計算思維與系統設計

顧學雍

benkoo@tsinghua.edu.dn

基礎工業訓練中心

清華大學

前言

網絡化、數字化的社會創造了許多全新的資源分享與配置的方式，讓懂得如何撬動資源的人群得到了大規模地利用網絡化的各種資源者的新機會。然而，這種機會是如何辨識、規劃、應用、並受到客觀而通用的指標檢驗，則尚未有一套簡要的指導原則，因此，我們還沒有找到一個眾所皆知的課程體系、或是項目孵化的“加速器”得以從創業達人的感性導引與重複試錯的硬漢模式，進入一套客觀的、科學化的認知與訓練內容，像訓練運動選手一般，明確地提升個人或是團隊的創新與創業能力。清華大學的基礎工業訓練中心的教學設計者，在跟資本市場的運營老手、國際化的孵化器操盤工作者、與政府的投資機構的負責人等不同領域的專家，經過多次的溝通，我們決定聯手開發一套由不同層次的系統設計工作者，一起用計算科學的底層認知規律，重新地整理並且紀錄多個不同層次的系統設計與檢驗方法，從而逐步探索出一套針對於調動應用技術、營運資金、政策與政府配套資源的科學論證方法。這一套論證方法，將三類不同的系統設計從業人員，從(A)宏觀的機會辨識、(B)中觀的技術架構、到(C)微觀的具體項目論證，圍繞著一個以4個月為週期的計算思維課程，配備一套移動互聯網採集的學習過程行為數據，把根據真實市場動態而反應的創造性行為，根據閉環控制與隨機過程規律的信息歸納理論，收納入通用的知識管理與出版的工具，並將這些學習的動態結果，向全球範圍內自願參與者，持續發佈與採集公開的回饋。而參與這一套知識管理系統開發的所有人員，將依託於一個定義“可計算詞彙” 的編譯程序與公開發佈系統，以及根據現有的“形式化語言”的知識編碼格式，提出不同具體項目和規模化經濟行為的應用場景，適用那一類“形式化語言”的分類標準。這A，B，C三類系統設計人員的學習任務，就是根據計算科學的邏輯論證方式，重新審視潛在機會的編碼與計算方法。三類人員的學習成果，就是持續地使用可超越時空界限的信息搜尋、編撰、出版工具，以他們本業的工作體驗，豐富如何辨識、規劃、應用網絡化、數字化資源的底層知識。此課程的內容累積方式，將引導創新與創業組織的從業人員，成為拓展信息科學的社會應用場景的知識工作者，對技術密集型、知識密集型社會的治理方法，將是一種開拓性的知識傳遞模式 。

Who should take this course?

Anyone who wants to become a full-stack digital engineer is welcome. We will guide a novice maker, from knowing almost absolutely nothing, to have a basic operational skill in creating a functional digital product and speak the global language of maker/hackers.

Content Outline

This course is an integrated web of knowledge that presents a bird’s-eye view of the maker technology and cultural landscape. The course contain the following knowledge components:

1. Collective Knowledge Management (Mediawiki, Git, Stack Overflow)
2. From Basic Digital Circuits to Software Applications (Nand2Tetris.org)
3. Key Personalities, Technologies, and Historical Institutions of the Maker/Hacker Community

We created the content outline of this course after a group of local makers took the famous online course: “Elements of Computing Systems”, (a.k.a. <http://nand2tetris.org>). We realized that to engage novice maker to better realize their objectives, just learning technical content or just knowing certain anecdotal stories of famous hackers is not enough. Effective makers must first learn to work together or utilize existing network tools to leverage technologies and resources developed by others, yet keep a good habit in using these tools. Then, they also need to have a good knowledge in mathematical/logical reasoning, and use or at least understand the various kinds of tools for different layers of abstraction. Finally, while they are learning about these ideas, makers need to know who and where these technologies or innovative devices came from. With these inter-related contextual knowledge, decision on what to make, and how to make, can be must better supported by a consistent intellectual framework. In other words, this is a digital technology literacy course designed for people who want to “make” in a networked society.

Infrastructure Preparation

This course first introduces popular Internet co-creation tools, such as Mediawiki and GitLab/GitHub services as a common protocol for sharing and managing collective knowledge. Novice makers would not just write code to simulate digital logic circuits, computer architectures, operating systems, programming languages, compilers, etc. they must also frequently store their incremental learning results using version control systems such as Git, and write appropriate learning reports in MediaWiki. This requirement is to build a habit of writing for novice makers. Clearly, to build this program, relevant software and hardware infrastructures must be prepared for all participating students. This include a website for Wiki, and a GitLab/GitHub compatible local file synchronization system, and other related project management software, and public social media services, such as WeChat and Youtube-like content sharing platforms. In this course, students are organized in teams to create documentation and digitized documentary of their own learning process using these popular digital knowledge management systems, and they will soon realize that these networked infrastructures are their venues to engage with the global marketplace.

Therefore, becoming to this class, students are encouraged to play with the basic tools and know about what will be taught and learned in this course. We will prepare material relevant to this class, and students should start experiencing these tools and concepts before starting this class.

|  |  |  |
| --- | --- | --- |
| Week | Content | Detail |
| -2 | Sign up to course explanation website | Read the syllabus carefully, and sign up to GitHub, GitLab, Wikipedia.com, and local MediaWiki website, submit an application to take this course. |
| -1 | Git and Wiki Worshop | Demonstrate the best practices of Git and Wiki. Give a short test to see if students know how to use Git |
| 0 | Models of Computation | An introductory lecture (maybe online) that explains the course, and outlines various models of computation and their applications. This includes Turing Computation, Symbolic Computation, Machine Learning Computation, and Networked Computation. Students needs to know about these terms, so that they can see why this is relevant to their daily lives in this digitized society. |

Technical and Mathematical Content

Makers are the modern alchemists. Modern alchemists; especially digital alchemists have their secret language. The language is a way to declare the testable perimeter of a designed artifact, also known as Design by Contract. To solidify the mindset of system design, Design by Contract (DbC) and automated testing tools are introduced in the very beginning of this course and will become the essential principle to control the quality of learning outcome assessment.

|  |  |  |
| --- | --- | --- |
| Week | Content | Detail |
| H1 | Design by Contract | Truth Table, Boolean Algebra, Digital Circuit Notation, Hardware Description Language |
| H2 | Numerical and Logic Operations | Number encoding system, Half/Full Adder Circuit, Combinatorial Logic, The Workhorse of CPU: Arithmetic Logic Unit |
| H3 | Memory, the basic unit of Space/ Time | Sequential Logic, Memory Circuitry, Program Counter, Flip-Flop and Clock Diagrams, Basic Instruments in Digital Electronics |
| H4 | Assembler : Symbolic Machine Language: | How to instruct a computer at the lowest level? What are the most primitive computer programs? How machine code is stored and run? |
| H5 | Computer System | Put together a General Purpose Turing Machine: Central Processing Unit, Human-Machine Interface, |
| H6 | Secure System | CPU with Write-Only Memory |
| H7 | Networked Systems | Networked Devices and Network Data Interface |

The course is intentionally divided into two parts. Hardware and software parts are taught in an interleaving manner. Conceptually, students can also learn both tracks separately, however, it would be ideal to learn both tracks in parallel.

|  |  |  |
| --- | --- | --- |
| Week | Content | Detail |
| S1 | Design by Contract | Computer as a data flow system. Use other metaphors, such as material flow to represent computation. Contracts before computation. Software installation, includes development tools, multi-media game API library, and manual installation |
| S2 | Program Control | How to write test cases, Unit Tests, System Tests, Stress Tests, use Video games as examples. |
| S3 | Virtual Machine I: Stack Arithmetic | Abstract Data Structures, Stack Arithmetic, What is Computation? Abstract Machine as a System Design Standardization technique |
| S4 | Virtual Machine II : Function Module | Functions as Subroutines, parameter passing, recursion, memory/resource allocation |
| S5.0 | High Level Programming Language | Language as a machine, Language Specification, Code examples, Software simulates physical applications |
| S5.1 | Language-specific Software Libraries | This is dedicated to the programming language of choice. We will primarily show how Python and its software libraries can be used in this course. Other tools and languages can be used here, too. |
| S6 | Compiler | Syntax Analysis: Parser Design, Grammar Specifications, Parse Tree, |
| S7 | Symbolic Computation | Symbolic Computation, Code Generation, Language-Oriented Engineering |
| S8 | Operating Systems | Data Representation, Resource Allocation, Input/Output Management, Device Drivers |

To further illustrate the power of digital contract, the universal expressive power of Turing Machine is presented in multiple layers of software/hardware systems, including the design and implementation of Assemblers, Compilers, Virtual Machines, Operating Systems, Arithmetic Logic Unit (ALU), and Central Processing Units (CPU).

All of these above systems have test cases, and relevant testing tools, to perform unit tests, system tests, and stress test. How and when these test cases are satisfied, can be an objective indicator for how well learning proceeds. These system design assignments are shown as recurring themes to demonstrate that complex engineered products can be decomposed into modularized, yet computationally testable design contracts.

After taking both hardware and software courses, students should have identified their project team members, and create a collaborative project as teams. The remaining weeks in a 16-week semester is divided as follow:

|  |  |  |
| --- | --- | --- |
| Week | Content | Detail |
| SD9 | System Design by Contract: build a networked system with modern Human-Computer Interfaces | Student team should propose a project that utilizes the above-mentioned computational tools and sensor technologies to build a human-computer interfacing system. This week is a project proposal week, every team should present a 5~10 minute project, and present their “System Design Contracts”. |
| SD10 | Project Control | Review and present their system design test cases, Unit Tests, System Tests, Stress Tests. In certain cases, the team could reverse engineer an existing Video game, and show its test cases. They should also show their program report and each teams’ code commit history. |
| SD11 | Functional Decomposition | Present the Software “Object Model” or “Function Model” of the desired system. Show a data/information flow |
| SD12 | Implementation Data Types and Function Calls | For the previously known system, identify Application Programing Interfaces (API), and tools that performs the specified tasks in the Object/Function Model |
| SD13 | Project Description Manual (Product Demo Material) | Present the Data Dictionary, Product Features, Resource Allocation, Main technical merits, and other product description content as a draft copy as website, digital content asset package. |
| SD14 | Knowledge Refinement | Each team should present their key findings in their respective projects, what can be done differently, what are their best practices. |
| SD15 | Knowledge Compilation: based on Git and Wiki data. | All teams should present their project finding as computational system, using a common set of vocabulary, and review other teams’ knowledge base on wiki and Git. Then, they should identify Design Patterns between different teams |
| SD16 | Final Presentation | Present and demonstrate the product with its digital documentation. |

Historical and Societal Contexts

To prolong students’ interest in absorbing these abstract and technically challenging concept, the entire course is gamified as a system of puzzles that leads participants into a stage-gated digital civilization. The history of computing science, the main personalities, key technologies, and influential institutions are presented to students during relevant course work. Students are required to know about these key personalities and institutions, so that they can see how breakthrough technologies are developed in certain social contexts.

In terms of personalities, we will introduce Alan Turing, Noam Chomsky, Claude Shannon, Richard Stallman, Linus Torvalds, the infamous Satoshi Nakamoto, and Jimmy Wales. These persons are directly responsible for the modern maker culture and how information/knowledge distribution can be democratized. As the course proceed, other personalities will also be included by instructors and students.

When we present technical terms, certain technologies, such as mechanical clock, mechanical calculators, integrated circuits, the Internet, HTTP/HTML, Bitcoin/Blockchain, Virtual Reality and Augmented Reality will be mentioned, and used in various homework assignments when appropriate. The goal is to make students aware that technologies are not just tools, they became popular because they all answer certain social needs. They are both the evidence and the enabler of social evolution. Finally, we will also introduce certain institutions, namely DARPA, Bell Lab, and the Free Software Foundation. These institutions have their social-economic agenda, and they all influenced the world in ways that went beyond the original expectation. Students are expected to incrementally and digitally present their essays on their interpretations of these historical/technical references. They are also welcome to identify other personalities, institutions, and technologies in their essays.

Time Allocation

We expect that learning this course would require 10 to 15 hours per week focusing on the given homework assignments. Weekly content presentation would range from 3 to 5 lecture hours. Some of the lectures can be delivered using MOOC platforms. To reduce the amount of trial and error time expended by novice makers, weekly even daily meet-up with teaching fellows and other students could be very helpful. Usually one hour tutorial or in-person technical per week can resolve a lot of problems. On the other hand, this course can be taught as a college level course with 3 to 5 credits in about 16 weeks.

Learner Feedback

At the point of this writing, the course development team has successfully delivered the hardware aspect of this course to a wide range of students, including 8-year-old boys and 40+ year-old housewives, mostly with little or no background in computation. The goal is to expand this computational approach to refine and deliver best learning practices for other foundational literacy skills, such as math, science and language learning. In the long run, creative making projects based on computational thinking and system design could be an integrative protocol to reveal the quality and quantity of collective intelligence in this increasingly digitized world.

Performance Assessment

The goal of this course is not to grade students, but to help them learn the secrets of great makers. Therefore, grading is not a focus. However, we will still provide feedback to students by observing whether they have frequently submitted their contributions to Wiki and Git. Moreover, as long as they committed their solutions to the testable assignments in specified time. They can receive full credit for each of the assignments. And the overall accumulated percentage of accomplished test cases is the final score. For team performance assessment, we plan to give certain prizes to various teams’ creative products. The quality of team work and collaborative effort can also be graded using Wiki and Git track records. However, at the time of this writing, a specific set of rules have not been implemented. The ideal way of doing this is to allow each class of students to write their own Design Contracts as their version of “constitution”.