Simple Questions

- How many cycles will it take to execute this code?

```
          lw $t2, 0($t3)
          lw $t3, 4($t3)
          beq $t2, $t3, Label      #assume not
          add $t5, $t2, $t3
          sw $t5, 8($t3)
Label:    ...
```

- What is going on during the 8th cycle of execution?
- In what cycle does the actual addition of `$t2` and `$t3` takes place?

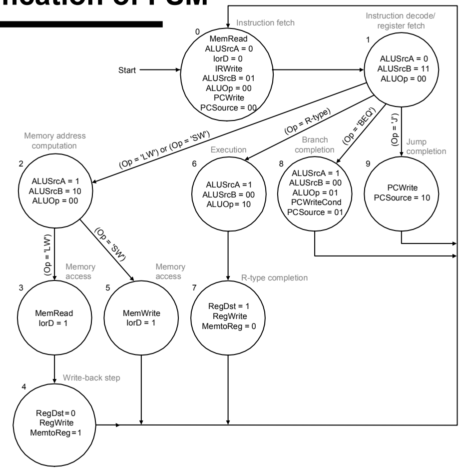
Implementing the Control

- Value of control signals is dependent upon:
  - what instruction is being executed
  - which step is being performed
- Use the information we've accumulated to specify a finite state machine
  - specify the finite state machine graphically, or
  - use micro-programming
- Implementation can be derived from specification

Deciding the Control

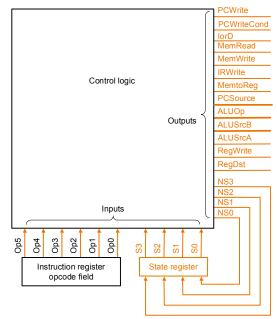
- In each clock cycle, decide all the action that needs to be taken
- The control signal can be 0 and 1 or x (don't care)
- Make a signal an x if you can to reduce control
- But any action that may destroy any useful value should not be allowed
- Control Signal required
  - ALU: SRC1 (1 bit), SRC2(2 bits), operation (Add, Sub, or from FC)
  - Memory: address (I or D), read, write, data clocked in IR or MDR
  - Register File: address (rt or rd), data (MDR or ALUOUT), read, write
  - PC: PCwrite, PCwrite-conditional, PC data (PC+4, branch, jump)
- Some of the control signal can be implied (register file read are values in A and B registers (actually A and B need not be registers at all)
- Explicit control vs indirect control (derived based on input like what instruction is being executed, or what function code field is) bits

Graphical Specification of FSM



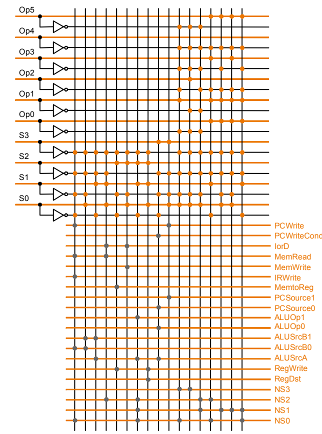
Finite State Machine for Control

- **Implementation:**



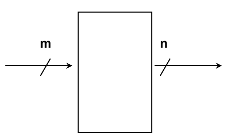
PLA Implementation

- **If I picked a horizontal or vertical line could you explain it?**



ROM Implementation

- **ROM = "Read Only Memory"**
  - **values of memory locations are fixed ahead of time**
- **A ROM can be used to implement a truth table**
  - **if the address is m-bits, we can address 2^m entries in the ROM.**
  - **our outputs are the bits of data that the address points to.**



0	0	0	0	0	1	1
0	0	1	1	1	0	0
0	1	0	1	1	0	0
0	1	1	1	0	0	0
1	0	0	0	0	0	0
1	0	1	0	0	0	1
1	1	0	0	1	1	0
1	1	1	0	1	1	1

m is the "height", and n is the "width"

ROM Implementation

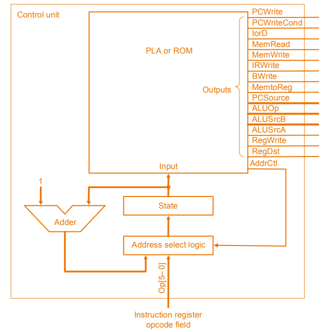
- **How many inputs are there?**
  6 bits for opcode, 4 bits for state = 10 address lines
  (i.e., 2^{10} = 1024 different addresses)
- **How many outputs are there?**
  16 datapath-control outputs, 4 state bits = 20 outputs
- **ROM is 2^{10} x 20 = 20K bits (and a rather unusual size)**
- **Rather wasteful, since for lots of the entries, the outputs are the same**
  — i.e., opcode is often ignored

ROM vs PLA

- Break up the table into two parts
  - — 4 state bits tell you the 16 outputs, 2^4 x 16 bits of ROM
  - — 10 bits tell you the 4 next state bits, 2^{10} x 4 bits of ROM
  - — Total: 4.3K bits of ROM
- PLA is much smaller
  - — can share product terms
  - — only need entries that produce an active output
  - — can take into account don't cares
- Size is (#inputs × #product-terms) + (#outputs × #product-terms)
  For this example = (10x17)+(20x17) = 460 PLA cells
- PLA cells usually about the size of a ROM cell (slightly bigger)

Another Implementation Style

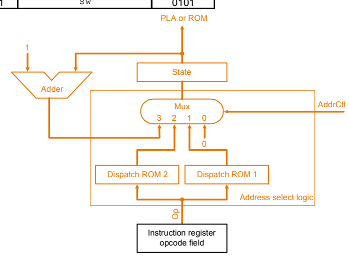
- Complex instructions: the "next state" is often current state + 1



Details

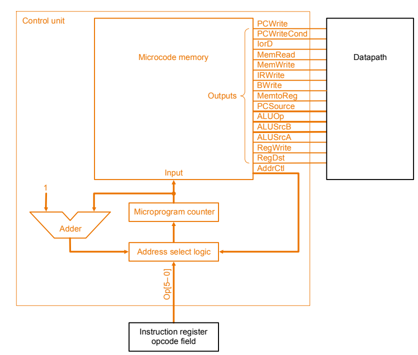
Dispatch ROM 1		
Op	Opcode name	Value
000000	R-format	0110
000010	jmp	1001
000100	beq	1000
100011	lw	0010
101011	sw	0010

Dispatch ROM 2		
Op	Opcode name	Value
100011	lw	0011
101011	sw	0101



State number	Address-control action	Value of AddrCtl
0	Use incremented state	3
1	Use dispatch ROM 1	1
2	Use dispatch ROM 2	2
3	Use incremented state	3
4	Replace state number by 0	0
5	Replace state number by 0	0
6	Use incremented state	3
7	Replace state number by 0	0
8	Replace state number by 0	0
9	Replace state number by 0	0

Microprogramming



Microprogramming

- A specification methodology
  - appropriate if hundreds of opcodes, modes, cycles, etc.
  - signals specified symbolically using microinstructions

Label	ALU control	SRC1	SRC2	Register control	Memory	PCWrite control	Sequencing
Fetch	Add	PC	4		Read PC	ALU	Seq
	Add	PC	Extshft	Read			Dispatch 1
Mem1	Add	A	Extend				Dispatch 2
LW2					Read ALU		Seq
				Write MDR			Fetch
SW2					Write ALU		Fetch
Rformat1	Func code	A	B				Seq
				Write ALU			Fetch
BEQ1	Subt	A	B			ALUOut-cond	Fetch
JUMP1						Jump address	Fetch

- *Will two implementations of the same architecture have the same microcode?*
- *What would a microassembler do?*

Microinstruction format

Field name	Value	Signals active	Comment
ALU control	Add	ALUOp = 00	Cause the ALU to add.
	Subt	ALUOp = 01	Cause the ALU to subtract; this implements the compare for branches.
	Func code	ALUOp = 10	Use the instruction's function code to determine ALU control.
SRC1	PC	ALUSrcA = 0	Use the PC as the first ALU input.
	A	ALUSrcA = 1	Register A is the first ALU input.
SRC2	B	ALUSrcB = 00	Register B is the second ALU input.
	4	ALUSrcB = 01	Use 4 as the second ALU input.
	Extend	ALUSrcB = 10	Use output of the sign extension unit as the second ALU input.
	Extshft	ALUSrcB = 11	Use the output of the shift-by-two unit as the second ALU input.
Register control	Read		Read two registers using the rs and rt fields of the IR as the register numbers and putting the data into registers A and B.
	Write ALU	RegWrite, RegDst = 1, MemtoReg = 0	Write a register using the rd field of the IR as the register number and the contents of the ALUOut as the data.
	Write MDR	RegWrite, RegDst = 0, MemtoReg = 1	Write a register using the rt field of the IR as the register number and the contents of the MDR as the data.
Memory	Read PC	MemRead, lorD = 0	Read memory using the PC as address; write result into IR (and the MDR).
	Read ALU	MemRead, lorD = 1	Read memory using the ALUOut as address; write result into MDR.
	Write ALU	MemWrite, lorD = 1	Write memory using the ALUOut as address, contents of B as the data.
PC write control	ALU	PCSource = 00 PCWrite	Write the output of the ALU into the PC.
	ALUOut-cond	PCSource = 01, PCWriteCond	If the Zero output of the ALU is active, write the PC with the contents of the register ALUOut.
	jump address	PCSource = 10, PCWrite	Write the PC with the jump address from the instruction.
Sequencing	Seq	AddrCtl = 11	Choose the next microinstruction sequentially.
	Fetch	AddrCtl = 00	Go to the first microinstruction to begin a new instruction.
	Dispatch 1	AddrCtl = 01	Dispatch using the ROM 1.
	Dispatch 2	AddrCtl = 10	Dispatch using the ROM 2.

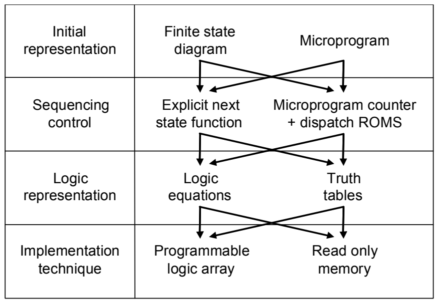
Maximally vs. Minimally Encoded

- No encoding:
  - 1 bit for each datapath operation
  - faster, requires more memory (logic)
  - used for Vax 780 — an astonishing 400K of memory!
- Lots of encoding:
  - send the microinstructions through logic to get control signals
  - uses less memory, slower
- Historical context of CISC:
  - Too much logic to put on a single chip with everything else
  - Use a ROM (or even RAM) to hold the microcode
  - It’s easy to add new instructions

Microcode: Trade-offs

- **Distinction between specification and implementation is sometimes blurred**
- Specification Advantages:
  - Easy to design and write
  - Design architecture and microcode in parallel
- Implementation (off-chip ROM) Advantages
  - Easy to change since values are in memory
  - Can emulate other architectures
  - Can make use of internal registers
- Implementation Disadvantages, SLOWER now that:
  - Control is implemented on same chip as processor
  - ROM is no longer faster than RAM
  - No need to go back and make changes

The Big Picture



Other Issues: Exception

- What should the machine do if there is a problem
- Two kinds of problems:
  - External condition: I/O interrupt, power failure, user wanting to stop the program, i.e. CTRL C
  - Internal condition: incorrect memory address for instruction read (branch or jump led to a non-existent memory location, data read or write in data memory, illegal operation code, arithmetic overflow and/or underflow
- Interrupts (external) and exception (internal) are handled similarly
- Control is transferred to an exception handling mechanism, stored at a pre-specified location
- Address of instruction is saved in a register called EPC

Vectored Interrupts/Exceptions

- Address of exception handler depends on the problem
  - Undefined Instruction C0 00 00 00
  - Arithmetic Overflow C0 00 00 20
  - Addresses are separated by a fixed amount, 32 bytes in MIPS
- PC is transferred to a register called EPC
- If interrupts are not vectored, then we need another register to store the cause of problem
- In what state what exception can occur?

Final Words on Single and Multi-Cycle Systems

- Single cycle implementation
  - Simpler but slowest
  - Require more hardware
- Multi-cycle
  - Faster clock
  - Amount of time it takes depends on instruction mix
  - Control more complicated
- Exceptions and Other conditions add a lot of compexity
- Other techniques to make it faster

Conclusions on Chapter 5

- Control is the most complex part
- Can be hard-wired, ROM-based, or micro-programmed
- Simpler instructions also lead to simple control
- Just because machine is micro-programmed, we should not add complicated instructions
- Sometimes simple instructions are more effective than a single complex instruction
- More complex instructions may have to be maintained for compatibility reasons