

深碗學習從下課後開始

Hsuan-Tien Lin (林軒田)
htlin@csie.ntu.edu.tw

臺灣大學資訊工程系

March 28, 2013

2013.01 接到邀請，但要講什麼呢？

- 課堂上可以怎麼惡搞？
 - 2012 ML: 不同語言的問候和笑話其實通常對培養過笑點的大學生研究生們來說**不是很好笑**
 - 2000 OOP: 謝謝老師還不忘變換身分請C學長**講冷笑話或唱歌**讓同學們醒過來
- 如何虐待學生？
 - 2012 DSA: 不教而殺為之虐，這學期感覺有點**被虐了**
 - 2012 ML: 謝謝老師帶給我們這麼精采又**慘無人道**的課
- 如何壓榨學生的課外時間？
 - 2012 DSA: **HW6**花了我大概**40**個小時卻也還是沒有全部完成
 - 2011 ML: 這是我第一次**熬夜熬這麼多次**的一學期

但是聽說會**錄影**……故先偽裝一下：

深碗學習從下課後開始

林軒田

- 臺灣大學資訊工程系學士(1997–2001)
- 加州理工學院電腦科學系博士(2003–2008)
- 臺灣大學資訊工程系助理教授(2008–2012)
臺灣大學資訊工程系副教授(2012–)

專業興趣：

- 機器學習(共同著作 *Learning from Data* 一書 )
- 資料探勘(共同指導臺大隊連三年獲全球資料探勘杯[KDDCup]冠軍)

- 臺灣大學：
 - 機器學習(Machine Learning): 2008–2012，研究所(大三以上)選修課程，英語授課，80-130 人
 - 物件導向程式設計(Object Oriented Programming): 2009, 2010, 2013, 大學部(大一以上)核心選修課程，中英混合授課，80-120 人
 - 資料結構與演算法(Data Structure and Algorithm): 2011, 2012, 2013, 大學部(大一)必修課程，中英混合授課，80-100 人
 - (合授) 機器學習理論與實務(Data Mining Case Study): 2010–2013, 參加 KDDCup 競賽訓練課程，約 20 人
- 加州理工學院：擔任數門選修課程助教

本演講：個人的成功與失敗經驗之(片斷)整理

關於主題

「臺灣的大學生對一門課程課後花的時間，平均不到三十分鐘，但國外知名大學學生對一小時的課，課外要花二到三小時」(蔣丙煌·聯合報·2009)

<http://catalog.caltech.edu/>

Each course at the Institute is assigned a number of units corresponding to the **total number of hours per week devoted to that subject**, including classwork, laboratory, and the normal outside preparation.

- CS 1. Introduction to Computer Programming. 9 units (3-4-2)
- CS 52. Microprocessor Systems Laboratory. 12 units (1-11-0)
- CS 150. Probability and Algorithms. 9 units (3-0-6)
- CS 151. Complexity Theory. 12 units (3-0-9)

- 學生：要學得深入，只靠課內三小時是不夠的
- 教師：要教得紮實，只靠課內三小時是不夠的

- 加深學習的作業
- 加深學習的專題
- 加深學習的助教
- 加深學習的課堂

加深學習的作業：公平的課堂規範(policy)

- 公平的規範：鼓勵認真，「打擊」投機份子
—讓遊戲、競賽、**學習**有了基本的邏輯與目標
- 學習的「功利主義」：成績、榮譽、滿足感……
- 公平 ≠ 齊頭式平等
—齊頭式平等其實是種不公平

我的想法：學習要成爲吸引學生的「遊戲」，首要有公平的規範

當 A 同學參考了過度認真的 C 同學的解答？

雙輸：C 同學沒有被獎勵學習；A 同學沒有學到更多東西

Course Principle (root from Caltech):

Taking any unfair advantages over other class members is not allowed.

It is everyone's responsibility to maximize the level of fairness.

- 加上明確的規範文件

www.csie.ntu.edu.tw/~htlin/course/dsa13spring/doc/policy.pdf

- 與課堂解說

Collaboration and Open-Book

- homework discussions: encouraged
- but fairness?
write the final solutions alone and understand them fully
- references (books, notes, Internet):
consulted, but **not copied from**
- no need to lend/borrow solutions

to maximize fairness (everyone's responsibility),
lending/borrowing/buying/selling not allowed

Collaboration and Open-Book

to maximize fairness (everyone's responsibility),
lending/borrowing/buying/selling not allowed

to maximize fairness (everyone's responsibility),
lending/borrowing/buying/selling not allowed

to maximize fairness (everyone's responsibility),
lending/borrowing/buying/selling not allowed

Deal? If your classmate wants to borrow homework from you, what do you say?

一些作法(續)

- no cheating, no lying, no plagiarism (抄襲)
- no individual score change: [白話] 要過關或高分，**要靠努力**，不能靠求情

收獲：在 95%+ 的同學身上看到良好的成效！(副作用?)

加深學習的作業：從作業中學新知識

- 作業不(只)是課後複習(深入舊知識)，也可以用來深入新知識
- 怎樣的新知識？要像「拼拼圖」
 - 有關的：講清楚和舊知識的關聯性
 - 重要的：學會了對整個課程會「有用」
 - 令人驚訝/啓發性的：創造額外的學習動機

我的想法：在作業中可以比課堂內學到更多更珍貴的知識

一些作法 @ 機器學習

- A方法(上課不教):
 - 作業：舉個理由在作業中由上課有教的P方法啓發A方法，請同學實作(發現方法的好處和潛在的問題)
 - 四週後：上課時用另一個推導方式，導得一個A方法的姐妹版(請用這個新方式解釋之前所發現的潛在問題)
- V性質：
 - 上課：V性質一般和P性質正相關，不過有一些反例
 - 作業：證明某個特別設計過的反例
- S方法、E方法與F工具：
 - 作業：運用上課教的F工具，和S方法結合，設計一系列的題目
 - 四週後：我的碩士論文的主要想法，來自把S方法和E方法利用F工具連結。
—嘖，你們作業已經導過了！

收穫：把作業當成**建構整門課的重要拼圖**，把握用作業延伸學習的機會

加深學習的作業：閱讀、觀看與搜尋

作業只能手寫/打字/寫程式嗎？

- 閱讀作業：課本、補充教材
- 觀看作業：相關影片、展示程式、線上課程
- 搜尋作業：相關課題、歷史

我的想法：這些作業是否較適合發展「自我學習」的能力？

一些作法

- 閱讀作業：

- 課本：每週指定閱讀作業，包含沒教的章節，並配合手寫作業及考試題目
 - 除「機器學習」課程外，**成效極差**

- 觀看作業：

- 線上課程：在上課前指定必須先去看完的線上課程，上課時直接接續下去，並配合手寫作業
 - 成效極差**

- 搜尋作業：

- 相關課題：在手寫作業中直接指定去找一個方法，搞懂以後用自己的話寫出來
 - 大部份同學都完成了任務**

現狀：還在實驗更適合臺灣學生的方式

加深學習的作業：丟到海裡游泳

作業只能出「符合程度」的嗎？

- 引導(逼迫)學生離開舒適圈
- 鼓勵學生發揮創造與想像力

我的想法：「超越程度」的作業，也是種可能的選擇

一些作法

- 超越程式程度的作業：
 - 資料結構與演算法(大一)：與 KDDCup 隊伍中的研究所學長一樣，處理 7.7GB 的資料
 - 在學校網路下載也要超過十分鐘
 - 用 Notepad 打開電腦會當掉
 - 程式沒寫好就會跑個幾天幾夜
- 超越數學程度的作業：
 - prove or disprove
 - 用不同角度的推導
 - 論文中經典結果的變化型式
- 每年把一些作業的難度調昇兩到三倍

收穫：刺激學生討論、參與助教時間並學到「一些」寶貴經驗

加深學習的作業：「無理數」作業

老師，可不可以告訴我哪裡有某某某定理的證明？

- 可以，但讀別人的證明，不如自己證一下
- 想出成常規的作業，卻覺得啓發性不夠……
- 轉念一想：「不用交的作業」最有趣

我的想法：用出作業的語言，引導學生自行導出經典的結果

一些作法

- 2012 機器學習作業 e/2：七小題證明 Hoeffding 不等式(機率中常見的不等式)
- 差一點就出了作業 $\sqrt{21}$ ：不過沒有力氣出完……

收獲：部份學生由此學到了更深入的知識

加深學習的作業：作業中的彩蛋(bonus)

這個題目好棒，可是這次作業已經太多題目了，怎麼辦？

- bonus 加分題：「騙」學生去做更多題目
- 做出加分題，更有成就感！
 - 要有點難度
 - 要跟「基本題」有點關聯性
 - 要有啟發性

我的想法：讓同學有「自我挑戰」及「破表」的快樂

一些作法

- 以機器學習的角度論「數列預測」的合理與不合理性(啓發)
- 證明某個上課教的不等式其實是等式(困難、啓發)
- 證明 KDDCup 競賽中可以「聰明的作弊」(困難、啓發、只要用到基本題的工具)

收獲：許多學生爲了「做出」加分題，纏著助教問個不停

加深學習的作業：「合法的」遲交

老師，我真的好想把這題加分題解出來，可是死線已經到了！如果再給我幾小時我一定會寫出來的。

- 作業的期限是爲了公平性，不是爲了「限制學習」
- 「彈性的延長死線」，搞不好可以學得更多？

我的想法：給同學「把作業做完」的一點點時間彈性

「金牌」

- no individual extension unless not violating the principle (e.g. institute-established cases of illness or emergency)
- late penalty:
90% of the value for 12-hour late, 80% of value for 24-hour late, ...
- **four penalty-free late half-days (金牌) per person**

收獲：許多好的副作用，如「找理由延交作業的同學變少了」

加深學習的專題：競賽式

- 全球資料探勘杯[KDDCup]: 全世界最重要的資料探勘大賽
—每年四到六月，我們課程裡的 20 名學生非常拼命！
- 加州理工學院開辦過類似課程
—學生爲了「百萬獎金」而拼命！
- 外在競賽使學生拼命
—那內在競賽(同儕競爭)呢？

我的觀察：內在競賽讓學生「更拼命」！

一些作法(機器學習)

仿照資料探勘大賽，做出最「準確」的預測方法

<http://main.learner.csie.ntu.edu.tw/php/ml12fall/main.php>

- 約 40 個隊伍，總共 13057 筆預測記錄，平均每隊做了**超過三百次**的嘗試
- 這些嘗試整合了整個學期的上課內容，**以及所有學生們搜尋得到或創造得出來的新方法**
- 課內頒獎：歡樂的教育機會

Award 1: First Submission Award

team	scoreboard	hidden	algorithm	time
Not Here~ Combo Three!!!	0.5018	0.4998	Random	2012/11/27 20:28:38

Award 4: Happy 2013 Award

team	scoreboard	hidden	algorithm	time
Minimaximizer	0.7632	0.7407	rwa	2013/01/01 00:00:08

Award 5: Goodbye 2012 Award

team	scoreboard	hidden	algorithm	time
anything	0.7704	0.7527	b	2012/12/31 23:59:24

Award 7-8: Hard Working Awards

team	submission count
A	1097
anything	1149

Award 13-15: Hidden Winner Awards

team	scoreboard	hidden	algorithm	time
r5082	0.7607	0.7560	test4	2012/12/15 10:59:07
super idol	0.7609	0.7562		2012/12/16 23:03:44
anything	0.7647	0.7612	LO	2012/12/18 13:44:15

一些作法(資料結構與演算法)

做出最「快」的程式

電子郵件的搜尋引擎：

<http://main.learner.csie.ntu.edu.tw/php/dsa12spring/main.php>

- 約 40 個隊伍，總共 5763 筆執行記錄，平均每隊做了超過一百次的嘗試
- 這些嘗試整合了整個學期的上課內容，以及所有學生們搜尋得到或創造得出來的新方法
- 課內頒獎：歡樂的教育機會

收獲：「追求卓越，互相學習」

加深學習的專題：開放式

- 緣起：「懶得」出期末專題的題目，覺得怎麼出都不夠嚴謹
- 轉念一想，換個方式吧！出不了題目，就「開放」一點
- 也像種內在競賽
—緣自「趨勢科技百萬獎金挑戰賽」的想法

我的觀察：開放式競賽讓學生「更有創意」！

一些作法(物件導向程式設計)

大富翁

只要是像大富翁的遊戲，什麼都行！

- 需要用到整個學期的上課內容，以及學生自己去學習的更深入知識+設計想像力
- 很多有創意的作品，如臺大選課大富翁、Pornopoly
- 成果展示(11小時連續無休息)：歡樂的教育機會

收獲：「追求卓越，互相學習」

加深學習的助教：公平與正確的評分

- 公平：源自課堂規範(policy)
- 正確：助「教」

我的想法：與助教約定時，一定要講清楚這兩點！

當 A 助教不公平的拒絕批改 B 同學的作業？

重點：助教的專業，來自公平與正確的評分

加深學習的助教：「鼓勵而不寵溺(spoil)」的迅速回答

- 鼓勵的回答：小心的維護學生「願意問問題」的火花
- 迅速的回答：不要讓火燒完了
- 不寵溺的回答：不要「洩答」、給學生思考空間

我的想法：助教是特殊的「客服人員」

一些作法

- 與助教一起維持(信件的)回答品質

FYI, I recently received a few complaints about the "cold" tone of the TAs, though. From my understanding, you guys are not cold and are far more friendly than the students expect. You are just under the pressure of replying lots of emails with frequent mistakes in a timely manner. So I don't intend to ask you to do anything more (in fact I've asked too much from you IMHO). I am just forwarding the complaints casually to you for your reference. You have my full support with whatever decision you believe to be fair for the class and to be helpful in learning.

- 以身作則！

經驗：優質的問題回答，可以造成正向的討論氣氛

加深學習的助教：主動出擊的助教時間

如果助教時間沒有人來，代表……

- 助教不夠有幫助？
- 學生不敢來？
- 作業不夠難？

我的想法：其實只是「不習慣」找助教、「不習慣」討論

一些作法

- 「地下室」的助教時間(2012 資料結構與演算法—**非常成功**)
- 「強制性」的線上討論—不回答郵件問題(2012 機器學習—**還算成功**)
- 「非正式」的線上討論：BBS? **還算成功** Facebook? **非常成功**

經驗：使用最能接觸「顧客」並強迫他們回應的方式

加深學習的課程：教與不教

- 什麼不教：閱讀作業、程式設計的細節程式碼、無法「闡述」的證明、程式執行環境的設定
—教學不是餵食
- 什麼可以點到為止：重覆的背景知識、一步之遙的推導、方法/理論背後的洞見、「教不完」的地方
—故意的留白，也是種美
- 什麼一定要教：脈絡、動機、思考

我的想法：教學效果不(夠)好者，不教！但應以教「完」(深入)為目標

一些作法

資料結構與演算法：選擇性教「完」(選擇重點，不要只選前面)

不講(鼓勵閱讀)、不講(閱讀作業)、略講、詳講、補充講

- 學生背景：懂C 程式設計
- 第一章：C++ 程式設計—與C 重覆部份點到為止；可自行閱讀處(例如：型別與型別轉換)不讀；有用處(運算子重載)略講並配合作業練習，極重要處(指標與參考)認真講。
- 第二章：物件導向程式設計—不教！短短的一章，講也講不好。鼓勵學生閱讀或未來修習更深入的課程。
- 第三章：陣列與串列—略講，但加上大量補充以避免淪為無聊之純複習。
- 第四章：數學分析工具—好好講，但控制在絕對必要之內容，以避免學生在此時被數學擊倒。

收穫：「若」其他作業成效良好，學生可從教「完」的全書學最多東西

加深學習的課程：著重思考與互動

- 上課要動腦，下課腦子才能繼續轉
- 同學的「冷漠」？「難以控制的風險」？

我的想法：只需要 10% 的人願意參與互動，就可以活絡 90% 的人的學習效果

一些作法

- 學生提問
 - 沒有不能問的問題
 - 讓學生回答學生的問題
- 亂數抽答
 - 「下一位」也是種答案
 - 「老師你覺得呢」也是種答案
 - 「答錯了，別急著糾正」
 - 把自己加入亂數吧
- 投票——手舉高
 - 「為什麼」？
- 無厘頭問題

收獲：思考與互動的機會，比起制式的知識傳授更珍貴

加深學習的課程：上課中的臭蟲(bug)

- 推導錯誤怎麼辦？投影片有錯怎麼辦？
—「化危機為轉機」，鼓勵「大家來找碴」！
- 深化學生的課堂參與感
- 教導學生「懷疑」的重要
- 「權威掃地」？拖慢進度？

我的想法：人非聖賢，有過則改，何須有懼？

一些作法

- 不小心的錯誤
- 故意的錯誤(要演得像，不可以偷笑)
 - 數學算式
 - 現場寫程式
- 大方的說自己想不清楚，找學生來推導

收獲：在所有與學生互動的方法中，「故意的錯誤」是最成功的一項

加深學習的課程：「試著突破」困難的課題

- 在「課堂中」直接提昇深度
- 老師不害怕教難的東西，學生也不害怕解決難的問題
- 冒著讓學生困惑、迷失甚至厭惡的風險

我的想法：不一定要詳細教，但要試著帶學生「突破」

機器學習：某個在書上寫了五頁的「經典」證明

- 2008：花兩個小時詳細教完，從學生反應看來，應該沒有人聽得懂
- 2009：幾乎沒教，附上教材讓學生自行閱讀
- 2010：教了證明的骨架及洞見，沒有教細節。有部份學生自行回去閱讀教材後，說「看懂了！！真是感動人心的證明」
- 2011, 2012：視課程進度及同學們在課程其他部份的興趣及反應，決定要教多少

收獲：很辛苦，但對自己對學生都很值得

加深學習的課程：關於搞笑……

- 降低距離，增進學生互動/發問的意願
- 調節沈重的課程(為學生也為老師)
- 不過，**總有人不喜歡**

我的想法：不需譁眾取寵，但**何妨偶爾瘋狂？**

一些作法

- 冷笑話、自嘲、「逼學生笑」
—時好時壞
- 演算法的內心戲
—還算成功
- 每次上課找一個不同的語言跟大家打招呼(2012機器學習)
—有點超過

警語：要有面對失敗的準備—臉皮要夠厚

- 加深學習的作業
- 加深學習的專題
- 加深學習的助教
- 加深學習的課堂

我的想法：每次課程，都是獨一無二的實驗機會；獨一無二的努力機會
—**bravely try everything creative without regret!**

- 葉丙成教授、陳雅雯、黃旨彥：演講邀請
- Prof. Yaser Abu-Mostafa & many others at Caltech
- 臺大資訊系師長們
- 我的父母—我的啓蒙老師
- 我的太太和小孩：讓我昨夜熬夜準備演講

Thank you!