

Summary of *The Ph.D. Grind*

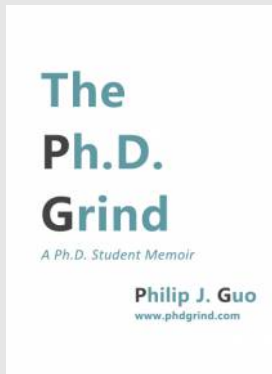
支部读书推介会

推介者：许利杰

中国科学院软件研究所
软工中心

2012 年 12 月 31 日

书名



Ph.D.

- Doctor of Philosophy
- 需要三名教授组成论文委员会并认可论文
- 计算机系大多需要4到8年时间毕业
- At least two respectable papers

Grind

- 碾，磨； 苦差事，苦活儿
- 研磨记：一名博士生的回忆录
- 6年，122页

○●



○○○○○○○○○○○○○○○○

-

要不要读

为啥读

- 受社会背景和父母影响，长大想找稳定又有点地位的工作，首选教授
- 本科暑期实习，觉得公司的活没有挑战
- 大三决定加入MIT的5年制的本硕连读项目，决心以后读PhD做研究

期间工作

- 硕士项目——对计算机编程语言C和C++进行运行时分析
- 读硕期间发了两篇二档的文章（ISSTA），其中一篇是一作
- 加上导师的推荐，顺利拿到心仪的Stanford的offer
- 还有两个奖学金，可以支持他五年的学费，并且衣食无忧

糟糕的第一年-研究方向

- 决定研究能够提高编程人员效率的工具
- 感觉本科实习和硕士期间编程繁琐，处理数据麻烦，工具不友好
- Stanford搞软件工程领的教授少，选了做系统程序分析的副教授



Dawson Engler

- MIT泰山北斗Frans Kaashoek教授的学生
- I now I have my own students, who sometimes listen to me.
- Junfeng Yang
- Best Paper, OSDI 2008/2004/2000

糟糕的第一年-具体工作

工作内容

- 测试Klee是否能找到Linux驱动程序的bug
- 配置Klee，并搭建好测试程序的环境需要n个小时，重复三个月
- bug没检测出来，反而发现这个组开发的Klee的工具具有bug
- 痛苦测试，直到paper deadline的前三天，Klee终于被初步改进
- 通宵三天写paper，被拒

心情

- 与自己的“用来提高编程人员效率的工具”目标差的很远
- 迫于规则苦干了三个月不需要太多思考的工作
- 导师加上学生需要1个月时间组织文章、写、融合各自的部分、润色
- Guo果断退出Klee项目，开始独自思考人生方向，读大量paper
- 去Google实习
- 思考新的研究方向应该是经验软件工程，确切说是软件质量测试

无奈的第二年

工作内容

- 观察程序员如何写程序，找规律，判定软件质量
- Google程序员不同意，Mozilla程序员不同意
- 与搞HCI（人机交互）方面的助理教授Scott合作
- 与博士生Joel组队，观察HCI本科生如何写一个课程作业程序
- 将各种有意思的发现写成paper，经过Scott的润色，投到了HCI上
- 可喜的是paper最后中了，可悲的是Guo不是一作
- 又回到了Dawson的研究组，回到了痛苦的Klee项目

心情

- 保密的原因，不想被一个站在旁边的人监视着
- 软件工程实战派遇到的最大问题啊，程序好找，人难找啊
- 慢慢体会到了做研究的感觉，比第一年痛苦干活几个月，毫无结果强
- 讲了很多论文撰写、投稿的一些门道，或者叫潜规则吧
- 分析了教授、学生、项目等等的利益关系和心理活动

悲剧的第三年-1

工作内容

- 仍然是低年级学生，只剩导师和马上就要毕业的博士生Cristi了
 - 做交叉检测的项目，目的是为了发现同一软件不同版本之间的不一致性
 - 与第一年生活一样，四个月过后，仍然没有可发表的结果
 - 得意门生Cristi毕业答辩了，想超越这个师兄太难了，毕业前景渺茫
-
- 导师说你要不改进下Klee的搜索算法吧，现有不高效
 - Klee搜索算法召回率提高到了94%，但之前已经是91%
 - 再一次捡起了之前的idea，新来PhD叫Peter，两人重新实现
 - 实现过程中，导师开始不耐烦，嫌他们工作进度太慢
 - 后来Guo和Peter感觉让导师失望了，终止了这个项目
 - Peter更是失望地换了导师，没多久也博士也退了

悲剧的第三年-2

心情

- 想要搞经验软件工程，帮助编程人员提高效率
- 导师却十分看好找软件bug的方向
- 工作状态跟第一年没什么实质区别
- Critsti最后拿到了伦敦帝国大学的教职

插曲

- 在国外，能拿到高水平大学教职对PhD毕业生而言是至高荣誉
- 能够在Stanford找到教职的人，平均顶尖大学的PhD毕业生75个才出一个
- 两年后，David 融合了Guo之前做过的两个idea，终于发表在CAV
- 但这代价太大了，前后经历了三次尝试，4个PhD，5年才得到一篇文章
- 只有一作幸存，Guo离开了Klee项目，另外两个最后连博士也一起退了

暑期实习插曲

工作内容

- 09年暑假申请去西雅图的微软总部研究院（MSR）实习
- MSR研究员看到Guo和Scott合作发的HCI论文后，发邮件邀请
- 推荐到了专门搞经验软件工程的学术新星Tom那
- 与Tom合作研究开始研究关于bug提交、分发、修复中的人为因素
- 依靠几个月的努力，加上微软庞大又规整的bug库，Tom高超的写作水平
- 投了ICSE，Guo是一作，该论文在10年被录用
- 决定回到Stanford，将失去微软强大的bug库，为研究带来巨大困难

心情

- 08年暑假投了MSR实习，结果被拒，原因是paper不够吸引人
- 原来MIT硕士导师也向这位研究员推荐过
- 之后的几年中，Tom将现有工作推进，又发了一篇CSCW11以及今年的ICSE

重生的第四年-1

工作内容

- 打算做IncPy，来减少程序员数据清洗、分析和可视化负担
 - Stanford无人支持，冒险实现自己的idea
 - 将原型设计投到一个USENIX的workshop
 - 认识了来自Harvard的Margo教授，让Guo去联系她学生Elaine
 - 继续IncPy项目，想投7个月后的顶会，要找人试用
-
- Berkeley神经学家Fernando要Guo做个报告，并安装一下IncPy
 - 安了软件，收到自己软件不能处理Python依赖包问题的邮件
 - 接下来几周，努力改进，但没有了回复
 - 放到网上，收到3个人的反馈，说实际工作中确实有用
 - 聊甚于无，paper老实地说软件经过3个人试用，提高了效率，被拒

重生的第四年-2

心情

- 对于大多数研究生来说，一般是跟着自己导师做导师申请的项目
- 当然最好是自己的idea并让导师认可，运气好的话能够发篇高水平文章
- 导师负责找另外两名教授组成指导委员会
- Stanford无人支持
- 求N多人试用其软件，大家都不愿意花时间去学一个PhD做的不成熟软件
- 是的，有时候，机会只有那么一次
- 相对于实际的软件开发，教授们更偏爱理论创新

高产的第五年-1

工作内容

- 下一个项目SlopPy，帮助用户在程序崩溃时分析查看部分计算结果
 - 花两个月写了个原型系统，投了篇二档会议，不认为是个很强的idea
 - 这个时候，有两篇正在被申，已经放弃做教授的梦想
 - 两篇都被拒，改了后将IncPy转投另一个二档会议，最后被接受
 - Sloppy转投USENIX的Workshop，被接受，参会时一块介绍了IncPy
-
- 又联系了Berkeley的Fernando，说现在计算神经学很多实验不能够重现
 - 跟导师讨论，说要做一个针对Python的可移植的执行结果可重现的工具
 - 大彻大悟，干脆做个通用的可移植的代码执行工具，简称CDE
 - 解决了困扰Linux十几年的不同发行版的软件安装、配置、执行问题
 - 还是半成品的时候，在网上发了一些demo，被评价是“U ROCK”
 - 将工作发了篇短文到USENIX ATC上，还有篇长文到USENIX的LISA上

高产的第五年-2

工作内容

- CDE被NASA, 开源组织, 欧洲网格项目, 各类研究人员, Google等采用
- 被邀请去Google做talk, Youtube直播, 只邀请著名教授或者杰出工程师
- 普通的PhD, 该转到下一个新项目, 学术界负责产生新idea和原型系统
- Guo个性显然与其他人不一样, 一定要做的满意为止

- 担心paper可能不中, 又联系了一名新的助理教授Jeff, 也是搞HCI的
- Guo给Jeff讲, 自己当前的想法就是做个新工具, 发顶级, 两人一拍即合
- 改进Wrangler, 两个半月实现系统的原型加评价, 并写了篇paper
- Jeff写了论文中的Introduction和实验评价与结果解释部分
- Jeff的强大论述能力和文章发表经验, 让这篇paper被顶会相中
- Jeff在自己读博时发了19篇高水平文章, 而且大部分是顶会
- 至此, Guo已经做了4个项目, 而且相应的文章都成功发表了

要结束的第六年-1

工作内容

- 重遇Margo，邀请他毕业后到Harvard做博后，与Elaine一块工作
- 做一个帮助研究人员组织记录实验笔记的辅助工具，称作BurritoBook
- 开始了最漫长的72天写代码阶段，每天工作10个小时以上，中间休息了5天
- 最后完成工具和一篇paper，回到Stanford，跟导师Dawson商量毕业的事情
- 创造了5个软件工具来帮助需要编程计算的人员提高生产效率
- 这5个项目发了1篇顶会，3篇二档会议，3个workshop
- 凑齐了3个教授：Dawson，Jeff和Margo，都有项目合作和文章发表
- 导师Dawson竖起大拇指，表示同意答辩

- 不只介绍5个工具，花了大量时间来组织和构思毕业论文（大概230页）
- 答辩包含1个小时的公开报告和1个小时的不公开问答
- Scott，Joel，Peter，Robert，Greg以及Fernando都来参加答辩，很感动
- 撒花，顺利通过，Guo拿到了Stanford的PhD学位

要结束的第六年-2

心情

- 经过几年的磨难与纠结，人生的重大决定：毕业后告别学术界
- 想做的东西一直被认为是非主流，paper也不够牛，对以后没有足够信心
- I just wanted to say that I really enjoyed it, partly because of the creativity of the work, partly because of the well-prepared talk, and partly because I had spent the previous year doing research programming.
- 那一年，哦，也就是今年，Stanford毕业了26个博士，自我感觉中等水平
- 横跨了编程语言、人机交互、操作系统三个计算机科学的子领域
- 最重要的是Guo已经选择好未来要走的方向：做一个专业技术人员

Words from the bottom of Guo's heart

I like these words

- ① I feel extremely lucky to have been able to take charge of my Ph.D. career in creative ways.
- ② It's been a long, circuitous, and unpredictable journey, but I'm incredibly grateful that I was able to turn this broad topic-one out of dozens that caught my interest over the years-into my Ph.D. dissertation.
- ③ This accomplishment wouldn't have been possible without a rare combination of great luck, personal initiative, insightful nudges from generous people, and nearly ten thousand hours of grinding.

Lessons

- 1 Results trump intentions
- 2 Outputs trump inputs
- 3 Find relevant information
- 4 Create lucky opportunities
- 5 Play the game
- 6 Lead from below
- 7 Professors are human
- 8 Be well-liked
- 9 Pay some dues
- 10 Reject bad defaults
- 11 Know when to quit
- 12 Recover from failures
- 13 Ally with insiders
- 14 Give many talks
- 15 Sell, sell, sell
- 16 Generously provide help
- 17 Ask for help
- 18 Express true gratitude
- 19 Ideas beget ideas
- 20 Grind hard and smart

感悟

人物方面

- ① 不喜欢按套路出牌
- ② 有想法、喜欢独立思考，“我的未来我做主”
- ③ 有完美主义倾向，认准的事一定要做，一定要做好
- ④ 内心强大，纵有再多失败，仍然可以调整自己，适应环境，最后努力获得成功
- ⑤ 善于学习，从各种事情中吸取教训，也能分析出和平衡各种利害关系
- ⑥ 教授也是人，名校教授在PhD时期相当于黄金圣斗士中处女座沙加

感悟

教育意义

- ① 从前到后没有任何炫耀，而是平铺直叙地讲出让人信服的道理
- ② 程老师的《逃离科研》，自己博士生最后选择去人大附中当中学老师
- ③ 其实那是在讨论一个高富帅是买房还是买车的问题
- ④ 受到了本科生、硕士生、博士生、毕业生、退学生、教授的关注
- ⑤ 大家都从中学到了一些东西，这就够了

感悟

学术方面

- ① 研究系统的人真心伤不起，USENIX ATC这样级别会议在Stanford只算二档
- ② 那一档的系统会议每年加起来也就不足百篇吧
- ③ Guo一开始就当起了自己的导师，过强的能力，过早的锻炼最后让他身心疲惫
- ④ 我喜欢他写的这个词 “emotionally and mentally all-consuming”

Thank you!

- 本书没有谈及Guo的感情经历
- 不过他的毕业论文中感谢了wife，也许一开始就有稳定的感情吧
- 书还提到自己经常回家，看了他在USENIX的报告视频，很帅的
- 《爱很简单》《Melody》《蝴蝶》《真的+猜不透》《后来》