Instruction Set Architecture

or

“How to talk to computers”

Crafting an ISA

• Designing an ISA is both an art and a science
• ISA design involves dealing in an extremely rare resource
  – instruction bits!
• Some things we want out of our ISA
  – completeness
  – orthogonality
  – regularity and simplicity
  – compactness
  – ease of programming
  – ease of implementation

Where are the instructions?

• Harvard architecture
  - “stored-program” computer
• Von Neumann architecture
Key ISA decisions

- **operations**
  - how many?
  - which ones
- **operands**
  - how many?
  - location
  - types
  - how to specify?
- **instruction format**
  - size
  - how many formats?

\[ y = x + b \]

operation

source operands

destination operand

How does the computer know what 0001 0100 1101 1111 means?

What enables performance in today's machines?

- This wasn’t true in the era in which most classical ISAs were defined...

- Parallelism!!
  - Superscalar
  - Pipelining
  - Multicore

Choice 1: Operand Location

- Accumulator
- Stack
- Registers
- Memory

We can classify most machines into 4 types: *accumulator, stack, register-memory* (most operands can be registers or memory), *load/store* (arithmetic operations must have register operands).

Choice 1B: How Many Operands?

**Basic ISA Classes**

<table>
<thead>
<tr>
<th>Accumulator:</th>
<th>Stack:</th>
<th>General Purpose Register:</th>
<th>Load/Store:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 address</td>
<td>0 address</td>
<td>2 address</td>
<td>3 address</td>
</tr>
<tr>
<td>add A</td>
<td>add</td>
<td>add A B</td>
<td>add Ra Rb</td>
</tr>
<tr>
<td>acc ← acc + mem[A]</td>
<td>tos ← tos + next</td>
<td>EA(A) ← EA(A) + EA(B)</td>
<td>Ra ← Rb + Rc</td>
</tr>
<tr>
<td>3 address</td>
<td>3 address</td>
<td>3 address</td>
<td>store Ra Rb</td>
</tr>
<tr>
<td>add A B C</td>
<td>add A B C</td>
<td>EA(A) ← EA(B) + EA(C)</td>
<td>mem[Rb] ← Ra</td>
</tr>
</tbody>
</table>

A load/store architecture has instructions that do either ALU operations or access memory, but never both.
Alternative ISA's

- \[ A = X \times Y - B \times C \]

<table>
<thead>
<tr>
<th>Stack Architecture</th>
<th>Accumulator</th>
<th>GPR</th>
<th>GPR (Load-store)</th>
</tr>
</thead>
</table>

Trade-offs

- Stack +
- Accumulator +
- GPR +
- Load-store +

Choice 2: Addressing Modes

- Register direct: \( R3 \) \( R6 = R5 + R3 \)
- Immediate (literal): \#25 \( R6 = R5 + 25 \)
- Direct (absolute): \[ M[10000] \] \( R6 = M[10000] \)
- Register indirect: \[ M[R3] \] (a.k.a register deferred) \( R6 = M[R3] \)
- Memory Indirect: \[ M[M[R3]] \]
- Displacement: \[ M[R3 + 10000] \] ...
- Index: \[ M[R3 + R4] \]
- Scaled: \[ M[R3 + R4 \times d + 10000] \]
- Autoincrement: \[ M[R3++] \]
- Autodecrement: \[ M[R3 - -] \]

Addressing Mode Utilization

Conclusion?
Displacement Size

• Conclusions – 16 bits is usually enough. If not, just use another instruction.

Choice 3: Which Operations?

• arithmetic
  – add, subtract, multiply, divide
• logical
  – and, or, shift left, shift right
• data transfer
  – load word, store word
• control flow

Does it make sense to have more complex instructions?
  – e.g., square root, mult-add, matrix multiply, cross product ...

Types of branches (control flow)

• conditional branch
  beq r1,r2, label
• jump
  jump label
• procedure call
  call label
• procedure return
  return

Conditional branch

• How do you specify the destination (target) of a branch/jump?
• How do we specify the condition of the branch?
Branch distance

- Average distance (in bits needed to specify) from branch to target.
- Conclusions!

Branch condition

**Condition Codes**

Processor status bits are set as a side-effect of arithmetic instructions or explicitly by compare or test instructions.

- sub r1, r2, r3
  - bz label

**Condition Register**

- cmp r1, r2, r3
  - bgt r1, label

**Compare and Branch**

- bgt r1, r2, label

Choice 4: Instruction Format

- Fixed (e.g., all RISC processors – SPARC, MIPS, Alpha)
  - opcode addr1 addr2 addr3

- Variable (VAX, ...)
  - opcode spec addr1 spec2 addr2 ... specn addrn

- Hybrid

- Tradeoffs?
- Conclusions?

The Customer is Always Right

- Compiler is primary customer of ISA
- Features the compiler doesn’t use are wasted
- Register allocation is a huge contributor to performance
- Compiler-writer’s job is made easier when ISA has
  - regularity
  - primitives, not solutions
  - simple trade-offs
- Summary -> simplicity over power
**Our desired ISA**

- Registers, Load-store
- Addressing modes
  - immediate (8-16 bits)
  - displacement (12-16 bits)
  - register deferred (register indirect)
- Support a reasonable number of operations
- Don’t use condition codes
- Fixed instruction encoding/length for performance
- regularity (several general-purpose registers)

**MIPS instruction set architecture**

- 32 32-bit general-purpose registers
  - R0 always equals zero
  - 32 or 16 FP registers
- 8-, 16-, and 32-bit integers, 32- and 64-bit fp data types
- immediate and displacement addressing modes
  - register deferred is a subset of displacement
- 32-bit fixed-length instruction encoding

---

**MIPS Instruction Format**

- R-type instruction
  - Opcode, rs, rt, Immediate
  - Encodes: Loads and stores of bytes, half words, words, double words. All immediates (rt = rs op immediate)
  - Conditional branch instructions (rs is register, rt unused)
  - Jump register, jump and link register
  - (rs = 0, rt = destination, immediate = 0)

- J-type instruction
  - Opcode
  - 26
  - Jump and jump and link
  - Trap and return from exception

---

**RISC vs CISC**

- MIPS is a classic RISC architectures (as are SPARC, Alpha, PowerPC, …)
- RISC stands for Reduced Instruction Set Computer. RISC architectures are load-store, few formats, minimal instruction sets.
- They were in contrast to the 70s and 80s which proliferated CISC ISAs (VAX, Intel x86, various IBM), which were characterized by complex and comprehensive instruction sets, and complex instruction decoding.
- RISC architectures thrived not because they supported fewer operations, but because they enabled parallelism.
**MIPS Operations and ISA**

- Read on your own!
- Get comfortable with MIPS instructions and formats

**A few sample instructions**

- `lw R1, 1000(R2)`
- `add R1, R2, R3`
- `addi R1, R2, #53`
- `jal label`
- `jr R3`
- `beq R1, R5, label`

**MIPS R2000 vs. VAX 8700**

Or “Why RISC?”

- ET = IC \* CPI \* CT

- \[ IC_{MIPS} = 2 \, IC_{VAX} \]
- \[ CPI_{VAX} = 6 \, CPI_{MIPS} \]

**ISA Key Points**

- Modern ISA’s typically sacrifice power and flexibility for regularity and simplicity; code density for parallelism and throughput.
- Instruction bits are extremely limited, particularly in a fixed-length instruction format.
- Registers are critical to performance – we want lots of them, and few strings attached.
- Displacement addressing mode handles the vast majority of memory reference needs.