公开课】

关于这个我可能可以提供一个新的视角，就是我认识一个上海交通大学的朋友，他在网上也有一些名气吧，他叫做迟先生。我跟他还算比较熟，打过一些交道，他在我的一个学习小组里，事实上他大一大二就开始学习外国名校的一些公开课，我听说他学校里的课都是能翘就翘，但是他花了大量时间在自学公开课，最后的结果是他的绩点是他们学院第一，因为他用这些资源学习好之后，来看学校里的课程基本就是**降维打击**。

所以我希望大家如果真的觉得冲突的话，大家可以眼界放宽一点，我只能这样说，尽量花时间去学一些更有价值的东西。如果学弟学妹真的觉得绩点很重要的话，我只能建议你们读研了之后，一定把这些基础的东西在找工作之前补回来。

**12.您在读研期间有多段实习经历，您觉得这些实习经历收获如何，对您的正式招聘有多大？**

实习在国内的来说对于找工作的帮助之大，几乎远超其他的东西的，就是你能够改变的东西中，影响最大的就是实习，哪怕是发论文都可能没有实习重要。实习不仅有转正的机会，而且你在实习中学到的东西，能够帮你解决面试时候的一些问题，另一个的话实习会有一定的镀金作用。我建议有志于就业的同学要抓住机会出去实习，我觉得实习还是非常重要的。

**13.在访谈的最后您还有什么想跟浙江大学的学弟学妹们说的话吗？**

我觉得对大家来说，还是我强调的那些，第一希望大家基础尽量扎实，第二就是大家如果对自己未来的职业发展有比较多想法和规划的，我建议大家保持终生学习的心态和状态，这个很重要我觉得。

第三点是我比较personal的一个想法，我觉得可能很多时候，大家可以考虑自己能为社会创造多大价值，如果你能为社会创造的价值越多，可能你自己收获的利益也就越多，当然这个也不绝对，我待过一些行业，虽然没有多大的社会价值但是也能赚很多钱。

当然我也希望我们浙江大学的一部分同学，可以真的成为竺可桢老校长说的，大学不仅要培养人才，还要能够培养能够主持国运的人才【据考证，竺老校长的原句应该是“大学教育的目标绝不仅是造就多少专家如工程师、医生之类，而尤在乎养成公忠坚毅，能担当大任、主持风气、转移国运的领导人才】，我相信浙江大学有同学能够做得到。

最后我想说，人生和职业生涯都很漫长，如果觉得自己的工作没有意思没意义的话也很难做好，祝大家都能找到自己觉得有意思也有意义的事业。