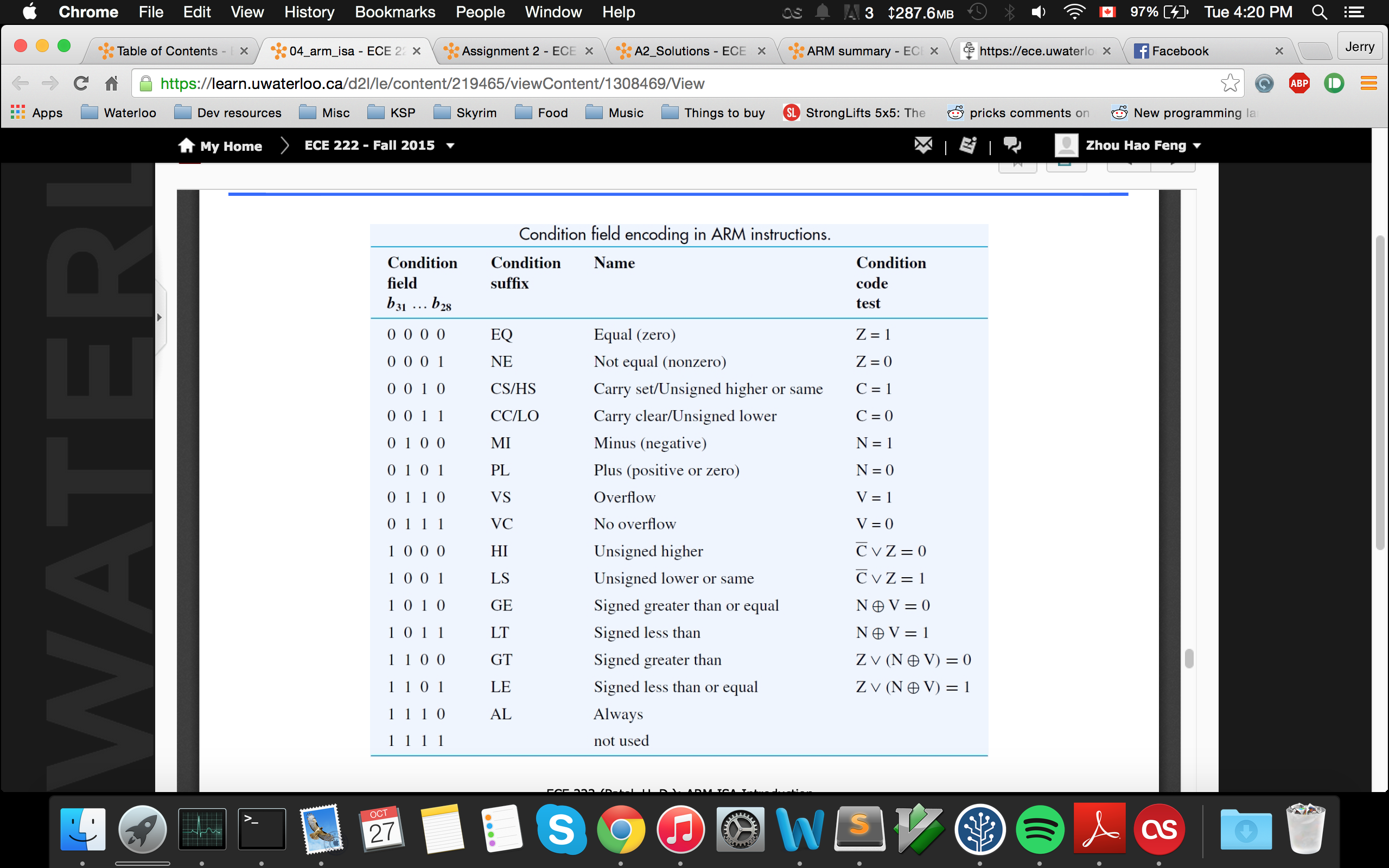
**ARM ISA**

* **Characteristics**
* RISC aspects
  + Instructions have fixed length (32 bits)
  + Only load/store instructions can access memory
  + Arithmetic/logic instructions only operate on registers
* CISC aspects
  + Auto-increment/-decrement and PC-relative addressing modes
  + Condition codes (N, Z, V, C) are used for conditional execution of instructions
  + Multiple registers can be loaded/stored from/to a block of memory
* Memory organization
  + Byte-addressable, 32-bit address space
  + Supports little/big-endian
  + Supports data sizes of words (32), half words (16), bytes (8)
* Registers
  + 16 32-bit registers R0 – R15
  + R15 = PC; R12 = frame pointer (by convention); R13 = stack pointer; R14 = link register
  + CPSR/PSR = status register – stores conditional code flags
    - N = 1 if result is negative; cleared otherwise
    - Z = 1 if result is 0; cleared otherwise
    - V = 1 if overflow occurs; cleared otherwise
    - C = 1 if carry-out occurs; cleared otherwise
* **Directives**
* Label/EQU
  + LABEL: MOV R1, #1 (instruction at address 100)
  + Equivalent to LABEL EQU 100
  + EQU directive tells assembler to replace LABEL with 100
* ORIGIN
  + ORIGIN 200
  + Tells assembler to store subsequent instructions starting at address 200
* RESERVE
  + RESERVE 100
  + Tells assembler to reserve 100 bytes for data – no data will be loaded into these memory locations
* DATAWORD
  + DATAWORD 25
  + Number of entries in the list whose memory is being reserved
* Memory operations
  + LDR/STR – load/store words
  + LDRH/STRH – half words
  + LDRB/STRB – bytes
  + LDRSH/LDRSB – sign-extended loads for half words/bytes
  + LDM/STM{FD/EA}{cond} Rn{!}, reglist – multi-word load/store
    - {} means optional
    - Address mode
      * FD – increment address before access
      * EA – decrement address after access
    - Cond – conditional execution
    - ! – load final address into Rn (i.e. writeback)
    - reglist – one or more registers
* PC-relative addressing
  + LDR Rd, LABEL
    - i.e. Rd ← [[PC] + offset(LABEL)]
  + PC += 8 when LDR executes
* Pre-index addressing
  + LDR Rd, [Rn, #8]
    - i.e. Rd ← [[Rn] + 8] **(indexed w/ offset)**
  + LDR Rd, [Rn, #8]! – with writeback
    - Additionally increments Rn: Rn ← [Rn] + 8
  + LDR Rd, [Rn, ±Rm, LSL #4]
    - Shifts [Rm] left by 4, then Rd ← [[Rn] ± [Rm]]
* Post-index addressing
  + LDR Rd, [Rn], #8
    - i.e. Rd ← [[Rn]] **(indexed w/o offset)**
    - Also increments Rn: Rn ← [Rn] + 8
  + LDR Rd, [Rn], ± Rm, LSR #4
    - Rd ← [[Rn]] then shifts [Rm] right by 4, and increments Rn: Rn ← [[Rn] ± [Rm]]
* Logical instructions
  + AND, ORR, EOR – bitwise and/or/xor
  + LSL/LSR/ASL/ASR Ri, Rj, #2
    - Ri ← [Rj] shifted by 2 bits
    - Arithmetic shift preserves the sign in the MSB
    - Arithmetic shift/logical shift left – add 0s at LSB, push everything left (multiply by 2)
    - Logical shift right – add 0s at MSB, push everything right (divide by 2)
    - Arithmetic shift right – add 0s or 1s at MSB, depending on the sign
  + ROR Ri, Rj, #2
    - Ri ← [Rj] rotated right by 2 bits
    - Rotate – copy bits from one end to another
    - C (carry) flag still retains the last bit being shifted out
    - E.g. C = 0, R1 = 0110…1011
      * ROR R1, R1, #2: C = 1, R1 = 110110…10
  + RRX Rj, Rj
    - Ri ← [Rj] rotated right using C flag by 1 bit
    - E.g. C = 0, R1 = 0110…1011
      * RLC R1, R1, #2: C = 1, R1 = 00110…101
* Arithmetic instructions
  + ADD R0, R2, R4
    - R0 ← [R2] + [R4]
  + SUB R0, R1, R2, LSL #4 (left shift operand)
    - R0 ← [R1] + [R2] \* 2^4
  + MUL R0, R1, R2
    - R0 ← [R1] \* [R2]
  + MLA R0, R4, R5, R6
    - R0 ← [R4] \* [R5] + [R6]
* Move instructions
  + MOV Rd, #value/Rm
    - Rd ← value/[Rm]
  + MOV Rd, #value/Rm, LSL #shift
    - Equivalent to LSL Rd, Rm, #shift
  + MVN Rd, #value/Rm
    - Moves the bit complement of #value/[Rm]
* Test/compare instructions
  + TST Rn, #value/Rm
    - Performs bitwise AND then sets condition code
  + TEQ Rn, #value/Rm
    - Performs bitwise XOR then sets condition code
  + CMP Rn, Rm
    - [Rn] – R[m] then sets condition code
  + These instructions always update condition code
* Append S to opcode to update condition code
  + E.g. ADDS, ORRS
* Branch instructions
  + B{condition} LOC
    - Branches to LOC if condition code flags satisfy {condition}
  + BL LABEL – store PC in R14 (link register), branch to LABEL
  + BX LR – set PC to address stored in link register, i.e. reset return address
* Instructions encoding
  + RISC – instructions are limited to single words
  + 32-bit immediate must be assigned by:
    - Load 16-bit, shift left 16 bits, load 16-bit
    - Or save immediate in memory, then load from memory into register
  + Register-operand format
    - MSB – Rsrc1 – Rsrc2 – Rdst – opcode – LSB
  + Immediate-operand format
    - MSB – Rsrc – Rdst – immediate – opcode – LSB
  + Call format
    - MSB – immediate – opcode – LSB
* Condition codes
* **The assembler**
* Assembler directives – instruct the assembler to perform certain tasks during translation from source program → object program
  + AREA – specifies start of CODE/DATA
  + ENTRY – specifies start of program execution
  + DCD – label and initialize data operands
  + EQU – declare symbolic names for constants
  + RN – ??
* Pseudo-instructions
  + LDR Ri, =X
  + Where X is an address value
* Two-pass assembly
  + Generate symbol table – label → address (position relative of start of program)
  + Go through table, substitute symbol names with addresses
* For multiple files/modules
  + Assemble each module as individual program
  + Linker adds appropriate offsets to reflect where each module is in memory
  + Position-independent code – relative addressing, instructions etc.
  + Position-dependent code – references to external labels