

## #Email

### #1.拜访

刚刚回国，没有直达西安的飞机，所以先在北京停靠。伯克利学期结束了，交大事情不多，所以这几天打算在北京休息一下，住在何俊逸屋里。

何俊逸第一年考研，今年申请出国，已经拿到一封来英国学校的录取信了，还在等其他学校消息。最近这一年他租住在学长的宿舍里，现在房间空了一张床，刚好可以让我过去住几天。这种感觉很好，虽然我大部分时间独来独往，这么多年了也有了一些好朋友，分散在全国各地。

突然想起了一个笑话。

“你有没有朋友啊？”

-- “这要看朋友的定义是什么”

“说这种话的人一般都没朋友”

我大多数时间独来独往，也不知道怎样的关系才算朋友。和有些人已经很久不联系了，但是一旦我遇到困难，他们一定提供力所能及的帮助。反之亦然。毕业生分散到了全国各地，这么想一下我在好多地方都有了落脚点，心里热乎乎的。

### #1.1 高中同学们

回到时间线上。下了飞机，换上出国前的手机卡，发现没信号，原来欠费太久，号码已经被回收了。再次来到中国农业大学餐厅(第三次了，和何俊逸，侯超一起吃了大餐，有烤鸭，烧鸡，梅菜扣肉，疙瘩汤，酸辣土豆丝。我那饱受思乡之苦的胃终于得到了满足。睡前同何俊逸聊到很晚很晚，天南海北无所不及，才知道牛犇去了澳大利亚读书，藏艺姝去了西班牙，王小萌刚结束了二次考研，于淼淼去了上海财经，李荣冠继续本校读研，琚旋在中国农业大学读研究生，董昭在北京一家很出名的会计师事务所工作..... 总的来说大多数同学都继续深造了。

第二天去拜访了罗兆勇。如今我们这群同学里，抱有学术理想的大概就是罗兆勇，侯超和我了。侯超希望在八年本硕博攻学习结束后去美国做博士后，并寻求教职(大学教授)。罗兆勇也希望去美国读博，可是担心自己的研究经历/成果不足，所以正在考虑读研期间延毕一年用来做研究。罗一直以来都是我非常敬佩的同学，相信今后是前程万里。即使未来依然动荡不定，我们仨依然约好他日异国相见。

为啥都选择去美国读博呢？唉，美国毕竟是学术的灯塔。

接下来几天还打算见一下来北京工作的大学同学，实习时的导师...

### #1.2 导师

今中午见到了微软实习时的导师 Houwen，聊了好久好久。当时我在 Jifeng 名下实习，可是 Jifeng 从未指导过我，最后也将我提前辞退了，那时的我经历了一段压抑的时光。今日重回公司，聊起来才发现 Jifeng 已经去了其他公司，(据说)因为他对实习生一直不太好，也因此

渐渐与组里的人矛盾加深，最后便离开了。这么想来，当时糟糕的实习经历也不全是我的责任，便有些释怀。一直以来，习惯于承担责任，从自己身上找错误，容易变得自卑。也可能错在别人身上呀，也可能俩人都没有错，只是不合适(匹配)。当时那么幼稚，轻易地就被他人否定了，现在我已经强壮了些。

这边实习时，一般每个导师带一两个学生，每日悉心指导，估计比大学里的博士生导师都要负责。我同学实习时每天与 mentor(导师)一起吃饭，争论科研话题，成长很快。而我当时则是一个没人养的小孩子，望着一摞论文，不知道怎么入门。Houwen 不是我的导师，但是机缘巧合相识后，他手把手指导我，那也是我成长最快得一段时间。

晚上去见在北京实习的大学同学，雨辰。当时我俩一起去伯克利读书，意趣相投。他做得一手好菜，每次我都在旁边打下手学习。自从他第一学期结束归国后，我就很久没吃到那种家乡的味道了，笑。

## #中西

直到飞机落地的时候，还没有切实感受到自己已经回到了祖国。  
身上依旧携带着伯克利生活留下的刻痕。数地铁站时自然而然的轻声说出了 one,two,three...而不是一二三四；推开门一定会停一会直到身后的人通过；习惯性点开国外的手机 app(应用)才想起在国内是打不开的。  
刚感慨完国内公共交通的便利，下一刻就被人潮挤到了角落里...  
时差还没倒换过来，白天眼皮突然变得沉重，看了手机才发现那是伯克利的睡觉时间。  
出去吃早饭，被驴肉火烧的香气吸引，看了一下价格，在心里默默除了一个 7(美元汇率)，哇，才 1 美元！狂喜。  
同学看我国外生活过的拮据，每次都点好多菜，可是这么多菜加起来还不到我在美国平常一顿饭的价格。这几天幸福地吃撑了。

## #圣诞

每逢佳节倍思亲，我赶在美国圣诞节之前离开，到达北京时是 25 号下午。  
在美国，圣诞便相当于国内的春节。圣诞前后大部分店铺都关门了，只有中餐厅还在勤恳地营业。圣诞前一个月便已经能感受到那种氛围。店铺里放起了圣诞主题曲，邻居家挂上了彩灯，学校食堂里摆上了圣诞树...上课地同学也越来越少，哈哈哈

## #未来

在伯克利最大的收获就是明确了自己下一阶段的方向。  
大一时订下的本科目标(马拉松，编程竞赛银奖，出国交流至少一学期)已经完成了，但是下一阶段要做什么，我想了很久也没想清楚，直到在伯克利这边，我发现了自己对超级计算机(比如神威，太湖之光)的研究兴趣。  
我已经申请了几所国外的学校，伯克利的教授也为我写了推荐信，申请结果来年春天，2，3 月份便会陆续公布。希望受到录取信。

我最想去的地方是 ETH(苏黎世联邦理工大学)，那是欧洲最好计算机学校吧，位于瑞士，是

爱因斯坦的母校。世界上超级计算机有一个排名榜单，ETH 便有一台排名前十的机器。而且那边学风严谨，注重理论研究，也非常契合我的性趣。我申请了 ETH 的硕士，希望以它为跳板，毕业后进一步申请更顶尖学校的博士。我也申请了美国一些学校的硕士或博士项目，就看谁要我了。

当然我最最想去的地方是伯克利，可是我没做好充分准备，错过了伯克利的申请截止日期。在伯克利的最后一段时间，我为自己正在进行的科研项目写了一篇小论文发给教授 James，James 便开始批改指导，并反馈给我。发回来的论文带着密密麻麻的批注，语法单词的使用也一一矫正，令我十分感动。那段时间，每天深夜我将论文改好发给他，早上便收到回复，我再改一整天，深夜发回去。一连两周，改了七八版，特别累，进步也特别快。

能够遇到这种愿意为学生一字一句改论文的导师，真的很奢侈。

我特别希望能成为 James 的学生，将来也把这种严谨的治学态度传承下去。

(附图)

The image shows a document page with mathematical proofs and a sidebar with sticky notes. The document is titled "Then we will prove a tradeoff between memory and communication for the simplified CNN model." and contains several lemmas and theorems. The sidebar contains several sticky notes with the text "demmel" and "requires additional load balance assumption".

**Document Content:**

Then we will prove a tradeoff between memory and communication for the simplified CNN model.

**Lemma 2.3** Consider the simplified CNN computation  $Out = Image \times Filter$ , where  $Image$  is  $N$ -by- $N$ ,  $Filter$  is  $L$ -by- $L$ , and  $Out$  is  $N$ -by- $N$ , on a  $P$ -processor distributed-memory parallel computer with  $M$  words of local memory per processor. Consider a processor that sends or receives  $T$  words of data during each phase except perhaps the last one, and performs  $N^2 L^2 / P$  scalar multiplications. The total amount of words that the processor must send or receive is at least

$$\left( \frac{9N^2 L^2}{P(M+T)^2} - 1 \right) T$$

**Proof.** The total number of scalar multiplications is  $N^2 L^2 / P$ , and from Lemma 2.2 we have that the maximum number of useful multiplications can be performed during one phase. Therefore, the number of phases is at least

$$\left\lceil \frac{9N^2 L^2}{P(M+T)^2} \right\rceil$$

Since the processor receives  $R$  words during each full phase, the total amount of communication cost is at least

$$\left\lceil \frac{9N^2 L^2}{P(M+T)^2} \right\rceil T$$

As  $\lceil x \rceil \geq x - 1$  for any positive  $x$ , we also have

$$\left\lceil \frac{9N^2 L^2}{P(M+T)^2} \right\rceil T \geq \left( \frac{9N^2 L^2}{P(M+T)^2} - 1 \right) T$$

which concludes the proof.

**Theorem 2.1.** Consider the simplified CNN computation  $Out = Image \times Filter$ , where  $Image$  is  $N$ -by- $N$ ,  $Filter$  is  $L$ -by- $L$ , and  $Out$  is  $N$ -by- $N$ . On a  $P$ -processor distributed-memory parallel computer with  $M$  words of local memory per processor. At least one processor must send or receive at least

$$\frac{9N^2 L^2}{4PM} - M$$

The total amount of communication is bounded by

$$\frac{9N^2 L^2}{4M} - PM$$

**Proof.** Notice that our lower bound in Lemma 2.3 is dependent on  $T$ , which is a free variable. In order to find the greatest lower bound for large problem sizes, we want the positive

**Sticky Notes:**

- Sticky Note 2019/11/26 3:53:17 demmel 选项
- Sticky Note 2019/11/26 3:53:59 demmel 选项
- Sticky Note 2019/11/26 3:54:24 demmel 选项
- Sticky Note 2019/11/26 3:54:47 demmel 选项
- Sticky Note 2019/11/26 3:55:21 demmel 选项
- Sticky Note 2019/11/26 3:56:23 demmel 选项

如果最后没能收到理想学校寄来的录取信，我便会工作一段时间，第二年继续申请，第三年继续申请。现在心态变得“佛系”，因为已经明确了自己的目标，知道了要学习哪些知识，无论身处何方，录取与否，都会一点点学习，向着目标前进。