Keith On...
Hardware

K.E. Schubert

Founder
Renaissance Research Labs

Professor
Department of Computer Science and Engineering
California State University, San Bernardino

Contents

Part I

Electronics

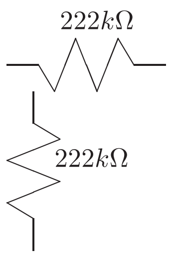
Chapter 1

Passive Components

1.1 Resistor

$$\begin{aligned} V &= IR && (1.1)\\ P &= VI && (1.2)\\ &= I^2R && (1.3)\\ &= \frac{V^2}{R} && (1.4) \end{aligned}$$

DC: $R = \frac{l \cdot \rho}{A}$, where l is the length in meters, A is the cross sectional area in square meters and ρ is the electric resistivity or specific electrical resistance in ohm-meters. This assumes the current density is uniform.

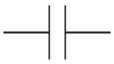


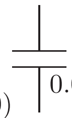
1.2 Capacitor

$$cV = q \tag{1.5}$$

When I took my physics E&M class my professor had an interesting way to remember equation 1.5. One of his friends in college used a beer slogan, “Canadian Velvet is the Queen of beers”, as a mnemonic.

$0.01\mu f$





put(70,0) $0.01\mu f$

1.3 Inductor

Symbol: L

Unit: Henry (volt sec/Amp)

$$\phi = LI \tag{1.6}$$

Note: ϕ is magnetic flux

1.4 Memristence

Memristors were first predicted in the 1970’s due to symmetry, but were first made in 2007. They have great potential to revolutionize memory.

$$\phi = M(q)q \tag{1.7}$$

Thus, at some instant in time

$$V(t) = M(q(t))I(t) \tag{1.8}$$

Note that M is not a constant, and in fact it is non-linear with hysteresis.

Chapter 2

Basic Laws

2.1 Coulomb's Law

$$F = \frac{kq_1q_2}{r^2} \tag{2.1}$$
$$= \frac{q_1q_2}{4\pi\varepsilon_0 r^2} \tag{2.2}$$

Note: Negative forces attract and positive repel.

Coulomb's constant, k is given by

$$k = \frac{1}{4\pi\varepsilon_0} \tag{2.3}$$
$$\approx 9 \times 10^9 \; [N \cdot m^2/C^2] \tag{2.4}$$

2.2 Maxwell's Laws

Maxwell's four Laws have individual names:

Gauss' Law of Electricity

Gauss' Law of Magnetism

Faraday's Law of Induction Basis of electrical generators and inductors

Ampere's Law

The symbols used are:

E Electric field

B Magnetic field

D Electric displacement

H Magnetic field strength

ρ charge density

ε permittivity

ε_0 permittivity of free space

μ permeability

μ_0 permeability of free space

M Magnetization

i electric current

J current density

c speed of light, $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$.

P Polarization

k Coulomb's constant, $k = \frac{1}{4\pi\varepsilon_0}$.

2.2.1 Gauss' Law for Electricity

Integral Form

$$\oint \vec{E} \cdot \vec{A} = \frac{q}{\varepsilon_0} \tag{2.5}$$

Differential Form

$$\nabla \cdot D = \rho \tag{2.6}$$

where D is

General Case $D = \varepsilon_0 E + P$

Free Space $D = \varepsilon_0 E$

Isotropic Linear Dielectric $D = \varepsilon E$

2.2.2 Gauss' Law for Magnetism

Integral Form

$$\oint \vec{B} \cdot \vec{A} = 0 \tag{2.7}$$

Differential Form

$$\nabla \cdot B = 0 \tag{2.8}$$

2.2.3 Faraday's Law of Induction

Simplified Form

$$V = -N\frac{\Delta(BA)}{\Delta t} \tag{2.9}$$

N number of turns in the coil

B Magnetic Field

A The cross sectional area (perpendicular to the magnetic field)

t Time

V Voltage or EMF (electro-motive force)

Integral Form

$$\oint \vec{E}d\vec{s} = -\frac{d\Phi_B}{dt} \tag{2.10}$$

$$= EMF \tag{2.11}$$

Differential Form

$$\nabla \times E = -\frac{\partial B}{\partial t} \tag{2.12}$$

2.2.4 Ampere's Law

Integral Form

$$\oint B \cdot ds = \mu_0 i + \frac{1}{c^2}\frac{\partial}{\partial t}\int E \cdot dA \tag{2.13}$$

Differential Form

$$\nabla \times H = J + \frac{\partial D}{\partial t} \tag{2.14}$$

where D is

General Case $D = \varepsilon_0 E + P$

Free Space $D = \varepsilon_0 E$

Isotropic Linear Dielectric $D = \varepsilon E$

and H is

General Case $B = \mu_0(H + M)$

Free Space $B = \mu_0 H$

Isotropic Linear Magnetic Medium $B = \mu H$

Chapter 3

Semiconductors

This will be a brief introduction to physical electronics. To properly study the field requires both quantum and statistical mechanics. Perhaps one day I will write up a book on these topics and will then have the background material available to show more of the why. For now I will attempt to give an understanding of the key concepts and an explanation of how to solve for key values.

3.1 Energy Levels

In electronics we are concerned about three basic types of materials: insulators, semiconductors, and conductors. Their properties come from the energy gap between their valence[1] (energy level of the valence band is denoted E_v) and conduction[2] (energy level of the conduction band is denoted E_c) bands, and where the Fermi energy[3] (denoted E_f) lies with respect to them.

Conductors have a small energy gap between the valence and conduction band and the Fermi energy is at or above the level of the conduction band. Charge carriers are readily available to carry current. This is how they conduct.

Semiconductors have a small to mid sized gap and the Fermi energy lies in this gap. Charge carriers are not readily available, but can be made to be available by other factors (temperature, electric field, photons, donors/acceptors, etc.). This ability to conduct or insulate is the source of the name. We break down semiconductors into different categories: intrinsic and extrinsic. Extrinsic is then broken down into p or n type.

Insulators have a large energy gap between the valence and conduction band and the Fermi energy is between them, though not close to the conduction band. This means charge carriers are very unlikely available to move and thus carry current. This is how they insulate.

Quantum theory tells us that the electrons around an atom are in shells that have quantized energy values. Further, due to the Pauli Exclusion Principle, no two electrons can have the same quantum numbers (n, l, m, s), which also applies in systems of multiple atoms. As atoms come closer together the shells split so the quantum numbers are unique between them.

[1] topmost filled band

[2] band above the valence band

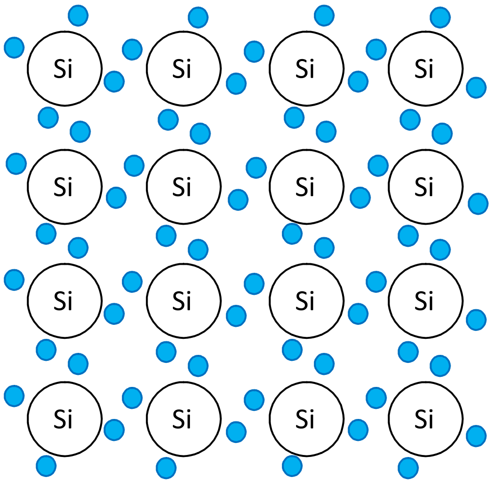
[3] A formal discussion is beyond the scope, so I will try to give a simple explanation. In thermal equilibrium the Fermi energy is the chemical potential, i.e. the amount the energy of the system changes when particles are added or subtracted from it. It is a crucial element in determining the probability a state contains an electron. The Fermi energy can also be thought of as the critical energy of the Fermi-Dirac distribution (the energy at which the probability is 0.5). Note the Fermi energy is always greater than the energy of the highest filled band.

Table 3.1: Semiconductors in the Periodic Table and Intrinsic Semiconductor Properties

IB	IIB	IIIA	IVA	VA	VIA
		5 B	6 C	7 N	8 O
		13 Al	14 Si	15 P	16 S
29 CU	30 Zn	31 Ga	32 Ge	33 As	34 Se
47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te
79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po

Material	Symbol	$E_g[eV]$	B$[cm^{-3}K^{-3/2}]$
Gallium Arsenide	GaAs	1.42	2.10×10^{14}
Germanium	Ge	0.66	1.66×10^{15}
Silicon	Si	1.12	5.23×10^{15}

Figure 3.1: Silicon crystal structure



3.2 Intrinsic Semiconductors

Intrinsic semiconductors are materials that semiconduct in and of themselves. They come in two varieties, elemental (Si, Ge) and compound (GaAs, InP, etc.), which simply tells you if the material is an element or a compound (made of several elements). Elemental semiconductors come from column IVa of the periodic table, see Table 3.1, as they have half their outer s and p sub-shells filled. Compound intrinsic semiconductors are compounds that behave like elemental intrinsic semiconductors, as they are formed by bonding elements on either side of column IVa. I will mostly discuss elemental, though the principles are the same for compound intrinsic semiconductors.

These are (and must be) pure materials, as even a small amount of impurities will cause them not to work. Most modern semiconductors are extrinsic[4], so I will just give a brief overview. At room temperature, Fermi-Dirac statistics shows the gap between valence and conduction band must be on the order of 1 electron volt or less. Several materials meet this requirement, such as gallium arsenide (GaAs, 1.42 eV), Silicon (Si, 1.12 eV), and Germanium (Ge, 0.66 eV). The electrons promoted to the conduction band leave behind holes in the valence band, and current is carried by both the flow of electrons in the conduction band and

[4]This is due to the excessive cost and difficulty of making a pure or nearly pure material.

holes in the valence band. The flow of the electron-hole pairs (in opposite directions) is the mechanism of current. Essentially the electron-hole movement is the same for extrinsic semiconductors also, but for extrinsic semiconductors these need not be equal (and are not). Since the number of electrons in the conduction band and the number of holes in the valence band are equal in an intrinsic semiconductor, we will just consider the density of the intrinsic carriers (either electron or hole), n_i.

$$n_i = BT^{\frac{3}{2}} e^{\left(\frac{-E_g}{2kT}\right)} \tag{3.1}$$

The temperature in Kelvin is T. The values of B and E_g are dependent on the material, and are provided for our top three intrinsic materials in Table 3.1.The Boltzmann constant, k, is $86 \times 10^{-6} [eV/K]$.

Example 1 *What is the carrier density of Germanium at room temperature?*

Answer:

Room temperature is not a well defined term, meaning it is not a set temperature. Roughly it could vary from $65°F$ *to* $85°F$*, or in kelvin, from 291K to 303K. For ease of calculation, we will choose 300K to be room temperature.*

$$n_i = BT^{\frac{3}{2}} e^{\left(\frac{-E_g}{2kT}\right)} \tag{3.2}$$

$$= 1.66 \times 10^{15} \cdot 300^{\frac{3}{2}} e^{\left(\frac{-0.66}{2 \cdot 86 \times 10^{-6} \cdot 300}\right)} \tag{3.3}$$

$$\approx 2.40 \times 10^{13} [cm^{-3}] \tag{3.4}$$

To make life easier in calculating this, I usually use a small SciLab program, see Code 3.1.

Listing 3.1: SciLab code to calculate intrinsic carrier density.

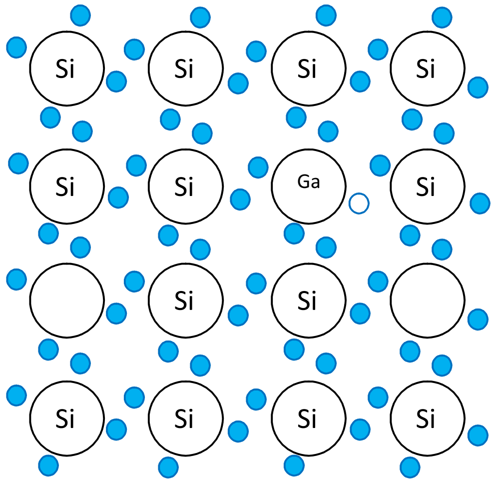
```
//Setup
GaAs=1;
Ge=2;
Si=3;
Eg = [1.42
      0.66
      1.12];//eV
B = [2.1E14
     1.66E15
     5.23E15];//cm^{-3}K^{-3/2}
k=86E-6;//eV/K

//user selections
T=300;//Kelvin
material = Ge;

ni = B(material) * T^1.5 * exp(-Eg(material) / (2*k*T)) // in cm^{-3}
```

3.3 Extrinsic Semiconductors

Extrinsic semiconductors are made by adding an impurity into the crystallin structure that is chosen to provide an extra electron above the valence band (a donor or n type), or to provide a deficiency of electrons (an acceptor or p type). Since the charge carriers (electrons and holes) are no longer balanced, we need to be able to calculate how many of them there are. A basic relationship is

$$n_e n_h = n_i^2 \tag{3.5}$$

Figure 3.2: P Type Silicon crystal structure



where n_e is the thermal equilibrium concentration of free electrons, n_h is the thermal equilibrium concentration of holes, and n_i is the intrinsic carrier density. Provided the concentration of the dopant is greater than the intrinsic carrier density, we can approximate the number of carriers of the type provided by the dopant by the concentration of the dopant. We will denote the dopant concentration by N_n for n type materials and N_p for p type materials.

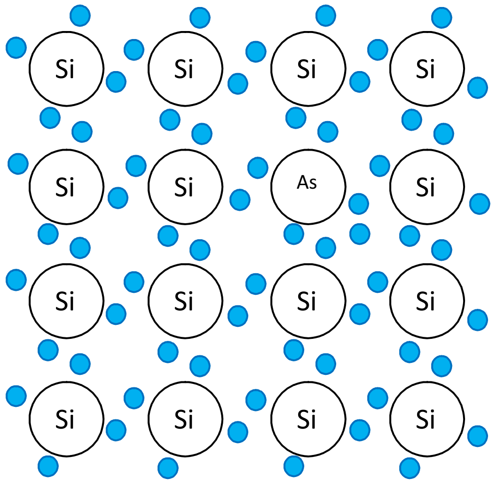
3.3.1 P Type Semiconductors

If an element like gallium or boron is doped, which has only 3 electrons in its outer shell of up to 8, into the crystallin structure of say silicon, which has 4 electrons in its outer shell of 8, a gap in the bonding is formed, see Figure 3.2. The outer shell of the gallium and the bonded silicon is just one shy of completing and thus it is already free to conduct a hole by stealing an electron from a neighbor. In terms of energy bands, the energy of this “hole” (usually denoted E_a) is very close to (but above) the energy of the valence band. The Fermi energy is thus close to the valence band, which means there will not be lots of free electrons, so the only free carriers are these holes[5].

3.3.2 N Type Semiconductors

[5]In truth the hole is the spot where the electron is missing (and thus not an actual thing), but since this means that the atom that is missing it is ionized (positive in the case of a lack of an electron), it appears that a positive charge is flowing. This is governed by statistical mechanics and so it is not possible to track the actual electron. Like it or not, holes are a reasonable description of how the charge is carried.

Figure 3.3: N Type Silicon crystal structure



Chapter 4

Diodes

Up till now we have considered individual semiconductors, now we want to consider what happens when we put two next to each other. By placing a p region next to an n region, holes start diffusing from the p region to the n region, and electrons start diffusing from the n region to the p region. This does several things.

- First, it locks up the charge carriers, so that there are not any available to carry current. For this reason it is called the depletion region.

- Second, it causes a charge differential across the boundary, which is called the potential barrier, V_{bi}. The potential barrier resists the further diffusion of charge carriers, because the depletion region in the n material is slightly positive, and the depletion region in the p region is slightly negative, resulting in an induced E-field from n to p. The potential barrier is given by

$$V_{bi} = \frac{kT}{e^-} \ln\left(\frac{N_n N_p}{n_i^2}\right), \tag{4.1}$$

  where, e^- is the electron charge[1]. We often lump $\frac{kT}{e^-}$ into a term called the thermal voltage, V_T, which is approximately $V_T \approx 0.026$ [V] at $T = 300$ [K].

- Third, the charge differential acts like a capacitor (it is storing charge). The nominal (or zero applied voltage) junction capacitance (or depletion layer capacitance), is given by, C_{j0} and is usually around a pico Farad (pF).

4.1 Reverse Bias

If we apply a voltage, such that the positive terminal is connected to the n material, and the negative terminal to the p material, the applied electric field, E_A, is in the same direction as the electric field of the potential barrier. This causes the depletion region to grow, because the free electrons in the n material are drawn to the positive terminal and the free holes in the p material are drawn to the negative terminal. The larger depletion region prevents charge from flowing so the diode is off. The reverse bias also effects the junction capacitance.

$$C_j = C_{j0}\left(1 + \frac{V_R}{V_b i}\right)^{-0.5} \tag{4.2}$$

[1] I am putting a minus sign in the exponent to distinguish the electron charge from the natural logarithm base. Thus I will put a plus in the exponent if I want to speak of the charge of a proton. The value of $e^- = 1.60217648710^{-19}$ [coulombs], which can also be calculated by $e^- = \frac{F}{N_A}$, where F is Faraday's constant ($9.64853399x10^4$ [C/mol]) and N_A is Avogadro's Number ($6.02213667x10^{23}$ [1/mol]).

with V_R the reverse bias voltage. Note the larger the larger the applied voltage the smaller the capacitance, which is due to the increased width of the depletion region (wider the region the lower the capacitance).

Chapter 5

Binary Junction Transistors

Characteristic	Common Base	Common Emitter	Common Collector
Input impedance	Low	Medium	High
Output impedance	Very High	High	Low
Phase Angle	$0°$	$180°$	$0°$
Voltage Gain	High	Medium	Low
Current Gain	Low	Medium	High
Power Gain	Low	Very High	Medium

Chapter 6

Field Effect Transistors

6.1 Ideal Behavior

A Field Effect Transistor (FET) is in one sense essentially a capacitor, and thus it is governed by

$$CV = Q \quad (6.1)$$

$$C_{gb}(V_{gc} - V_t) = Q_{channel} \quad (6.2)$$

where,

- C_{gb} is the capacitance between the gate and the body, this is often just called the gate capacitance.
- V_{gc} is the Voltage between the gate and the channel. Note that $V_c = V_{ds}/2$, so $V_{gc} = V_g s - V_{ds}/2$.
- V_t is the threshold voltage, i.e. the minimum voltage to cause an inversion layer to form.
- $Q_{channel}$ is the charge carries available in the channel to conduct.

The more charge carriers, $Q_{channel}$, the easier the current will flow, so calculating this is an essential step to quantitatively analyzing a FET. First we need to find out what our capacitance, C_{gb} is, this is done by

$$C_{gb} = \varepsilon_{ox}\frac{WL}{t_{ox}} \quad (6.3)$$

$$= 3.9\varepsilon_0\frac{WL}{t_{ox}} \quad (6.4)$$

- ε_{ox} is the permittivity[1] of the insulating oxide layer. Note: the 3.9 is the relative permittivity of silicon dioxide compared the the permittivity of free space. Relative permittivity is denoted ε_r, and varies by material, frequency, temperature, and sometimes even direction. We will treat it as a constant, which is ok for our operating situation.
- ε_0 is the permittivity of free space (8.85×10^{-14} [F/cm]).
- W is the width of the gate (along source and drain).
- L is the length under the gate (between source and drain).

[1] Permittivity is the resistance to forming an electric field.

Now we want to get the current flowing in the channel, but that means we need to know how fast they are moving. In a semiconductor the velocity of the charge carrier, v_c, is given by

$$\begin{aligned} v_c &= \mu_c E && (6.5) \\ &= \mu_c \frac{V_{ds}}{L} && (6.6) \end{aligned}$$

where μ_c is the mobility of the charge carrier. The length of the channel divided by the velocity of the carriers, gives us the time for a charge to cross the channel, $T_{channel}$. The total charge in the channel divided by this time is then the current.

$$\begin{aligned} i_{ds} &= \frac{Q_{channel}}{T_{channel}} && (6.7) \\ &= \frac{C_{gb}(V_{gc} - V_t)}{\frac{L}{v_c}} && (6.8) \\ &= \frac{3.9\varepsilon_0 \frac{WL}{t_{ox}}(V_{gc} - V_t)}{\frac{L}{\mu_c \frac{V_{ds}}{L}}} && (6.9) \\ &= 3.9\varepsilon_0 \frac{W}{t_{ox}}(V_{gs} - V_t)\mu_c \frac{V_{ds}}{L} && (6.10) \\ &= \frac{3.9\varepsilon_0}{t_{ox}}\mu_c \frac{W}{L}(V_{gc} - V_t)V_{ds} && (6.11) \\ &= \frac{3.9\varepsilon_0}{t_{ox}}\mu_c \frac{W}{L}(V_{gs} - V_t - V_{ds}/2)V_{ds} && (6.12) \end{aligned}$$

6.2 Amplification

$$i_{DS} = \frac{k}{2}(V_{GS} - V_T)^2 \qquad (6.13)$$

if $V_{DS} \geqslant V_{GS} - V_T \geqslant 0$.

Chapter 7

Logic Families

There are a great many logic families in use today. Probably the most famous is the TTL family, though it has largely been replaced by CMOS families. Even so, there are reasons for using different families (power, current, voltage, static, noise rejection, bus design, etc.). In the following sections we will examine some of the more well known families, their advantages, and how to interface them.

7.1 Diode Logic

Diode Logic (DL) uses diodes and resistors to implement logic gates. DL is a simple but old technology not used in integrated circuits. They are helpful to understand, as they are similar in some ways to later families. DL only has **and** and **or** gates.

Consider the circuit in Figure 7.1. If either in_1 or in_2 is high, the corresponding diodes (D_1 or D_2 respectively) turns on, making the output high. Since the output will be about 0.6v less than the input you can't put too many of these in series before the logic level drops below useful levels. If both inputs are off then both diodes don't conduct and the resistor to ground (R_1) pulls the output down to a low output (hence the name pull down resistor). The circuit is thus an **or** gate.

Now consider the circuit in Figure 7.1. If either in_1 or in_2 is low, the corresponding diodes (D_1 or D_2 respectively) turns on, making the output low, though it will be about 0.6v higher than the inputs so just like with the **or** gate, you can't do too many of these in series. If by inputs are high, then both diodes are off and the output is isolated from the input. The resistor to V_{cc} pulls the output up (hence the name, pull up resistor). The gate is thus an **and** gate.

Figure 7.1: Diode Logic (a) **Or** Gate and (b) **And** Gate.

(a) (b)

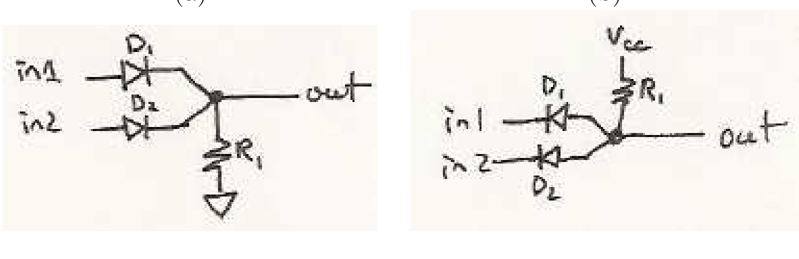


Figure 7.2: Diode Transistor Logic **Nand** Gate.

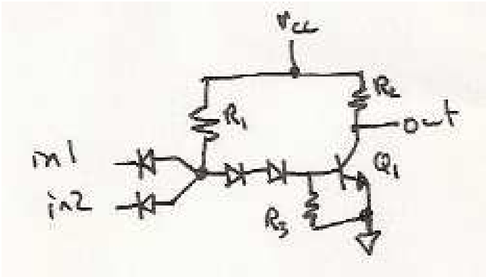
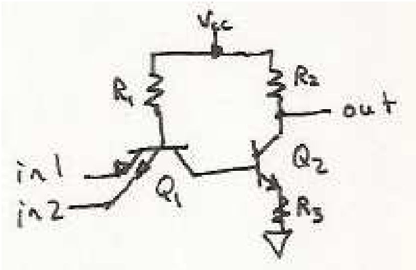


Figure 7.3: Transistor Transistor Logic **Nand** Gate.



7.2 Resistor Transistor Logic

Resistor Transistor Logic, (RTL) replaces the diodes of DL with transistors, which allowed for negation. This is thus the first full logic family.

7.3 Diode Transistor Logic

7.4 Transistor Transistor Logic

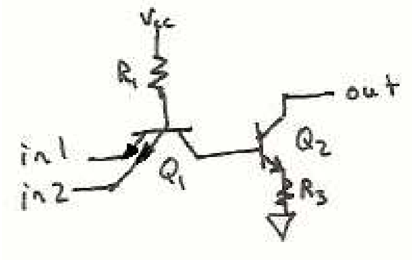
Transistor transistor logic (TTL) is arguably the most famous logic family. It has been used for around 40 years, and can still be purchased today. It is reasonably fast, has good noise rejection, and has good protection from static. A large number of interfaces are TTL compatible, so even when the components are not used, its design implications are still felt. The venerable 7400 and 5400 (milspec[1]) series are the most famous TTL components, and they have been used widely in engineering labs since I was in school way back when... If you want to see what components were in the series, see Appendix G.

7.4.1 Open Collector Outputs

If you noticed in earlier logic families, typically the collector of the output transistor is connected to power by a pull-up transistor, and without it the "high" output would not function correctly (it would be a weak, floating high). Since all the outputs use one, you could omit the resistor, and then wire the outputs together, and put an external pull-up. Such an output is "open collector" and they allow you to do active-low wired-OR

[1]Milspec means it is built to military specification, which would be enough to be noteworthy, but milspec parts are useful in hazardous environments, such as space, marine (water and saline), industrial fabrication environments, extreme temperature ranges, etc.

Figure 7.4: Transistor Transistor Logic **Nand** Gate with Open Collector Output.



and active-high wired-AND functionality. Wired gates are "gates" formed by wiring the outputs together, so you get a free gate. If you were to try this with driven outputs, you would get a short as both high and low are typically driven. Open collector outputs are generally slow, but proper resistor selection can improve things, and you can get two levels of logic for one level, which also saves time. I would advise against them unless you really know what you are doing and why. If you need to use them, the pull-up resistor is generally sized by calculating the minimum and maximum values per below.

$$R_{max} = \frac{\min(V_{cc}) - V_{OH}}{\sum \max(I_{OH}) + \sum \max(I_{IH})} \tag{7.1}$$

$$R_{min} = \frac{\max(V_{cc}) - V_{OL}}{I_{OL} - \sum \max(I_{IL})} \tag{7.2}$$

where, V_{cc} is the supply voltage, V_{OH} is the high output voltage of the gate, I_{OH} is the high output current for every gate whose outputs are connected to the pull-up resistor, I_{IH} is the high input current of every gate whose input is connected to the pull-up resistor, I_{OL} is the low output current of the gate, and I_{IL} is the low input current for every gate hooked to the pull-up resistor. This might look fancy, but it is actually just Ohm's law ($R = V/I$ in this case), where the voltage is the difference from the output to the supply, and the current must consider all the possible currents. We then maximize the top and minimize the bottom and vice versa to get our extreme cases. Memorizing this would be tough, understanding it is easy and from this it can be easily recreated.

7.4.2 Totem Pole Outputs

Instead of trying to take advantage of wired logic through a pull-up, we could consider how can we make the outputs switch as fast as possible. To do this we would need to pull up and down with separate transistors, so that we could quickly drive the output high or low. This is what totem pole outputs does. It is called totem pole outputs because the output transistors are on top of each other like a totem pole.

7.4.3 Tristate Outputs

7.5 CMOS Families

7.6 Interfacing

There are a lot of logic families out there so I will only cover the most common ones in this section, the rest can be handled in similar ways, you just need to know the standard considerations of voltages (not just min and max but also the logic levels), current (input and output at high and low levels), impedance (not always

Figure 7.5: Transistor Transistor Logic **Nand** Gate with Totem Pole Output.

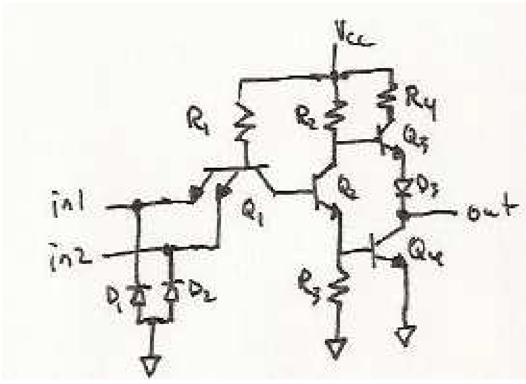
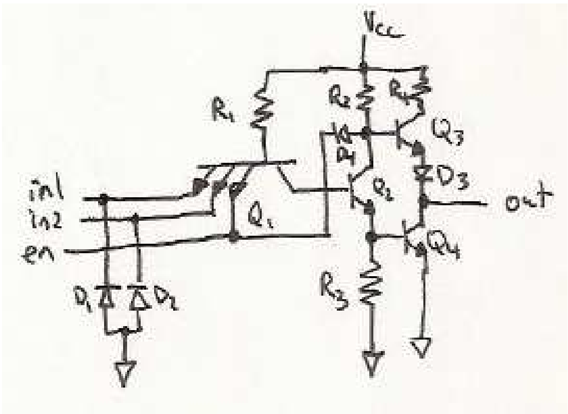


Figure 7.6: Transistor Transistor Logic **Nand** Gate with Tristate Output.



needed but can be important on bus lines and such), floating outputs (high, low, or not at all), timing (rise time, fall time, latency, etc.), and data rates (really this is an implication of timing, but it is big enough to be mentioned separately). It is helpful to know if the circuits you are interfacing are switching ground or power, as you can make a more reliable circuit using this information (i.e. you do the same in your circuit, which will be more compatible and thus also more reliable as it will not run into as many glitches caused by misreading a voltage level.). The basic requirements to call two chips compatible are:

Driver Output		Load Input
$V_{OH_{max}}$	$<$	$V_{IH_{max}}$
$V_{OH_{min}}$	$>$	$V_{IH_{min}}$
$V_{OL_{max}}$	$<$	$V_{IL_{max}}$
$V_{OL_{min}}$	$>$	$V_{I_{min}}$
$-I_{OH_{max}}$	$>$	$I_{IH_{max}}$
$I_{OL_{max}}$	$>$	$-I_{IL_{max}}$

The first two rows require that the high (true in positive logic) output voltage range must be contained in the high input voltage range, so that any H produced is correctly received. The next two lines do the same thing for the low values. The last two are to ensure that the driving chip can supply the needed current for a H and sink the needed current for a L. Note the negative signs are present because they flow out of the corresponding terminals rather than in. Level shifters can be placed between to meet voltage requirements, or current amplifiers/buffer stages can be used to meet current requirements. Typically, the first requirement is met by keeping the supply voltage the same, provided they can both take the same supply voltage. Similarly the fourth requirement is met by providing both the same ground. The last two requirements need to be verified for the entire load they are driving (fan-out and fan-in problems).

Before you design a circuit to interface, you should check if there is a device that already does the interfacing. In many cases there are devices designed for interfacing. For instance, between the old TTL family and the newer CMOS families (C, HC, AC, ACH, etc.), there are T versions (CT, HCT, ACT, ACHT, etc.) that can drop in replace the old TTL components, or one of the T devices can sit between the families and convert (say a buffer or two inverters). This is by far the easiest way to do the conversion, and I would do it this way unless forced to do otherwise. It is useful to know how to convert, should you ever have to, so below are the basics.

If you are straight converting signals, you often want to go through two inverters[2], as these devices as they are often used for this purpose, they are frequently designed to handle input and output. Note that you do not have to use inverters, they are a protection layer. The inverters serve as sacrificial elements (one for each logic family) to protect the circuits they are interfacing. Frequently you put them and any other interfacing hardware on a shim board, then if anything gets damaged it is the shim, not the original circuit.

Let's pick up the case mentioned above, where we wanted to go from an old TTL output to a newer CMOS input (say from a 74LS to a 74HC), but assume for some reason, we didn't just want to us a 74HCT to interface. We would need three circuit elements: a but would also need a pull up resistor of about 1k between them for two reasons. First, the voltage levels are incompatible, particularly at the high range, and the pull up solves this. Second, the pull up resistor is used to guarantee the input does not stay in the

[2]The inverters buffer the input. Don't use an actual buffer because a buffer is slower than an inverter, so buffers should be avoided unless absolutely needed. Basically, you put an inverter of the same logic type of the output immediately after the output and an inverter of the same logic type as the input immediately before the input.

dangerous 0.8v to 2.0v range[3]. To calculate the pull-up resistor more precisely you can use

$$R_{pull-up} \geqslant \frac{V_{cc}max - V_{TTL\ Low\ Output}}{I_{TTL\ Low\ Output} + (num\ inputs)I_{CMOS\ Low\ Input}} \tag{7.3}$$

$$R_{pull-up} \leqslant \frac{V_{cc} - V_{CMOS\ Input\ High}}{(num\ inputs)I_{CMOS\ Input\ High}} \tag{7.4}$$

Usually the minimum value is in the mid to high hundreds so a 1k resistor is a good guess up till about $num\ inputs = 8$, past that I would guess 2k till about $num\ inputs = 16$, I would not drive more than 16 gates directly with anything at the moment (theoretically you can but current, heat, and transients become big problems). The input current of CMOS devices is very small, so it can usually be ignored in the lower bound, and will cause the upper bound to be large (but finite). As the number of devices grows it cannot be ignored and at around $num\ inputs = 18$ there is a crossover, thus no pull-up resistor will work. You should calculate the minimum and maximum for any problem to ensure there is a feasible region and you are in it.

Depending on the circuit, you might also have to guarantee a particular rise time (the second reason we wanted a pull-up). In this case we have a simple first order[4] equation

$$V_{CMOS\ Input\ High} = V_{CC}\left(1 - e^{-\frac{t}{C \cdot R_{pull-up}}}\right) \tag{7.5}$$

where t is the desired rise time, and C is the total capacitance of the circuit, which is the sum of the output capacitance of the driving circuit, the input capacitance of the receiving circuits, and the capacitance of the line (often negligible but not always, it is probably about 1pF/cm). You can solve this for the value of $R_{pull-up}$. All of this assumed open-collector output (nothing driving high). If something is driving the circuit high, the pull-up resistor only has to account for the missing voltage. For instance totem pole output (typical for many TTL such as the LS family logic gates) is driven to at least 2.7 volts in around 10ns. Given some driven output voltage the equation for the pull-up resistor is

$$V_{CMOS\ Input\ High} - V_{Driven\ Output} = (V_{CC} - V_{Driven\ Output})\left(1 - e^{-\frac{t}{C \cdot R_{pull-up}}}\right). \tag{7.6}$$

Again solve for $R_{pull-up}$ and then ensure it falls between the minimum (Eq. 7.3) and maximum (Eq. 7.4).

If we were going the other way, we could just directly connect them, as long as we were in the fan-out restrictions, which is at least 10 gates for LS-TTL or 2-4 gates for TTL. Often designers put an additional CMOS buffer, say a 4096, for timing input to the slower TTL circuits. Note the data rates must be compatible, or no buffer will be able to solve incompatible data rates for continuous data streams.

As an interesting side note TTL at 5v is directly compatible, both ways with CMOS at 3v. Fan in and out restrictions still apply.

It is worth noting that CMOS has a wide range of voltage operation, so it is not uncommon to have to convert voltage ranges. Resistor voltage dividers are common for going down, and amplifier circuits, such as an open drain CMOS device with a pull-up resistor. As a particularly interesting case is ECL, which is usually run between 0v and -5.2v so the voltage differences and potential logic inversions need to be handled, or you can just run the CMOS from the same supply, and then just use diodes on the interface for protection.

---

[3]In the transition voltage range the P-channel to Vcc and the N-channel to ground can both be open, creating a path from Vcc to ground. This causes a current spike which can damage circuits if it is around too long. This happens every transition between high and low, but in the new CMOS families transitions are so fast the time of current spike is thus so short it does not effect things. The TTL output is slower and thus can allow significant damage. A pull up (or pull down if you want the default low) resistor solves this problem.

[4]This is the solution of a first order derivative equation for the rise time of a driven RC circuit. Theoretically they covered a lot of this in your physics sequence.

Part II

Digital Logic

Chapter 8

Boolean Algebra

Our goal is to design circuits, to do this we need to understand how the different circuit elements interact together to produce an output. A function can be described in three basic ways: algebraically, graphically, and by a table of values. Algebraically is usually thought of as the preeminent method as it covers every value precisely. While graphs are theoretically precise it is difficult to do it in practice. While tables are precise they are not exhaustive. When we deal with boolean values as opposed to numbers, graphs make no sense but tables are now exhaustive as there are only finitely many values. Proof by table is thus a legitimate technique in Boolean algebra. As a side note Boolean algebra derives its name from its systematizer, George Boole.

8.1 Postulates and Theorems

Boolean algebra has many similarities with the regular algebra you are used to. In fact all the usual properties like commutativity, distributivity, and associativity are present, and a few new ones to boot. Some important notes are in order before we get in too far.

- Primal refers to the properties of "or", which is the analog of addition hence the "+".

- Dual refers to the properties of "and", which is the analog of multiplication hence"$\cdot$".

- You will notice there are "primal" and "dual" versions of all the properties, which is different than with regular algebra. For instance if the primal (+) distributive property was true for regular algebra and say $a = 5$, $b = 2$, and $c = 3$ then $5 + 2 \cdot 3 = 11 = (5 + 2) \cdot (5 + 3) = 56$. The dual distributive property is the one you are used to.

- Some properties exist as only special cases in regular algebra. For instance, the primal idempotent property works in regular algebra only if a=0, and the dual idempotent property works in regular algebra for 0 and 1.

- Some properties are gone, for instance both the inverses (additive, $-a$, and multiplicative, $\frac{1}{a}$) don't exist. They are replaced by the concept of a complement, which does not exist in regular algebra.

Name	Primal	Dual
Commutativity	$a + b = b + a$	$a \cdot b = b \cdot a$
Distributivity	$a + (b \cdot c) = (a + b) \cdot (a + c)$	$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$
Identity	$0 + a = a + 0 = a$	$1 \cdot a = a \cdot 1 = a$
Complement	$a + a' = 1$	$a \cdot a' = 0$
Associativity	$(a + b) + c = a + (b + c)$	$(a \cdot b) \cdot c = a \cdot (b \cdot c)$
Idempotent	$a + a = a$	$a \cdot a = a$
Involution	$(a')' = a$	
Absorbtion	$a + a \cdot b = a$	$a \cdot (a + b) = a$
(special case)	$1 + b = 1$	$0 \cdot b = 0$
Simplification	$a + a' \cdot b = a + b$	$a \cdot (a' + b) = a \cdot b$
DeMorgan's Law	$(a + b)' = a' \cdot b'$	$(a \cdot b)' = a' + b'$

8.2 DeMorgan's Law

DeMorgan's Law is probably the most useful theorem in the table. DeMorgan's Law is the basis of our use of only one gate (either "nand" or "nor" can be that one gate) to design actual circuits. I don't want to make my notes purely a mathematical proof record, but it is important to be able to prove things. If you can't prove something, you don't understand it. Note that knowing a proof is also insufficient in and of itself, you need to know how to prove it and how to use it. I will prove DeMorgan's algebraically as I want to do the general statement, which has arbitrary numbers of variables which can't be represented simply in a table.

The most general statement of DeMorgan's Law is

$$(a_1 + a_2 + a_3 + \ldots + a_n)' \quad = \quad a_1' \cdot a_2' \cdot a_3' \cdot \ldots \cdot a_n' \tag{8.1}$$

and

$$(a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_n)' \quad = \quad a_1' + a_2' + a_3' + \ldots + a_n' \tag{8.2}$$

Proof:

The proof will be by induction on 8.1.

1. (Basis) Show that $(a_1 + a_2)' = a_1' \cdot a_2'$.

   By definition of complement, $a + a' = 1$ and $a \cdot a' = 0$. DeMorgan's Theorem states that the complement of $(a_1 + a_2)$ is $(a_1' \cdot a_2')$ so

   (a) First requirement: $a + a' = 1$

$$\begin{aligned}
(a_1 + a_2) + (a_1' \cdot a_2') &= (a_1' + (a_1 + a_2)) \cdot (a_2' + (a_1 + a_2)) \qquad \text{distributivity} \\
&= (a_1' + a_1 + a_2) \cdot (a_2' + a_1 + a_2) \qquad \text{associativity} \\
&= (a_1' + a_1 + a_2) \cdot (a_2' + a_2 + a_1) \qquad \text{commutativity} \\
&= ((a_1' + a_1) + a_2) \cdot ((a_2' + a_2) + a_1) \qquad \text{associativity} \\
&= (1 + a_2) \cdot (1 + a_1) \qquad \text{definition of complement} \\
&= 1 \cdot 1 \qquad \text{Absorbtion special case} \\
&= 1 \qquad \text{Idempotent}
\end{aligned}$$

   Thus they satisfy the first part of the definition.

(b) Second requirement: $a \cdot a' = 0$

$$\begin{aligned}(a_1 + a_2) \cdot (a_1' \cdot a_2') &= (a_1 \cdot (a_1' \cdot a_2')) + (a_2 \cdot (a_1' \cdot a_2')) \quad \text{distributivity} \\ &= (a_1 \cdot a_1' \cdot a_2') + (a_2 \cdot a_1' \cdot a_2') \quad \text{associativity} \\ &= (a_1 \cdot a_1' \cdot a_2') + (a_2 \cdot a_2' \cdot a_1') \quad \text{commutativity} \\ &= ((a_1 \cdot a_1') \cdot a_2') + ((a_2 \cdot a_2') \cdot a_1') \quad \text{associativity} \\ &= (0 \cdot a_2') + (0 \cdot a_1') \quad \text{complement} \\ &= 0 + 0 \quad \text{Absorbtion special case} \\ &= 0 \quad \text{Absorbtion special case}\end{aligned}$$

Thus they satisfy the second part of the definition and are therefore complements of each other.

2. (Inductive Step) Assume it works for $(a_1 + a_2 + a_3 + \ldots + a_{n-1})' = a_1' \cdot a_2' \cdot a_3' \cdot \ldots \cdot a_{n-1}'$ and show it thus works for $(a_1 + a_2 + a_3 + \ldots + a_n)' = a_1' \cdot a_2' \cdot a_3' \cdot \ldots \cdot a_n'$

$$\begin{aligned}(a_1 + a_2 + a_3 + \ldots + a_{n-1} + a_n)' &= ((a_1 + a_2 + a_3 + \ldots + a_{n-1}) + a_n)' \quad \text{associativity} \\ &= (a_1 + a_2 + a_3 + \ldots + a_{n-1})' \cdot a_n' \quad \text{basis} \\ &= a_1' \cdot a_2' \cdot a_3' \cdot \ldots \cdot a_{n-1}' \cdot a_n' \quad \text{induction hypothesis}\end{aligned}$$

♢ SDG ♢

Example

Verify the following by both algebra and truth tables.

$$A + A' \cdot B = B + B' \cdot A$$

Sol:

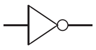
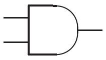
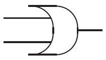
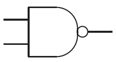
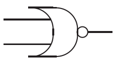
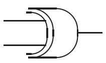
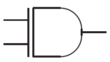
$$\begin{aligned}A + A' \cdot B &= A \cdot 1 + A' \cdot B \\ &= A \cdot (B' + B) + A' \cdot B \\ &= A \cdot B' + A \cdot B + A' \cdot B \\ &= A \cdot B' + (A + A') \cdot B \\ &= A \cdot B' + 1 \cdot B \\ &= A \cdot B' + B \\ &= B + B' \cdot A\end{aligned}$$

A	B	$A' \cdot B$	$A + A' \cdot B$
0	0	0	0
0	1	1	1
1	0	0	1
1	1	0	1

A	B	$B' \cdot A$	$B + B' \cdot A$
0	0	0	0
0	1	0	1
1	0	1	1
1	1	0	1

Notice the truth tables are the same for $A + A' \cdot B$ and $B + B' \cdot A$, so they are equal.

8.3 Gates

<table>
<tr><th>Name</th><th>Expression</th><th>Symbol</th><th>Truth Table</th></tr>
<tr><td>Not</td><td>$z = x'$</td><td></td><td><table><tr><th>x</th><th>x'</th></tr><tr><td>0</td><td>1</td></tr><tr><td>1</td><td>0</td></tr></table></td></tr>
<tr><td>And</td><td>$z = x \cdot y$</td><td></td><td><table><tr><th>x</th><th>y</th><th>$x \cdot y$</th></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table></td></tr>
<tr><td>Or</td><td>$z = x + y$</td><td></td><td><table><tr><th>x</th><th>y</th><th>$x + y$</th></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>1</td><td>0</td><td>1</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table></td></tr>
<tr><td>Nand</td><td>$z = (x \cdot y)'$</td><td></td><td><table><tr><th>x</th><th>y</th><th>$(x \cdot y)'$</th></tr><tr><td>0</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>1</td><td>0</td><td>1</td></tr><tr><td>1</td><td>1</td><td>0</td></tr></table></td></tr>
<tr><td>Nor</td><td>$z = (x + y)'$</td><td></td><td><table><tr><th>x</th><th>y</th><th>$(x + y)'$</th></tr><tr><td>0</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>0</td></tr></table></td></tr>
<tr><td>Xor</td><td>$z = x \oplus y$</td><td></td><td><table><tr><th>x</th><th>y</th><th>$x \oplus y$</th></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>1</td><td>0</td><td>1</td></tr><tr><td>1</td><td>1</td><td>0</td></tr></table></td></tr>
<tr><td>Xnor</td><td>$z = x \odot y$</td><td></td><td><table><tr><th>x</th><th>y</th><th>$x \odot y$</th></tr><tr><td>0</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table></td></tr>
</table>

A	B	False, Ground	$A \cdot B$, And	$A \nRightarrow B$, Negated Implication, andn	A	$B \nRightarrow A$, Negated Implication, norn	B	$A \oplus B$, Xor	$A + B$, Or	$\overline{A + B}$, Nor	$A \odot B$, Equiv, Xnor	$\overline{B}$, Not	$B \implies A$, Implication, orn	$\overline{A}$, Not	$A \implies B$, Implication, nandn	$\overline{A \cdot B}$, Nand	True, Power
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

Chapter 9

Logic Conventions

In a standard digital logic course, a usual starting point is to associate a high voltage, say 5v or 3.3v, with true (1), and a low voltage, usually ground, with false (0). The purpose of this chapter is to blow that assumption out of the water. We really have two completely different things we need to associate some way. One is a system of logic, composed of truth (1), falsehood (0)[1], and logical gates. The other system is a physical one of high voltages (Vcc or Vdd), low voltages (ground), and hardware devices that operate off these voltage values.

Ideally we would like a system that allows us to look at the logic without having to think about the hardware, or to look at the hardware without thinking about the logic. Mixed logic allows us to do this. To use it we will design the logic as we normally would, without any thought of the hardware that we will use to implement. When we go to select the devices to implement the logic gates, we will use mixed logic to give us flexibility in the selection of devices, by strategically and consistently changing the logic convention in place at different locations in our design. As long as we do this we will not change the logic of our design.

I need to take a little pedantic aside, because the term logic convention, which is standardly used to refer to the association of logic values to voltage values has an unintended implication that seems to confuse people. Logic convention causes people to think that the voltages are preserved but the logic is changing, which is no problem for analysis but causes unnecessary confusion in design. We could have used the term voltage convention to refer to the same think because it is a design oriented term, i.e. it does not suggest you are changing your logic, you are changing your voltage associations, but then people would get confused in design. Once you become used to a term you are fine, but it is the learning I care about, and it is for this reason I suggest logic-voltage convention (LVC). LVC does not have any connotation toward design or analysis, and thus I hope will cause people to understand it better and use it more.

9.1 Logic-Voltage Conventions

The first LVC is positive logic, which is what most digital logic students think is the only one there is. Positive logic is also called active-high, which is more in keeping with my pedantic aside from the introduction. See Table 9.1.

The second LVC is negative logic, which becomes important in developing the other two canonical forms in Section 9.2. Negative logic is also called active-low, which is more in keeping with my pedantic aside from the introduction. See Table 9.2.

The final LVC is mixed logic, which uses either positive or negative logic rules on a wire by wire basis. The key to using this is to have a system of marking the wires and the signal names so you can tell which

[1]These associations of true with 1 and false with 0 are conventions also, and we could play with them also, but we will leave that off to a discussion of math for another book.

Table 9.1: Positive Logic/Active-High

Logic	Voltage
F	L
T	H

Table 9.2: Negative Logic/Active-Low

Logic	Voltage
F	H
T	L

convention is in place.

- The traditional way to mark wires for positive logic (active-high) was to do nothing, i.e. just draw the wire. The new (from 1984) IEEE standard has us put a flag on top of the wire at each end that points in the direction of flow (into the device for inputs, out of the device for outputs), and is associated with the term active-high. The flags look like a small right triangle with the base formed by the wire and the hypotenuse pointing in the direction of the flow.
- Positive logic wire/signal names have a '.H' or '.h' appended to it. Some people append either a '+' or a '↑' but I find this more tedious. A final convention puts a lower case p in front. In other cases nothing is added to the name for these, and the absence lets you know the convention, though this is error prone as you can't tell if the signal was just missed in the naming. I suggest you use the '.h' to be clear.
- Wires that use the negative logic convention have an open circle on all ends with the classic logic shapes. Active-low flags (open arrow, only on the lower part, pointing in the direction of signal travel) are used interchangeably with negative logic bubbles, though you should pick a convention and stick with it. The flags look like a small right triangle with the base formed by the wire and the hypotenuse pointing in the direction of the flow.
- Negative logic (active-low) name/signal has a '.L' appended. Some people also append: a '−', a '↓', a '#'. Other notations put leading symbols of a 'n' or a slash, '/', which is designed to look like an bar over the signal, which is the final way. I don't like the overbar or slash as it is easily confused with not, though it is the most common[2]. The '`_B`' notation is confusing as it could mean byte in other contexts.

In general the bubbles go with the classic logic shapes, and the arrows go with the new IEEE 91-1984 standard, which calls for boxes with symbols. Mixed logic is much more flexible in the ability to use other hardware devices to implement gates. One way of thinking of this is that we can implement a logic function with a variety of hardware devices, or put the other way one hardware device can implement a variety of logic gates. This is easiest to explain by an example.

Example 2 *Say we need a 'not' gate. With either positive or negative logic we have only one choice, but consider mixed logic. We could have the input as either positive or negative and a similar but independent choice for the output. This means we have four possible mixed conventions. But how many devices? Is this only an illusion of choice?*

[2]It is an inconsistent use of the bubbles and slashes that causes so much confusion in digital logic students, so I will avoid them. Hopefully when you feel comfortable with the conventions you will then have no problem reading the highly overloaded syntax that is commonly used.

1. *Positive logic to positive logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*L*	*H*	*T*
T	*H*	*L*	*F*

This requires a voltage inverter, which is what most people think a 'not' gate is.

2. *Negative logic to negative logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*H*	*L*	*T*
T	*L*	*H*	*F*

This also requires a voltage inverter, so no new requirement is added.

3. *Positive logic to negative logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*L*	*L*	*T*
T	*H*	*H*	*F*

The voltage is already correct so only a wire is needed to connect them. We now have something new, a 'bare wire not'. Think about this for a second, we have a wire that can do logical negation. That is pretty cool.

4. *Negative logic to positive logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*H*	*H*	*T*
T	*L*	*L*	*F*

Again the voltages are correct so only a wire is needed.

Our four logic combinations gave us two different devices (inverter or bare wire) that could fulfil our needs, depending on the convention picked. That is one more than with either straight positive logic or straight negative logic, which yielded the same one possibility (inverter) as each other. The increased design flexibility is important in a real design situation.

Now lets try from a different perspective. The last example started with a requirement on the logic and found what devices could work, now let's start with the device and find out what it can do for our logic.

Example 3 *The voltage characteristics of an inverter is*

Voltage In	*Voltage Out*
L	*H*
H	*L*

Now we just have to add the interpretation, i.e. the logic convention. We have four possibilities for a single input, single output.

1. *Positive logic to positive logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*L*	*H*	*T*
T	*H*	*L*	*F*

This is 'not', and as we noted in the last example this is why most people think an inverter is 'not'.

2. *Negative logic to negative logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*H*	*L*	*T*
T	*L*	*H*	*F*

This also is 'not'.

3. *Positive logic to negative logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*L*	*H*	*F*
T	*H*	*L*	*T*

This is a logic convention changer. It preserves the interpretation (logic value) but switches conventions.

4. *Negative logic to positive logic*

Logic In	*Voltage In*	*Voltage Out*	*Logic Out*
F	*H*	*H*	*T*
T	*L*	*L*	*F*

Again the we see the inverter also 'inverts' the convention.

We thus have that an inverter can serve one of two purposes: logic value inversion (not) or logic convention inversion (converter).

The options are even larger with two input gates.

Example 4 *Consider an '**andn**' gate in positive logic. What could we use to make it with standard TTL Gates (**nand**, **nor**, **and**, **or**, **not**, **xor**, **xnor**)?*

Answer:

*Let's start by looking at the logic table of an **andn** gate.*

A	*B*	*A andn B*
0	*0*	*0*
0	*1*	*0*
1	*0*	*1*
1	*1*	*0*

Since this is positive logic, we have to find a device or devices that give us the voltage pattern below.

In 1	*In 2*	*Out*
L	*L*	*L*
L	*H*	*L*
H	*L*	*H*
H	*H*	*L*

1. *If we used positive logic everywhere, we would have an **andn** gate, which we have no implementation for directly, so we could use an **not** on in2(B) and then **and** the result with in1(A).*

A	$\bar{B}$	A *and* $\bar{B}$
0	*1*	*0*
0	*0*	*0*
1	*1*	*1*
1	*0*	*0*

2. *If we used negative logic on in2, we would have an **and** gate, and we could use an inverter (**not**) to take the initial positive logic system on in2 to make it negative logic without negating the input.*

A	*B.L*	*A and B.L*
0	*1*	*0*
0	*0*	*0*
1	*1*	*1*
1	*0*	*0*

3. *If we used negative logic on in1, we would have a **nor** gate, and we could use an inverter (**not**) to take the initial positive logic system on in1 to make it negative logic without negating the input.*

A.L	B	A.L nor B
1	0	0
1	1	0
0	0	1
0	1	0

4. *If we used negative logic on both inputs, we would have a norn gate, which is not implemented. The negation of B.L can be handled by a bare wire not, which will also work to go from B.h to B.L. Going from A.h to A.L can be handled by an inverter (**not**). The rest can be handled by a **nor** gate, so that this is the same as the last case.*

A.L	B.L	A.L norn B.L
1	1	0
1	0	0
0	1	1
0	0	0

5. *If we used negative logic only on the output we would have nandn, which is not implemented. We need a **not** on in2, and a bare wire not on the output handling the logic level and not of the output, leaving the main gate as an **and**.*

A	B	(A nandn B).L
0	0	1
0	1	1
1	0	0
1	1	1

6. *If we used negative logic on in2 and the output, we would have an **nand** gate, and we could use an inverter (**not**) to take the initial positive logic system on in2 to make it negative logic without negating the input and similarly an inverter (**not**) could be used to take the initial negative logic output and convert to positive logic.*

A	B.L	(A nand B.L).L
0	1	1
0	0	1
1	1	0
1	0	1

7. *If we used negative logic on in1 and the output, we would have a **or** gate, and we could use an inverter (**not**) to take the initial positive logic system on in1 to make it negative logic without negating the input and similarly an inverter (**not**) could be used to take the initial negative logic output and convert to positive logic.*

A.L	B	(A.L or B).L
1	0	1
1	1	1
0	0	0
0	1	1

8. *If we used negative logic on both inputs and the output, we would have **orn**, which does not exist. We could make it by using a bare wire not on in2 and inverters (**not**) on in1 and the output. The main gate is now an (or).*

A.L	*B.L*	*(A.L or B.L).L*
1	*1*	*1*
1	*0*	*1*
0	*1*	*0*
0	*0*	*1*

The above cases reduce to four possibilities:

1. *an* ***and*** *with a* ***not*** *on in2,*
2. *an* ***nor*** *with a* ***not*** *on in1,*
3. *an* ***nand*** *with a* ***not*** *on in2 and output,*
4. *an* ***or*** *with a* ***not*** *on in1 and output.*

It is straightforward to show they are equivalent, the nice thing is that mixed logic can generate them all.

It is this flexibility is one great reason that mixed logic so popular.

9.2 Canonical Forms

Only in rare cases are problems easy enough to reduce to a single gate we can recognize. In most cases we need to design more complicated circuits to achieve the desired result. An important result in boolean logic is that every possible output pattern can be realized from input signals in two levels of logic if each gate can have as many inputs as you need. The practical use of this is that we can create canonical forms. Two main canonical forms are used, Sum-of-Products (SOP) and Product-of-Sums (POS). Each canonical is made up of terms, and each term corresponds to one row of a truth table. Since each term corresponds to a row in the truth table the terms can be referenced by the row number or the actual equation for the term (I will show how to get the equations below) The names were designed to be descriptive, as follows.

9.2.1 Sum of Products

A sum is a series of terms connected by "+", which is **or** in our case. A product is a series of terms connected by "·", which is **and** in our case, thus SOP is bunch of terms that only use **and** in them that are connected together by **or**. We call each term in a SOP a Miniterm because it is only true for one combination of inputs (since **and** is only true for one combination of inputs this follows directly). In essence each Miniterm places one true (1) value in the output, and thus can be thought of as tracking the 1's. Each Miniterm's equation is written such that it will be true for that row only of the truth table. Consider the following.

x	y	z	Row	Miniterm
0	0	0	0	$x' \cdot y' \cdot z'$
0	0	1	1	$x' \cdot y' \cdot z$
0	1	0	2	$x' \cdot y \cdot z'$
0	1	1	3	$x' \cdot y \cdot z$
1	0	0	4	$x \cdot y' \cdot z'$
1	0	1	5	$x \cdot y' \cdot z$
1	1	0	6	$x \cdot y \cdot z'$
1	1	1	7	$x \cdot y \cdot z$

Notice that the row number is just the decimal value of binary number (xyz). Also note that the Miniterm is formed by placing complements where the corresponding variable is zero, this forces all the variables (or complements) to be true for the equation on that row. To get a better appreciation of what it means for a Miniterm to be adding or tracking the 1's consider a series of truth tables.

$a_1 = x \cdot y \cdot z$

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

$a_2 = x \cdot y \cdot z + x \cdot y' \cdot z$

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

$a_3 = x \cdot y \cdot z + x \cdot y' \cdot z + x \cdot y' \cdot z'$

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1

$a_4 = x \cdot y \cdot z + x \cdot y' \cdot z + x \cdot y' \cdot z' + x' \cdot y \cdot z'$

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1

Each time a term is added the truth table shows the output in the corresponding row to the new term becomes 1. It is thus evident that we can go the reverse direction. We always want a shorter way to write things, so since Miniterm implies small we can also represent the term by a small "m" followed by the row number, so $m_5 = x \cdot y' \cdot z$. Using this notation the designs above are $a_1 = \{m_7\}$, $a_2 = \{m_5, m_7\}$, $a_3 = \{m_4, m_5, m_7\}$, and $a_4 = \{m_2, m_4, m_5, m_7\}$.

This is nice and short but we want even shorter so we abbreviate the list of Miniterms by $\sum$ followed by a list of the numbers of the terms (you might recall that $\sum$ means a series of '+' in math). While it is not a general rule, I list the inputs as subscripts of $\sum$, it makes it easier to tell the sequence and what signals (wires) to connect. Thus our summation notation for the designs would be, $a_1 = \sum_{x,y,z}(7)$, $a_2 = \sum_{x,y,z}(5,7)$, $a_3 = \sum_{x,y,z}(4,5,7)$, and $a_4 = \sum_{x,y,z}(2,4,5,7)$. Note the listing of inputs as subscripts can be done with the listing of Miniterms, $a_1 = \{m_7\}_{x,y,z}$, $a_2 = \{m_5, m_7\}_{x,y,z}$, $a_3 = \{m_4, m_5, m_7\}_{x,y,z}$, and $a_4 = \{m_2, m_4, m_5, m_7\}_{x,y,z}$.

9.2.2 Product of Sums

By the colloquial descriptions above for sum and product, POS is a bunch of terms that only use **or** gates internally and are connect by **and** gates. We call each term in a POS a Maxiterm because it is true for every input combination but one (since it is made of **or** gates). A Maxiterm is thus false for only one combination of the inputs. In essence each Maxiterm places one false (0) value in the output, so it can be thought of as tracking the 0's. Each Maxiterm's equation is written such that it will be true for that row only of the truth table. Consider the following.

x	y	z	Row	Maxiterm
0	0	0	0	$x + y + z$
0	0	1	1	$x + y + z'$
0	1	0	2	$x + y' + z$
0	1	1	3	$x + y' + z'$
1	0	0	4	$x' + y + z$
1	0	1	5	$x' + y + z'$
1	1	0	6	$x' + y' + z$
1	1	1	7	$x' + y' + z'$

Notice that the row number is just the decimal value of binary number (xyz). Also note that the Maxiterm is formed by placing complements where the corresponding variable is one, this forces all the variables (or complements) to be false for the equation on that row. To get a better appreciation of what it means for a Maxiterm to be adding or tracking the 0's consider a series of truth tables.

$\bar{a}_1 = (x + y + z)$

x	y	z	a
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

$\bar{a}_2 = (x+y+z)\cdot(x+y+z')$

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

$\bar{a}_3 = (x + y + z) \cdot (x + y + z') \cdot (x + y' + z')$

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

$a_4 = (x + y + z) \cdot (x + y + z') \cdot (x + y' + z') \cdot (x' + y' + z)$

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1

Notice that each time a term is added the truth table shows the output in the corresponding row to the new term becomes 0. It is thus evident that we can go the reverse direction. We always want a shorter way to write things, so since Maxiterm implies large we can also represent the term by a capital "M" followed by the row number, so $M_5 = x' + y + z'$. Using this notation the designs above are $\bar{a}_1 = \{M_0\}$, $\bar{a}_2 = \{M_0,\, M_1\}$, $\bar{a}_3 = \{M_0,\, M_1,\, M_3\}$, and $a_4 = \{M_0,\, M_1,\, M_3,\, M_6\}$.

This is nice and short but we want even shorter so we abbreviate the list of Maxiterms by $\prod$ followed by a list of the numbers of the terms (you might recall that $\prod$ means product in math). While it is not a general rule, I list the inputs as subscripts of $\prod$, it makes it easier to tell the sequence and what signals (wires) to connect. Thus our product notation for the designs would be, $\bar{a}_1 = \prod_{x,y,z}(0)$, $\bar{a}_2 = \prod_{x,y,z}(0,1)$, $\bar{a}_3 = \prod_{x,y,z}(0,1,3)$, and $a_4 = \prod_{x,y,z}(0,1,3,6)$. Note the listing of inputs as subscripts can be done with the listing of Maxiterms, $\bar{a}_1 = \{M_0\}_{x,y,z}$, $\bar{a}_2 = \{M_0,\, M_1\}_{x,y,z}$, $\bar{a}_3 = \{M_0,\, M_1,\, M_3\}_{x,y,z}$, and $a_4 = \{M_0,\, M_1,\, M_3,\, M_6\}_{x,y,z}$.

As a final note, the last problem in this section is the same as the last one in the SOP section and so the designs must be equivalent. We thus have $a_4 = \prod_{x,y,z}(0,1,3,6) = \sum_{x,y,z}(2,4,5,7)$, from which we can note that the if we take all the numbers from the truth table and remove the ones from the $\prod$ list, we have the $\sum$ list and vice versa. This gives us a nice way to switch between the two forms provided we know how many rows are in the table, which you can know from counting the number of inputs in our subscript (another good reason for listing them).

Example

Obtain the sum of products form by algebra and the product of sums form by truth table for $A + B \cdot (C + A) \cdot (B' + A' \cdot B)$.

$$
\begin{aligned}
A + B \cdot (C + A) \cdot (B' + A' \cdot B) &= A + B \cdot (B' + A' \cdot B) \cdot (C + A) \\
&= A + (B \cdot B' + B \cdot A' \cdot B) \cdot (C + A) \\
&= A + A' \cdot B \cdot (C + A) \\
&= A + A' \cdot B \cdot C + A' \cdot B \cdot A \\
&= A + A' \cdot B \cdot C \\
&= A \cdot (B + B') \cdot (C + C') + A' \cdot B \cdot C \\
&= A \cdot B' \cdot C' + A \cdot B' \cdot C + A \cdot B \cdot C' + A \cdot B \cdot C + A' \cdot B \cdot C \\
&= m_4 + m_5 + m_6 + m_7 + m_3 \\
&= \Sigma(3,4,5,6,7)_{A,B,C}
\end{aligned}
$$

A	B	C	$A + B \cdot (C + A) \cdot (B' + A' \cdot B)$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

The three terms with 0's are thus M_0, M_1, and M_2, yielding $\Pi(0, 1, 2)_{A,B,C}$.

Chapter 10

Combinational Circuits

Combinational circuits are the most basic type of circuit that can be designed in that they have no memory input. In a combinational circuit the outputs are completely determined by the inputs.

Consider a simple example. Say I have two hard disks on my computer and I want to hook up a light that shows when either is accessed. Hard disks have an output line that signals when they are accessed so we have two variables d_1 and d_2 which are the disk access output signals from the drives. Since I want the light to come on when either drive is accessed the truth table describing this is

d_1	d_2	light	comments
0	0	0	neither disk accessed, no light
0	1	1	second disk being accessed, light on
1	0	1	first disk being accessed, light on
1	1	1	both disks being accessed, light on

This table is identical to the definition of "or" so we have that light $= d_1 + d_2$. Thus by connecting the signals from the disks to an "or" gate and using the output of the gate to drive the light. This is a combinational circuit because it does not matter what happened in the past or what some variable's value is (the value of variables in a circuit with memory is known as state).

Combinational circuits are the foundation of digital design, as sequential circuits (the circuits with memory) can be handled as a combinational circuits driven by inputs and memory and the outputs not only drive other circuits, they modify the memory that drives the input. You can thus consider all sequential circuits as combinational ones with feedback.

10.1 Designing: Tables

If there are only a few trues (or falses) that need to be generated and a small number of input variables, then it is easy to do the design off a truth table by reading the canonical terms. Even complex problems can be designed with the use of decoders or multiplexors (mux).

10.1.1 Implementing With Sum of Products

Sum of Products design rules:

- For each row in the table where the output is a "1", connect the inputs to a **nand** gate (or an **and** gate) being sure to invert any input line that has a "0" in that row.
- Connect the outputs of the previous gates into another **nand** gate (or an **or** gate if you used **and** gates in the previous step).

- The output of the last gate is the desired output.

10.1.2 Implementing With Product of Sums

Sum of Products design rules:

- For each row in the table where the output is a "0", connect the inputs to a **nor** gate (or an **or** gate) being sure to invert any input line that has a "1" in that row.
- Connect the outputs of the previous gates into another **nor** gate (or an **and** gate if you used **or** gates in the previous step).
- The output of the last gate is the desired output.

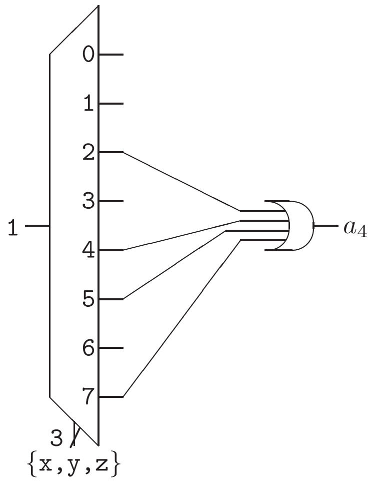
10.1.3 Implementing With Decoders

Decoders have an enable input, n address lines, and 2^n output lines that are true if the decoder is enabled and the address on the address line is their line on the decoder. Decoder designs have a few simple rules:

- Enable the decoder.
- Connect the inputs to the address lines in the sequence of the table.
- Connect the decoder outputs that correspond to 1's in the table to an **or** gate.
- The output of the **or** gate is the desired output.

It is easiest to see this by an example. Consider the following table from our canonical term section.

x	y	z	a_4
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1



The technique can be seen to be fairly straightforward. It can require a large decoder if the 1's and 0's are mixed, but it can be done with a small decoder if there are few tightly grouped 1's or 0's.

10.1.4 Implementing With Multiplexors

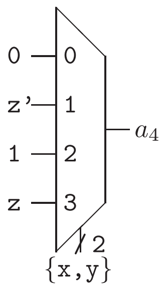
A mux has 2^n input lines, n address select lines, and 1 output. The input line that corresponds to the address is passed to the output. The design technique is a little tricky.

- All but one of the inputs are connected to the address select lines.

- The remaining input is used to divide the table into pairs, each pair corresponding to one of the 2^n input lines.
  - If both outputs in a pair in the table are both zero then the corresponding input line is grounded.
  - If both outputs in a pair in the table are both one then the corresponding input line is set high.
  - If both outputs in a pair in the table are the same as the one unconnected input, then that input is connected to the corresponding input to the mux.
  - If both outputs in a pair in the table are the opposite of the one unconnected input, then the inverse of that input is connected to the corresponding input to the mux.
- The desired function is on the output of the mux.

Let us one more time examine our example from our canonical term section.

x	y	z	a_4	Mux Input
0	0	0	0	0
0	0	1	0	
0	1	0	1	z’
0	1	1	0	
1	0	0	1	1
1	0	1	1	
1	1	0	0	z
1	1	1	1	



This one is hard to see at first but it is easy and compact once understood.

10.2 Designing: Karnaugh Maps

Karnaugh maps are a nice visual way to handle small design problems, i.e. those with less than 6-8 inputs. Karnaugh maps are formed by making a table indexed in both rows and columns by the inputs which are arranged in Grey code order (00,01,11,10)[1].

Basic rules for all encirclements:

1. Always encircle only a number of items equal to a power of 2 (1, 2, 4, 8, 16, etc.).
2. Only encircle either 0’s or 1’s, but not a mixture. Since don’t cares, x, could be either a 0 or 1, you can mix and match them.
3. Make only the largest encirclements possible.
4. Overlap encirclements (partial due to above rule) whenever possible to remove errors of type 1. Diagonal overlaps will take care of errors of type 2.

Rules for encircling with **and** gates for SOP:

1. Only encircle 1’s.
2. Encirclements must be horizontally or vertically aligned rectangles.

Rules for encircling with **or** gates for POS:

[1]Veitch diagrams are just like Karnaugh maps, but they are in normal binary order. This makes them a pain to recognize patterns and so they are rarely used.

1. Only encircle 0's.
2. Encirclements must be horizontally or vertically aligned rectangles.

Rules for encircling with **xor** or **equiv** gates:

1. Only encircle 1's.
2. Encirclements must be checkerboard patterned (diagonal).

Example 5 *Make a Karnaugh map for $\Pi(1, 2, 3, 6, 7, 9, 11, 15)_{A,B,C,D}$ and use it to simplify the expression. Implement your result using And, Or, and inverter gates in a HDL module to describe the circuit.*

Sol:

			A	A	
	1	1	1	1	
	0	1	1	0	D
C	0	0	0	0	D
C	0	0	1	1	
		B	B		

Three 4 entry encirclements of zeros (two squares and a row). This gives the simplification as:

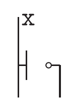
$$(C' + D') \cdot (A + C') \cdot (B + D')$$

Alternately you could make three 4 entry encirclements of ones (two squares and a row). The simplification would then be:

$$C' \cdot D' + B \cdot C' + A \cdot D'$$

A HDL implementation of the simplified sum of products form is:

```
module my_circ(F,A,B,C,D);
  input A,B,C,D;
  output F;

  wire c,d,e,x,y,z;

  not g1(c,C);
  not g2(d,D);
  and g3(x,c,d);
  and g4(y,B,c);
  and g5(z,A,d);
  or g6(e,x,y);
  or g7(F,e,z);
endmodule
```



Example 6 *We wanted to design a system to check three lines, say A, B, C. If only one line is active we want to receive a signal. We are also interested in knowing if lines A and C are active, and we want no errors of type-1. The design is small, so we start with a Karnaugh map.*

			A	A
	0	1	0	1
C	1	0	1	1
		B	B	

10.3 Quine-McCluskey

Originally proposed by Quine and then modified by McCluskey, this method provides an symbolic tabular way to minimize a boolean algebraic function. Graphical methods like Karnaugh maps are great for up to about 6 variables, but then they bog down really badly.

The idea of this method is to combine terms using the rule $xy + xy' = x$, where x represents multiple variables, but y is only one variable.

$$\sum_{abcd}(1, 3, 4, 5, 6, 7, 10, 14, 15) \tag{10.1}$$

We begin by writing the minterms in binary. We then sort them so they will be in order of increasing index. Index is defined to be the number of 1's in the expression. In order to combine terms they may differ by only 1 value, so we only need compare each index group with the group above it. Here is the term by term combination to generate the 2-term implicants from the minterms. When a minterm is used it is checked to note it cannot be a prime implicant, though we continue to use it to generate other terms as a minterm can be in multiple groupings.

Minterms		2-terms
0001		
0100		
0011		
0101		
0110		
1010		
0111		
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100		
0011	✓	
0101		
0110		
1010		
0111		
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100		0-01
0011	✓	
0101	✓	
0110		
1010		
0111		
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	
0110		
1010		
0111		
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	
1010		
0111		
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	0-11
1010		
0111	✓	
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	0-11
1010		01-1
0111	✓	
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	0-11
1010		01-1
0111	✓	011-
1110		
1111		

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	0-11
1010		01-1
0111	✓	011-
1110	✓	-110
1111		

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	0-11
1010	✓	01-1
0111	✓	011-
1110	✓	-110
1111		1-10

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	0-11
1010	✓	01-1
0111	✓	011-
1110	✓	-110
1111	✓	1-10
		-111

Minterms		2-terms
0001	✓	00-1
0100	✓	0-01
0011	✓	010-
0101	✓	01-0
0110	✓	0-11
1010	✓	01-1
0111	✓	011-
1110	✓	-110
1111	✓	1-10
		-111
		111-

We now move on to generating the 4-term implicants from the 2-term implicants. We do it in the exact same way as the preceding development.

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0- -1
0100	✓	0-01		
0011	✓	010-		
0101	✓	01-0		
0110	✓	0-11		
1010	✓	01-1	✓	
0111	✓	011-		
1110	✓	-110		
1111	✓	1-10		
		-111		
		111-		

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0- -1
0100	✓	0-01	✓	0- -1
0011	✓	010-		
0101	✓	01-0		
0110	✓	0-11	✓	
1010	✓	01-1	✓	
0111	✓	011-		
1110	✓	-110		
1111	✓	1-10		
		-111		
		111-		

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0- -1
0100	✓	0-01	✓	0- -1
0011	✓	010-	✓	01- -
0101	✓	01-0		
0110	✓	0-11	✓	
1010	✓	01-1	✓	
0111	✓	011-	✓	
1110	✓	-110		
1111	✓	1-10		
		-111		
		111-		

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0- -1
0100	✓	0-01	✓	0- -1
0011	✓	010-	✓	01- -
0101	✓	01-0	✓	01- -
0110	✓	0-11	✓	
1010	✓	01-1	✓	
0111	✓	011-	✓	
1110	✓	-110		
1111	✓	1-10		
		-111		
		111-		

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0--1
0100	✓	0-01	✓	0--1
0011	✓	010-	✓	01--
0101	✓	01-0	✓	01--
0110	✓	0-11	✓	-11-
1010	✓	01-1	✓	
0111	✓	011-	✓	
1110	✓	-110		
1111	✓	1-10		
		-111		
		111-	✓	

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0--1
0100	✓	0-01	✓	0--1
0011	✓	010-	✓	01--
0101	✓	01-0	✓	01--
0110	✓	0-11	✓	-11-
1010	✓	01-1	✓	-11-
0111	✓	011-	✓	
1110	✓	-110	✓	
1111	✓	1-10		
		-111	✓	
		111-	✓	

The prime implicants are thus:

0--1	01--	-11-	1-10
A'D	A'B	BC	ACD'

Now let's add some don't care conditions

$$\sum_{abcd}(1,3,4,5,6,7,10,14,15) + DC(2,9,11) \tag{10.2}$$

I will keep the old chart and just add in the new terms to save space.

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0--1
0010		0-01	✓	01--
0100	✓	010-	✓	-11-
0011	✓	01-0	✓	
0101	✓	0-11	✓	
0110	✓	01-1	✓	
1001		011-	✓	
1010	✓	-110	✓	
0111	✓	1-10		
1011		-111	✓	
1110	✓	111-	✓	
1111	✓			

Minterms		2-terms		4-terms
0001	✓	00-1	✓	0--1
0010		0-01	✓	01--
0100	✓	010-	✓	-11-
0011	✓	01-0	✓	
0101	✓	0-11	✓	
0110	✓	01-1	✓	
1001		011-	✓	
1010	✓	-110	✓	
0111	✓	1-10		
1011		-111	✓	
1110	✓	111-	✓	
1111	✓			

Chapter 11

Synchronous Circuits

SR latches and flip-flops are the fastest, as they are just the latch with possible clocking. Use them when you need high speed.

Characteristic

S	R	Q
0	0	q
0	1	0
1	0	1

Excitation

q	Q	S	R
0	0	0	x
0	1	1	0
1	0	0	1
1	1	x	0

D latches and flip-flops are primarily used in memory applications. The design process is simple because the simplicity of the excitation table.

Characteristic

D	Q
0	0
1	1

Excitation

Q	D
0	0
1	1

T latches and flip-flops are usually used for counters and dividers.

Characteristic

T	Q
0	q
1	q’

Excitation

q	Q	T
0	0	0
0	1	1
1	0	1
1	1	0

JK latches and flip-flops give such easy designs that they are preferred for most every design. Usually one of the others is only used in the special cases mentioned above.

Characteristic

J	K	Q
0	0	q
0	1	0
1	0	1
1	1	q’

Excitation

q	Q	J	K
0	0	0	x
0	1	1	x
1	0	x	1
1	1	x	0

Typically, the characteristic tables are used when doing analysis, and the excitation table is used for design.

11.1 Counters

11.2 General Design

Give the logic diagram using any flip-flops you want and a PAL for the state diagram in below.

Any undesignated states will go to 111/0, which will be our garbage state. You could also decide to send it to 000/1, but since this state machine looks for words of the pattern ((01*01*0)*(10*10*1)*)*[1], having the undesignated states go to 000/1 would violate the pattern. Additionally I will use D flip-flops. I am doing this at home, so I don't have the drawing program, so I will leave the equations for the PAL, the connection is straightforward.

[1]this is a regular expression that is equivalent to the state machine, regular expressions and their relation to FA/FSM is covered in formal languages and automata theory. This is included for your information and is thus not expected for you to know for the test.

Present State	Input	Next State	Output
000	0	010	1
	1	100	1
001	0	111	0
	1	111	0
010	0	011	0
	1	010	0
011	0	000	0
	1	011	0
100	0	100	0
	1	101	0
101	0	101	0
	1	000	0
110	0	111	0
	1	111	0
111	0	111	0
	1	111	0

Most Significant Bit (S2)

S0,In\S2,S1	00	01	11	10
00	0	0	1	1
01	1	0	1	1
11	1	0	1	0
10	1	0	1	1

I will use SOP on the zeros then complement (three encirclements)

$D2 = ((S2' \cdot S1) + (S2' \cdot S0' \cdot In')$

$+(S2 \cdot S1' \cdot S0 \cdot In))'$

Middle Bit (S1)

S0,In\S2,S1	00	01	11	10
00	1	1	1	0
01	0	1	1	0
11	1	1	1	0
10	1	0	1	0

I will again use SOP on the zeros then complement (three encirclements)

$D1 = ((S2 \cdot S1') + (S1' \cdot S0' \cdot In)$

$+(S2' \cdot S1 \cdot S0 \cdot In'))'$

Least Significant Bit (S0)

S0,In\S2,S1	00	01	11	10
00	0	1	1	0
01	0	0	1	1
11	1	1	1	0
10	1	0	1	1

I will again use SOP on the zeros then complement (five encirclements)

$D0 = ((S1' \cdot S0' \cdot In') + (S2' \cdot S1' \cdot S0')$

$+(S2' \cdot S0' \cdot In) + (S2' \cdot S1 \cdot S0 \cdot In')$

$+(S2 \cdot S1' \cdot S0 \cdot In))'$

Output

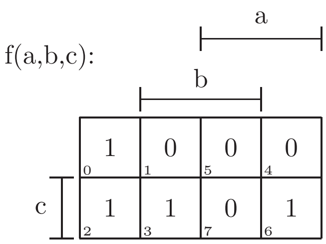
This can be read off the table trivially:

$Out = S2' \cdot S1' \cdot S0'$

Chapter 12

Timing

12.1 Combinational Circuits



12.2 Sequential Circuits

The timing on sequential circuits revolves around ensuring that the setup and hold times of a flip flop are met in the circuit. We will be using a bunch of different measurements of a circuit so we will begin by defining them.

Trigger The event which is used to start a sequential circuit, usually the rising or falling edge of a clock.

Setup time (T_s) The minimum time the inputs must be stable before a trigger so the correct value is latched. Failing to do so is a setup violation.

Hold time (T_h) The minimum time the inputs must be stable after a trigger so the correct value is latched. Failing to do so is a hold violation.

Clock period (T_{clk}) The time between successive rising (or falling) edges in the clock signal.

Clock skew (T_{skew}) The propagation time difference between furthest components, which thus is the time difference of them reading the same clock. You can think of it as the time error range.

Flip Flop Clock propagation ($T_{clk-xmit}$) The time from when a flip flop receives the trigger till when the data is transmitted from it. This is sometimes referred to as the time from clock to q.

Combinational Logic Delay (T_{comb}) Time for a signal to pass through the combinational circuit. Sometimes called **propagation delay**.

Now to ensure there is no problem in a sequential circuit, we must verify two conditions are met: the loop time in Eq. 12.1, and the arrival time in Eq. 12.2.

$$T_{clk} \geqslant T_s + T_{comb} + T_{clk-xmit} + T_{skew} \tag{12.1}$$
$$T_h \leqslant T_{comb} + T_{clk-xmit} + T_{skew} \tag{12.2}$$

Note that the loop time constrains the setup time, while the arrival time is a constraint on the hold time.

- In a new design, you use the arrival time equation to determine the flip flop to use, and the loop time equation to determine the clock.
- In an FPGA, you are stuck with the logic, flip flops, and the clock, so those parameters are fixed. The skew depends on position of the circuit elements (layout) is design dependent, so the equations are checked to verify a design. If the design does not meet the clock timing an excessive skew warning is issued.

12.3 Flip Flops and Hazards

In Table 12.1, I list the setup, hold, and the sum, which is the metastable interval or window.

12.4 How Often?

Since the primary failure mode for entering metastability is a data change during setup and hold, the smaller these times the better, which means faster logic families. The equations for calculating mean time between failures (MTBF) are

$$MTBF = \frac{e^{\frac{T_r}{T_\gamma}}}{F_d F_c T_p} \tag{12.3}$$
$$= \frac{e^{\frac{T_r + \frac{1}{F_c} - T_s}{T_\gamma}}}{F_d F_c^2 T_p^2} \tag{12.4}$$

F_d Data Frequency

F_c Clock Frequency

T_p Propagation delay of the flip flop

T_s Setup Time

T_r Resolve time (clock time minus the path time)

T_γ Resolution time of flip flop

Table 12.1: Interval when Metastability is most likely to occur

Device	T_s[ns]	T_h[ns]	T_{ms}
SN74LS74A	20	5	25
SN74ALS74A	15	0	15
SN74AS74A	4.5	0	4.5
SN74F74	3	1	4
CD74ACT74-Q1	4	0	4
SN54AHC74	5	0.5	5.5
SN54AHCT74	5	0	5
SN54LVC74A-SP	3	1	4
SN74AC74	3	0.5	3.5
SN74AC74-EP	3	0.5	3.5
SN74ACT74	3.5	1	4.5
SN74ACT74-EP	4	1	5
SN74AHC74	5	0.5	5.5
SN74AHC74-EP	5	0.5	5.5
SN74AHC74Q-Q1	5	0.5	5.5
SN74AHCT74	5	0	5
SN74AHCT74-EP	5	0	5
SN74AHCT74Q-Q1	5	0	5
SN74AUC74	0.7	0.3	1
SN74HC74	21	0	21
SN74HC74-EP	150	0	150
SN74HC74-Q1	17	0	17
SN74HCT74	14	0	14
SN74LV74A-EP	3	2.15	5.15
SN74LV74A-Q1	5	0.5	5.5
SN74LVC74A	3	0	3
SN74LVC74A-EP	3	1	4
SN74LVC74A-Q1	3	1	4
SN74S74	3	2	5

Part III

Data Representation and Manipulation

Chapter 13

Codes

Codes are used to represent members of a set by a sequence of symbols. For our purposes, the sequence of symbols will always be a sequence of $\{0, 1\}$. Codes have an encoding for each member to be represented. Codes can be fixed or variable in length. Fixed length codes like ascii have the same number of symbols in every encoding of the code. Variable length codes use different numbers of symbols to represent the encodings. For instance if '1' is 'a', '01' is 'b', and '00' is 'c', then the code is variable length. The major trouble with variable length codes is splitting the message up into the individual encodings. If the code is prefix (postfix) then the code can be directly read from left to right (right to left).

13.1 Standard Codes

13.1.1 Unsigned

decimal	Binary	Gray	BCD	2421	Residue(5,3)	Residue(7,2)
0	0000	0000	0000	0000	000,00	000,0
1	0001	0001	0001	0001	001,01	001,1
2	0010	0011	0010	0010	010,10	010,0
3	0011	0010	0011	0011	011,00	011,1
4	0100	0110	0100	0100	100,01	100,0
5	0101	0111	0101	1011	000,10	101,1
6	0110	0101	0110	1100	001,00	110,0
7	0111	0100	0111	1101	010,01	000,1
8	1000	1100	1000	1110	011,10	001,0
9	1001	1101	1001	1111	100,00	010,1
10	1010	1111			000,01	011,0
11	1011	1110			001,10	100,1
12	1100	1010			010,00	101,0
13	1101	1011			011,01	110,1
14	1110	1001			100,10	-
15	1111	1000			-	-

BCD is a decimal code designed to be compatible with standard binary numbers. It is sometimes called 8421 code due to the weights on the columns. The 2421 code was designed to be the same as BCD for 0-4 and make the 9's complement, which is important for easy subtraction, of 0-4 (i.e. 9-5 respectively) be easy to take because you can simply flip the bits.

Gray code is an alternate to binary. It is not a decimal code, and hence does not waste 6 codes for every four bits. Gray code was designed to have only one bit flip at any given time. This is helpful in systems which have analog components and need to count. For instance in an NC drill, we might want to encode the shaft position and hence put gray code bars on the shaft and have an ir sensor read them. Since only one bit flips between each consecutive number, it is easy to verify if we are reading correctly and thus get a good idea of how fast the shaft is spinning and where the shaft is. Gray code is also useful to us in Karnaugh maps and code maps because the one bit flipping property lets us find errors of type one easily (Karnaugh maps) and measure Hamming distance easily (code maps). Notice that the first bit of a gray code is just like binary (all 0's first then 1's), while the rest follow a 0110 pattern on reducing scales.

The easiest way to read grey code is to start from the left and just copy the first bit. From then on if the next digit to the right is 0 then repeat the last digit you wrote, if it is 1 flip the last digit you wrote.

Example 7 *What is the value of* 101111_{gray}*?*

Starting at the left copy the first bit:

Gray	*1*	*0*	*1*	*1*	*1*	*1*
Binary	*1*					

The next bit is a 0 so repeat the last bit you wrote (in this case a 1):

Gray	*1*	*0*	*1*	*1*	*1*	*1*
Binary	*1*	*1*				

The next bit is a 1 so flip the last bit you wrote (in this case 1 flips to 0):

Gray	*1*	*0*	*1*	*1*	*1*	*1*
Binary	*1*	*1*	*0*			

The next bit is a 1 so flip the last bit you wrote (in this case 0 flips to 1):

Gray	*1*	*0*	*1*	*1*	*1*	*1*
Binary	*1*	*1*	*0*	*1*		

The next bit is a 1 so flip the last bit you wrote (in this case 1 flips to 0):

Gray	*1*	*0*	*1*	*1*	*1*	*1*
Binary	*1*	*1*	*0*	*1*	*0*	

The next bit is a 1 so flip the last bit you wrote (in this case 0 flips to 1):

Gray	*1*	*0*	*1*	*1*	*1*	*1*
Binary	*1*	*1*	*0*	*1*	*0*	*1*

Binary 110101 is 53, so gray 101111 is 53.

Residue number systems (residue codes) are fun though rarely used because of the difficulty in converting back from them to binary. Residue codes are specified by a series of remainders, taken to relatively prime bases (listed parenthesis and separated by commas). The remainders are in the same order as the specified bases and also separated by commas. The advantage of this system is you can perform fast addition, multiplication, and subtraction (if the divisor is not zero in any of the residues you can also do division efficiently), extremely fast, as the modulo terms are independently calculated by the modulo of the arithmetic operation being performed.

Example 8 *Calculate* $7+3$*,* $3*4$*,* $14-8$*, and* $14/7$ *in Modulo(5,3). Note we can do division because* $7 \mod 5 = 2 > 0$ *and* $7 \mod 3 = 1 > 0$*.*

$7+3 = (010, 01) + (011, 00) = (010 + 011 \mod 5, 01 + 00 \mod 3) = (000, 01) = 10$

$3*4 = (011, 00) * (100, 01) = (011 * 100 \mod 5, 00 * 01 \mod 3) = (010, 00) = 12$

$14-8 = (100, 10) - (011, 10) = (100 - 011 \mod 5, 10 - 10 \mod 3) = (001, 00) = 6$

$14/7 = (100, 10) - (010, 01) = (100/010 \mod 5, 10/01 \mod 3) = (010, 10) = 2$

13.1.2 Signed

decimal	Signed Binary	1's Comp	2's Comp	Excess-7	Excess 8
8	-	-	-	1111	-
7	0111	0111	0111	1110	1111
6	0110	0110	0110	1101	1110
5	0101	0101	0101	1100	1101
4	0100	0100	0100	1011	1100
3	0011	0011	0011	1010	1011
2	0010	0010	0010	1001	1010
1	0001	0001	0001	1000	1001
0	0000,1000	0000,1111	0000	0111	1000
-1	1001	1110	1111	0110	0111
-2	1010	1101	1110	0101	0110
-3	1011	1100	1101	0100	0101
-4	1100	1011	1100	0011	0100
-5	1101	1010	1011	0010	0011
-6	1110	1001	1010	0001	0010
-7	1111	1000	1001	0000	0001
-8	-	-	1000	-	0000

Note that both signed binary and 1's compliment have a positive and negative 0. Signed binary was an early development, but is not that useful because you can't use a standard adder/subtractor.

1's compliment is easy to calculate (flip the bits to convert from positive to negative), and is useful in turning an adder into a subtractor (the number to be subtracted is turned into the 2's complement, by finding the 1's complement, then setting the carry-in bit of the adder to do the +1).

2's compliment is the standard form for storing negative numbers in computers because you can easily convert (either by flipping bits and adding 1, or by starting on the right and copying bits up to and including the first 1, then flipping the remaining bits), and standard adder/subtractor circuits can be used.

Excess codes are most commonly used in floating point number exponents, as they preserve the numeric order of greatness (you can use standard compare circuits to check size). The excess is either half the total numbers ($16/2 = 8$ for excess 8) or half the total numbers minus 1 ($16/2 - 1 = 7$ for excess 7).

13.2 Huffman Codes

Huffman codes are variable length codes that produce optimal expected code lengths.

$$ecl = \sum_{l \in C} (freq(l) \times length(l))$$

Example:

Consider the string "adabaabcaabacadaccac" that we want to encode. There are four members of the set (a, b, c, d) which means the members can be represented by a two bit fixed code. But consider the following encoding (a=1, b=001, c=01, d=000). The frequencies of the members are (a=10/20=.5, b= 3/20=.15, c=5/20=.25, d =2/20=.1). The ecl of the variable code is

$$\begin{aligned} ecl &= .5 * 1 + .15 * 3 + .25 * 2 + .1 * 2 \\ &= 1.65 \end{aligned}$$

The expected code length is only 1.65 bits/character.

13.2.1 Huffman Algorithm

1. Calculate the frequencies of each member

$$\frac{\#\, occurrences\, of\, member}{Total\, occurrences}$$

2. Form decode tree from forest

   (a) make 1 node tree for each member with frequency and member name

   (b) join two trees with the smallest frequency on root node by making them branches of a new root node and giving the new root node the sum of the frequencies of the old root nodes

   (c) put new tree in forest and repeat joining till only one tree remains (the answer)

3. encode or decode message

13.3 Error Detection and Correction

Errors can happen in a variety of ways. Bits can be added, deleted, or flipped. Errors can happen in fixed or variable codes. For simplicity we will consider only bit flips in fixed codes. Note that variable codes can be packed into fixed length blocks for transmission and storage, so this is not as restrictive as it might sound at first.

The Hamming distance (d_H) between two codewords is the number of bit flips to turn one codeword into the other codeword. It can also be thought of as the number of bits that are different between two codewords. The Hamming distance can be extended to a set, by defining it as the minimum distance between any two codewords in the set. The Hamming distance is useful in codes because it tells us how many errors can be detected (E_d) and how many errors can be corrected (E_c) The relations are given by

$$\begin{aligned} d_H &\geqslant 1 + E_d \\ d_H &\geqslant 1 + 2 \times E_c \end{aligned}$$

Example

Consider the codes $(00001, 01100)$.

1. What is the Hamming distance?

   3

2. How many errors can be detected? How many can be corrected?

   $3 \geqslant 1 + d$ thus detect 2

   and

   $3 \geqslant 1 + 2c$ thus correct 1

3. It is desired to add another codeword without reducing the Hamming distance. What codeword do you suggest?

   any of the following will work:

   - 10010
   - 10110

- 10111
- 11010
- 11011
- 11111

13.3.1 Hamming Code

To detect and/or correct errors, two pieces of information must be sent, the original data (D_i) and check bits (C_j). Consider numbering in binary each position in an array of bits to be sent starting at 1, and positioning the check bits at the powers of two.

Address	0	0	0	0	0	0	0	1	1	1
	0	0	0	1	1	1	1	0	0	0
	0	1	1	0	0	1	1	0	0	1
	1	0	1	0	1	0	1	0	1	0
Code	C_0	C_1	D_1	C_2	D_2	D_3	D_4	C_3	D_5	D_6

The check bits are then calculated by taking the exclusive-or (xor) of all the data bits (D_i), whose address contains a 1 in the same place as the check bit. Thus,

Address	0	0	0	0	0	0	0	1	1	1
	0	0	0	1	1	1	1	0	0	0
	0	1	1	0	0	1	1	0	0	1
	1	0	1	0	1	0	1	0	1	0
Code	C_0	C_1	D_1	C_2	D_2	D_3	D_4	C_3	D_5	D_6

$$C_0 = D_1 \oplus D_2 \oplus D_4 \oplus D_5$$

Address	0	0	0	0	0	0	0	1	1	1
	0	0	0	1	1	1	1	0	0	0
	0	1	1	0	0	1	1	0	0	1
	1	0	1	0	1	0	1	0	1	0
Code	C_0	C_1	D_1	C_2	D_2	D_3	D_4	C_3	D_5	D_6

$$C_1 = D_1 \oplus D_3 \oplus D_4 \oplus D_6$$

And so on.

The Hamming distance is three, which will be proved in three cases.

1. If the data portion of two codewords differs by only one bit, then note that the address of each data bit has at least two ones in it. This means that the data bit that is different will cause at least two check bits to be different, yielding a Hamming distance of three.

2. If the data portion of two codewords differs by two bits, then note that no two data bits affect all the same check bits. Thus, there exists at least one check bit that is affected by only one of the two data bits that differs, and will thus be different between the two codewords, yielding a Hamming distance of three.

3. If the data portion of two codewords differs by more than two bits the result is trivial.

Q.E.D.

A Hamming distance of three means

$$\begin{aligned} 3 &\geqslant 1 + E_d \\ 2 &\geqslant E_d \\ 3 &\geqslant 1 + 2 \times E_c \\ 2 &\geqslant 2 \times E_c \\ 1 &\geqslant E_c. \end{aligned}$$

One error can be corrected or two detected. To find the error for correction you create its address by taking the exclusive-or of the check bits and the data that created them. A 1 will result only if an odd number of errors happened in the subset checked. The address that results is the address of the error, which is fixed by toggling.

Example

the data "1010" is to be sent by Hamming Code. Since there are only four bits of data, only three check bits are needed. The data is put in place.

Address	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
Code	C_0	C_1	1	C_2	0	1	1

Next the check bits are calculated and

$$\begin{aligned} C_0 &= D_1 \oplus D_2 \oplus D_4 \\ &= 1 \oplus 0 \oplus 1 \\ &= 0 \\ C_1 &= D_1 \oplus D_3 \oplus D_4 \\ &= 1 \oplus 1 \oplus 1 \\ &= 1 \\ C_2 &= D_2 \oplus D_3 \oplus D_4 \\ &= 0 \oplus 1 \oplus 1 \\ &= 0 \end{aligned}$$

Thus,

Address	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
Code	0	1	1	0	0	1	1

Now, assume an error happens. It could be anywhere, but for this example assume that the bit in position 6 is toggled.

Address	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
Code	0	1	1	0	0	0	1

To find it get the address by

$$\begin{aligned}
A_0 &= C_0 \oplus D_1 \oplus D_2 \oplus D_4 \\
&= 0 \oplus 1 \oplus 0 \oplus 1 \\
&= 0, \\
A_1 &= C_1 \oplus D_1 \oplus D_3 \oplus D_4 \\
&= 1 \oplus 1 \oplus 0 \oplus 1 \\
&= 1, \\
A_2 &= C_2 \oplus D_2 \oplus D_3 \oplus D_4 \\
&= 0 \oplus 0 \oplus 0 \oplus 1 \\
&= 1.
\end{aligned}$$

Yielding the address, $A_2A_1A_0 = 110 = 6$, which is the error.

Example: Hello There

Compress "hello there" using a Huffman code designed off it. Then use a Hamming code on 11 bit blocks of the compressed message. How does the overall message size compare to the original? I will just list the code, the tree is obvious from it. Note that other trees are possible.

letter	frequency	code
h	$\frac{2}{11}$	100
e	$\frac{3}{11}$	11
l	$\frac{2}{11}$	101
o	$\frac{1}{11}$	011
sp	$\frac{1}{11}$	010
t	$\frac{1}{11}$	001
r	$\frac{1}{11}$	000

Huffman code: 10011101101 01101000110 01100011

Hamming Code

Since I don't have enough bits to do 3 groups of 11, I could pad with 0's or 1's or I could make the last packet shorter. Alternately I could have made an EOF code in my Huffman code. In this case I will just skip them so you see how that works. You should mention the problem and what you will do along with the solution.

Data Section	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
First	c_0	c_1	1	c_2	0	0	1	c_3	1	1	0	1	1	0	1
Second	c_0	c_1	0	c_2	1	1	0	c_3	1	0	0	0	1	1	0
Third	c_0	c_1	0	c_2	1	1	0	c_3	0	0	1	1			

Data Section	1	2	**3**	4	**5**	6	**7**	8	**9**	10	**11**	12	**13**	14	**15**
First	1	c_1	1	c_2	0	0	1	c_3	1	1	0	1	1	0	1
Second	1	c_1	0	c_2	1	1	0	c_3	1	0	0	0	1	1	0
Third	0	c_1	0	c_2	1	1	0	c_3	0	0	1	1			

Data Section	1	2	**3**	4	5	**6**	**7**	8	9	**10**	**11**	12	13	**14**	**15**
First	1	0	1	c_2	0	0	1	c_3	1	1	0	1	1	0	1
Second	1	0	0	c_2	1	1	0	c_3	1	0	0	0	1	1	0
Third	0	0	0	c_2	1	1	0	c_3	0	0	1	1			

Data Section	1	2	3	4	**5**	**6**	**7**	8	9	10	11	**12**	**13**	**14**	**15**
First	1	0	1	0	0	0	1	c_3	1	1	0	1	1	0	1
Second	1	0	0	0	1	1	0	c_3	1	0	0	0	1	1	0
Third	0	0	0	1	1	1	0	c_3	0	0	1	1			

Data Section	1	2	3	4	5	6	7	8	**9**	**10**	**11**	**12**	**13**	**14**	**15**
First	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1
Second	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0
Third	0	0	0	1	1	1	0	0	0	0	1	1			

The length is thus 42 bits for the compressed code with error correction. The original message was 11 chars $\times$ 7 bits/char $=$ 77 bits. The new message is much smaller (less than 4/7).

Chapter 14

Integers

14.1 Integer numbers

unsigned All the bits are used for the magnitude of the number. (0 to $2^n - 1$)

signed int The first bit indicates the sign (1 is negative), the remaining $n-1$ bits are used for magnitude. ($-2^{n-1} + 1$ to $2^{n-1} - 1$)

1's complement Positive numbers are the same as signed int, but negative are found by inverting each bit of the positive number with the same magnitude. ($-2^{n-1} + 1$ to $2^{n-1} - 1$)

2's complement As 1's complement, but negative numbers have 1 added to them after the bitwise inversion. This removes a -0 code, so the extra code is assigned to -2^{n-1}. This is the natural way to handle numbers if addition and subtraction of mixed sign numbers are needed. (-2^{n-1} to $2^{n-1} - 1$)

2^{n-1} **excess** The code is found by adding 2^{n-1} to the value (hence the name). This gives a slightly larger negative range. (-2^{n-1} to $2^{n-1} - 1$)

$2^{n-1} - 1$ **excess** The code is found by adding $2^{n-1} - 1$ to the value (hence the name). This gives a slightly larger positive range. ($-2^{n-1} + 1$ to 2^{n-1})

Example 9 *Convert the following*

1. *-39 to 8 bit 2's complement*

39	
19	1
9	1
4	1
2	0
1	0
0	1

$39_{10} = 100111_2 = 00100111_2$

$-39_{10} = 11011001_2$

2. *234 to 8 bit unsigned*

234	
117	0
58	1
29	0
14	1
7	0
3	1
1	1
0	1

$234_{10} = 11101010_2$

14.2 Addition

The basic addition routines can be modified to work for any of the codes as well as subtraction for the codes. The special customizations will be considered later. Right now, the typical techniques for addition are considered.

Example 10 *Calculate the following in binary using 8 bits.*

1. $42 - 51$
2. $51 - 42$

Sol:

	42	51
+	00101010	00110011
-	11010110	11001101

42	00101010	51	00110011
-51	11001101	-42	11010110
	11110111		100001001
-9	-00001001	9	00001001

14.2.1 Ripple Adders

This is the technique that is covered in CSCI 310. Basically, full bit adders, see Figure 14.1, are created and cascaded together. The carry bit from the previous full adder must arrive before the result is added. The resulting valid carries thus ripple down to the most significant bit (hence the name). Adding n bit numbers, thus takes the propagation time of $n + 1$ levels of logic, i.e. it is O(n) in time to calculate addition. Thus if 32 bit numbers are added on fast logic (1ns per stage/gate) the process would take 33ns. This is way too slow. On the bright side, none of the gates take more than 2 inputs so the size of the gates is O(1).

14.2.2 Conditional Sum

Conditional sum is a divide and conquer algorithm, and hence exploits binary tree parallelism. The algorithm works by calculating both possible results for each bit (if carry in was 1 or 0), then performing paired conditional concatenation using the actual carry bit of the lower number, see Figure 14.2.

1. form conditional terms for each digit in summation → (digit with carry, digit without carry) = $(x_i + y_i + 1, x_i + y_i)$
2. group by twos from right and for both conditional values in the right parenthesis form the result as follows:

Figure 14.1: (left) Half Adder, (right) Full Adder

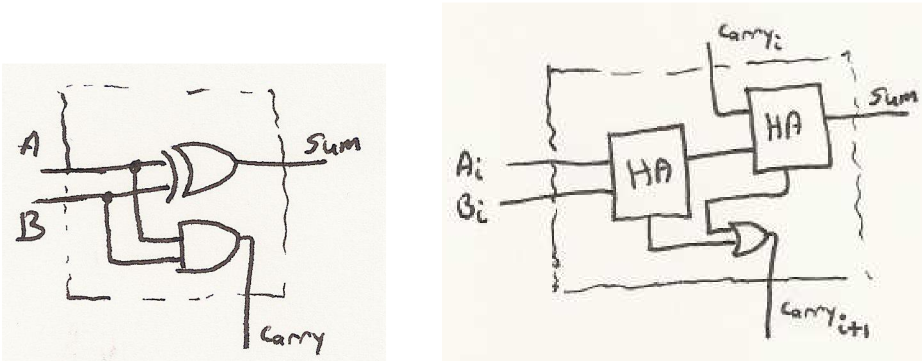
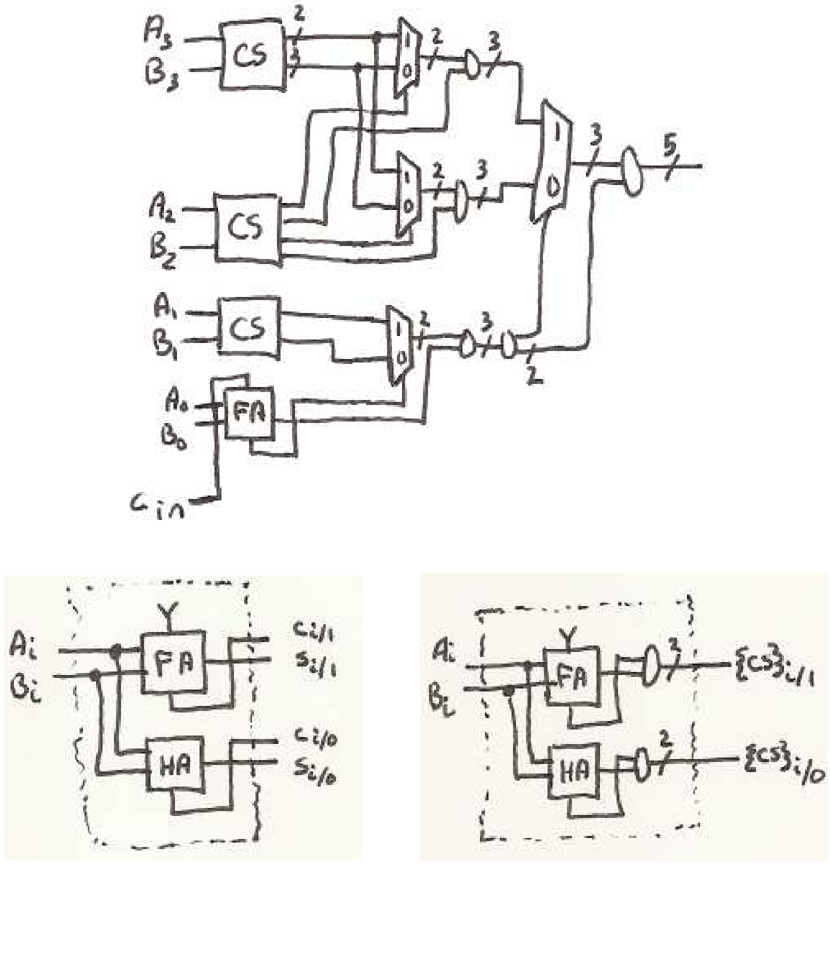


Figure 14.2: Conditional Sum Adder (above), and its sub-blocks (below, left and right).



(a) the leftmost bit of the two terms on the right are the carry bits used to select the term on the left

(b) concatenate the appropriate term on the left (picked by carry) with each term on right after removing the parity bits of the right terms

3. continue pairings until only 1 term remains. pick right number if $c_{in} = 0$ else pick left.

Example 11 *Add $x = 0110$ and $y = 1111$ by conditional sum and indicate if overflow occurred.*

0+1 1+1 1+1 0+1
↓ ↓ ↓ ↓
(10,01) (11,10) (11,10) (10,01)
↘ ↙ ↘ ↙
(101,100 (110,101)
↘ ↙
(10110,10101)
1 0101

No overflow occurred (added a positive and negative number).

Example 12 *Calculate $7 - 8$ by conditional sum.*

$7 = 0111$ *and* $-8 = 1000$

0 1 1 1
+1 0 0 0
(10,01) (10,01) (10,01) (10,01)
↘ ↙ ↘ ↙
(100,011) (100,011)
↘ ↙
(10000,01111)

Since this was done as addition no carry-in was set so the solution is 0 | 1111 *or* -1 *in signed base ten.*

Example 13 *Add by conditional sum $x = 01100110$ and $y = 00110011$.*

$0+0$ $1+0$ $1+1$ $0+1$ $0+0$ $1+0$ $1+1$ $0+1$
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
(01, 00) (10, 01) (11, 10) (10, 01) (01, 00) (10, 01) (11, 10) (10, 01)
↘ ↓ ↘ ↓ ↘ ↓ ↘ ↓
(010, 001) (110, 101) (010, 001) (110, 101)
↘ ↙ ↘ ↙
(01010, 01001) (01010, 01001)
↘ ↙
(010011010, 010011001)
0 10011001

Why go through this? First, by a folk theorem of Dr. Alan Laub, "*What is hard for us tends to be easy for computers (and vice versa).*" In reality this process is really easy for a computer to do. Second, the process is highly parallel, so it can be done very fast. If the numbers to be added are n bits long this takes $2(\log_2(n) + 1)$ levels of logic, much better than the $n + 1$ levels of logic required by ripple calculations. Thus it is O(log(n)) in time complexity. For example, for adding the 32 bit numbers considered already, conditional sum would take $2(\log_2(32) + 1) = 12$ levels of logic, so on the fast logic described it would be 12ns, a huge improvement.

14.2.3 Carry-Lookahead

This is also referred to as lookahead carry. Assume $x + y = z$. Pre-generate all carries with 2-level logic. Usually form (g,p,c) generate, propagate, carry.

$$\begin{aligned}
G_i &= x_i \cdot y_i \\
P_i &= x_i + y_i \\
C_i &= G_i + P_i \cdot C_{i-1} \\
&= G_i + P_i \cdot (G_{i-1} + P_{i-1} \cdot C_{i-2}) \\
&= G_i + P_i \cdot G_{i-1} + P_i \cdot P_{i-1} \cdot C_{i-2} \\
&= G_i + P_i \cdot G_{i-1} + P_i \cdot P_{i-1} \cdot G_{i-2} + \ldots + P_i \cdot P_{i-1} \cdot \ldots \cdot P_0 \cdot C_{in}
\end{aligned}$$

This method is very fast (regardless of size it take 5 levels of logic) but requires large gates for problems of reasonable size (even 16 or 32 bit numbers) and thus has problems with fan-in, fan-out, and size.

Often blocks of a number are handled with lookahead, and the blocks are connected in some fashion (for example ripple) to get the net result (i.e. just like single bit adds from a full adder are connected to propagate the carry bit, blocks or 4, 8, or more could be handled lookahead then connected to propagate the carry bit between them to handle a larger number, say 32 bits). Even better than cascading (ripple connection) the adders, is to us group carry-lookahead, in which each of the carry-lookahead adders output their group propagate and group generate variables to a circuit that generates the carry-in bits for each group. It takes 5 logic levels to generate the carries to each individual carry-lookahead adder, and each adder then takes 5 levels of logic to get the result, for a total of 10 levels of logic. For the example of adding 32 bit numbers with fast logic, it would take 10ns with group carry-lookahead adders (probably four or eight bits in a group).

Example 14 *Specify the equations of a two bit binary adder with carry in (i.e.: one equation for each of the sum bits and one equation for the carry out). Put the equations in sum of products form.*

Sol: Let the two numbers to be added be A_1A_0 and B_1B_0. Let the resulting sum be S_1S_0. Let the carries be C_{in} and C_{out}. Finally, let C_0 be the carry from the first bit added (saves writing).

$$\begin{aligned}
S_0 &= A_0 \oplus B_0 \oplus C_{in} \\
C_0 &= A_0 \cdot B_0 + A_0 \cdot C_{in} + B_0 \cdot C_{in} \\
S_1 &= A_1 \oplus B_1 \oplus C_0 \\
C_{out} &= A_1 \cdot B_1 + A_1 \cdot C_0 + B_1 \cdot C_0
\end{aligned}$$

Putting this in sum of products form yields

$$
\begin{aligned}
S_0 &= A_0' \cdot B_0' \cdot C_{in} + A_0' \cdot B_0 \cdot C_{in}' + A_0 \cdot B_0' \cdot C_{in}' + A_0 \cdot B_0 \cdot C_{in} \\
S_1 &= A_1' \cdot B_1' \cdot (A_0 \cdot B_0 + A_0 \cdot C_{in} + B_0 \cdot C_{in}) + A_1' \cdot B_1 \cdot (A_0 \cdot B_0 + A_0 \cdot C_{in} + B_0 \cdot C_{in})' + \\
&\quad A_1 \cdot B_1' \cdot (A_0 \cdot B_0 + A_0 \cdot C_{in} + B_0 \cdot C_{in})' + A_1 \cdot B_1 \cdot (A_0 \cdot B_0 + A_0 \cdot C_{in} + B_0 \cdot C_{in}) \\
&= A_1' \cdot B_1' \cdot A_0 \cdot B_0 + A_1' \cdot B_1' \cdot A_0 \cdot C_{in} + A_1' \cdot B_1' \cdot B_0 \cdot C_{in} \\
&\quad + A_1' \cdot B_1 \cdot (A_0' \cdot B_0' + A_0' \cdot C_{in}' + B_0' \cdot C_{in}') \\
&\quad + A_1 \cdot B_1' \cdot (A_0' \cdot B_0' + A_0' \cdot C_{in}' + B_0' \cdot C_{in}') \\
&\quad + A_1 \cdot B_1 \cdot A_0 \cdot B_0 + A_1 \cdot B_1 \cdot A_0 \cdot C_{in} + A_1 \cdot B_1 \cdot B_0 \cdot C_{in} \\
&= A_1' \cdot B_1' \cdot A_0 \cdot B_0 + A_1' \cdot B_1' \cdot A_0 \cdot C_{in} + A_1' \cdot B_1' \cdot B_0 \cdot C_{in} \\
&\quad + A_1' \cdot B_1 \cdot A_0' \cdot B_0' + A_1' \cdot B_1 \cdot A_0' \cdot C_{in}' + A_1' \cdot B_1 \cdot B_0' \cdot C_{in}' \\
&\quad + A_1 \cdot B_1' \cdot A_0' \cdot B_0' + A_1 \cdot B_1' \cdot A_0' \cdot C_{in}' + A_1 \cdot B_1' \cdot B_0' \cdot C_{in}' \\
&\quad + A_1 \cdot B_1 \cdot A_0 \cdot B_0 + A_1 \cdot B_1 \cdot A_0 \cdot C_{in} + A_1 \cdot B_1 \cdot B_0 \cdot C_{in} \\
C_{out} &= A_1 \cdot B_1 + A_1 \cdot (A_0 \cdot B_0 + A_0 \cdot C_{in} + B_0 \cdot C_{in}) \\
&\quad + B_1 \cdot (A_0 \cdot B_0 + A_0 \cdot C_{in} + B_0 \cdot C_{in}) \\
&= A_1 \cdot B_1 + A_1 \cdot A_0 \cdot B_0 + A_1 \cdot A_0 \cdot C_{in} + A_1 \cdot B_0 \cdot C_{in} \\
&\quad + B_1 \cdot A_0 \cdot B_0 + B_1 \cdot A_0 \cdot C_{in} + B_1 \cdot B_0 \cdot C_{in}
\end{aligned}
$$

14.2.4 Other notes

Integer numbers larger than the word size of the computer can be handled by chaining. Two special assembly commands are often available to aid in chaining: addc, subb. Normally when you add the first carry in is zero, but for blocks of bits after the first block, the lower block might need to carry up. Addc uses the carry bit as c_{in} rather than assuming $c_{in} = 0$.

Two different signals are used to warn that the integer result might not be valid[1] : carry (c) and overflow (v). Carry is used for unsigned integers, and overflow is used for two's complement. Since both carry and overflow bits are both calculated at the same time[2] it is important to know what they mean, when they are relevant, and how they are calculated.

Overflow set if last two carries are different.

14.2.5 Signed Int

Addition

- if signs are same then add two $n-1$ digit numbers and keep sign
- else flip sign of second term and subtract (subtracting with same signs).

Subtraction ($S_1 - S_2$)

- if $S_1 \geqslant S_2 \geqslant 0$ or $S_1 \leqslant S_2 < 0$ then preserve sign and subtract absolute magnitudes,

[1] Overflow and carry are two of the typical condition codes. It is possible for a condition code to be set but the result is still valid. For instance carry could be set and overflow could be unset after an operation with 2's complement numbers. In this case the number is still valid since overflow is the signal for 2's complement.

[2] On some machines every arithmetic operation generates the condition codes, on other machines, like the SPARC, the condition codes are set only when special versions of the arithmetic commands that end in *cc* are used.

- if $S_2 > S_1 \geqslant 0$ or $S_2 < S_1 < 0$ then flip sign and subtract absolute magnitudes reversed,
- else flip sign of second term and add (adding with same signs).

14.2.6 2's Comp

For addition you just add the numbers normally with $c_{in} = 0$(no special cases).

For subtraction you take the 1's complement of the second number and add with $c_{in} = 1$(no special cases, note 1's complement +1 is 2's complement).

14.2.7 Excess

For addition, you need to carry extra bits while calculating, because you have to subtract the excess number after adding. This is needed because the excess was in each of the numbers added, so an extra excess is present which must be removed.

For subtraction, the excess gets removed in the process so it must be added back in after subtraction. Note the subtraction can result in an intermediate negative number, so extra bits are needed during calculation.

14.3 Multiplication

14.3.1 unsigned

Algorithm 1

1. set v to 0
2. for each digit do:
   (a) if lsb of x is 1, add y to v
   (b) left shift y
   (c) right shift x

This basically only handles numbers whose product fits in 1 register. In general multiplication could take up to 2 registers.

Algorithm 2

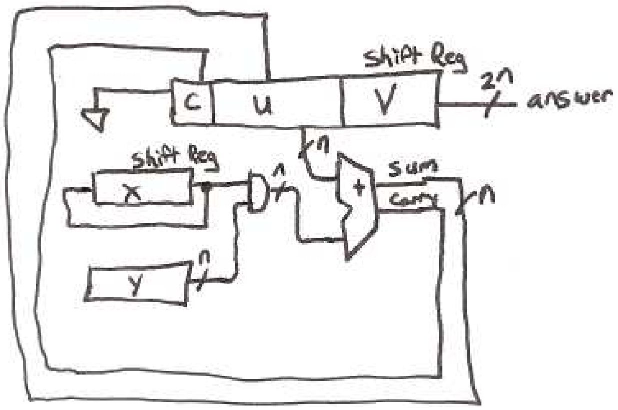
1. group two regs (u,v) for product, set to 0
2. for each digit do:
   (a) add (y and lsb(x)) to u hold carry in c
   (b) right shift (c,u,v)
   (c) circulant right shift x

Right shifting the product with carry is the same as left shifting (y_{hi},y), but without the need for a second register to hold the high order bits. The algorithm can be implemented in a circuit as is done in Figure 14.3.

Example 15 *Multiply 10 and 12 in binary using algorithm 2*

First we need to convert our numbers to binary: $x = 10_{10} = 1010_2$ and $y = 12_{10} = 1100_2$.

Figure 14.3: Unsigned Multiplier of Algorithm 2



c	u	v	x	*Comments*
0	*0000*	*0000*	*1010*	*Setup (Step 1)*
				Round 1
0	*0000*			*Step 2a: add $y \cdot 0$ to u (0+0=0)*
0	*0000*	*0000*		*Step 2b: rotate right cuv*
			0101	*Step 2c: circulant right shift x*
0	*0000*	*0000*	*0101*	*End of round 1*
				Round 2
0	*1100*			*Step 2a: add $y \cdot 1$ to u (0+12=12)*
0	*0110*	*0000*		*Step 2b: rotate right cuv*
			1010	*Step 2c: circulant right shift x*
0	*0110*	*0000*	*1010*	*End of round 2*
				Round 3
0	*0110*			*Step 2a: add $y \cdot 0$ to u (6+0=6)*
0	*0011*	*0000*		*Step 2b: rotate right cuv*
			0101	*Step 2c: circulant right shift x*
0	*0011*	*0000*	*0101*	*End of round 3*
				Round 4
0	*1111*			*Step 2a: add $y \cdot 1$ to u (3+12=15)*
0	*0111*	*1000*		*Step 2b: rotate right cuv*
			1010	*Step 2c: circulant right shift x*
0	*0111*	*1000*	*1010*	*End of round 4*

Note x is returned to its original value and $uv = 01111000_2 = 120_{10}$.

Example 16 *Multiply 14 and 7 in binary using algorithm 2*

First we need to convert our numbers to binary: $x = 14_{10} = 1110_2$ and $y = 7_{10} = 0111_2$.

c	u	v	x	*Comments*
0	*0000*	*0000*	*1110*	*Setup (Step 1)*
				Round 1
0	*0000*			*Step 2a: add* $y \cdot 0$ *to* u *(0+0=0)*
0	*0000*	*0000*		*Step 2b: rotate right cuv*
			0111	*Step 2c: circulant right shift* x
0	*0000*	*0000*	*0111*	*End of round 1*
				Round 2
0	*0111*			*Step 2a: add* $y \cdot 1$ *to* u *(0+7=7)*
0	*0011*	*1000*		*Step 2b: rotate right cuv*
			1011	*Step 2c: circulant right shift* x
0	*0011*	*1000*	*1011*	*End of round 2*
				Round 3
0	*1010*			*Step 2a: add* $y \cdot 1$ *to* u *(3+7=10)*
0	*0101*	*0100*		*Step 2b: rotate right cuv*
			1101	*Step 2c: circulant right shift* x
0	*0101*	*0100*	*1101*	*End of round 3*
				Round 4
0	*1100*			*Step 2a: add* $y \cdot 1$ *to* u *(5+7=12)*
0	*0110*	*0010*		*Step 2b: rotate right cuv*
			1110	*Step 2c: circulant right shift* x
0	*0110*	*0010*	*1110*	*End of round 4*

Note x *is returned to its original value and* $uv = 01100100_2 = 98_{10}$.

14.3.2 2's complement

Booth's Algorithm

1. group two regs (u,v) for product, set to 0
2. set x_{-1} to 0 (this is a single bit)
3. for each digit do:
   (a) if (lsb of x is 1,) and (x_{-1}=0), subtract y from u
   (b) if (lsb of x is 0) and (x_{-1}=1), add y to u
   (c) arithmetic right shift (u,v)
   (d) circular right shift x

Booth's algorithm can be implemented in a circuit as is done in Figure 14.4.

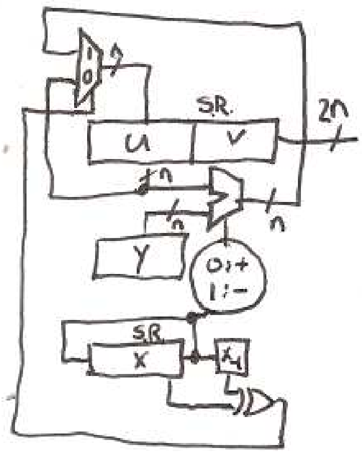
Example 17 *Multiply* 6 *(*$x = 0110$*) and* -1 *(*$y = 1111$*) using Booth's algorithm. Show the values at each stage in a table.*

Booth's

u	v	x	x_{-1}
0000	*0000*	*1111*	*0*
1010	*0000*		
1101	*0000*	*1111*	*1*
1110	*1000*	*1111*	*1*
1111	*0100*	*1111*	*1*
1111	*1010*	*1111*	*1*

Note the answer is 11111010, which is -6 *in 2's complement.*

Figure 14.4: Booth's Algorithm



Example 18 *Multiply -3 and 5 using Booth's algorithm and 4 bit numbers. Perform the indicated calculations showing all steps.*

$y = 5 = 0101$

$-y = -5 = 1011$

u	v	x	x_{-1}
0000	0000	1101	0
1011			
1011	0000	1101	0
1101	1000	1110	1
0101			
0010	1000	1110	1
0001	0100	0111	0
1011			
1100	0100	0111	0
1110	0010	1011	1
1111	0001	1101	1

The result is 11110001, which is -15 *in 2's complement.*

Example 19 *Multiply -3 and -6 using Booth's algorithm and 4 bit numbers. Perform the indicated calculations showing all steps.*

$x = -3 = 1101$, $y = -6 = 1010$ *and* $-y = 6 = 0110$.

U	V	X	X_{-1}
0000	0000	1101	0
0110	0000	1101	0
0011	0000	1110	1
1101	0000	1110	1
1110	1000	0111	0
0100	1000	0111	0
0010	0100	1011	1
0001	0010	1101	1

$00010010 = 18$

14.3.3 Systolic Array

The preceding algorithms are $O(n^2)$ if implemented with ripple adders, $O(n\log(n))$ if implemented with conditional sum adders, or $O(n)$ if implemented with look-ahead adders. The look-ahead adders have a large constant, so the $O(n)$ is not a perfect indicator of performance, and they are currently not practical beyond about 8 bits. It would be nice to find a way to multiply that has $O(n)$ and a small constant multiplier. Systolic arrays are $O(n)$, and have a constant multiplier of about 6 depending on your hardware, which is about half what it takes with even block (group) carry look-ahead adders using serial routines.

14.4 Integrated Examples

Example 20 *Calculate the following expression in binary using 2's complement and 8 bits total. Show all work.*

$$(9 * 9 - 24)/3$$

Sol:

$9_{10} = 00001001_2$ *and* $3_{10} = 00000011_2$

$$\begin{array}{r|l} 24 & \\ \hline 12 & 0 \\ 6 & 0 \\ 3 & 0 \\ 1 & 1 \\ 0 & 1 \end{array}$$

$24_{10} = 00011000_2$ *thus* $-24_{10} = 11101000_2$. *Thus* $9 * 9$,

$$\begin{array}{ccccccccccc}
 & & & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
 & & & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ \hline
 & & & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & & & \\ \hline
 & & & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1
\end{array}$$

Then (subtracting 24),

$$\begin{array}{ccccccccc}
 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\
 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ \hline
1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\
 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1
\end{array}$$

Now perform the division:

$$\begin{array}{cc|cccccc}
\multicolumn{2}{c}{} & & 1 & 0 & 0 & 1 & 1 \\ \cline{3-8}
1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\
 & & 1 & 1 & & & & \\ \cline{3-4}
 & & & 0 & 1 & 0 & 0 & \\
 & & & & & 1 & 1 & \\ \cline{6-7}
 & & & & & & 1 & 1 \\
 & & & & & & 1 & 1 \\ \cline{7-8}
 & & & & & & & 0
\end{array}$$

The answer is thus $00010011_2 = 19_{10}$.

14.5 Residue Arithmetic

We have shown different ways of calculating the sum and product of binary numbers. In this section we will examine a different way to represent numbers and thus to calculate. In residue arithmetic numbers are

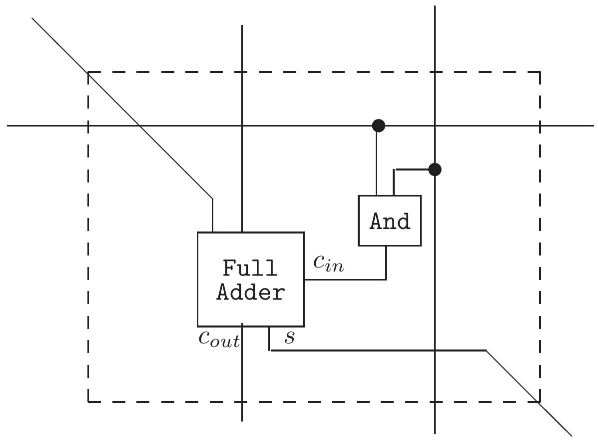


Figure 14.5: Individual Cell of Systolic Array

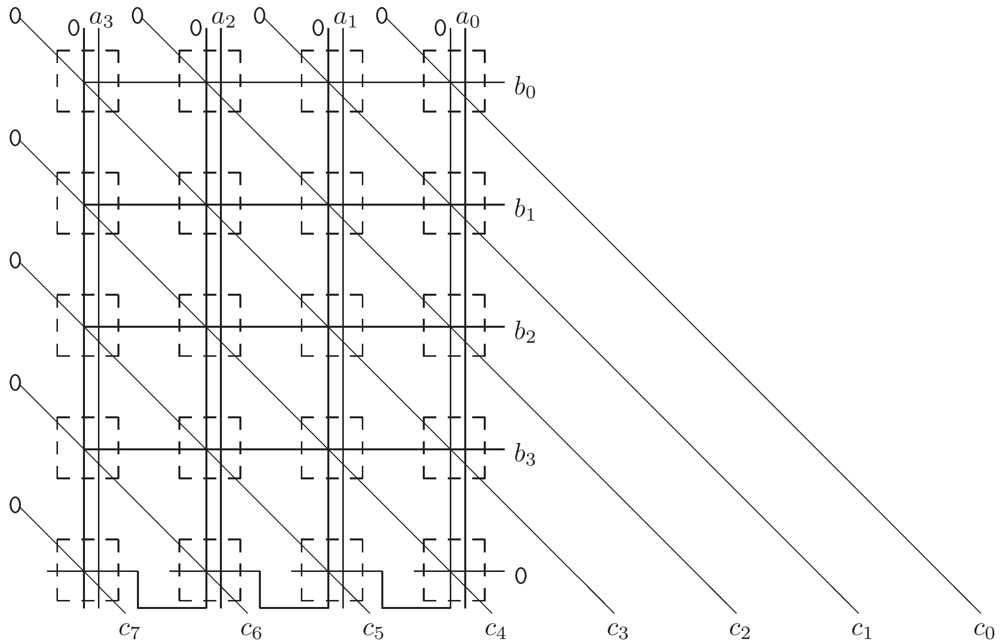


Figure 14.6: Systolic Array For 4 Bit Numbers

represented by their remainders of a group of numbers that constitute the basis of the representation. Let's consider a simple example of how numbers can be represented in this method.

Number	%2	%3
0	0	0
1	1	1
2	0	2
3	1	0
4	0	1
5	1	2

Note that each of the numbers from 0 through 5 can be represented uniquely by their remainders. Note that the number 6 would be 0,0 and thus not distinguishable from 0. You can represent six numbers (1-5) because the product of the basis numbers is $2 \times 3 = 6$. That we can represent the numbers is one thing, being able to calculate easily is another. Lets consider addition first:

1=1,1	2=0,2
2=0,2	3=1,0
3=(0+1)%2,(1+2)%3	5=(0+1)%2,(2+0)%3
=1,0	=1,2

If you look up (1,0) in our table you will find it corresponds to 3, similarly (1,2) corresponds to 5. Now lets try some multiplication problems:

2=0,2	1=1,1
2=0,2	3=1,0
4=(0 × 0)%2,(2 × 2)%3	3=(1 × 1)%2,(1 × 0)%3
=0,1	=1,0

If you look up (0,1) in our table you will find it corresponds to 4, similarly (1,0) corresponds to 3. Subtraction is slightly more complex, similar to the 2's complement[3] an inverse of each remainder (the representation) must be found. This is done by subtracting each remainder from the number it was modulused from. This is easiest to see in an example.

Example 21 *First, let's get a table of the numbers and their negatives (additive inverses):*

Number Decimal	*Residue %2,%3*	*Negative %2,%3*	*Negative Decimal*
0	*0,0*	*(2-0)%2=0,(3-0)%3=0*	*0*
1	*1,1*	*(2-1)%2=1,(3-1)%3=2*	*5*
2	*0,2*	*(2-0)%2=0,(3-2)%3=1*	*4*
3	*1,0*	*(2-1)%2=1,(3-0)%3=0*	*3*
4	*0,1*	*(2-0)%2=0,(3-1)%3=2*	*2*
5	*1,2*	*(2-1)%2=1,(3-2)%3=1*	*1*

Now let's do some calculations.

$$\begin{aligned} 5-2 &= (1,2)-(0,2) \\ &= (1,2)+(0,1) \\ &= (1+0,2+1) \\ &= (1,0) \\ &= 3 \end{aligned}$$

[3]In fact it is a radix complement, in particular since for our example their are 6 numbers in our example, we will be calculating the 6's complement and then finding its residue.

$$\begin{aligned}
4-4 &= (0,1)-(0,1)\\
&= (0,1)+(0,2)\\
&= (0+0,1+2)\\
&= (0,0)\\
&= 0
\end{aligned}$$

$$\begin{aligned}
2-1 &= (0,2)-(1,1)\\
&= (0,2)+(1,2)\\
&= (0+1,2+2)\\
&= (1,1)\\
&= 1
\end{aligned}$$

The basis of the representation must be relatively prime, that is they must have unique prime factors (they cannot share prime factors with other basis numbers). This means that you can have a number like 4 (2×2) as long as no other basis had 2 as a factor, but you could not have 9 (3×3) and 12 ($2 \times 2 \times 3$), or 6 (2×3) and 10 (2×5) in the same basis. To see why consider the basis (4,6), it should give unique representations for $4 \times 6 = 24$ numbers (0-23).

Number	%4	%6	Number	%4	%6
0	0	0	12	0	0
1	1	1	13	1	1
2	2	2	14	2	2
3	3	3	15	3	3
4	0	4	16	0	4
5	1	5	17	1	5
6	2	0	18	2	0
7	3	1	19	3	1
8	0	2	20	0	2
9	1	3	21	1	3
10	2	4	22	2	4
11	3	5	23	3	5

Notice the first and second column are the same, and thus do not give us the full range we wanted.

Chapter 15

Floating Point

The main goal of this chapter is to introduce floating point numbers and the issues around their use and misuse. Toward that end, we will first cover fixed point numbers.

15.1 Fixed Point Numbers

Example:

Convert π to binary and hexadecimal. Assume you have four bits before the radix point and 8 bits after the radix point.

Sol:

before the decimal we have $3 = 0011$

after the decimal

0.1415926 . . .	
0.2831852	0
0.5663704	0
1.1327408	1
0.2654816	0
0.5309632	0
1.0619264	1
0.1238528	0
0.2477056	0

combining gives 0011.00100100

To convert to hexadecimal we group the digits together in groups of four starting at the radix point, thus we are forcing the hexadecimal digits to represent either integer or fractional portions.

0011	0010	0100
3	2	4

Thus the answer is $0x3.24$.

Example:

Convert 25.6875 to binary.

25	/2	.	*2	.6875
12	1		1	.375
6	0		0	.75
3	0		1	.5
1	1		1	0
0	1			

11001.1011

15.2 Floating Point Numbers

I came up with the following program in my doctoral work at UCSB.

```
#include <iostream>
#include <iomanip>
#include <cmath>

using namespace std;

int main(){
    double pi, e, result;
    int i;

    e=exp(1);

    pi=atan(1)*4;

    result=pi;

    for(i=1;i<53;i++){
        result=sqrt(result);
    }

    for(i=1;i<53;i++){
        result=result*result;
    }

    cout << setiosflags(ios::showpoint | ios::fixed) << setprecision(16);
    cout << "Pi     = " << pi << endl;
    cout << "Result = " << result << endl;
    cout << "e      = " << e << endl;

    return 0;
}
```

The results are

```
Pi     = 3.1415926535897931
Result = 2.7182818081824731
e      = 2.7182818284590451
Press any key to continue
```

Notice that Result is e to 7 significant digits, but it should be π. This underscores the importance of being numerically aware when writing programs.

15.3 IEEE 754

Floating point numbers are based off scientific notation. Consider a typical number in base 10 scientific notation,

$$-1.23 \times 10^3.$$

The number is composed of five pieces of information,

1. sign of the number (-),
2. significant or mantissa (1.23),
3. base (10),
4. sign of the exponent (+),
5. magnitude of the exponent (3).

There are two basic number formats called out in IEEE 754, single precision (float in c/c++), and double precision (double in c/c++). In addition there are two extended formats, which are only used as intermediate results while calculating.

e	f	Category	Interpretation
$1 \ldots 11$	$1 \ldots 11$ $\vdots$ $0 \ldots 01$	NaN	See Codes
$1 \ldots 11$	$0 \ldots 00$	$\pm\infty$	$\pm\infty$
$1 \ldots 10$ $\vdots$ $0 \ldots 01$	$1 \ldots 11$ $\vdots$ $0 \ldots 00$	Numbers	$(-1)^s \times 1.f \times 2^{(e-127)}$
$0 \ldots 00$	$1 \ldots 11$ $\vdots$ $0 \ldots 00$	Denormals	$(-1)^s \times 0.f \times 2^{(-126)}$
$0 \ldots 00$	$0 \ldots 00$	± 0	± 0

NaN codes:

Dec	Meaning	Example
1	invalid square root	$\sqrt{-1}$
2	invalid addition	$\infty + -\infty$
4	invalid division	$\frac{0}{0}$
8	invalid multiplication	$0 \times \infty$
9	invalid modulo	$x mod 0$

For this discussion, the notation $fl(x)$ will be used to mean the number x as it is represented in floating point on a computer.

$$(-1)^s \cdot 1.f \times 2^{e-127}$$

<table>
<tr><td>01</td><td>02</td><td>03</td><td>04</td><td>05</td><td>06</td><td>07</td><td>08</td><td>09</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td><td>32</td></tr>
<tr><td>s</td><td colspan="8">e</td><td colspan="23">f</td></tr>
</table>

This is equivalent to saying

$$(-1)^s \cdot 1.f \times 2^E$$

0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3
1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2
s	e=E+127								f																						

They are the same because $e - 127 = E$ is the same equation as $e = E + 127$. I think the latter is easier to use because you read E from the number and want e. The first form (standard for most texts) involves you guessing what number produced what you are seeing (rather than calculating it). It is like trying to solve $y = mx + b$ for y given x but using the form $\frac{(y-b)}{m} = x$ to do it. It works, just not well. In any case, consider some examples.

Example:

Convert 7.892 to single precision IEEE.

Step 1: Convert 7.892 to binary

$7.892 = 111.11100100010110100000111$

Step 2: Normalize and note sign

$7.892 = (-1)^0 1.1111100100010110100000111 \times 2^2$

Step 3: Calculate Excess 127 code for exponent

$e = 2 + 127 = 129 = 10000001$

Step 4:Round $1.f$ to 24 digits

$fl(1.1111100100010110100000111) = 1.11111001000101101000100$

Step 5: Assemble

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

Example:

Calculate 3.75×29.625 in IEEE-754 single precision floating point.

Convert:

$3.75 = 11.11 = 1.111 \times 2^1$

$29.625 = 11101.101 = 1.1101101 \times 2^4$

Multiply Significants:

	1.	1	1	0	1	1	0	1			
×	1.	1	1	1							
	1.	1	1	0	1	1	0	1			
	0.	1	1	1	0	1	1	0	1		
	0.	0	1	1	1	0	1	1	0	1	
	0.	0	0	1	1	1	0	1	1	0	1
1	1.	0	1	1	1	1	0	0	0	1	1

1.10111100011×2^1

Add exponents to normalization exponent and put in excess 127:

$1 + 4 + 1 + 127 = 133 = 10000101$

Write in single precision:

0	10000101	1011 1100 0110 0000 0000 000

Example:

Perform the following for IEEE-754, single precision

1. Show the representation of $x = 93.3125$

   $x = 93.125_{10} = 1011101.001_2 = 1.011101001 \times 2^6$

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

2. calculate $x * y$ for y equal to

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

   exponent: 128+133-127=134

   float: shortcut, note that y only has two 1's in the expansion (hidden and near end) and they are farther apart than the length of the significant portion of x. This will cause the x float to be placed starting at these locations. The comma below notes where the last bit of precision lies.

   $$z_{fl} = 1.0111010010000000000000101, 1101001$$

   Note that the first bit after the comma is a 1 so the number gets rounded up.

   z is

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

Example:

Convert 3.03125 to IEEE single precision

3		.	03125
1	1	0	0625
0	1	0	125
		0	25
		0	5
		1	0

$3.03125_{10} = 11.00001_2 = 1.100001_2 \times 2^1$

$1 + 127 = 128$

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

Now perform the following on your result and

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

1. Addition

   $x = 1.0000000100000001_2 \times 2^5$

   $y = 1.100001_2 \times 2^1 = 0.0001100001_2 \times 2^5$

   $$\begin{aligned} x + y &= 1.0000000100000001_2 \times 2^5 + 0.0001100001_2 \times 2^5 \\ &= (1.0000000100000001_2 + 0.0001100001_2) \times 2^5 \\ &= (1.0001100101000001_2) \times 2^5 \end{aligned}$$

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

2. Multiplication

   exponent is $132 + 128 - 127 = 133$

   significant is $1.0000000100000001 \times 1.100001 = 1.1000010110000101100001$

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

Example:

Perform the following for IEEE-754, single precision

1. Show the representation of $x = 0.8125$

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

2. calculate (show steps) $x * y$ for x from above and

y is

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

Exponent: $(10000001 + 01111110) - 01111111 = 11111111 - 01111111 = 1000000$

float= $1.101 * 1.11 = 10.11011 = 1.011011 \times 2^1$, so add 1 to exponent

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

3. Perform the multiplication above in decimal and verify the answer.

$.8125 * (-7) = -5.6875 = -101.1011_2$

15.4 Rounding versus Chopping

Rounding is almost always used because of two reasons. To see both, let the interval between two numbers in the representation is 2δ then for rounding $x - fl(x) \in [-\delta, \delta)$, while for chopping it is $x - fl(x) \in [0, 2\delta)$. The first problem is that the error magnitude is up to twice as large for chopping. This is obviously bad, but it is not as bad as the second problem. The second problem is that all the errors of chopping have the same sign, so no error cancellation is possible when calculations are done. To see why this is bad, consider the following.

Example:

Find out the error in calculating $\sum_{i=1}^{n} x_i$ on a computer. First note that what you actually calculate is $\sum_{i=1}^{n} fl(x_i)$. The error (actual minus calculated) is thus $Err = |(\sum_{i=1}^{n} x_i) - (\sum_{i=1}^{n} fl(x_i))|$. Also let $fl(x_i) = x_i + \gamma_i$ for γ_i in the error interval of your method.

$$\begin{aligned}
Err &= \left|(\sum_{i=1}^{n} x_i) - (\sum_{i=1}^{n}(x_i + \gamma_i))\right| \\
&= \left|(\sum_{i=1}^{n} x_i) - (\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\gamma_i)\right| \\
&= \left|\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\gamma_i\right| \\
&= \left|\sum_{i=1}^{n}\gamma_i\right| \\
&\leqslant \sum_{i=1}^{n}|\gamma_i|
\end{aligned}$$

For chopping the last inequality is actually an equality, i.e. chopping always has the worst case error. For a typical case on rounding the errors are distributed with some positive and some negative, thus cancellation can occur. For large sums (many terms) the law of large numbers and an assumed uniform distribution of γ_i indicates that the error for rounding will go to 0! This is a great result.

Example

Write C/C++ code to sum the following $\sum_{i=1}^{100} \frac{1}{i}$. Make sure you do it in the right order.

```
double sum=0;
int i;

for(i=100;i>=0;i--){
    sum+=1.0/i;}
```

15.5 Evaluating a Polynomial

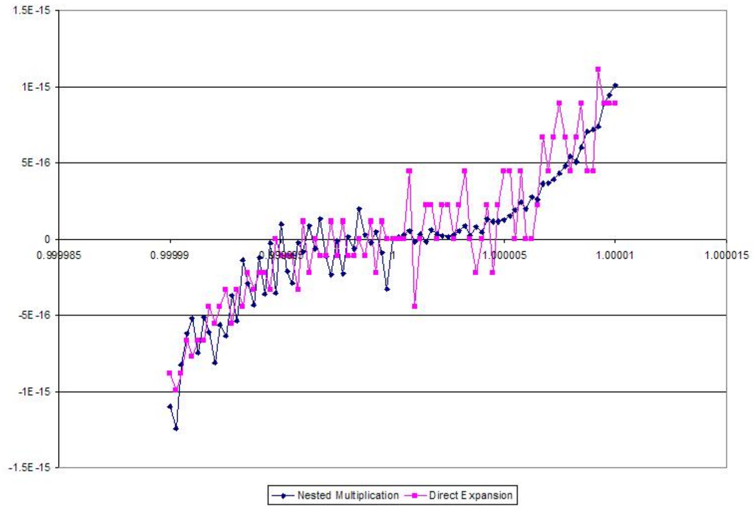


Figure 15.1: Close-up Look at Resulting Values of Two Evaluation Methods for $y = x^3 - 3x^2 + 3x - 1$

Part IV

Organization

Chapter 16

Arithmetic Operations

We have looked at number representation and calculation techniques, now we will look at how to specify the operations to a computer. In order to do an arithmetic operation, we need to know where the two operands (sources) are located and where the result should be placed (destination). Computers are classified by how many of the addresses must be explicitly stated and how many are implicit.

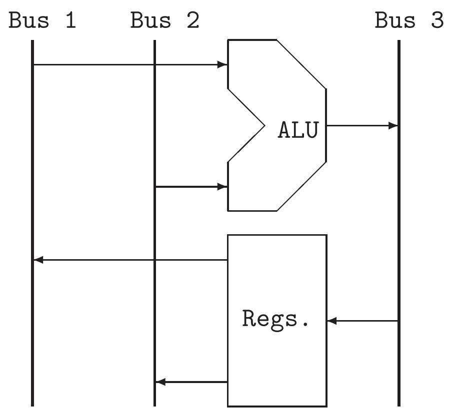
16.1 Three Address Machines

This is the most flexible form. Each address can be specified by the user. The commands are of the form

command source1, source2, destination

or

command destination, source1, source2



16.2 Two Address Machines

The destination is also a source in this case. The commands are of the form

command destination, source

16.3 One Address Machines

A special register, called the accumulator, is designated to be a source and destination. The accumulator has two special instructions, load accumulator and store accumulator. Accumulator machines rarely use additional registers, though it is not technically required. The arithmetic commands are of the form
command source

16.4 Zero Address Machines

The internal registers are arranged as a stack. The source operands are taken from the stack in order (first operand on top, second operand below). The result is pushed on the stack. These are often called stack machines. The arithmetic commands are of the form
command

16.5 Comparison Code

Consider the following equation:

$$\begin{aligned} y &= x^2 + 2x + 3 \\ &= (x + 2) * x + 3 \end{aligned}$$

Assume x is at 100, 2 is at 104, 3 is at 108, and y is at 112. The following uses a three address scheme with destination first.

version 1 $y = x^2 + 2x + 3$	version 2 $y = (x + 2) * x + 3$
mpy 112,100,100	add 112,100,104
mpy 116,100,104	mpy 112,112,100
add 112,112,116	add 112,112,108
add 112,112,108	

The following shows the second version on different machines.

3 address	2 address	1 address	0 address
add 112,100,104	move 112,100	load 100	push 100
mpy 112,112,100	add 112,104	add 104	push 104
add 112,112,108	mpy 112,100	mpy 100	add
	add 112,108	add 108	push 100
		store 112	mpy
			push 108
			add
			pop 112

Assume x is in R_1, 2 is in R_2, 3 is in R_3, and y is in R_4.

Chapter 17

Stack Machines

Stack machines are also known as 0-address machines, because no address must be specified for arithmetic operations. The most common example of a stack machine is an HP calculator. The application "Toy Stack" is an executable for Windows XP, which is available at the website. It has 64 bytes of memory split into 32 for instructions and 32 for data. All variables are 1 byte long and stored in 2's complement or unsigned form. Instructions are 1 byte long, but can have two commands in it in some cases. There is no branch delay slot. The commands are

Memory

0	0	P	Addr

where,

$$\begin{aligned} P &= \begin{cases} 0, & \text{Push;} \\ 1, & \text{Pop.} \end{cases} \\ Addr &= \text{5-bit address in memory.} \end{aligned}$$

Branching

0	1	C	Addr

where,

$$\begin{aligned} C &= \begin{cases} 0, & \text{Always;} \\ 1, & \text{Less (i.e. the top number on the stack is negative).} \end{cases} \\ Addr &= \text{5-bit address in memory to branch to.} \end{aligned}$$

Note: branch less is also branch bit set, for the most significant bit on the top of the stack.

Arithmetic

1	0	Op_1	Op_2

where,

$$Op_i = \begin{cases} 000, & \text{halt } (Op_1) \text{ or nop } (Op_2); \\ 001, & \text{addition;} \\ 010, & \text{subtraction;} \\ 011, & \text{negation;} \\ 100, & \text{unsigned multiplication;} \\ 101, & \text{signed multiplication;} \\ 110, & \text{unsigned division;} \\ 111, & \text{signed division.} \end{cases}$$

Note: Nop is no operation, and is used to allow, just one arithmetic command to execute rather than two. Halt is used to terminate the program run. If something other than nop is in Op_2 after a halt then that command is executed before termination.

Shifting

1	1	0	L/R	mode	times

where,

$$L/R = \begin{cases} 0, & \text{left shift;} \\ 1, & \text{right shift.} \end{cases}$$
$$mode = \begin{cases} 00, & \text{fill with 0's;} \\ 01, & \text{fill with 1's;} \\ 10, & \text{arithmetic shift;} \\ 11, & \text{circulant shift.} \end{cases}$$
$$times = \text{shift (1+times) bits (times is a two bit number).}$$

Push Signed Constant

1	1	1	0	Const

where, Const is a four bit number that is sign extended to eight bits and pushed on the stack.

Logic

1	1	1	1	0	Op

where,

$$Op = \begin{cases} 000, & \text{or;} \\ 001, & \text{nor;} \\ 010, & \text{orn;} \\ 011, & \text{xor;} \\ 100, & \text{and;} \\ 101, & \text{nand;} \\ 110, & \text{andn;} \\ 111, & \text{equivalence.} \end{cases}$$

Note: all logic functions are bitwise.

Undefined

1	1	1	1	1	Op

where, Op is a three bit operand. This operation is left undefined.

At the moment you have to enter your programs and data values manually, sorry I just started writing this. A load and save feature has been added which saves the memory to a file in encrypted format. You can only load programs that were encrypted with your exact name (spelling and caps count). Essentially this removes sharing data files as you need to submit your solutions electronically to me, with the exact spelling of your name (so I can load them). I will not give credit to you unless the name is yours.

17.1 Affine Encryption Program

Affine encryption is one of the simplest methods for doing encryption. Let P_i be the i^{th} character in the plain text message, and let C_i be the corresponding encoded character. Let there be n possible characters to encode, then the basic idea is to pick two numbers (a, b) to encode a message such that $\gcd(a, n) = 1$ (so a has an inverse). No requirement on b is needed if your modulus function has been encoded correctly. The encoded character can then be found by

$$a \times P_i + b = C_i \mod n.$$

Note that the " mod n" at the end says the equation holds in $\mathbb{Z}_n$, the set of integers mod n with appropriately defined arithmetic.

To decrypt the message, the equation

$$\bar{a} \times (C_i + d) \quad = \quad P_i \mod n$$

is used. The term $\bar{a}$ is the inverse of a in $\mathbb{Z}_n$, which is found by solving

$$a \times \bar{a} \quad = \quad 1 \mod n$$

or

$$a \times \bar{a} \quad = \quad m \times n + 1.$$

Note that m is any whole number. The term d is the additive inverse of b in $\mathbb{Z}_n$, which is found by solving

$$d = n - (b \mod n).$$

We can summarize this by saying an affine cipher is an encryption technique that encodes using three integers: a, b, and n. If *plain* is the character to be encoded (with 'A'=0 and 'Z'=25) then $code = (a*plain+b)$ mod n. Decoding is also done using three integers: c, d, and n. If *code* is the character to be encoded (with 'A'=0 and 'Z'=25) then $plain = (c * (code + d)) \mod n$. The requirements on (a, b, c, d, n) are:

- $\gcd(a, n) = 1$
- $(ac) \mod n = 1$
- $(b + d) \mod n = 0$

Below is C code to implement a particular case of affine cyphers.

```
char affine_encode(char plain){
    // affine codes capital letter in plain using a=5, b=12 thus this is modulo 26
    int iCode, iPlain, a=3,b=0;

    // convert char to integer and shift so A=0
    iPlain=int(plain)-65;

    // do the encoding
    iCode = (a*iPlain+b)%26;

    // return the result as a char
    return char(iCode+65);
}

char affine_decode(char code){
    // affine decodes capital letter in plain using c=21, d=8 thus this is modulo 26
    int iCode, iPlain, c=9, d=0;

    // convert char to integer and shift so A=0
    iCode=int(code)-65;

    // do the decoding
```

```
    iPlain = (c*(iCode+d))%26;

    // return the result as a char
    return char(iPlain+65);
}
```

Using this we consider affine encryption for standard ASCII including the control codes. In this case $n = 2^7 = 128$. Note that the standard arithmetic on our stack machine is $\mathbb{Z}_{2^8}$ so we can calculate normally then drop the leading bit to get $\mathbb{Z}_{2^7}$. As long as a does not have 2 as a factor it will meet the requirement $\gcd(a, n) = 1$. Let $a = 3$ then $3 \times \bar{a} = m \times n + 1$ for some $m \in \{1, 2, \ldots\}$. Start with $m = 1$, then $\bar{a} = 129/3 = 43$. Since the result is an integer, it is an inverse. If the result was not an integer, m would be incremented and the process would continue. Finally, let $b = 57$ then $d = 128 - 57 = 71$.

Let the memory locations of the variables be:

Variable	Address	Value
P	00000	your choice
C	00001	per calculation
P(calc)	10000	per calculation
a	11100	00000011
$\bar{a}$	11101	00101011
b	11110	00111001
d	11111	01000111

The variable P(calc) was added so the decoded plain text would not overwrite the original. The program to encode is thus:

Machine	Assembly	;Comment
00011110	push b	;load data
00011100	push a	;
00000000	push P	;
	unsigned multiply	;aP+b
10100001	add	;
11000000	shl0 1	;drop leading bit
11010000	shr0 1	;
00100001	pop C	;store
10000000	halt	;done

The program to decode is thus:

Machine	Assembly	;Comment
00011101	push $\bar{a}$	;load data
00011111	push d	;
00000001	push C	;
	add	;$\bar{a}(C + d)$
10001100	unsigned multiply	;
11000000	shl0 1	;drop leading bit
11010000	shr0 1	;
00110000	pop P(calc)	;store
10000000	halt	;done

17.2 Babylonian Algorithm

Implement the following Babylonian algorithm to find Pythagorean Triples[1] on the Toy Stack.

- Start with 2 (unsigned) integers p, q with $p > q$ (assume these are present)

[1] The algorithm actually predates Pythagoras.

- calculate the three numbers by: $n_1 = 2pq$, $n_2 = p^2 - q^2$, $n_3 = p^2 + q^2$

To understand how this works note that

$$\begin{aligned} n_1^2 &= (2pq)^2 \\ &= 4p^2q^2 \end{aligned}$$

and

$$\begin{aligned} n_2^2 &= (p^2 - q^2)^2 \\ &= p^4 - 2p^2q^2 + q^4 \end{aligned}$$

and

$$\begin{aligned} n_3^2 &= (p^2 + q^2)^2 \\ &= p^4 + 2p^2q^2 + q^4 \end{aligned}$$

thus

$$\begin{aligned} n_1^2 + n_2^2 &= (4p^2q^2) + (p^4 - 2p^2q^2 + q^4) \\ &= p^4 + 2p^2q^2 + q^4 \\ &= n_3^2 \end{aligned}$$

The assembly is

```
push 0   ! calculate 2pq
push 1
push #2
umul
umul
pop 16   ! 2pq stored in 16
push 0   ! calculate p^2
push 0
umul
pop 2    ! p^2 stored in 2
push 1   ! calculate q^2
push 1
umul
pop 3    ! q^2 stored in 3
push 3   ! calculate p^2 - q^2
push 2
sub
pop      ! p^2 - q^2 stored in 17
push 3   ! calculate p^2 + q^2
push 2
add
pop      ! p^2 + q^2 stored in 17
```

For the machine code see the website.

Chapter 18

Instruction Set Architecture

18.1 RISC vs. CISC

RISC reduced instruction set computer- For high level language programmers (reduces time for each instruction)

CISC complex instruction set computer- For assembly programmers (reduces instructions for same program)

	RISC	CISC
Number of addressing modes	few	many
Access to main memory	Only in loads and stores (hence load-store architecture)	One or more operands in most instructions can access
Size of instruction set	small	large
Complexity of each instruction	small	large

RISC is currently and has been more efficient.

18.2 Memory Access

Most machines are byte addressable (i.e. each byte in memory has an address). Memory access typically come in three sizes and are often distinguished by the operand suffix .b (byte), .h (halfword), .w (word).

18.3 Branching

Conditional branching
Three ways: compare two, compare to zero, condition registers
cmp
Branch delay and pipelining
short circuit (positional) put in sum of expressions form and then do a series of conditional branches
Bitwise (and,or,xor,andn,orn)
bb (bitbranch reg,bit,targ)
bset
bclr
shift L/R

zero fill
one fill
rotate
usually to carry

Chapter 19

Addressing

- .bss
- .data
- .text

.bss (block started by symbol) memory, reserved only

.data memory, predefined values

.text instructions

.reserve val (alternately ".skip val") sets aside val bytes of memory

.equate name, val (alternately ".set name, val") makes name a constant with value val

.byte val (alternately .b, ub, sb) specifies the operation to be on a byte

.half val (alternately .h, uh, sh) specifies the operation to be on a half word (2 bytes)

.word val (alternately .w) specifies the operation to be on a word (4 bytes)

.align val aligns the memory location counter

Note that val may be a constant expression for readability.

Name	Generic	Sparc	Uses
memory direct	mX	[%r0+X]	
register direct	rX	%rX	
immediate	#X	X	
memory indirect	@mX	-	pointers
register indirect	@rX	[%rX]	pointers
memory indexed	label[mX]	-	arrays
register indexed	label[rX]	[%rY + %rX]	arrays
			(note %rY is loaded with label)
pre-increment	+[rX]	-	increments by size (stride) each time
post-increment	[rX]+	-	increments by size (stride) each time
pre-decrement	-[rX]	-	decrements by size (stride) each time
post-decrement	[rX]-	-	decrements by size (stride) each time
memory displaced	mX → label	-	struct
register displaced	rX → label	[%rX + label]	struct

m0	0x00	0x00	0x00	0x12
m4	0x00	0x00	0x00	0x08
m8	0x01	0x23	0x45	0x67
m12	0x89	0xAB	0xCD	0xEF
m16	0x12	0x34	0x56	0x78
m20	0x9A	0xBC	0xDE	0xF0
m24	0x11	0x11	0x11	0x11

r0	0x00	0x00	0x00	0x00
r1	0x00	0x00	0x00	0x08
r2	0x00	0x00	0x00	0x0C
r3	0x00	0x00	0x00	0x04
r4	0x00	0x00	0x00	0x10

Let var1 be a label for the value 8.

Representation	X=4	Effective Address	Expression
mX	m4	0x00000004	0x00000008
rX	r4	-	0x00000010
#X	#4	-	0x00000004
@mX	@m4	0x00000008	0x01234567
@rX	@r4	0x00000010	0x12345678
var1[mX]	8[m4]	0x00000010 (i.e.: 8+8)	0x12345678
var1[rX]	8[r4]	0x00000018 (i.e.: 8+16)	0x11111111
+[rX]	+[r4]	0x00000014	0x9ABCDEF0
			r4 ← 0x00000014 before
[rX]+	[r4]+	0x00000010	0x12345678
			r4 ← 0x00000014 after
-[rX]	-[r4]	0x0000000C	0x89ABCDEF
			r4 ← 0x0000000C before
[rX]-	[r4]-	0x00000010	0x12345678
			r4 ← 0x0000000C after
mX → var1	m4 → 8	0x00000010 (i.e.: 8+8)	0x12345678
rX → var1	r4 → 8	0x00000018 (i.e.: 8+16)	0x11111111

19.0.1 Arrays

For instance consider an array of 10 integers.

```
int my_int[10];
```

This creates both the array of integers and a pointer to the first element. The elements are numbered 0 to 9 and are accessed by $my_int[i]$ for $i \in \{0, 1, \ldots, 9\}$. They can also be accessed by $*(my_int + i)$. In assembly we would have:

```
my_int:  .skip 10*4    ; each int is 4 bytes
```

The contents can be accessed by:

```
set i, %r2
ld  [%r2], %r2
umul %r2, 4, %r3
set my_int, %r4

ld [%r4 + %r3], %r5
```

or if my_int (the address) fits in a 13 bit signed constant:

```
  set i, %r2
  ld  [%r2], %r2
  umul %r2, 4, %r3

  ld [%r3+my_int], %r5
```

Essentially the address is my_int + i*4, but this assumes that start of my array is zero. How about a language like Pascal or VB which allows other starting values? Consider defining an array (-m,-m+1,. . . , -1, 0, 1, . . . , n). To use the address my_int + i*size we have

```
        .skip m*size        ! negatives
        .skip (n+1)*size    ! zero and positives
```

Alternately,

```
        .skip (m+n+1)*size  ! whole thing
```

This causes the address to be my_int + (i+m)*size. Now you might think this will be longer, but note that it can be rewritten as

my_int + (i+m)*size
my_int + i*size + m*size
(my_int + m*size) + i*size

That is, rather than constantly biasing the index, it makes more sense to bias the base. Essentially it makes the second method look like the first, but it works for a positive starting number (by making m a negative). Since it is more general the later form is what is used in practice.

19.0.2 String Storage

string256 (aka length plus value) length of string in first byte, string following

NULL terminated string followed by 0

19.0.3 Structs

```
struct book{
  int pages;
  float price;
  char title[20];
}library[100];
```

Would be implemented:

```
        .set pages, 0
        .set price, 4
        .set title, 8
        .set book_size, 28

.bss

library: .skip 100*book_size
```

.bss is done in .data on some assemblers or machines

Chapter 20

Subroutines

20.1 Basic Overview

Before we get into this, let's establish some basic definitions.

Caller the section of code that initiates the call

Callee the section of code that is called

Return Address The address of the instruction to be executed after the call is done (usually the one following the branch or jump)

Subroutine Linkage data structure used to share information between caller and callee

20.1.1 What needs to be passed?

A subroutine can be called from different sections of code and with different parameters. The subroutine needs to know what data it must operate on and where to resume execution when it finishes. Additionally the subroutine usually must return some data, and thus it must place the data in an easy to locate area. The basic data that must be exchanged is thus,

- return address
- return value
- parameters

20.1.2 General Call Sequence

Caller		Callee
Startup Sequence	→	Prologue
		Body
Cleanup Sequence		Body
	←	Epilogue

20.2 Return Addresses in Leaf and Non-Leaf Subroutines

For the moment we will look only at the issues surrounding return addresses. The following distinctions must be made:

> Leaf subroutines do not make subroutine calls, where as non-leaf subroutines call at least one subroutine (itself or another subroutine).

The most basic leaf subroutine call looks like:

Caller		Callee
address+(4 or 8) to r31	→	none
branch sub		Body
none		Body
none	←	branch @r31

The basic leaf routine is quick and easy, but it cannot be used on non-leaf procedures as the return address would be lost. Consider the following subroutine to calculate x^n:

Code

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!! name: pow
!! desc: calculates x^n
!! meth: recursive function call
!!            x*(x^{n-1})
!! parm: x in r8
!!       n in r9
!! pre : nothing in r16, it is used as
!!       a temporary variable
!! post:
!! ret : x^n in r8
!! date: 20 May 2003
!! rev : 1.0
!! revh:
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
pow:        cmp r9,r0         ! see if x^0
            breq,a pow_done   !  if n=0
            add r0,1,r8       !  then ans=1

            cmp r9,1          ! see if x^1
            breq pow_done     ! if n=1
            nop               ! then ans=x

            mv r8,r16         ! else n>1
            call pow          ! calc r8=x^{n-1}
            sub r9,1,r9       !
pow_r:      smul r16,r8,r8    ! ans = x*x^{n-1}
pow_done:   retl
            nop
```

Sample run

Assume the call was to calculate 5^2 and return to the label "retn". For our machine the return address is stored in r31. We will assume that annulled commands become nop's (they really do, the results are just sent to r0 and the condition codes are not set).

Instruction	r8	r9	r16	r31
cmp r9,r0	5	2	-	retn
breq,a pow_done	5	2	-	retn
nop	5	2	-	retn
cmp r9,1	5	2	-	retn
breq pow_done	5	2	-	retn
nop	5	2	-	retn
mv r8,r16	5	2	5	retn
call pow	5	2	5	pow_r

Notice at this point we lost the return address!

Instruction	r8	r9	r16	r31
sub r9,1,r9	5	1	5	pow_r
cmp r9,r0	5	1	5	pow_r
breq,a pow_done	5	1	5	pow_r
nop	5	1	5	pow_r
cmp r9,1	5	1	5	pow_r
breq pow_done	5	1	5	pow_r
nop	5	1	5	pow_r
retl	5	1	5	pow_r
nop	5	1	5	pow_r
smul r16,r8,r8	25	1	5	pow_r
retl	25	1	5	pow_r
nop	25	1	5	pow_r

At this point it should have gone back to "retn" but since that address was lost it will loop endlessly.

If the subroutine is non-leaf and not part of a cycle (recursive or otherwise) then the following modification will work nicely.

Caller		Callee
address+(4 or 8) to r31	→	r31 to mem
branch sub		Body
none		mem to r31
none	←	branch @r31

the two versions can be combined as:

Caller		Callee
address+(4 or 8) to r31	→	if nonleaf r31 to mem
branch sub		Body
none		if nonleaf mem to r31
none	←	branch @r31

20.3 Parameter Passing

We now turn our attention on the parameters. First we need to consider how to represent the data. For instance if you just need to send an integer to do a calculation but you don't want it modified then you would pass by value. If on the other hand you need to pass an instance of a class you must pass by reference. The three ways data may be handled are

1. pass by value (not returned)
2. pass by value/result (modify and return)
3. pass by ref (pointer to actual object)

Beyond these basic considerations, there is a question as to where to locate the data for the subroutine call. The information could be located in the registers for speed, or in static variables in RAM (parameter block). Neither of the options discussed so far will handle cyclic subroutines or dynamic local variables. If either cyclic subroutines or dynamic local variables are needed the information must be passed on the stack (dynamic variables in RAM). The methods are:

1. register
   - fast
   - leaf subroutine
2. parameter block
   - larger data
   - non-leaf and non-cyclic subroutines
3. stack
   - larger data
   - (dynamic) local variables
   - cyclic and recursive calls

20.4 Register

Caller		Callee
mv params into r8 to r13		
address+(4 or 8) to r31		none
branch sub	$\nearrow$	Body
none		mv result to r8
none	$\nwarrow$	branch @r31

Example

We have discussed affine ciphers already. You might have noticed that the equation for encoding and decoding is very similar. We can combine them with only a small alteration to the decoding formula and one of the requirements. Decoding is still done using three integers: c, d, and n. If *code* is the character to be decoded (with 'A'=0 and 'Z'=25) then $plain = (c * code + d) \mod n$. The requirements on (a, b, c, d, n) are:

- $\gcd(a, n) = 1$
- $(ac) \mod n = 1$
- $(cb + d) \mod n = 0$

Below is C code to implement a particular case of affine cyphers.

```
char affine(char letter, int scale, int offset){
    // affine codes capital letter in 'letter' thus this is modulo 26
    int iCode, iLetter;

    // convert char to integer and shift so A=0
    iLetter=int(plain)-65;

    // do the encoding
    iCode = (scale*iLetter+offset)%26;

    // return the result as a char
    return char(iCode+65);
}
```

The SPARC syntax is then
affine

```
        ! calculates affine encryption:
        !    crypt = (a*(orig-off)+b) mod p + off
        ! a     is passed   in r8
        ! b     is passed   in r9
        ! n     is passed   in r10
        ! off   is passed   in r11
        ! orig  is passed   in r12
```

```
        ! crypt is returned in r
        .text
affine: sub  r12, r12, r11  ! orig-off
        mult r8,  r12, r8   ! a*(orig-off)
        add  r8,  r8,  r9   ! a*(orig-off)+b
        div  r9,  r8,  r10  ! x= y mod z = y - y/z*z
        mult r9,  r9,  r10
        sub  r8,  r8,  r9   ! (a*(orig-off)+b) mod n
        add  r8,  r8,  r11  ! done
        retl
```

encrypt call

```
        ! affine encrypt
        ! a is passed in r8
        ! b is passed in r9
        ! n     is passed   in r10
        ! off   is passed   in r11
        ! orig  is passed   in r12
        ! crypt is returned in r8
        .text
        set r8, 3            ! given
        set r9, 0            ! given
        set r10, 26          ! letters in alphabet
        set r11, 65          ! A in ascii
        call affine          ! call and link
        ld.b r12, add_plain  ! assume have label add_plain
                             !   where plain text is stored
        st.b r8,  add_code   ! assume have label add_code where
                             !   cypher text is to be stored
```

decrypt call

```
        ! affine decrypt
        ! a is passed in r8
        ! b is passed in r9
        ! n     is passed   in r10
        ! off   is passed   in r11
        ! orig  is passed   in r12
        ! crypt is returned in r8
        .text
        set r8, 9            ! given
        set r9, 0            ! given
        set r10, 26          ! letters in alphabet
        set r11, 65          ! A in ascii
        call affine          ! call and link
        ld.b r12, add_code   ! assume have label add_code
                             !   where cypher text is stored
```

```
            st.b r8,  add_plain  ! assume have label add_code where
                                 !   plain text is to be stored
```

20.5 Parameter Block

Caller		Callee
store params into block using labels		allocate block and labels in .data
store address+(4 or 8) to block		none
branch sub	↗	Body
load result to desired register		store result to block
none	↖	ld return address to r31
none		branch @r31

20.6 Stack

The stack is a large block of RAM which data is pushed onto. Any piece of information can be pushed onto the stack. All the data passed to and from the subroutine with all the local variables composes a block of information on the stack called the frame. The frame is created in the startup and prologue and removed in the epilogue and cleanup. The startup allocates space for all the information that must be passed (return address, parameters, and return values), and the cleanup removes it. The prologue allocates any local variables or storage to protect registers and the epilogue removes this local information.

Caller		Callee
push params		
allocate return value		push registers to protect
push return address	↗	push local variables
branch sub		Body
pop result to desired register		store result to stack at offset
pop params	↖	pop locals (remove)
		pop register back
		pop return address to r31
		branch @r31

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!! name: pow
!! desc: calculates x^n
!! meth: recursive function call
!!           x*(x^{n-1})
!! parm: stack passing:
!!       x              at fp+20
!!       n              at fp+16
!!       return value   at fp+12
!!       return address at fp+8
!! pre :
!! post:
!! ret : x^n at fp+12
```

```
!! date: 22 May 2003
!! rev : 1.1
!! revh:
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
            .set s16,0       ! offset to save r16
            .set s17,4       ! offset to save r17
            .set ra,8        ! offset to ret add
            .set rv,12       ! offset to ret val
            .set n,16        ! offset to n
            .set x,20        ! offset to x
pow:        sub sp,8,sp      ! allocate save space
            mv sp,fp         ! set frame
            st r16,[fp+s16]  ! save r16
            st r17,[fp+s17]  ! save r17

            ld [fp+n],r17    ! load n
            cmp r17,r0       ! see if x^0
            breq,a pow_done  !  if n=0
            add r0,1,r16     !  then ans=1

            cmp r17,1        ! see if x^1
            breq pow_done    !  if n=1
            ld [fp+x],r16    !  then ans=x

                             ! else n>1
            sub sp,4,sp      !  decrement pointer
            st r16,[sp]      !  push x
            sub r17,1,r17    !  calc n-1
            sub sp,4,sp      !  decrement pointer
            st r17,[sp]      !  push n-1
            sub sp,8,sp      !  decrement pointer
                             !  for return value
                             !  and address
            call pow         !  calc r8=x^{n-1}
            st r31,[sp]      !  push return address

            ld [sp],r16      !  get x^{n-1}
            add sp,12,sp     !  deallocate
            mv sp,fp         !  restore frame

            ld [fp+x],r17    !  get x
            smul r16,r17,r16 !  ans = x*x^{n-1}

pow_done:   st r16,[fp+rv]   ! store return value
            ld [fp+s16],r16  ! restore r16
            ld [fp+s17],r17  ! restore r17
            ld [fp+ra],r31   ! get return address
            retl
```

```
add sp,12,sp     !  deallocate ra, s16, s17
```

20.7 Temperature Conversion

Write a function that converts Fahrenheit to Celsius by following the steps below. A C/C++ command to do the conversion is:

```
celsius = ((fahrenheit - 32)* 5) / 9;
```

Note: I added an extra set of parenthesis to let you know you must do the multiplication first! Why does the multiplication have to be done first? Include an example.

If you do not multiply first, you can loose precision. ex: 2/9*5=0, while 2*5/9=1 (in integer math).

1. State the passing convention you will use (include what needs to be passed and where you will pass it) and any other reasonable assumptions on the machine.

   I will use register passing and will use register r8 to pass both the parameter and the result. Since this is a leaf procedure and I do not need other registers, I will use the book's leaf procedure (return address in r31). I will further assume that my machine has call and retl that automatically store and access the return address. Finally, I will assume there is a branch delay slot, the destination is always the first location, and I have all addressing modes. (your choices may be different).

2. Write the function.

   ```
   fahr_2_cels:   sub r8, r8, 32
                  mpy r8, r8, 5
                  retl
                  div r8, r8, 9
   ```

3. Show how it would be called. Assume that the Fahrenheit temperature is stored in a memory location specified by the label "fahr_temp". The result should be stored at the memory location specified by the label "cels_temp".

   ```
   set r1, fahr_temp
   call fahr_2_cels
   ld.w r8, @r1
   set r1, cels_temp
   st.w @r1, r8
   ```

Chapter 21

MIPS Assembly

R-Format	op	rs	rt	rd	shamt	funct
Bits	6	5	5	5	5	6
add $r1,$r2,$r3	0	$r2	$r3	$r1	0	32
addu $r1,$r2,$r3	0	$r2	$r3	$r1	0	33
sub $r1,$r2,$r3	0	$r2	$r3	$r1	0	34
subu $r1,$r2,$r3	0	$r2	$r3	$r1	0	35

I-Format	op	rs	rt	address
Bits	6	5	5	16
lw $r1,off($r2)	35	$r2	$r1	off
sw $r1,off($r2)	43	$r2	$r1	off

21.1 Registers

Number	Name	Use
0	$zero	0
1	$at	assembler use
2	$v0	return value (value)
3	$v1	return value (value)
4	$a1	parameters (arguments)
5	$a2	parameters (arguments)
6	$a3	parameters (arguments)
7	$a4	parameters (arguments)
8	$t0	temp (not saved)
9	$t1	temp (not saved)
10	$t2	temp (not saved)
11	$t3	temp (not saved)
12	$t4	temp (not saved)
13	$t5	temp (not saved)
14	$t6	temp (not saved)
15	$t7	temp (not saved)
16	$s0	saved temp
17	$s1	saved temp
18	$s2	saved temp
19	$s3	saved temp
20	$s4	saved temp
21	$s5	saved temp
22	$s6	saved temp
23	$s7	saved temp
24	$t8	temp (not saved)
25	$t9	temp (not saved)
26	$k0	OS
27	$k1	OS
28	$gp	global pointer (0x10008000) points to middle of 64k block
29	$sp	stack pointer
30	$fp	frame pointer
31	$ra	return address

21.2 Keeping Your Ends Straight

Big (LR) and little (RL) endian

Consistent (same for bits)

Sparc is inconsistent big-endian.

Endian	Consistent				Inconsistent			
Big	0	1	. . .	n	0	1	. . .	n
	0. . . 7	0. . . 7	. . .	0. . . 7	7. . . 0	7. . . 0	. . .	7. . . 0
Little	n	. . .	1	0	n	. . .	1	0
	7. . . 0	7. . . 0	. . .	7. . . 0	0. . . 7	0. . . 7	. . .	0. . . 7

21.3 Data Structures

Implement the following data structure in assembly then write a MIPS function to calculate $mykey.block = mykey.p \times mykey.q$.

```
struct keys{
   int p;
   int q;
   int public;
   int private;
   int block;
};
```

```
.data
mykey:
mykey_p: .word 0
mykey_q: .word 0
mykey_public: .word 0
mykey_private: .word 0
mykey_block: .word 0
.set mykey_off_p=mykey_p - mykey
.set mykey_off_q=mykey_q - mykey
.set mykey_off_public=mykey_public - mykey
.set mykey_off_private=mykey_private - mykey
.set mykey_off_block=mykey_block - mykey

.text
! Since this operates on data we know the location of,
!  we don't need to pass anything
la $t1, mykey
lw $t2, mykey_off_p($t1)
lw $t0, mykey_off_q($t1)
mul $t0,$t2
mflo $t0
sw $t0,mykey_off_block($t1)
```

21.4 Register Passing

21.4.1 Exponentiation by Multiplication

Write code to calculate n^m for n a non-zero finite integer and m a non-negative integer.

```
# n^m by loop
# n !=0 finite in a0
# m >=0 finite in a1
# n^m          in v0
# 0 in
pow_by_loop:
```

```
  # ensure arguments are ok
  mov  $v0,$zero
  beqz $a0,pow_done
  bltz $a1,pow_done
  # m=0 and setup
  addi $v0,$v0,1
  beqz $a1,pow_done
  # m>0, loop
pow_loop:
  mul  $v0,$a0
  mflo $v0
  subi $a1,$a1,1
  bgtz $a1,pow_loop
pow_done:
  jr $ra
```

Now how do we call it? Assume that n is in $s0 and m is at address "int_m" and we want the result in $s1.

```
mov $a0,$s0
la  $t1,int_m  # note I use $t1 for address scrap space
lw  $a1,0($t1)
jal pow_by_loop:
mv $s1, $v0
```

21.4.2 Polynomial Evaluation

Write the MIPS assembly code for the following function. Assume the array a has been defined as size n+1. You do not need to write the code to call the function but you need to state where you assume the parameters and return address will be.

```
int poly_eval(int* a, int n, int x){
  y=a[n];
  for(i=n-1;i>=0;i--){
    y=y*x+a[i];
  }
  return y;
}
```

```
###########################################
# poly_eval
# leaf procedure to evaluate polynomials
# parameters:
# a1 : pointer to array of coefficients
# a2 : largest index in array
# a3 : point to evaluate polynomial
# return value:
# v0 : value of polynomial
# temporary values:
```

```
# t0 : offset in array
# t1 : address in array
poly_eval: add  $t0, $a2, $a2         # four bytes per integer
           add  $t0, $t0, $t0
           add  $t1, $t0, $a1         # address of element to get
           lw   $v0,0($t1)            # initialize the answer
           beq  $t0,$zero, poly_done  # if only one element then done
poly_do:   mul  $v0, $v0, $a3         # y=y*x
           subi $t0, $t0, 4           # next coefficient is four bytes down
           add  $t1, $t0, $a1         # next coefficient's address
           lw   $t2, 0($t1)           # next coefficient
           add  $v0, $v0, $t2         # add next coefficient
           bne  $t0,$zero, poly_do    # more coefficients left
poly_done: jr   $ra                   # return
```

21.4.3 Xor Encryption

Consider the problem of xor encryption. The i^{th} cipher text character, C_i is given by

$$C_i = P_i \oplus K_i$$

where P_i is the i^{th} plain text character and K_i is the i^{th} key character. The decryption is then given by

$$P_i = C_i \oplus K_i.$$

This encryption method is thus symmetric.

```
#
#  xor
#
# $a0 contains plaintext
# $a1 contains key
# $a2 contains ciphertext
xor:
  mov $t3,$a1
  lb $t0,0($a0)
  lb $t1,0($a1)
xor_loop:
  xor $t2,$t0,$t1
  sb  $t2,0($a2)
  addi $a0,$a0,1
  addi $a1,$a1,1
  addi $a2,$a2,1
  lb $t0,0($a0)
  beqz $t0, xor_done
xor_load:
  lb $t1,0($a1)
  bgtz $t1, xor_loop
  mov $a1,$t3
```

```
  j xor_load
xor_done:
  jr $ra
```

21.4.4 Bubble Sort

```
#
#  Bubble Sort
#
# $a0 points to start of array
# $a1 points to last element in array
         move $t0, $a0
         move $t1, $a1
outter:  move $t4, $0           # 1
         lw   $t2, 0($t0)       # 2
inner:   lw   $t3, 4($t0)       # 3
         addi $t0, $t0, 4       # 4
         ble  $t2, $t3, no_swap # 5
swap:    sw   $t2, 0($t0)       # 6
         sw   $t3, -4($t0)      # 7
         ori  $t4, $0, 1        # 8
         blt  $t0, $t1, inner   # 9
         subi $t1, $t1, 4       # 10
         move $t0, $a0          # 11
         b    outter            # 12
no_swap: move $t2, $t3          # 13
         blt  $t0, $t1, inner   # 14
         subi $t1, $t1, 4       # 15
         move $t0, $a0          # 16
         bnez $t4, outter       # 17
```

21.5 Block Passing

Let us reconsider affine encryption as outlined in Section 17.1

We will be passed a pointer to a string of plaintext, *P, and the length of the string, len. Additionally we need the affine parameters a, b, and n. Five parameters cannot be passed in registers, as we only have four, so we will use a block. Modulus is handled nicely by div in mips so we have no problems there. To be really careful I will use divu (unsigned division).

If an error is detected I will use `break $zero` to halt execution. You could also write your own error handler but that did not seem reasonable given the length of the code already (3 pages). I have tried to exhibit good commenting techniques. They greatly simplify others reading and editing.

```
##############################################
# _affine_encrypt
#
#
# Author: Keith Schubert
# Date  : Nov 4, 2005
# Desc  : Affine encryption of a string
# Method: calculate then modulus.
# BlkPtr: _affine_encrypt_block_pointer
#         var  contents     offset
# Return:
# RetAdd:                        _affine_encrypt_off_ra
# Params: *P   plaintext         _affine_encrypt_off_p
#         len  plaintext.length  _affine_encrypt_off_len
#         *C   ciphertext        _affine_encrypt_off_c
#         a    affine scale      _affine_encrypt_off_a
#         b    affine shift      _affine_encrypt_off_b
#         n    # of code chars   _affine_encrypt_off_n
# Pre   :
# Post  : contents of $t0-$t8 changed, $ra changed
#
##############################################
.data
_affine_encrypt_block_pointer:
_affine_encrypt_base_ra:
    .word 0
_affine_encrypt_base_p:
    .word 0
_affine_encrypt_base_c:
    .word 0
_affine_encrypt_base_len:
    .word 0
_affine_encrypt_base_a:
    .word 0
_affine_encrypt_base_b:
    .word 0
_affine_encrypt_base_n:
    .word 0
_affine_encrypt_block_bottom:

    .set _affine_encrypt_off_ra =
         _affine_encrypt_base_ra - _affine_encrypt_block_pointer
    .set _affine_encrypt_off_p =
         _affine_encrypt_base_p - _affine_encrypt_block_pointer
    .set _affine_encrypt_off_c =
         _affine_encrypt_base_c - _affine_encrypt_block_pointer
    .set _affine_encrypt_off_len =
         _affine_encrypt_base_len - _affine_encrypt_block_pointer
```

```
    .set _affine_encrypt_off_a =
         _affine_encrypt_base_a - _affine_encrypt_block_pointer
    .set _affine_encrypt_off_b =
         _affine_encrypt_base_b - _affine_encrypt_block_pointer
    .set _affine_encrypt_off_n =
         _affine_encrypt_base_n - _affine_encrypt_block_pointer
    .set _affine_encrypt_block_size =
         _affine_encrypt_block_bottom - _affine_encrypt_block_pointer
.text
_affine_encrypt:

#
# Setup
#
# t0 = current char index
# t1 = *p
# t2 = *c
# t3 = len
# t4 = a
# t5 = b
# t6 = n
# t7 = current char
# t8 = effective address
#
la  $t1, _affine_encrypt_block_pointer
lw  $t2, _affine_encrypt_off_c($t1)
lw  $t3, _affine_encrypt_off_len($t1)
bgtz $t3,_affine_encrypt_len_ok
break $zero  #error stop execution
_affine_encrypt_len_ok
lw  $t4, _affine_encrypt_off_a($t1)
lw  $t6, _affine_encrypt_off_n($t1)

#
# Data validity
#
# see if gcd(a,n)=1
mov  $t5, $t4
mov  $t0, $t6
break $zero  # MIPS error
break $zero
# Euclid's alg
_affine_encrypt_Euclid:
divu $t5,$t0
mov  $t5,$t0
mfhi $t0
bgez $t0,_affine_encrypt_Euclid
subi $t5,$t5,1
```

```
beqz $t5,_affine_encrypt_ab_ok
break $zero
_affine_encrypt_ab_ok:

#
# Finish loads
#
lw  $t5, _affine_encrypt_off_b($t1)
lw  $t1, _affine_encrypt_off_p($t1)
mov $t0,$zero

#
# main loop
#
# get char, scale, shift, mod, then store
#
_affine_encrypt_loop:
add  $t8,$t0,$t1
lbu  $t7,0($t8)
mulu $t7,$t4
mflo $t7
add  $t7,$t7,$t5
divu $t7,$t6
mfhi $t7
add  $t8,$t0,$t2
sb   $t7,0($t8)
addi $t0,$t0,1
sle  $t8,$t0,$t3
beqz $t8,_affine_encrypt_loop

#
# Return
#
la $t1,_affine_encrypt_block_pointer
lw $ra,_affine_encrypt_off_ra($t1)
jr $ra
```

21.6 Stack Passing

On some machines you can/must manually allocate your own stack using .bss and .skip. On MIPS the stack is predefined and the OS initializes the stack pointer for you. We are going to define two macros, push and pop. To define a macro we use .macro and .endmacro.

```
.macro push arg1
  addui $sp,$sp,-4  # allocate space
  sw    arg1,0($sp) # place contents
.endmacro

.macro pop arg1
```

```
  lw    arg1,0($sp) # get contents
  addui $sp,$sp,4   # deallocate space
.endmacro
```

Let's consider Euclid's algorithm for finding the GCD of two numbers

1. Let a,b be positive numbers
2. a=b and b=a mod b
3. repeat 2 until b=0
4. gcd=a

iteration	a	b
1	15	12
2	12	3
3	3	0

iteration	a	b
1	49	84
2	84	49
3	49	35
4	35	14
5	14	7
6	7	0

```
###############################################
#
# _euclid_alg_gcd
#
# Author: Keith Schubert
# Date  : Nov 4, 2005
# Desc  : greatest common divisor
# Method: Euclid's Algorithm, recursive
#         var       offset
# Return: gcd       _euclid_alg_gcd_off_gcd
# RetAdd: ra        _euclid_alg_gcd_off_ra
# Params: a         _euclid_alg_gcd_off_a
#         b         _euclid_alg_gcd_off_b
# Pre   :
# Post  : contents of $t0-$t8 changed, $ra changed
#
##############################################
_euclid_alg_gcd:
```

21.6.1 Towers of Hanoi

Implement a recursive function to solve the towers of Hanoi in MIPS.

```
#
# hanoi
#
# Frame: Return address
#        *Answer
#        Answer Size
```

```
#        Number of disks
#        Free
#        Destination
#        Source
.set hanoi_off_ra=0
.set hanoi_off_ans=4
.set hanoi_off_size=8
.set hanoi_off_num=12
.set hanoi_off_free=16
.set hanoi_off_dest=20
.set hanoi_off_source=24
.set hanoi_allocate=-28
.set hanoi_deallocate=28
.set newline="\n"
.set arrow=">"
hanoi:
  lw $t0,hanoi_off_num($t0)
  subi $t0,$t0,1
  blez $t0,done
  #
  # move stack-1 to free
  mov $fp,$sp
  addiu $sp,$sp,hanoi_allocate
  sw $t0,hanoi_off_num($sp)    # num-1
  lw $t0,hanoi_off_ans($fp)    # same string
  sw $t0,hanoi_off_ans($sp)
  lw $t0,hanoi_off_size($fp)   # same size
  sw $t0,hanoi_off_size($sp)
  lw $t0,hanoi_off_free($fp)   # new dest=free
  sw $t0,hanoi_off_dest($sp)
  lw $t0,hanoi_off_dest($fp)   # new free=dest
  sw $t0,hanoi_off_free($sp)
  lw $t0,hanoi_off_source($fp) # source same
  sw $t0,hanoi_off_source($sp)
  la $t0,back1                 # return address
  sw $t0,hanoi_off_ra($sp)
  j hanoi
  back1:
  #don't deallocate yet, we are calling another in a sec
  #
  # store "source>dest\nNull"
  lw $t1,hanoi_off_ans($sp)
  lw $t0,hanoi_off_size($sp)
  add $t1,$t1,$t0
  lw $t2,hanoi_off_source($sp)
  sb $t2,0($t1)
  li $t2,arrow
  sb $t2,1($t1)
```

```
  lw $t2,hanoi_off_dest($sp)
  sb $t2,2($t1)
  li $t2,newline
  sb $t2,3($t1)
  sb $zero,4($t1)
  addi $t0,$t0,4
  sw $t0,hanoi_off_size($sp)
  #
  # move stack-1 to dest
  lw $t0,hanoi_off_dest($fp)     # same dest
  sw $t0,hanoi_off_dest($sp)
  lw $t0,hanoi_off_source($fp)   # new free=source
  sw $t0,hanoi_off_free($sp)
  lw $t0,hanoi_off_free($fp)     # new source=free
  sw $t0,hanoi_off_source($sp)
  la $t0,back2                   # return address
  sw $t0,hanoi_off_ra($sp)
  j hanoi
  back2:
  addiu $sp,$sp,hanoi_deallocate
  lw $ra,hanoi_off_ra($sp)
  jr $ra

done:
  # store "source>dest\nNull"
  lw $t1,hanoi_off_ans($sp)
  lw $t0,hanoi_off_size($sp)
  add $t1,$t1,$t0
  lw $t2,hanoi_off_source($sp)
  sb $t2,0($t1)
  li $t2,arrow
  sb $t2,1($t1)
  lw $t2,hanoi_off_dest($sp)
  sb $t2,2($t1)
  li $t2,newline
  sb $t2,3($t1)
  sb $zero,4($t1)
  addi $t0,$t0,4
  sw $t0,hanoi_off_size($sp)
  lw $ra,hanoi_off_ra($sp)
  jr $ra
```

21.6.2 Tracing Code

The code that follows, implements the algorithm

$$n_{k+1} = \begin{cases} 3n_k + 1 & \text{if } n_k \text{ is odd} \\ \frac{n_k}{2} & \text{if } n_k \text{ is even} \end{cases}$$

in MIPS. Trace the code by showing how the register values change. What is the value that is returned? Note: this code is a somewhat famous problem in number theory. The problem is to prove that starting at any number, the algorithm will bring you to 1.

```
!   code                    $t0  |  $a0   |  $v0
!                                |   3    |
!---------------------------------------------
secret:                     !    |        |
  bgtz $a0, ok              !    |        |
  break $zero               !    |        |
ok:                         !    |        |
  addi $v0,$zero,1          !    |        |
  subi $t0,$a0,1            !    |        |
  beqz $t0, end             !    |        |
loop:                       !    |        |
  addi $v0,$v0,1            !    |        |
  andi $t0,$a0,1            !    |        |
  beqz $t0, even            !    |        |
  sll  $t0,$a0,1            !    |        |
  add  $a0,$a0,$t0          !    |        |
  addi $a0,$a0,1            !    |        |
  b loop                    !    |        |
even:                       !    |        |
  sra  $a0,$a0,1            !    |        |
  subi $t0,$a0,1            !    |        |
  bgtz $t0, loop            !    |        |
end:
```

I will show changes on successive loops by placing a comma and then the new value

```
#   code                    $t0               |  $a0              |  $v0
#                                             |   3               |
#-------------------------------------------------------------------------------
secret: bgtz $a0, ok      #                  | 3                 |
        break $zero       #                  |                   |
ok:     addi $v0,$zero,1  #                  | 3                 | 1
        subi $t0,$a0,1    # 2                | 3                 | 1
        beqz $t0, end     # 2                | 3                 | 1
loop:   addi $v0,$v0,1    # 2,6,4 ,10,7,3,1| 3,10,5 ,16,8,4,2| 2,3,4,5,6,7,8
        andi $t0,$a0,1    # 1,0,1 ,0 ,0,0,0| 3,10,5 ,16,8,4,2| 2,3,4,5,6,7,8
        beqz $t0, even    # 1,0,1 ,0 ,0,0,0| 3,10,5 ,16,8,4,2| 2,3,4,5,6,7,8
        sll  $t0,$a0,1    # 6  ,10           | 3    ,5           | 2   ,4
        add  $a0,$a0,$t0  # 6  ,10           | 9    ,15          | 2   ,4
        addi $a0,$a0,1    # 6  ,10           | 10   ,16          | 2   ,4
        b loop            # 6  ,10           | 10   ,16          | 2   ,4
even:   sra  $a0,$a0,1    #   0    ,0 ,0,0,0|   5     ,8 ,4,2,1|   3   ,5,6,7,8
        subi $t0,$a0,1    #   4    ,7 ,3,1,0|   5     ,8 ,4,2,1|   3   ,5,6,7,8
        bgtz $t0, loop    #   4    ,7 ,3,1,0|   5     ,8 ,4,2,1|   3   ,5,6,7,8
end:
```

Returns 8.

Chapter 22

Data Transfer

22.1 I/O

Transmission of data from one device to another is the essence of I/O. Usually, I/O is accomplished by defining registers to hold the information necessary to transmit the data. The registers that handle the transmission are called the I/O port. At least three registers are used, one for the data, one for the control, and one for the Status.

Data the codes to be transmitted. These can be traditional codes, such as ASCII, or even an address of data being requested.

Control the commands specifying what is to be done.

Status a series of bits specifying what is going on with the bus and the current transaction.

Accessing the registers (reading from or writing to) can be accomplished in two ways.

Memory Mapped the registers of the I/O port, have addresses in regular memory, and thus can be treated as a regular memory location for access purposes.

Isolated the registers are in a separate (isolated) memory address scheme, and thus the memory must be access through special commands.

22.2 Busses

Internal vs. External (relative to cpu)
Master/Slave (initiator/target)

(Transaction) Master the initiator of a transaction.

(Transaction) Slave the target of a transaction.

Bus Master any device that can be a (transaction) master.

Burst Mode Transaction transaction which transmits several values.

Bus Transaction data transfer on an external bus.

Synchronous Bus Lines

Line/Signal	Num	Owner
Clock	1	Bus
Start	1	Master
Address	k	Master
$R/\overline{W}$	1	Master
Data	n	Master/Slave
Done	1	Slave

Arbitration is usually overlapped

22.2.1 Synchronous/Asynchronous Transfer

Busses have to have a way to specify when to transfer and if data has been received. The two basic schemes for transfer is synchronous and asynchronous.

Synchronous transfers uses a clock signal to coordinate communication, and is thus very fast. For a data request, we only need to spend one bus cycle to sent the request, the access time to find the data, and one bus cycle to send the answer. The time to transmit the data is thus

$$T_{transmit} = \frac{2}{f_{bus}} + T_a,$$

were T_a is the time to access the data, and f_{bus} is the bus clock rate[1]. The faster the clock the less time to transmit the data. The bandwidth of the bus in terms of transactions is

$$BW_{transaction} = \frac{W_{bus}}{T_{transmit}},$$

where W_{bus} is the width of the bus[2]. Frequently however, buses are measured not by an actual transaction but by what a one way message would be

$$\begin{aligned} BW &= \frac{W_{bus}}{T_{bus}} \\ &= W_{bus} f_{bus}. \end{aligned}$$

Let's consider a few examples. Note that we will be reporting bandwidth in megabytes per second (MB/s). A byte is 8 bits, and a megabyte is 2^{20} bytes. Bus frequencies (sometimes called speeds) are reported in megaHertz (MHz), but here mega is in base 10 not base 2, so it is 10^6 Hertz. Recall a Hertz is a reciprocal second. Sometimes this distinction is ignored to simplify calculations.

Example 22 (PCI) *A basic PCI bus is 32 bits wide (4 bytes) and runs at 33.3 MHz. Thus the bandwidth is*

$$\begin{aligned} BW &= W_{bus} f_{bus} && (22.1) \\ &= \left(4[B]\frac{1[MB]}{2^{20}[B]}\right)\left(33.3[MHz]\frac{10^6[Hz]}{1[MHz]}\right) && (22.2) \\ &= \left(\frac{1}{2^{18}[MB]}\right)(3.33 \times 10^7[Hz]) && (22.3) \\ &= \left(\frac{1}{2^{18}[MB]}\right)(3.33 \times 10^7[Hz]) && (22.4) \\ &\approx 127[MB/s] && (22.5) \end{aligned}$$

[1] A one way transmission must finish in this time.

[2] How much data can be sent simultaneously, i.e. the number of wires measured in bits or bytes. A bus that has 32 data wires is 32 bits wide or 4 bytes wide.

Clock signals take time to transfer down the wire and thus is subject to clock skew. To understand clock skew, consider a simple example of two clocks 3 kilometers apart. The clocks are synchronized by a beam of light, which travels at 3×10^5 km/s, and thus it takes $10\mu s$ for the synchronization pulse to arrive from the master clock. If the clocks were only synchronized once per second the fraction of the synchronization time used to transmit the pulse would be $\frac{10\mu s}{1s} = .001\%$, which is basically insignificant. What if we wanted to synchronize the clocks every tenth of a milisecond (.1ms)? The fraction of time to transfer now is $\frac{10\mu s}{.1ms} = 10\%$, which is very significant. When the clock pulse arrives it is off by 10%! That is called clock skew, when the transmission time of the clock pulse takes a significant portion of the clock frequency. Clock skew is effected by the distance (d) and the clock rate (f). If the clock skew is some fraction (s) and we assume that the clock signal is carried at the speed of light (c) then the relation between the variables is

$$\frac{d}{c} = \frac{s}{f}$$

Assuming we want the skew to be less than a third ($s = .33\ldots$), the distance is measured in meters and the bus clock will be measured in megahertz, then

$$df = 100.$$

In other words a 100MHz bus (f=100) can only be 1 meter long (d=1) to keep clock skew under 33.3%! Given that bus speeds of 400MHz are very reasonable, this would limit bus length to about 9in. Thus we see that clock skew limits bus length, and thus synchronous buses are fast but short.

Asynchronous transfers get around the problem of clock skew by doing a procedure called handshaking. Basically two units that want to talk send messages back and forth letting each other know what is going on. A basic handshaking protocol between a sender (S) and a receiver (R) to request data from R is

1. S to R: Here is the address of the data I want.
2. R to S: I got your request and will look it up.
3. S: Drop request when recieve
4. R: looking up data.
5. R to S: Here is your data.
6. S to R: I got it.
7. S: Wait till see data signal drop then drop acknowledgement.

Call the time for the signal to travel from sender to receiver or vice versa T_h (for handshake time), and the time to get the data as T_a(for access time). If we are clever we can overlap items 2,3 with item 4, so that we will only take the longer of $2T_h$ or T_a rather than $2T_h + T_a$. The total time for one transfer is thus

$$T_{transfer} = 4T_h + \max(2T_h, T_a).$$

The bandwidth of the bus is the rate at which data can be sent, and thus

$$BW = \frac{W_{bus}}{T_{transfer}},$$

where W_{bus} is the width of the bus.

22.2.2 Polling and Interrupts

There are two basic ways to handle bus communication with the CPU: polling, interrupts. Direct Memory Access (DMA) is a special case of interrupts.

Polling - CPU Controlled Data Transfer

$$\begin{aligned}
\text{Fraction of CPU Time} &= \frac{\text{Cycles Per Second used on Polls}}{\text{Clock Frequency}} \\
&= \frac{\frac{\text{Polls}}{\text{Sec}}\frac{\text{Cycles}}{\text{Poll}}}{\text{Clock Frequency}} \\
&= \frac{\frac{\text{Data Rate}}{\text{Poll Size}}\frac{\text{Cycles}}{\text{Poll}}}{\text{Clock Frequency}}
\end{aligned}$$

Interrupt Driven - CPU Controlled Data Transfer

$$\begin{aligned}
\text{Fraction of CPU Time} &= \frac{\text{Cycles Per Second used on Interrupts}}{\text{Clock Frequency}} \\
&= \frac{\frac{\text{Interrupts}}{\text{Sec}}\frac{\text{Cycles}}{\text{Interrupts}}}{\text{Clock Frequency}} \\
&= \frac{\frac{\text{Data Rate}}{\text{Packet Size}}\frac{\text{Cycles}}{\text{Interrupt}}}{\text{Clock Frequency}}
\end{aligned}$$

Interrupt Driven - Direct Memory Access (DMA)

$$\begin{aligned}
T_{\text{Transfer}} &= \frac{\text{Size Transfer}}{\text{Speed Transfer}} \\
&= \frac{\text{Data Size}}{\text{Data Rate}} \\
\text{Cycles to Handle} &= C_h \\
&= \frac{\text{Cycles to Start} + \text{Cycles to Complete} + f_e \times \text{Cycles to handle errors}}{1 - f_e} \\
\text{Fraction of CPU Time} &= \frac{\text{Cycles Per Second used to handle DMA}}{\text{Clock Frequency}} \\
&= \frac{\frac{C_h}{T_{\text{Transfer}}}}{\text{Clock Frequency}} \\
&= \frac{C_h}{T_{\text{Transfer}}\text{Clock Frequency}}
\end{aligned}$$

Example

You are given a 32-bit **asynchronous** bus with a handshaking time of 15 ns. Your computer has the following equipment attached:

Hard Drive	RAM
Total Latency: 7.2 ms	Access Time: 40ns
Disk Transfer Rate: 10MB/s	No Burst Mode
Number of Disks: 4	

Showing all work calculate the following:

1. the band width of the bus,
2. the percent of the bus utilized by continuous paging of a virtual memory system with 32KB pages,
3. the number of cache to RAM transfers that can occur if: The bus is continuously paging and 10% of the bandwidth must be left for other transactions (Hint: calculate the available bandwidth for the RAM transactions and use the size of the transactions).

The bandwidth of the bus is:

$$\begin{aligned}
\text{BW} &= \frac{\text{Data Transfered}}{\text{Time to Transfer}} \\
&= \frac{\text{Bus Width}}{4T_{Hand} + \max 2T_{Hand}, T_{RAM}} \\
&= \frac{4B}{4(15ns) + \max 2(15ns), 40ns} \\
&= \frac{4B}{100ns} \\
&= 40MB/s
\end{aligned}$$

The effective transfer rate of the pages from the disks is:

$$\begin{aligned}
\text{Rate}_{\text{Disk}} &= \frac{\text{Data Transfered}}{\text{Time to Transfer}} \\
&= \frac{\text{Data Transfered}}{\text{Total Latency} + \frac{\text{Data Transfered}}{\text{Combined Disk Transfer Rate}}} \\
&= \frac{32KB}{7.2ms + \frac{32KB}{4\times 10MB/s}} \\
&= \frac{32KB}{7.2ms + .8ms} \\
&= 4MB/s
\end{aligned}$$

Thus the bandwidth available to RAM is $40 - 4 - 4 = 32$ MB/s. Since each transfer is 4 B, the transfers per second is 8×10^6 transfers/sec or 1 cache miss every 125 ns.

Chapter 23

Memory and Cache

23.1 Memory

2D

2.5D

A synchronous memory bus for a system with 2^k addresses of n bit words would require at least:

- k address lines
- n data lines
- 4+ control lines

or a total of $k + n + 4$ parallel lines. See Section 22.2

Memory is usually byte-addressable, but I don't just load it one byte at a time. In a typical 2D or 2.5D RAM configuration though, if I had all of memory in one large module/array, I would only be able to access one byte at a time. To allow access to more than one byte at a time, memory is interleaved: the first byte is stored in the first location of the first module/array, the second byte in the first location of the second module/array, and so on. When all the module/arrays have their first location addressed, the second locations are specified, see Table 23.1.

Module Address	Module 1	Module 2	...	Module N
0	0	1	...	$N-1$
1	N	$N+1$	...	$2N-1$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
2^k-1	$(2^k-1)N$	$(2^k-1)N+1$	...	2^kN-1

Table 23.1: Mapping Memory Module's Addresses to the Computer's Memory Addresses

A number of potential problems can arise. Consider the four byte integer, $0x12345678$, stored starting in address 2 on a machine with four modules. In the easiest and fastest way to implement the hardware, the first byte of the returned number comes from the first module, the second byte from the second module and so on. By examining Table 23.2 you will notice that this means the value sent back is $0x56781234$ or even $0xABCD1234$ depending on how the addresses are selected!

To prevent such problems, systems adopt standards of how memory must be stored. The simplest method is justified, in which the first byte of any new memory item must start in the first module. Justified can obviously lead to some inefficiencies in memory utilization. A more sophisticated method is aligned, in which

First Byte Address	Module 1	Module 2	Module 3	Module N
0	$0xAB$	$0xCD$	$0x12$	$0x34$
1	$0x56$	$0x78$	$0x00$	$0x00$

Table 23.2: Memory Contents of Non-Aligned Integer

a new memory item must start at an address that is divisible by the number of bytes in the memory item (e.g.: a 4 byte integer can start at any address that can be expressed as $4i$ for i a non-negative integer).

23.1.1 Endian

Big (LR) and little (RL) endian

Consistent (same for bits)

Sparc is inconsistent big-endian.

Endian: Consistent

Endian				
Big	0	1	...	n
	0...7	0...7	...	0...7
Little	n	...	1	0
	7...0	7...0	...	7...0

Endian: Inconsistent

Endian				
Big	0	1	...	n
	7...0	7...0	...	7...0
Little	n	...	1	0
	0...7	0...7	...	0...7

23.2 Cache Design

In general DRAM has a cycle-time of about 50ns to 80ns, and SRAM has a cycle-time of 5ns to 20ns. Main memory is almost exclusively DRAM due to size and cost, so access will be slow. Strategies must be used to speed up access to main memory. Several common techniques are:

Wide Memory memory that passes multiple words at a time.

Interleaving memory that has successive addresses stored in different components that can be accessed simultaneously.

Prefetching buffer that fetches most likely instructions (or sometimes data) when memory is idle.

Cache data and instructions that have been accessed are stored in fast memory (SRAM) that is close to the CPU often as well as in main memory.

Usually, a variety of techniques are used, and often multiple levels of cache (l1, l2, and even l3).

Cache can be:

fully associative any main memory location can be stored in any cache location.

2^k-way set associative each main memory location must be stored in one of n prescribed cache locations. Usually, $16 \geqslant k \geqslant 1$.

direct mapping each main memory location must be stored in a particular cache location. This is the same as 1-way set associative.

Let's introduce some formalisms. Let 2^k be the associativity of the cache, 2^l be the size of a cache location (block size, usually less than 16 words), 2^m be the number of cache locations, and 2^n be the size of main memory.

Then

$$\begin{aligned}
\text{number of sets} &= m-k \\
\text{size of the cache} &= 2^{(l\times m)} \\
\#\text{ address bits inferred by location} &= m-k+l \\
\#\text{ tag address bits} &= n-(m-k+l)
\end{aligned}$$

n-(m-k+l)	m-k	l
tag address bits	set address bits	offset in block

Example: Cache for Toy Stack

Design a 4 way associative, 8 byte cache for a 64 byte system (i.e.: the Toy Stack). Show an example of how your system would do a cache lookup (ie: through all the steps for a lookup, you may pick memory and cache to have any values you want)

The numbers of our design are as follows.

- 64 bytes means 6 bit addresses
- 8 byte cache means 3 bit addresses
- 4 way associative means the high two bits of each cache address do not need to match the corresponding bits in main memory, but the least bit does.
- 5 bits of address from main memory need to be identified for each cache location, with the valid bit, this makes 6 tag bits for each cache location.
- the least significant bit of the main memory address to be checked for is used as a lookup on the cache to provided the 4 specific locations in cache that must be checked
- the 5 address tag bits of each of the 4 cache locations is compared with the high 5 bits of the main memory address.
- if any of them match and the corresponding valid bit is set then we have a cache hit and the data is sent
- if there is no match or the match is not valid main memory is accessed.

lookup

Let the address to be checked for be 010111, and let the cache be

Tag Bits						Address			Contents
High Address					Valid Bit				
0	0	1	1	0	1	0	0	0	11011101
0	1	0	1	0	1	0	0	1	11010110
0	0	0	0	0	0	0	1	0	00011100
1	0	0	0	0	0	0	1	1	10010100
1	1	0	1	1	1	1	0	0	11101101
1	0	1	0	1	0	1	0	1	11011110
1	0	0	0	0	1	1	1	0	11111111
0	1	0	1	1	1	1	1	1	11010000

First, the low bit (a 1) of the address tells us to look at the 4 odd addresses in cache:

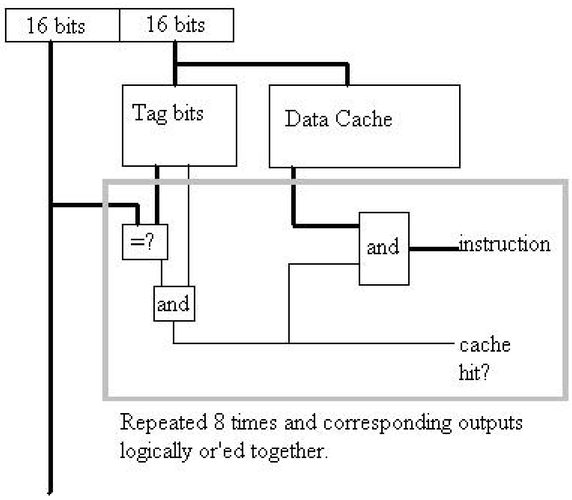


Figure 23.1: 8-Way Set Associative Cache

Tag Bits						Address			Contents
High Address					Valid Bit				
0	1	0	1	0	1	0	0	1	11010110
1	0	0	0	0	0	0	1	1	10010100
1	0	1	0	1	0	1	0	1	11011110
0	1	0	1	1	1	1	1	1	11010000

The 5 address tag bits are checked against the high five bits of the address (01011):

Tag Bits						Address			Contents
High Address					Valid Bit				
0	1	0	1	1	1	1	1	1	11010000

The address matches and the valid bit is set so 11010000 is sent as the contents.

Example: 8-way set associative

Consider a machine with 32 bit addressing (up to 4GB of RAM) and 512k (2^{19}) of data cache with 1 byte blocks. To define the 8-way set association, it will be required that main memory addresses must have the same last 16 bits (19-3=16) as a cache location to be stored in that cache location. Every cache location has 17 extra bits, 16 for addressing, and one for validity. Eight location in cache must be checked for each main memory access (it is 8-way for a reason). The main memory address to be checked is split into the upper and lower 16 bits. The lower 16 bits are used to identify the eight cache locations, whose 16 address tag bits are then compared to the 16 high bits of the main memory address, see Figure 23.1. This generates eight signals (true if match was found) that are then logically and'ed together with the corresponding 8 validity bits (might have the same address but might not be current). If any generates a hit (is true) then its contents are sent as the data.

Replacement policies

LRU Least Recently Used

FIFO First-in First-out

LFU Least Frequently Used

Random Random

23.2.1 Neat Little LRU Algorithm

Let the number of cache slots (locations) be 2^k, then we create a matrix of bits that is $2^k \times 2^k$ (so we can associate the cache address with both a row and column). Initially they are all cleared. When a cache slot, say address p, is accessed:

1. 1's are placed in every bit of the matrix row p,
2. 0's are placed in every bit of the matrix column p.

Note that the second step will delete one of the 1's you placed in the first step.

The the address that was least recently used corresponds to the number of the row that has a sum of zero. Equivalently, the address that was least recently used corresponds to the number of the column with the largest sum.

Example: Fully Associative Cache With 4 Slots

For simplicity we will assume main memory has 256 (2^8) bytes, and the data length is 1 byte. The cache starts empty.

NLLRU				Tag Bits			Data
0	1	2	3	V	D	Address	
0	0	0	0	0	0	0x00	0x00
0	0	0	0	0	0	0x00	0x00
0	0	0	0	0	0	0x00	0x00
0	0	0	0	0	0	0x00	0x00

Address 0x1A, which contains 0x49, is accessed.

NLLRU				Tag Bits			Data
0	1	2	3	V	D	Address	
0	1	1	1	1	0	0x1A	0x49
0	0	0	0	0	0	0x00	0x00
0	0	0	0	0	0	0x00	0x00
0	0	0	0	0	0	0x00	0x00

Address 0x05, which contains 0x11, is accessed.

NLLRU				Tag Bits			Data
0	1	2	3	V	D	Address	
0	0	1	1	1	0	0x1A	0x49
1	0	1	1	1	0	0x05	0x11
0	0	0	0	0	0	0x00	0x00
0	0	0	0	0	0	0x00	0x00

Address 0x25, which contains 0xFF, is accessed.

NLLRU				Tag Bits			Data
0	1	2	3	V	D	Address	
0	0	0	1	1	0	0x1A	0x49
1	0	0	1	1	0	0x05	0x11
1	1	0	1	0	0	0x25	0xFF
0	0	0	0	0	0	0x00	0x00

The value 0x33 is stored to address 0x05.

NLLRU				Tag Bits			Data
0	1	2	3	V	D	Address	
0	0	0	1	1	0	0x1A	0x49
1	0	1	1	1	1	0x05	0x33
1	0	0	1	0	0	0x25	0xFF
0	0	0	0	0	0	0x00	0x00

The value 0xF5 is stored to address 0x06.

NLLRU				Tag Bits			Data
0	1	2	3	V	D	Address	
0	0	0	0	1	0	0x1A	0x49
1	0	1	0	1	1	0x05	0x33
1	0	0	0	0	0	0x25	0xFF
1	1	1	0	1	1	0x06	0xF5

The value 0x07 is stored to address 0x07.

NLLRU				Tag Bits			Data
0	1	2	3	V	D	Address	
0	1	1	1	1	1	0x07	0x07
0	0	1	0	1	1	0x05	0x33
0	0	0	0	0	0	0x25	0xFF
0	1	1	0	1	1	0x06	0xF5

23.2.2 Cache Performance

We will be concerned with some basic numbers

Hit Ratio (HR) The number of cache hits over the number of lookups.

Miss Ratio (MR) The number of cache misses over the number of lookups.

Effective Access Time (EAT or T_{eff}) The average time spent in a memory access.

First let us consider the hit and miss ratios. For a series of lookups, the number of hits was "Hit" and the number of misses was "$Miss$", thus $Hit + Miss = lookups$. Given this,

$$\begin{aligned} HR &= \frac{Hit}{Hit + Miss} \\ MR &= \frac{Miss}{Hit + Miss} \\ 1 &= HR + MR \end{aligned}$$

thus,

$$\begin{aligned} T_{eff} &= \frac{Hit \times T_{Hit} + Miss \times T_{Miss}}{Hit + Miss} \\ &= HR \times T_{Hit} + MR \times T_{Miss}. \end{aligned}$$

Usually, the miss time is the access time (T_{Hit}), plus a miss penalty (say $T_{Penalty}$).

$$
\begin{aligned}
T_{Miss} &= T_{Hit} + T_{Penalty} \\
T_{eff} &= HR \times T_{Hit} + MR \times T_{Miss} \\
&= HR \times T_{Hit} + MR \times (T_{Hit} + T_{Penalty}) \\
&= (HR + MR) \times T_{Hit} + MR \times T_{Penalty} \\
&= T_{Hit} + MR \times T_{Penalty}
\end{aligned}
$$

Example

Use the following chart to show the state of a 4 location, 2-Way associative cache, that uses LRU. If a location has a number printed in it, the address is valid, if no number appears the contents are invalid. For simplicity the computer only has 16 locations in memory. If the cache takes 5ns to access and RAM takes 60ns, what is the effective access time given the sequence?

Time	0	1	2	3	4	5	6	7	8	9	10
Lookup Address	-	2	5	6	B	5	2	2	B	C	5
Cache location 00	A										
Cache location 01	B										
Cache location 10											
Cache location 11											

Time	0	1	2	3	4	5	6	7	8	9	10
Lookup Address	-	2	5	6	B	5	2	2	B	C	5
Cache location 00	A	A	A	6	6	6	6	6	6	C	C
Cache location 01	B	B	B	B	B	B	B	B	B	B	B
Cache location 10		2	2	2	2	2	2	2	2	2	2
Cache location 11			5	5	5	5	5	5	5	5	5

MR=.4

$$
\begin{aligned}
T_{eff} &= T_{cache} + MR(T_{RAM}) \\
&= 5ns + .4(60ns) \\
&= 29ns
\end{aligned}
$$

23.3 Virtual Memory

A 32-bit virtual memory system has a 64KB page size, and 1 GB of RAM. How large is the physical page number in bits? Assuming that the each entry in the table is word aligned, how large is the lookup table in bytes?

64KB = 2^{16}

1GB = 2^{30}

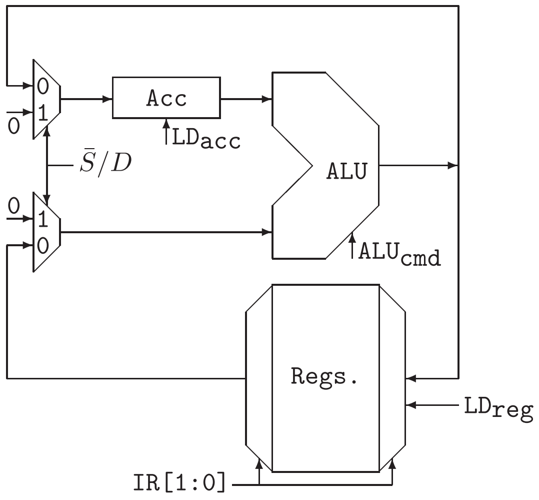
So the physical page number takes 30-16=14 bits or almost 2B to store in the table. We also need to add memory protection, ownership, validity, location, etc. I will assume that I can fit all this in 4B.

The table size is $2^{(32-16)} \times 4B = 2^{18}B = 256KB$

Chapter 24

CPU Control

24.1 Tiny Accumulator



The tiny accumulator has four commands

Mach. Code	Assem. Lang.	Description.
00MN	STC MN	Store Acc to location MN and clear Acc
01MN	ADD MN	Add Acc and location MN placing result in Acc
10MN	SUB MN	Sub location MN from Acc, placing result in Acc
11MN	BRL MN	if Acc is negative, Branch to $nPC + MN\bar{N}$

STC MN The store and clear command not only allows storage, but due to the clear, allows a load if it is followed by adding the desired value to load. The instruction is implemented as follows. The signal $\bar{S}/D$ is set to 1, which puts a zero both on the accumulator and the second input of the ALU. The ALUop is set to add, which thus does ACC plus zero, and so the value of the ACC is placed on the answer line. Both the ACC and the register file is told to read, which results in the ACC loading zero, and register M loading the value that had been in the ACC.

ADD MN This instruction makes it easy to load the ACC as mentioned in STC MN, as well as providing an arithmetic command. The instruction is implemented as follows. The signal $\bar{S}/D$ is set to 0, which allows the selected register to go to the second input of the ALU and allows the result of the ALU to go to the ACC input. the ALUop is set to add, and finally the ACC is told to load, so the result becomes stored.

SUB MN This instruction is very similar to ADD. The instruction is implemented as follows. The signal $\bar{S}/D$ is set to 0, which allows the selected register to go to the second input of the ALU and allows the result of the ALU to go to the ACC input. the ALUop is set to sub, and finally the ACC is told to load, so the result becomes stored.

BRL MN This instruction allows loops and conditional executions to be handled. The offset is taken to be a three bit, two's compliment number, of which the first two are MN and the last bit is the flip of N. While this may sound strange it makes the displacements to be

MN	$MN\bar{N}$	displacement
11	110	-2
10	101	-3
01	010	2
00	001	1

The negative numbers allow loops which include one or two instructions besides the branch, and the positive numbers allow for conditional statements of one or two instructions. Note the negative numbers are larger in magnitude by one to include the branch statement.

This gives us a full architecture that can be programmed, but is small enough to be built by hand.

24.2 GST ISA

Gomez-Schubert-Tafas Instruction Set Architecture.

My thought is to implement 1k-word of memory for each processor, and to do memory mapped IO so we don't need special commands. The word size is 16 bits and this is the smallest addressable size, again for simplicity. The "network" port should have a buffer of, say, 16 words. Initially there will not be a cache because since this will be a SOC there is no access time advantage.

The ISA is load-store. I have broken the 16 bit instruction into 4 nibbles for different purposes as seen below. I have tried to pair commands by opcode to make for easier control. I left two unused in case there is anything you want to add.

We only use register, immediate, and indexed addressing, to keep things simple and still provide flexibility. These three modes allow us to do anything.

I am only considering two's complement numbers, so no unsigned numbers. While this is a limitation for real computers, I don't think it will matter for this test architecture.

24.2.1 R Type Commands

FEDC	BA98	7654	3210
Opcode	RD	RS1	RS2

or

FEDC	BA98	7654	3210
Opcode	RD	RS1	Imm1

24.2.2 I Type commands

FEDC	BA98	76543210
Opcode	RD	Imm2

24.2.3 B Type commands

FEDC	BA9876543210
Opcode	Imm3

24.2.4 Commands

Opcode	Assembly	Comments
0000	load RD(RS1+RS2)	$RD \leftarrow M[RS1 + RS2]$
0001	store RD(RS1+RS2)	$RD \rightarrow M[RS1 + RS2]$
0010	ldi RD,Imm2	$RD[F : 8] \leftarrow Imm2$
0011		
0100	add RD,RS1,RS2	$RD \leftarrow RS1 + RS2$
0101	sub RD,RS1,RS2	$RD \leftarrow RS1 - RS2$
0110		
0111		
1000	sll RD,RS1,Imm1	$RD \leftarrow RS1 << Imm$
1001	sra RD,RS1,Imm1	$RD \leftarrow RS1 >> Imm$
1010	nand RD,RS1,RS2	$RD \leftarrow (RS1 \cdot RS2)'$
1011	nor RD,RS1,RS2	$RD \leftarrow (RS1 + RS2)'$
1100	brlt RD,RS1,Imm1	$(RD < RS1) \Rightarrow (PC \leftarrow nPC + \{Imm1[3:0], \not{I}mm1[0]\})$
1101	brle RD,RS1,Imm1	$(RD \leqslant RS1) \Rightarrow (PC \leftarrow nPC + \{Imm1[3:0], \not{I}mm1[0]\})$
1110	br Imm3	$PC \leftarrow PC + Imm3$
1111	j RD	$PC \leftarrow PC + RD$

Note: SE is sign extend.

24.2.5 Registers

0	R0	Zero	8	L0	Local Register 0
1	R1	General Purpose Register 1	9	L1	Local Register 1
2	R2	General Purpose Register 2	10	L2	Local Register 2
3	R3	General Purpose Register 3	11	L3	Local Register 3
4	R4	General Purpose Register 4	12	L4	Local Register 4
5	R5	General Purpose Register 5	13	L5	Local Register 5
6	R6	General Purpose Register 6	14	SP	Stack Pointer
7	R7	General Purpose Register 7	15	RA	Return Address

Part V

Performance

Chapter 25

Performance

25.1 Cost

$$\text{Cost of IC} = \frac{\text{Cost of die} + \text{Cost of Testing} + \text{Cost of Packaging}}{\text{Final Yield}}$$

$$\text{Cost of Die} = \frac{\text{Cost of Wafer}}{\text{Dies per Wafer} \times \text{Die Yield}}$$

$$\text{Die Yield} = \frac{WaferYield}{\left(1 + \frac{\text{Defects per Area} \times \text{Die Area}}{\alpha}\right)^{\alpha}}$$

$$\begin{aligned}
\text{List Price} &= \frac{4}{3}\text{Average Selling Price} \\
&= \frac{4}{3}\frac{4}{3}\text{Production Cost} \\
&= \frac{4}{3}\frac{4}{3}\frac{6}{5}\text{Component Cost} \\
&= \frac{32}{15}\text{Component Cost} \\
&\approx 2\text{Component Cost}
\end{aligned}$$

25.2 Power, Energy, and Heat

These are probably the most misused terms in computers (and many other fields as well). They are not synonyms and should not be used as such.

Work Electrical work is electrical force applied on a charge over a distance. Usually Electrical force is calculated by the the charge times the electrical field. For computers a computation involves moving charges from one place to another by applying a voltage, i.e.: electrical work. The work done does not change with the time it takes to do the computation. Think of it as this is what you want to do.

Energy The ability to do work. You can also consider this the cost of doing work. In a computer Energy use is primarily due to dynamic operations (switching transistors), so $E_d = C \cdot V^2$, where E_d is the dynamic energy, C is the capacitive load of the computer (consider it constant for a computer design), and V is the voltage of the computer. Energy for laptops are stored in batteries, and since this is a fixed source energy is a major issue to laptops (i.e. we care about the work done which is proportional to the computations we do).

Power The rate at which energy is used (and thus work done). We are primarily concerned with dynamic power (again from switching transistors), so $P_d = \frac{1}{2}C \cdot V^2 \cdot f_s$, where P_d is dynamic power, C is capacitive load, V is voltage, and f_s is the switching frequency, which is proportional to the clock frequency of the processor. Static power loss is caused primarily from leakage current in the transistors and thus is constant even for inactive circuits (the computer must be on of course though). Static power, P_s is given by $P_s = i_s \cdot V$, where i_c is the static current (leakage current in one transistor time the number of transistors), and V is still voltage. Static power accounts for more than 25% in current computers. Computers that have a continuous power source are more concerned with power, as power also tells us the rate of heat production. We are at the limits of air cooling, so this is a major issue.

25.3 Performance

Response Time (aka execution time) the time between the start and completion of a task.

Throughput The number of task completed in a period of time.

There are four tasks (a, b, c, and d) which are composed of four subparts (1, 2, 3, 4 for each of a, b, c, and d) that are independent (i.e. you can do a1 and a2 simultaneously). You are to run them on a four processor machine. Ignoring memory and overhead, we can schedule the processes as:

Time

Processor

	1	2	3	4
1	a1	a2	a3	a4
2	b1	b2	b3	b4
3	c1	c2	c3	c4
4	d1	d2	d3	d4

or

Time

Processor

	1	2	3	4
1	a1	b1	c1	d1
2	a2	b2	c2	d2
3	a3	b3	c3	d3
4	a4	b4	c4	d4

25.4 Time

Time can be different things. There is time that we exist in, sometimes called "wall time" due to measurements by wall clocks. There is the CPU time of the program, but even here do we mean the total time from start to finish, or just the time spent on the program without counting system functions or other programs (execution time). We will in general speak of only the execution time or CPU Time (CPUT, T_{CPU}) of the program, for simplicity.

The longer a process takes to run the worse the performance, this should be obvious as who wants a slower machine. We could also say, the less time a process takes the better the performance. Execution time and performance are thus inversely related:

$$\text{Perf} = \frac{1}{\text{Execution Time}}$$

If the performance of system A is n times better than system B then

$$\begin{aligned}\text{Perf}_A &= n\text{Perf}_B \\ \frac{\text{Perf}_A}{\text{Perf}_B} &= n.\end{aligned}$$

Alternately we note

$$\begin{aligned}\text{Perf}_A &= n\text{Perf}_B \\ \frac{1}{\text{Execution Time}_A} &= n\frac{1}{\text{Execution Time}_B} \\ \frac{\text{Execution Time}_B}{\text{Execution Time}_A} &= n.\end{aligned}$$

Putting all this together we obtain:

$$\frac{\text{Perf}_A}{\text{Perf}_B} = \frac{\text{Execution Time}_B}{\text{Execution Time}_A}.$$

25.5 Measuring CPU Time

$$\begin{aligned}CPUT &= \#\text{ cycles} \times \text{cycle time} \\ &= \#\text{ cycles} \times \frac{1}{\text{cycle rate}}\end{aligned}$$

Cycle rate is easily known for a machine so only the # cycles is needed.

25.5.1 First Approximation

$$\begin{aligned}\#\text{ cycles} &= \#\text{ instruct} \times \frac{\#\text{ cycles}}{\#\text{ instruct}} \\ &= IC \times \text{CPI}\end{aligned}$$

CPI is the cycles per instruction, and IC is the instruction count. It can be measured on average for a running program, and theoretical predictions of it can be made fairly easily.

25.5.2 Second Approximation

CPI for different types of instructions are different. For instance, arithmetic instructions like addition are usually much faster than memory access instructions.

$$\begin{aligned}
\#\text{ cycles} &= IC_{total} CPI_{avg} \\
&= IC_{total} \sum_{i=1}^{n} f_i \times \text{CPI}_i \\
&= IC_{total} \sum_{i=1}^{n} \frac{IC_i}{IC_{total}} \times \text{CPI}_i \\
&= \sum_{i=1}^{n} IC_i \times \text{CPI}_i
\end{aligned}$$

where f_i is the frequency of instruction type i. These frequencies can be measured for a large number of software packages to give typical results.

Consider, for example, a program that executes 50,000 instructions running on a machine that is typified by

	ALU	Branch	Memory
CPI	1	3	4
freq	0.5	0.2	0.3

In this case the average CPI of the machine would be given by

$$\begin{aligned}
CPI_{avg} &= \sum_{i=1}^{n} f_i \times \text{CPI}_i \\
&= .5 \times 1 + .2 \times 3 + .3 \times 4 \\
&= .5 + .6 + 1.2 \\
&= 2.3
\end{aligned}$$

It is interesting to note that memory accounts for more of the CPI than the other two combined, and branching accounts for more than ALU operations even though there are over twice as many ALU operations.

25.6 Amdahl's Law

The performance difference between two machines, or two configurations of the same machine for that matter, can be compared by setting them as a ratio as we have seen. Let's refer to the performance difference of the two machines as the speedup (S). From what we have seen we can write for two machines a and b that

$$\begin{aligned}
S &= \frac{P_a}{P_b} \\
&= \frac{T_b}{T_a} \\
&= \frac{IC_b CPI_b \frac{1}{\text{cycle rate}_b}}{IC_a CPI_a \frac{1}{\text{cycle rate}_a}} \\
&= \frac{IC_b CPI_b \text{cycle rate}_a}{IC_a CPI_a \text{cycle rate}_b}
\end{aligned}$$

Now, let's assume that we are dealing with two versions of the same machine, one enhanced and one not enhanced. If the time of the original code was $T_{original}$, and the instructions that would be speed up by the enhancement took up a fraction, f of the original time and resulted in that portion be completed in $\frac{1}{S_{enhanced}}$ the time, then

$$T_{enhanced} = T_{original}\left((1-f) + f\frac{1}{S_{enhanced}}\right).$$

The speedup, per the second form above is

$$\begin{aligned}
S_{overall} &= \frac{T_{original}}{T_{enhanced}} \\
&= \frac{T_{original}}{T_{original}\left((1-f) + f\frac{1}{S_{enhanced}}\right)} \\
&= \frac{1}{(1-f) + \frac{f}{S_{enhanced}}}
\end{aligned}$$

This result can be extended to cover many enhancements, say n of them.

$$S = \frac{1}{(1 - \sum_{i=1}^{n} f_i) + \sum_{i=1}^{n} \frac{f_i}{S_i}}$$

25.6.1 Alternate Approach

We could have assumed that the enhanced time took $T_{enhanced}$, and that the instructions using the enhanced mode took up a fraction g of the enhanced time. If the speedup of the enhanced mode was still $S_{enhanced}$ then

$$T_{original} = T_{enhanced}\left((1-g) + gS_{enhanced}\right)$$

We can relate f and g by noting that

$$\begin{aligned}
T_{enhanced}gS_{enhanced} &= T_{original}f \\
gS_{enhanced} &= fS_{overall}
\end{aligned}$$

By observing that $S_{overall} \leqslant S_{enhanced}$, with strict inequality if $S_{enhanced} > 1$, we find that $g \leqslant f$, with strict inequality for the same condition. Alternately, we could note that

$$\begin{aligned}
T_{enhanced}(1-g) &= T_{original}(1-f) \\
1-g &= (1-f)S_{overall} \\
1-g &= S_{overall} - gS_{enhanced} \\
S_{overall} &= (1-g) + gS_{enhanced}
\end{aligned}$$

An alternate way of finding the overall speedup is by using the formula for speedup directly.

$$\begin{aligned}
S_{overall} &= \frac{T_{original}}{T_{enhanced}} \\
&= \frac{T_{enhanced}\left((1-g) + gS_{enhanced}\right)}{T_{enhanced}} \\
&= (1-g) + gS_{enhanced}
\end{aligned}$$

Since the speedup must be the same, we can also find a formula to calculate the speedup for the enhanced portion in terms of just f and g.

$$\begin{aligned}
(1-g) + gS_{enhanced} &= \frac{1}{(1-f) + \frac{f}{S_{enhanced}}} \\
((1-g) + gS_{enhanced})\left((1-f) + \frac{f}{S_{enhanced}}\right) &= 1 \\
1 - g - f + fg + (1-g)\frac{f}{S_{enhanced}} + (1-f)gS_{enhanced} + fg &= 1 \\
g(S_{enhanced} - 1) + f\left(\frac{1}{S_{enhanced}} - 1\right) &= fg\left(S_{enhanced} - 1 + \frac{1}{S_{enhanced}} - 1\right) \\
&= fg(S_{enhanced} - 1) + fg\left(\frac{1}{S_{enhanced}} - 1\right) \\
g(1-f)(S_{enhanced} - 1) &= f(1-g)\left(1 - \frac{1}{S_{enhanced}}\right) \\
g(1-f)(S_{enhanced} - 1) &= f(1-g)\frac{S_{enhanced} - 1}{S_{enhanced}} \\
g(1-f)S_{enhanced} &= f(1-g) \\
S_{enhanced} &= \frac{f}{1-f}\frac{1-g}{g} \\
S_{enhanced} &= \frac{f}{g}S_{overall}
\end{aligned}$$

We can thus calculate the overall speedup a number of ways

$$\begin{aligned}
S_{overall} &= S_{enhanced}\frac{g}{f} \\
&= \frac{1-g}{1-f} \\
&= (1-g) + gS_{enhanced} \\
&= \frac{1}{(1-f) + \frac{f}{S_{enhanced}}}
\end{aligned}$$

Consider, for example, that on an unenhanced machine a piece of code runs in 10 seconds, and the instructions that could have used the enhanced mode (were it available) took up 6 seconds of that time. On an enhanced machine the same code uses the enhanced mode for a total of 1 second of the time. What is f and g? What is the speedup of the enhancement and the overall system?

We can find f directly.

$$\begin{aligned}
f &= \frac{6sec}{10sec} \\
&= 0.6
\end{aligned}$$

We can find g by noting that the original code has 4 seconds that are not speed up, so the total time after must be 5 seconds.

$$\begin{aligned}
g &= \frac{1sec}{5sec} \\
&= 0.2
\end{aligned}$$

If you did not make this observation you could have first found the speedup of the enhanced mode and used it to find g. The speedup of the enhancement is simple, given this information.

$$\begin{aligned} S_{enhanced} &= \frac{6sec}{1sec} \\ &= 6 \end{aligned}$$

Using this, we could have found

$$\begin{aligned} S_{enhanced} &= \frac{f}{1-f}\frac{1-g}{g} \\ 6 &= \frac{0.6}{0.4}\frac{1-g}{g} \\ 4 &= \frac{1-g}{g} \\ 5g &= 1 \\ g &= 0.2 \end{aligned}$$

The same we found before. The overall speedup is equally easy to get, by a bunch of ways.

$$\begin{aligned} S_{overall} &= \frac{T_{original}}{T_{enhanced}} \\ &= \frac{10sec}{5sec} \\ &= 2 \end{aligned}$$

Or

$$\begin{aligned} S_{overall} &= S_{enhanced}\frac{g}{f} \\ &= 6\frac{.2}{.6} \\ &= 2 \end{aligned}$$

Or

$$\begin{aligned} S_{overall} &= \frac{1-g}{1-f} \\ &= \frac{1-.2}{1-.6} \\ &= \frac{.8}{.4} \\ &= 2 \end{aligned}$$

Or

$$\begin{aligned} S_{overall} &= (1-g) + gS_{enhanced} \\ &= (1-0.2) + 0.2 \times 6 \\ &= 0.8 + 1.2 \\ &= 2 \end{aligned}$$

Or

$$\begin{aligned}
S_{overall} &= \frac{1}{(1-f)+\frac{f}{S_{enhanced}}} \\
&= \frac{1}{(1-0.6)+\frac{0.6}{6}} \\
&= \frac{1}{0.4+0.1} \\
&= \frac{1}{0.5} \\
&= 2
\end{aligned}$$

As you can see, it doesn't matter which formula you use, they all give the same answer. You should also notice that if you improve the enhanced mode more, you will gain almost nothing in the overall speedup. For example consider allowing $S_{enhanced} = \infty$, then

$$\begin{aligned}
S_{overall} &= \frac{1}{(1-f)+\frac{f}{S_{enhanced}}} \\
&= \lim_{x\to\infty} \frac{1}{(1-0.6)+\frac{0.6}{x}} \\
&= \frac{1}{0.4} \\
&= 2.5
\end{aligned}$$

In this case $g = 0$ so some of the equations have the indeterminate form $0 \times \infty$, which we avoid by using a form that does not have this problem. The really big thing to see though is that even a huge increase in the speedup of the enhanced mode made little difference, because the non-enhanced portions are dominating. This brings up one of the most basic interpretations of Amdahl's Law, always improve the most common case.

25.6.2 Relating the CPIs

Assuming we are dealing with enhancements to a machine, it is thus reasonable that the code length would not change, so $IC_a = IC_b$. Additionally we will assume it is not a trivial improvement of increasing the clock speed, so cycle rate_a = cycle rate_b. Thus

$$\begin{aligned}
S &= \frac{CPI_{original}}{CPI_{enhanced}} \\
CPI_{enhanced} &= CPI_{original}\left((1-\sum_{i=1}^{n} f_i)+\sum_{i=1}^{n}\frac{f_i}{S_i}\right)
\end{aligned}$$

Without changing the clock or reducing instructions, we can then find that the maximum speedup possible for a single issue system is $CPI_{original}$, since the ideal CPI for a single issue system is 1.

25.7 Putting It All Together

Example

You are to select a compiler to develop applications for a company with two types of computers. The company wants the best average performance with both machines. Assume all the machines are 1GHz machines.

Type	CPI 1	CPI 2	Compiler 1	Compiler 2
Arithmetic	1	1	35%	30%
Branch	6	3	25%	20%
Memory	3	5	40%	50%

If the code is 10000 lines (for either compiler) when assembled how long does it take to run on each machine?

	Compiler 1	Compiler 2
Machine 1	$1 \times .35 + 6 \times .25 + 3 \times .4 = 3.05$	$1 \times .3 + 6 \times .2 + 3 \times .5 = 3$
Machine 2	$1 \times .35 + 3 \times .25 + 5 \times .4 = 3.1$	$1 \times .3 + 3 \times .2 + 5 \times .5 = 3.4$
Average	3.075	3.2

Since time is the inverse of performance, we want the lowest average and ergo pick compiler 1. If each command runs only once (a bad assumption in reality but we will use it for now), the code will run in:

machine 1: $\frac{10000 \times 3.05}{10^9} = 3.05 \times 10^{-4}$ seconds.

machine 2: $\frac{10000 \times 3.1}{10^9} = 3.1 \times 10^{-4}$ seconds.

Chapter 26

Instruction Level Parallelism

26.1 Trouble In Paradise

There are three types of hazards we can encounter.

Structural hardware cannot support the instruction combo. Big problem in multi-cycle execution, out of order execution, and superscalar, but it can also happen in simple pipelines with things like memory access. Fixing this requires hardware design.

Data data is not available to proceed. Typical solutions fall into two categories, wait till the answer is here or send the answer from where it is now. These are discussed more below.

Control at branch, which do I take and how can I rearrange code around branches in dynamic execution?

26.1.1 Data Hazards

Dependence	Hazard	Example	When
True (data)	RAW	add r2,r3,r4 add r5,r2,r6	When: read happens before the write can finish Requires: pipelining (without forwarding), multi-cycle units, out of order execution, etc.

Output
(name)
	WAW	add r2,r3,r4
		brgtz r7, label
		add r2,r5,r6
			When: instructions finish out of order.
			Requires: out of order execution or multiple can multi-cycle execution units.

Antidependence (name)	WAR	add r3,r2,r4 add r2,r5,r6	When: instructions start out of order. Requires: out of order execution
None	RAR	add r3,r2,r4 add r5,r2,r6	There is no problem here, and it is not a hazard. I put it in because people kept asking.

Read after write (RAW) data hazards are also called true dependence or data dependence, because the second instruction actually needs the result from the first. It is the strongest dependence in the sense that it cannot be broken - the second instruction must have the result of the first instruction. Since it is so fundamental, it is the easiest to have happen. RAW occurs when the second instruction tries to access a result before it has been written by the first instruction. This commonly occurs in pipelines, as there are typically multiple cycles after the execute cycle completes till the result is updated in the registers. Each cycle of delay till the update could cause an instruction being decoded to access the wrong value. The two most common solutions to this problem are slips and register forwarding, though register renaming will also handle it (explained in subsection 26.1.2.

Write after write (WAW) hazards is the second most easy data hazard to generate, but the last most people think about. Usually people look at this and wonder if this can ever be a problem. This is actually the most dangerous data hazard in terms of potential to harm your results. Most machines today allow instructions to finish out of order, either by starting out of order, or because some instructions are slower and the fast ones are allowed to pass. If two instructions finish out of order and are writing to the same register, then we have a WAW hazard. The severity of the problem is caused by the number of instructions that are impacted. Normally, the first instruction would finish and its result would be available for use till the second one finished in which case the second answer would be available from then on. When a WAW hazard occurs, the second one finishes first and its result is available in the intermediate time, then the first ones result is available from then on. Unlike a RAW hazard which impacts one instruction (and those dependent on it), WAW can effect many instructions (and those dependent on them). The entire problem is based on the output so it is often called an output dependence. The problem is also due to the reuse of a register for different values, so it is called name dependence (it depends on the register name you picked). It can be fixed by a reorder buffer or register renaming.

Write after read (WAR) hazards are the hardest to occur, and have a small impact, but seem to make reasonable sense to most people. They occur when instructions start out of order causing one instruction to read the result of an instruction that was supposed to happen after it. It can only happen with out of order execution units, and it only effects the instruction that did the read (and those which use its results - but this is true of all data hazards). The dependence is in reverse order so it is sometimes called anti-dependence, but it is also based on reuse of a register so it is also considered a name dependence as WAW is. Both reorder buffers and register renaming will work to solve WAR hazards. The most commonly known algorithm for solving this problem is by Tomasulo and is covered in chapter 28.

26.1.2 Hazard Solutions

What can we do with data hazards. Remove all performance measures and execute single instructions slowly. I'm not kidding, it will work for all problems. The problems are challenges that come from performance improvements, so if you are willing to run non-pipelined, single threaded, non-superscalar processors at a few hundred megahertz you will never hit one of these problems. Your performance will stink, you won't be able to play modern games or movies, but you won't have any problems. Most people want speed, and so we have to come up with other solutions. Here are some of the most famous.

- register interlocking

  This is basically a stop until the data is available. Two variety exists

  Stall Entire processor is held for an instruction (or more), particularly important for structural hazards such as multi-cycle units or memory operations, since the units between the pipeline buffer registers keep running, and thus can finish what they are doing. Essentially this is like slowing the clock down when you need to. This tends to kill performance, but it avoids errors. Stalling will not solve the problems register forwarding will. It is the easiest method to implement.

  Slip only the held-up instruction and those after must wait, others can proceed. Note it could be one of these that produces the desired answer, so this handles the same problems as forwarding, and can handle the problems that stalling does. Overall it is the most versatile (it handles everything stalling and forwarding does), but it is not the fastest solution (same as stalling on performance). It is the second easiest to implement.

- register forwarding

  Often the value exists, it is just not in the final destination yet. This technique sends the value that is missing, to the execution unit. There is no delay if you can do it. It cannot handle multi-cycle execution or memory accesses, and it adds cost and complexity to the design (though not bad for what you get). This is straightforward to implement, but does add several multiplexors, wires, and control circuits to track where the result is (comparators or counters are common).

- register renaming

  Used to solve WAR and WAW hazards. Register renaming adds a status field to each register, which contains the address of the instruction that is calculating its current value or 0, which means it has the correct value. Instructions are fetched and issued in order, so the registers have the correct values in the status field, but are then buffered and executed when the system is ready (kind of like giving them a number and sticking them in a waiting room). It can do almost anything (it can't handle control hazards). The most basic (and famous) of these algorithms is Tomasulo's algorithm, see chapter 28.

- reorder buffer

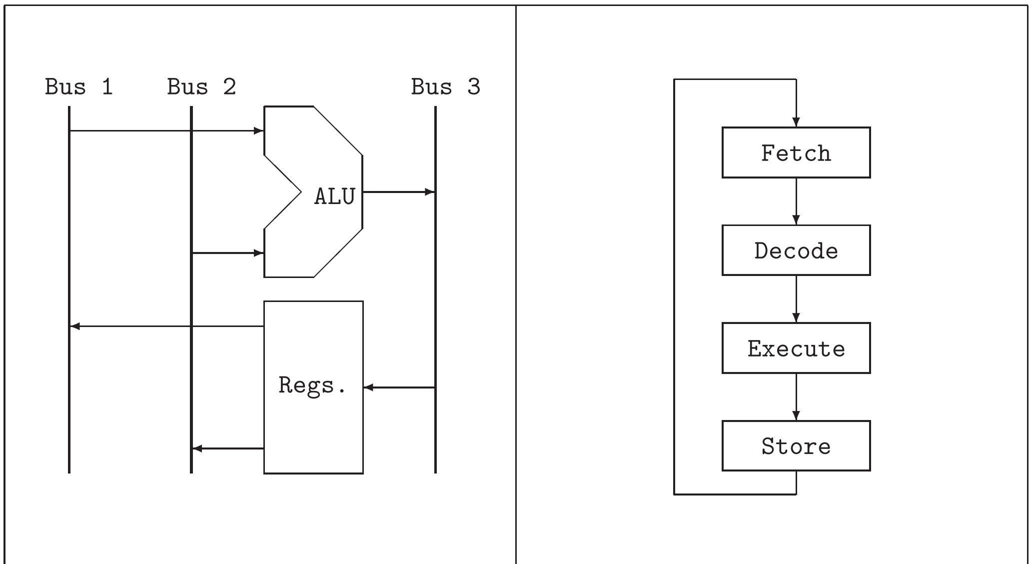
  Instructions are held in a buffer for writing to the register files, then they are written in the order of the original code. These are different buffers than the pipeline buffers. This preserves the order of the writes and thus solves WAR and WAW hazards, but increases the latency of the instruction execution. On the bright side it can handle control hazards (the only one listed that can).

Chapter 27

Pipelining

27.1 Basic Architecture

Consider the following architecture.



Architecture Fetch/Execute

The architecture and Fetch/Execute loop, lend themselves to a four stage pipeline. We will make each of the stages in the Fetch/Execute loop to be a stage in our pipeline.

Use registers at boundaries of hardware portions that do the stages of the IFetch (more fully to separate the clock cycles).

27.1.1 Calculating efficiency

Our basic equations of pipeline performance are

$$
\begin{aligned}
\text{speedup} &= \frac{T_{\text{original}}}{T_{\text{modified}}} \\
\text{efficiency} &= \frac{\text{actual speedup}}{\text{ideal speedup}}
\end{aligned}
$$

Consider m instructions running on a computer with n stages. If this is not pipelined then the time of execution will take $T_{nopipe} = m \times n \times T_{clock}$. To get this we just used that $T = \#\, cycles \times T_{clock}$. If it is pipelined then the execution will take $T_{pipe} = (m + n - 1) \times T_{clock}$. To see why consider this for m¿¿n (the usual case)

	Cycle								
Instruction	0	1	...	n-1	n	...	m-1	...	m+n-1
Inst 1	x	x	...	x					
Inst 2		x	...	x	x				
$\vdots$			$\ddots$	$\ddots$	$\ddots$	$\ddots$			
$\vdots$						$\ddots$	$\ddots$	$\ddots$	
Inst m							x	...	x

Using this we can find that as the speedup of pipelining for m instructions in an n stage machine as m gets very large (long program run) is

$$
\begin{aligned}
\text{speedup} &= \frac{T_{nopipe}}{T_{pipe}} \\
&= \lim_{m\to\infty} \frac{mnT_{clock}}{(m+n-1)T_{clock}} \\
&= \lim_{m\to\infty} \frac{mn}{m+n-1} \\
&= n
\end{aligned}
$$

Yielding the famous result that the ideal speedup is the number of stages in a pipeline. If a stall were to happen a finite number of times it would not effect the asymptotic speedup, however if a stall happened a fraction of the time that is a different matter. For instance, assume the pipeline stalls P_{err} cycles in $f_{T,err}$ of all instructions of type T ($m \times fT$ total instructions) then the time of the pipelined machine would be $T_{pipe} = (m + n - 1 + mf_T f_{err} P_{err}) \times T_{clock}$. The non-ideal speedup would be

$$
\begin{aligned}
\text{speedup} &= \frac{T_{nopipe}}{T_{pipe}} \\
&= \lim_{m\to\infty} \frac{mnT_{clock}}{(m+n-1+mf_T f_{err} P_{err})T_{clock}} \\
&= \lim_{m\to\infty} \frac{mn}{m+n-1+mf_T f_{T,err} P_{err}} \\
&= \frac{n}{1+f_T f_{T,err} P_{err}} \\
&= \frac{n}{1+f_{err} P_{err}}
\end{aligned}
$$

where $f_{err} = f_T f_{T,err}$. Note that the numerator is the CPI of the non-pipelined machine and the denominator is the CPI of the non-ideal pipelined machine. Thus we see that CPI for a pipelined machine is

$$CPI = 1 + \sum_{i=1}^{n} f_i P_i.$$

If there are no errors the ideal CPI is thus 1. Consider an example of this with branches incurring a penalty when they taken (i.e. the machine assumes branch not taken).

$$CPI_{avg} = (1 - P_b)CPI_{no\,branch} + P_b((1 - P_{take})CPI_{no\,branch} + P_{take}(1 + b))$$

CPI Cycles per instruction. The smaller the better. Nominally for a RISC machine this will be 1, but bubbles will increase it and pipelining will decrease it.

P Probability that something will happen (the event is indicated by the subscript).

b Branch penalty, which indicates how large the bubble in the pipeline is, that is caused by taking a branch.

27.1.2 Branch Prediction

Normally branches are assumed to be not taken but this is a simplistic assumption. A more sophisticated choice is to do what was done most recently. So for instance if the second instruction is a branch, and last time I was there I took it, I would have:

Address	Taken
0	0
1	0
2	1
3	0

This would require an extra bit for every memory location, most of which would be unused.

Performance

A pipelined RISC computer has 8 stages, and runs at 1.25 GHz. The cache has a miss rate of 1% for data and instructions, and a miss penalty of 24 ns. The system has a dynamic branch predictor that is wrong only 10% of the time. Branch errors cost 5 cycles.

1. What is the ideal (no stalls) speedup over a non-pipelined machine?
2. What is the impact to the CPI due to cache misses on a non-memory operation?
3. What is the impact to the CPI due to cache misses on a memory operation?
4. What is the impact to the CPI due to branch errors on branching instructions?
5. If memory operation make up 20% of the commands in a typical program and branching make up 15% of the commands, what is the average CPI?

1. $n = \frac{\text{Time Without Pipeline}}{\text{Time With Pipeline}} = \frac{I \times 8}{I + 8} \approx 8$ for large I (number of instructions).
2. $\Delta CPI = \text{Miss Rate} \times \text{Miss Penalty} \times \text{Clock Frequency} = (.01)(24ns)(1.25GHz) = .3$
3. Twice above or (0.6).
4. $\Delta CPI = \text{Branch Error Rate} \times (BranchPenalty) = .1 \times 5 = .5$
5. $CPI_{avg} = .2(1 + .6) + .15(1 + .3 + .5) + .65(1 + .3) = .32 + .27 + .845 = 1.435$

Superscalar

Superscalar pipelines have multiple pipelines to execute commands on (for example the latest pentium has 2). The advantage is that a machine with n pipelines could have a CPI of $\frac{1}{n}$. They have their own challenges in programming though.

Consider the following section of a program:

```
loop:  lw $t3,0($t1)       # first data
       add $t5, $t5, $t3   # running sum
       addi $t1, $t1, 4    # increment counter
       brne $t0, $t1, loop # check if done
exit:
```

And place the commands to be scheduled on two pipelines in the most obvious way.

Pipeline 1	Pipeline 2
lw \$t3,0(\$t1)	Nop
add \$t5, \$t5, \$t3	addi \$t1, \$t1, 4
brne \$t0, \$t1, loop	Nop

Granting myself a perfect branch predictor, so I have no stalls due to branching (in class we considered stalls), I still only get:

$$CPI = \frac{3}{4} = .75$$

Now consider a clever rearrangement:

Pipeline 1	Pipeline 2
lw \$t3,0(\$t1)	addi \$t1, \$t1, 4
add \$t5, \$t5, \$t3	brne \$t0, \$t1, loop

Granting myself a perfect branch predictor, I get:

$$CPI = \frac{2}{4} = .5$$

Can I always do such a rearrangement? Sorry but no. Consider the following:

```
loop:  lw $t3,0($t1)        # first data
       mult $t3, $t1        # multiplication
       mflo $t3             # get the product
       add $t5, $t5, $t3    # running sum
       addi $t1, $t1, 4     # increment counter
       brne $t0, $t1, loop  # check if done
exit:
```

And place the commands to be scheduled on two pipelines in the most obvious way.

Pipeline 1	Pipeline 2
lw \$t3,0(\$t1)	Nop
mult \$t3, \$t1	addi \$t1, \$t1, 4
mflo \$t3	Nop
add \$t5, \$t5, \$t3	brne \$t0, \$t1, loop

Granting myself a perfect branch predictor, so I have no stalls due to branching, I still only get:

$$CPI = \frac{4}{6} = .66$$

And note that the second pipeline is only half used.

27.2 Unrolling

Now let us unroll the loop, by considering two runs through at once. Note that on the second run through the data accessed is at four bytes higher than the first run.

```
loop:  lw $t3,0($t1)        # first data
       lw $t4,4($t1)        # second data
       mult $t3, $t1        # multiplication
       mflo $t3             # get the product
       add $t5, $t5, $t3    # running sum
       addi $t1, $t1, 4     # increment counter
       breq $t0, $t1, exit  # check if done
       mult $t4, $t1        # multiplication
       mflo $t4             # get the product
       add $t5, $t5, $t4    # running sum
       addi $t1, $t1, 4     # increment counter
       brne $t0, $t1, loop  # check if done
exit:
```

Pipeline 1	Pipeline 2
lw $t3,0($t1)	lw $t4,4($t1)
mult $t3, $t1	addi $t1, $t1, 4
mflo $t3	mult $t4, $t1
add $t5, $t5, $t3	breq $t0, $t1, exit
mflo $t4	addi $t1, $t1, 4
add $t5, $t5, $t4	brne $t0, $t1, loop

Granting myself a perfect branch predictor, so I have no stalls due to branching, I now get:

$$CPI = \frac{6}{12} = .5$$

As a general rule you unroll n copies of the loop for a machine with n pipelines. In this case I unrolled 2 copies because I had two pipes to fill.

27.3 Unrolling, Part II

Consider the following code to calculate the Fibonacci numbers.

```
top: add r4, r3, r2
     mov r2, r3
     mov r3, r4
     addi r1, r1, -1
     brgtz r1, top
```

The first three instructions are the data manipulations, and the last two are loop overhead (indexing and branching). There is a large amount of wasted effort spent in moving data around. Consider two loops worth of just the data manipulation portions.

```
add r4, r3, r2
mov r2, r3
mov r3, r4
add r4, r3, r2
mov r2, r3
mov r3, r4
```

Note that the "mov" commands are only to set up the problem for the next loop. In particular the contents of r2 are removed and the contents of r3 and r4 are shuffled. Consider the following change.

```
add r2, r3, r2
add r4, r3, r2
mov r3, r4
```

The contents of the registers are the same at the end of the loop, as the original, but considerable savings have been achieved. by noting the last mov command only shifts the results of the second add, we note that it is equivalent to the following

```
add r2, r3, r2
add r3, r3, r2
```

Thus by unrolling we can see the loop is equivalent to

```
 top: add r2, r3, r2
      add r3, r3, r2
      addi r1, r1, -2
      brgtz r1, top
      mov r4, r3
      breqz r1, exit
      mov r4, r2
exit:
```

Note the last three commands are cleanup only, so two iterations of the original loop can be done in less instructions than the unoptimized code. The loop can be scheduled efficiently on a two pipeline machine as

top:	add r2, r3, r2	addi r1, r1, -2
	add r3, r3, r2	bgtqz r1, top
	mov r4, r3	breqz r1, exit
	mov r4, r2	
exit:		

27.4 Software Pipelining

Returning to the original code

```
top: add r4, r3, r2
     mov r2, r3
     mov r3, r4
     addi r1, r1, -1
     brgtz r1, top
```

And let us again consider two iterations of the Fibonacci number loop.

```
add r4, r3, r2
mov r2, r3
mov r3, r4
add r4, r3, r2
mov r2, r3
mov r3, r4
```

First note that each pair of moves can be done simultaneously.

```
add r4, r3, r2
mov r2, r3         mov r3, r4
add r4, r3, r2
mov r2, r3         mov r3, r4
```

Now we will move the second add ahead in the scheduling so it is simultaneous with the first moves.

```
add r4, r3, r2
mov r2, r3         mov r3, r4         add r4, r4, r3
mov r2, r3         mov r3, r4
```

Now note that the `mov r2, r3` commands are useless and can be dropped.

```
add r4, r3, r2
mov r3, r4         add r4, r4, r3
mov r3, r4
```

This suggests the following parallel execuation

mov r2, r3	add r3, r3, r2	addi r1, r1, -1	brgtz r1, top

time	r3	r2	r1
0	1	1	3
1	2	1	2
2	3	2	1
3	5	3	0

27.4.1 Example

Consider the following code

```
top: ld r2, 0(r1)
     addi r3, r2, 1
     st r3, 0(r1)
     addi r1, r1, 4
     brlt r1, r4, top
```

st r3, 0(r1)	addi r3, r2, 1	ld r2, 8(r1)

Chapter 28

Tomasulo

28.1 Multiple Issue Tomasulo

To illustrate the method we will consider a simple piece of code.

```
loop:
  mul  $t4,$t2
  mflo $t4
  subi $t3,$t3,1
  bgtz $t3,loop
```

This code will calculate $\$t4 = \$t2^{\$t3}$, assuming $\$t4 = 1$ initially and $\$t2 > 0$ and $\$t3 > 1$.

Further lets assume add/sub/move takes 1 cycle of execution, multiply takes 2 cycles, and branches take 2 cycle. The branch predictor will always predict branch taken in this example. Let's schedule this for our machine.

Cycle 1

Reorder Buffer

Entry	Busy	Instruction	State	Destination	Value
1	yes	mul $t4,$t2	Issue	$Hi, $Lo	
2	yes	mflo $t4	Issue	$t4	
3					
4					
5					
6					
7					
8					
9					
10					

Registers

Field	Data	Reorder	Busy
$t0			
$t1			
$t2	5		
$t3	2		
$t4	1	#2	yes
$t5			
$t6			
$t7			
$t8			
$t9			

Reservation Station

Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1		mflo			#1		#2	
Add2								
Add3								
Add4								
Mul1		mul	1	5			#1	
Mul2								
Br1								
Br2								

Cycle 2

Reorder Buffer

Entry	Busy	Instruction	State	Destination	Value
1	yes	mul \$t4,\$t2	Exec	\$Hi, \$Lo	
2	yes	mflo \$t4	Issue	\$t4	
3	yes	subi \$t3,\$t3,1	Issue	\$t3	
4	yes	bgtz \$t3,loop	Issue		
5					
6					
7					
8					
9					
10					

Registers

Field	Data	Reorder	Busy
\$t0			
\$t1			
\$t2	5		
\$t3	2	#3	yes
\$t4	1	#2	yes
\$t5			
\$t6			
\$t7			
\$t8			
\$t9			

Reservation Station

Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1		mflo			#1		#2	
Add2		subi	2	1			#3	
Add3								
Add4								
Mul1	yes	mul	1	5			#1	
Mul2								
Br1		bgtz			#3		#4	
Br2								

Cycle 3

Reorder Buffer

Entry	Busy	Instruction	State	Destination	Value
1	yes	mul \$t4,\$t2	Exec	\$Hi, \$Lo	
2	yes	mflo \$t4	Issue	\$t4	
3	yes	subi \$t3,\$t3,1	Exec	\$t3	
4	yes	bgtz \$t3,loop	Issue		
5	yes	mul \$t4,\$t2	Issue	\$Hi, \$Lo	
6	yes	mflo \$t4	Issue	\$t4	
7					
8					
9					
10					

Registers

Field	Data	Reorder	Busy
\$t0			
\$t1			
\$t2	5		
\$t3	2	#3	yes
\$t4	1	#6	yes
\$t5			
\$t6			
\$t7			
\$t8			
\$t9			

Reservation Station

Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1		mflo			#1		#2	
Add2	yes	subi	2	1			#3	
Add3		mflo			#5		#6	
Add4								
Mul1	yes	mul	1	5			#1	
Mul2		mul		5	#2		#5	
Br1		bgtz			#3		#4	
Br2								

Cycle 4

Reorder Buffer

Entry	Busy	Instruction	State	Destination	Value
1	no	mul $t4,$t2	Commit	$Hi, $Lo	5
2	yes	mflo $t4	Exec	$t4	
3	no	subi $t3,$t3,1	done	$t3	1
4	yes	bgtz $t3,loop	Exec		
5	yes	mul $t4,$t2	Issue	$Hi, $Lo	
6	yes	mflo $t4	Issue	$t4	
7	yes	subi $t3,$t3,1	Issue	$t3	
8	yes	bgtz $t3,loop	Issue		
9					
10					

Registers

Field	Data	Reorder	Busy
$t0			
$t1			
$t2	5		
$t3	1	#7	yes
$t4	1	#6	yes
$t5			
$t6			
$t7			
$t8			
$t9			

Reservation Station

Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1	yes	mflo	5				#2	
Add2		subi	1	1			#7	
Add3		mflo			#5		#6	
Add4								
Mul1								
Mul2		mul		5	#2		#5	
Br1	yes	bgtz	1				#4	
Br2		bgtz		#7			#8	

Cycle 5

Reorder Buffer

Entry	Busy	Instruction	State	Destination	Value
1					
2	no	mflo $t4	Commit	$t4	5
3	no	subi $t3,$t3,1	Commit	$t3	1
4	yes	bgtz $t3,loop	Exec		
5	yes	mul $t4,$t2	Exec	$Hi, $Lo	
6	yes	mflo $t4	Issue	$t4	
7	yes	subi $t3,$t3,1	Exec	$t3	
8	yes	bgtz $t3,loop	Issue		
9	yes	mul $t4,$t2	Issue	$Hi, $Lo	
10	yes	mflo $t4	Issue	$t4	

Registers

Field	Data	Reorder	Busy
$t0			
$t1			
$t2	5		
$t3	1	#7	yes
$t4	5	#10	yes
$t5			
$t6			
$t7			
$t8			
$t9			

Reservation Station

Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1		mflo			#9		#10	
Add2	yes	subi	1	1			#7	
Add3		mflo			#5		#6	
Add4								
Mul1		mul		5	#6		#9	
Mul2	yes	mul	5	5			#5	
Br1	yes	bgtz	1				#4	
Br2		bgtz		#7			#8	

Cycle 6

Reorder Buffer						Registers			
Entry	Busy	Instruction	State	Destination	Value	Field	Data	Reorder	Busy
1	yes	subi $t3,$t3,1	Issue	$t3		$t0			
2	yes	bgtz $t3,loop	Issue			$t1			
3						$t2	5		
4	no	bgtz $t3,loop	Commit			$t3	1	#1	yes
5	yes	mul $t4,$t2	Exec	$Hi, $Lo		$t4	5	#10	yes
6	yes	mflo $t4	Issue	$t4		$t5			
7	no	subi $t3,$t3,1	Done	$t3	0	$t6			
8	yes	bgtz $t3,loop	Issue			$t7			
9	yes	mul $t4,$t2	Issue	$Hi, $Lo		$t8			
10	yes	mflo $t4	Issue	$t4		$t9			

Reservation Station								
Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1		mflo			#9		#10	
Add2		subi	0	1			#1	
Add3		mflo			#5		#6	
Add4								
Mul1		mul		5	#6		#9	
Mul2	yes	mul	5	5			#5	
Br1		bgtz			#2		#2	
Br2	yes	bgtz	0				#8	

Cycle 7

Reorder Buffer						Registers			
Entry	Busy	Instruction	State	Destination	Value	Field	Data	Reorder	Busy
1	yes	subi $t3,$t3,1	Issue	$t3		$t0			
2	yes	bgtz $t3,loop	Issue			$t1			
3						$t2	5		
4						$t3	1	#1	yes
5	no	mul $t4,$t2	Commit	$Hi, $Lo	25	$t4	5	#10	yes
6	yes	mflo $t4	Exec	$t4		$t5			
7	no	subi $t3,$t3,1	Done	$t3	0	$t6			
8	yes	bgtz $t3,loop	Exec			$t7			
9	yes	mul $t4,$t2	Issue	$Hi, $Lo		$t8			
10	yes	mflo $t4	Issue	$t4		$t9			

Reservation Station								
Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1		mflo			#9		#10	
Add2		subi	0	1			#1	
Add3	yes	mflo	25				#6	
Add4								
Mul1		mul		5	#6		#9	
Mul2								
Br1		bgtz			#2		#2	
Br2	yes	bgtz	0				#8	

Cycle 8

Reorder Buffer					
Entry	Busy	Instruction	State	Destination	Value
1	yes	subi $t3,$t3,1	Exec	$t3	
2	yes	bgtz $t3,loop	Issue		
3	yes	mul $t4,$t2	Issue		
4	yes	mflo $t4	Issue		
5					
6					
7					
8	no	bgtz $t3,loop	Flush		
9	yes	mul $t4,$t2	Exec	$Hi, $Lo	
10	yes	mflo $t4	Issue	$t4	

Registers			
Field	Data	Reorder	Busy
$t0			
$t1			
$t2	5		
$t3	0	#1	yes
$t4	25	#4	yes
$t5			
$t6			
$t7			
$t8			
$t9			

Reservation Station								
Name	Busy	Op	V_1	V_2	S_1	S_2	Dest	A
Add1		mflo			#9		#10	
Add2	yes	subi	0	1			#1	
Add3		mflo			#3		#4	
Add4								
Mul1	yes	mul	25	5			#9	
Mul2		mul		5	#10		#3	
Br1		bgtz			#2		#2	
Br2								

At this point the buffers and stations will be flushed, the executions cancelled, and the registers not updated (they are at the right point). New commands will be loaded from after the branch, and execution proceeds normally.

Chapter 29

Thread Level Parallelism

29.1 Taxonomy

Flynn

SISD Single Instruction Single Data (Modern uniprocessors)

SIMD Single Instruction Multiple Data (Vector machines, and some multimedia)

MISD Multiple Instruction Single Data (No commercial, possible in special applications)

MIMD Multiple Instruction Multiple Data (Modern multiprocessors)

MIMD is broken into two groups based on memory configuration. Memory is either shared equally by all processors or distributed among the processors.

29.2 Shared Memory

The first group centralizes the memory and has each processor with its cache connect via a shared memory bus.

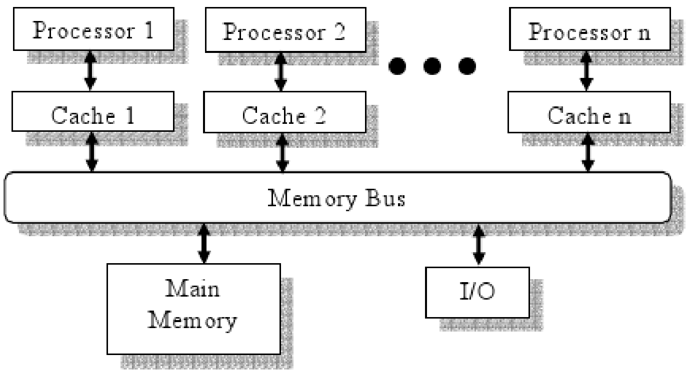


Figure 29.1: Centralized shared memory multiprocessor

The first group is also referred to by

- Centralized Shared Memory
- Symmetric Multiprocessors (SMP)
- Uniform Memory Access (UMA)

These alternate titles are used since the the memory is central and shared, it is thus symmetric to all, and thus the access for each processor is uniform. The main problem here is that as the number of processors grows, the need for memory bandwidth grows. Without the needed bandwidth, requests will have to be scheduled resulting in increased latency.

Example 23 *Using Figure 6.10 in the book, fill in the table, assuming all events are for an address relative to a cache in a SMP system.*

Event	*Source*	*State*
Startup	-	*Invalid*
Read Miss	*CPU*	
Read Miss	*Bus*	
Write Hit	*CPU*	
Write Miss	*Bus*	
Write Miss	*CPU*	
Read Miss	*Bus*	

Event	*Source*	*State*
Startup	-	*Invalid*
Read Miss	*CPU*	*Shared*
Read Miss	*Bus*	*Shared*
Write Hit	*CPU*	*Exclusive*
Write Miss	*Bus*	*Invalid*
Write Miss	*CPU*	*Exclusive*
Read Miss	*Bus*	*Shared*

29.3 Distributed Memory

The second group distributes the memory to each processor so the memory bandwidth grows with the need. This results in the problem of data sharing and communications between the nodes. We could just treat the distributed memories like one big memory, giving each an address (shared address space). This would allow the memories to be shared. Access to different parts of memory is no longer uniform (addresses corresponding to "local" memory will be fast and the addresses corresponding to "remote" memory will be slow). This scheme is referred to as

- Distributed Shared Memory (DSM)
- Nonuniform Memory Access (NUMA)

Alternately we could keep each address space separate (local addresses) and pass messages between nodes containing the data or communications. This scheme makes each machine look like an individual computer (multi-computers) and often each processor is a separate machine (clusters).

Shades of grey exist between the two, for instance a network OS can use message passing to pass a page of memory and implement what looks like shared address space by utilizing paging capabilities.

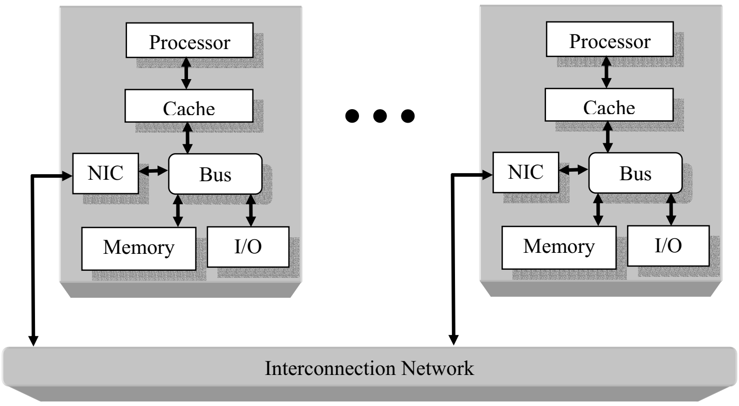


Figure 29.2: Distributed memory multiprocessor

29.4 Performance

Amdahl's Law, for n processors is

$$S = \frac{1}{\sum_{i=1}^{n}\left(\frac{f_i}{i}\right)}, \tag{29.1}$$

where f_i is the fraction of time when i processors are busy. Note that

$$\sum_{i=1}^{n} f_i = 1. \tag{29.2}$$

Example 24 *Consider a 4 processor machine. What must the fractions be to ensure a speedup of at least 3.*

$$\begin{aligned} 3 &= \frac{1}{\frac{f_1}{1}+\frac{f_2}{2}+\frac{f_3}{3}+\frac{f_4}{4}} \\ 1 &= 3\left(\frac{f_1}{1}+\frac{f_2}{2}+\frac{f_3}{3}+\frac{f_4}{4}\right) \\ 4 &= 12f_1+6f_2+4f_3+3f_4 \end{aligned}$$

Note that if the least common multiple of the numbers 1 *through* n *is denoted* LCM*, then for an* n *processor system trying to achieve a speedup of* s *we can say*

$$\frac{LCM}{s} = \sum_{i=1}^{n} \frac{LCM}{i} f_i$$

is the equation describing this situation that has integer coefficients. We also know

$$1 = f_1+f_2+f_3+f_4.$$

Combining yields

$$\begin{bmatrix}4\\1\end{bmatrix} = \begin{bmatrix}12 & 6 & 4 & 3\\ 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix}f_1\\f_2\\f_3\\f_4\end{bmatrix}.$$

This is indefinite (more unknowns than equations), but we can solve for the fractions in terms of f_1 and f_2.

$$\begin{aligned} 8f_1 + 2f_2 &= f_4 \\ 1 - 9f_1 - 3f_2 &= f_3 \end{aligned}$$

The second equation implies that individually $f_1 < \frac{1}{9} \approx .11$ and $f_2 < \frac{1}{3} \approx .33$ and together $3f_1 + f_2 \leqslant \frac{1}{3}$. Further, if f_3 is negligible then $.67 \leqslant f_4 \leqslant .88$ is the minimum range to ensure a speedup of 3.

The last example shows how great the required thread level parallelism is to achieve a reasonable speedup. The lack of thread level parallelism is one of the two great problems/challenges in multiprocessing. The other great problem/challenge is the latency of remote accesses, which effectively adds a fixed penalty to the CPI of each processor thereby limiting performance.

The efficiency is given by

$$E = \sum_{i=1}^{n} \left(f_i \frac{i}{n} \right) \leqslant 1 \tag{29.3}$$

Scientific programs are often used to benchmark multiprocessor performance. For the following table n is the problem size, p is the number of processors, and the α numbers are the scaling factors.

Application	$\alpha_{compute}$	$\alpha_{communicate}$
FFT	$\frac{n \log n}{p}$	$\frac{n}{p}$
LU/Ocean	$\frac{n}{p}$	$\sqrt{\frac{n}{p}}$
Barnes-Hut	$\frac{n \log n}{p}$	$\sqrt{\frac{n}{p}} \log n$

Example 25 *An Ocean application takes 1 hour to run on a uniprocessor, and 33 minutes on a dual processor. How long will it take to run the Ocean application on a problem 16 times the original size on a 128 processor machine?*

Processors(p)	*Size(n)*	*Time*	*Computation*	*Communication*
1	*1*	*1 hour*	*1*	*0*
2	*1*	*33 min*	$\frac{1}{2}$	$\frac{\sqrt{1}}{\sqrt{2}}$
128	*16*		$\frac{16}{128}$ $\frac{1}{8}$	$\frac{\sqrt{16}}{\sqrt{128}}$ $\frac{1}{2\sqrt{2}}$

From the table, if we call the base time of computation x, and the base time of communication y then we get (using minutes to be consistent):

$$60 = 1x + 0y \tag{29.4}$$

$$33 = .5x + \frac{1}{\sqrt{2}}y \tag{29.5}$$

from the first equation, $x = 60min$ and from the second equation $y = 3\sqrt{2}min$. Thus the time for the third case (problem solution) is

$$\begin{aligned} T &= \frac{1}{8}x + \frac{1}{2\sqrt{2}}y \\ &= \frac{1}{8}60 + \frac{1}{2\sqrt{2}}3\sqrt{2} \\ &= 7.5 + 1.5 \\ &= 9min \end{aligned}$$

Example 26 *An LU application runs in 4000 seconds on a uniprocessor. The same problem runs in 1020 seconds on a four processor machine. How long will it take to run on a 16 processor machine? How long will it take to run on a 64 processor machine?*

$$\begin{aligned} 4000 &= 1x + 0y \\ 1020 &= .25x + .5y \end{aligned}$$

So $x = 4000$ *and* $y = 40$*. Using this,*

$$\begin{aligned} \frac{1}{16}4000 + \frac{1}{4}40 &= 250 + 10 \\ &= 260 \end{aligned}$$

So a 16 processor machine runs it in 260 seconds.

$$\begin{aligned} \frac{1}{64}4000 + \frac{1}{8}40 &= 62.5 + 5 \\ &= 67.5 \end{aligned}$$

Note that communication takes up $\frac{20}{1020} \approx .0196$ *or just under 2% of the time on a four processor. When we have a sixteen processor machine it takes up* $\frac{10}{260} \approx .0385$ *or about 3.85% of the time. On the 64 processor machine the communication took up* $\frac{5}{67.5} \approx .0385$ *or about 7.41% of the time. Communication takes ever increasing fraction of the time, and becomes a limit to performance. Consider running this on a 4096 processor machine. It would take*

$$\begin{aligned} \frac{1}{4096}4000 + \frac{1}{64}40 &\approx .977 + .625 \\ &\approx 1.6 \end{aligned}$$

Thus communication takes $\frac{.625}{1.6} \approx .391$ *or almost 40% of the time!*

Part VI

Appendices

Appendix A

Sample Computers

A.1 32 Bit Pipelined Computer

Consider a 32 bit pipelined computer with a 1.0 GHz clock and an ISA that has three categories of commands:

	Freq
Branch	.2
Memory	.3
Other	.5

The computer has a 64 bit memory bus that operates at 500 MHz. The bus requires that requests and responses take 1 cycle. The memory takes 40ns to respond to a request and can do burst sends with a delay of 10ns. The bus requires 3 cycles to initiate a request and 2 cycles to transmit the response. The bus is DMA and requires 710 CPU cycles to set up a transfer, 275 cycles to complete, 500 cycles to handle errors (1% of the time).

The machine has two disks that have a combined transfer rate of 20MB/s, and a total latency of 6.8 ms. The computer has virtual memory with a page size of 64KB.

1. What is the bandwidth of the bus?

   We have been assuming the installed RAM to be integral in the bus design, so the answer would be:

$$\begin{aligned}
\text{Bandwidth} &= \frac{\text{Data Transferred}}{\text{Time of Transfer}} \\
&= \frac{\text{Data Transferred}}{\text{Number of Cycles} \times \text{Time of 1 Cycle}} \\
&= \frac{\text{Data Transferred} \times \text{Bus Clock Frequency}}{\text{Cycles to Initiate} + \text{Cycles to Respond} + \text{Cycles to Get Data}} \\
&= \frac{\text{8Bytes} \times \text{500MHz}}{3 + 2 + (\text{40ns} \times \text{500MHz})} \\
&= \frac{\text{4000MB/s}}{25} \\
&= \text{160MB/s}
\end{aligned}$$

   You might have noted that be RAM supports a burst transfer mode. As the size of the burst increases the effective time to get the data approaches the burst time of 10 ns (down from 40 ns). If you assumed this you would have found the bandwidth to be 400 MB/s.

2. If the computer had to continually page, how much of the CPU's time and the bus's bandwidth would it use?

   Note that in memory KB = 2^{10} bytes, but in networks KB = 10^3 bytes. As they are similar, we will ignore the difference as the book does. Additionally, we will assume the pages are spread across both disks so as to maximize the transfer.

   The time it takes to transfer one page is given by:

$$\begin{aligned} T_{\text{transfer}} &= \text{time to get the data} + \text{time to send the data} \\ &= \text{total latency} + \frac{\text{Data Sent}}{\text{Transmission Rate}} \\ &= 6.8\text{ms} + \frac{64\text{KB}}{20\text{MB/s}} \\ &= 6.8\text{ms} + 3.2\text{ms} \\ &= 10\text{ms}. \end{aligned}$$

   The data rate for the transfer is:

$$\begin{aligned} R_{\text{Data}} &= \frac{\text{Data Sent}}{T_{\text{transfer}}} \\ &= \frac{64\text{KB}}{10\text{ms}} \\ &= 6.4\text{MB/s}. \end{aligned}$$

   Using the figure of 160 MB/s for the bus's bandwidth we find:

$$\begin{aligned} \text{Percent Utilization of Bus} &= \frac{\text{Bandwidth Used}}{\text{Bandwidth Available}} \times 100\% \\ &= \frac{6.4\text{MB/s}}{160\text{MB/s}} \times 100\% \\ &= 4\%. \end{aligned}$$

   Now let's look at the impact on the CPU. We need to find the number of cycles the CPU must use to handle the transfer.

$$\begin{aligned} \text{Cycles Per Transfer} &= \frac{\text{Cycles to Set Up} + \text{Cycles to Finish} + \text{error rate} \times \text{Cycle to Handle Errors}}{1 - \text{error rate}} \\ &= \frac{710 + 275 + .01 \times 500}{1 - .01} \\ &= \frac{990}{.99} \\ &= 1000 \end{aligned}$$

   The utilization of the CPU is thus:

$$\begin{aligned} \text{Percent Utilization of CPU} &= \frac{\frac{\text{Cycles Per Transfer}}{T_{\text{transfer}}}}{\text{CPU Clock Frequency}} \times 100\% \\ &= \frac{\frac{1000}{10\text{ms}}}{1\text{GHz}} \times 100\% \\ &= 0.01\% \end{aligned}$$

   Thus we have a negligible impact.

3. What block size of the cache would cause the least impact on the CPI of the computer due to misses, assuming the instruction and data miss rate are equal?

Block Size	2 words	4 words	8 words	16 words
Miss Rate	4%	2%	1.2%	1%

The average increase to a command's CPI due to cache misses depends on if the command accesses memory just for the instruction fetch or also for the commands implementation. We will therefor assess the impact to memory commands separate from branch and other commands. The average increase for branch and other commands is given by:

$$\begin{aligned}\Delta\text{CPI} &= \text{miss rate} \times \text{Bus Cycles to Transfer} \times \frac{\text{CPU Clock Rate}}{\text{Bus Clock Rate}} \\ &= \text{miss rate} \times \text{Bus Cycles to Transfer} \times 2.\end{aligned}$$

The average impact for branch commands is twice the increase of branch and other commands.

The bus cycles to transfer $2N$ words is given by:

$$\begin{aligned}\text{Cycle to Transfer} &= \text{Cycles to Initiate} + \text{Cycles to Get First 2 Words} \\ &\quad +(N-1) \times \text{Cycle to Burst Get 2 Words} \\ &\quad +N \times \text{Cycles to Send 2 Words} \\ &= 3 + (40\text{ns} \times 500\text{MHz}) + (N-1) \times (10\text{ns} \times 500\text{MHz}) + N \times 2 \\ &= 18 + 7N\end{aligned}$$

Block Size	2 words	4 words
Miss Rate	4%	2%
Bus Cycles to Transfer	18+7(1)=25	18+7(2)=32
ΔCPI Not Memory	(.04)(25)(2)=2	(.02)(32)(2)=1.28
ΔCPI Memory	(2)(2)=4	(2)(1.28)=2.56

Block Size	8 words	16 words
Miss Rate	1.2%	1%
Bus Cycles to Transfer	18+7(4)=46	18+7(8)=74
ΔCPI Not Memory	(.012)(46)(2)=1.104	(.01)(74)(2)=1.48
ΔCPI Memory	(2)(1.104)=2.208	(2)(1.48)=2.96

The least impact is given by a cache with blocks of 8 words in this case.

4. Design a dynamic branch predictor for the computer.

A good estimate of whether a branch will be taken is to remember whether it was taken last time. Remembering if a branch was taken or not requires 1 bit per instruction tracked. To keep the problem realistic we will add an additional 32-bit register. Each bit in the register will indicate if the branch was taken for the last instruction whose address modulo 32 corresponds to the bit's location. An easy way to implement this would be to take the outputs of the 32 bits and pass them into a 32×1 MUX, whose address select lines are given the last five bits of the command's address (from PC for instance). The branch taken signal could be sent from the control to the particular bit by using a 1×32 DeMUX.

5. For this system, 60% of the branch instructions make loops and the rest are for conditional execution. On average, the code in a loop is executed 10 times. What fraction of the time does your dynamic branch predictor, correctly predict the branch taken?

Loops occur 60% of the time, conditional execution occurs 40% of the time. The dynamic branch selected above does not likely do anything for conditional execution branches, so it is most likely correct on 50% of the conditional execution branch instructions. In the loops, the method designed would be correct in all but the first and last execution of the loop, so 80% on loops.

The net effect is $(.4)(.5) + (.6)(.8) = .68$ or 68% of the time it is right.

6. Using the best cache and your dynamic branch predictor, calculate the average CPI and the performance of the computer in MIPS.

   I forgot to give you base CPI and the penalty to CPI for missing a branch. I wanted the base for all instructions to be 1 (ideal for piplined) and the penalty to be 3. Sorry about that.

   Average CPI is given by:

$$
\begin{aligned}
\mathrm{CPI}_{\mathrm{avg}} &= \sum_i (\mathrm{CPI}_i \times \mathrm{frequency}_i) \\
&= \mathrm{CPI}_{\mathrm{Memory}} \times \mathrm{freq}_{\mathrm{Memory}} + \mathrm{CPI}_{\mathrm{Correct\ Branch}} \times \mathrm{freq}_{\mathrm{Correct\ Branch}} \\
&\quad + \mathrm{CPI}_{\mathrm{Incorrect\ Branch}} \times \mathrm{freq}_{\mathrm{Incorrect\ Branch}} + \mathrm{CPI}_{\mathrm{Other}} \times \mathrm{freq}_{\mathrm{Other}} \\
&= (1 + 2.208)(.2) + (1 + 1.104)(.3 \times .68) \\
&\quad + (1 + 1.104 + 3)(.3 \times .32) + (1 + 1.104)(.5) \\
&= 2.6128
\end{aligned}
$$

   MIPS is given by:

$$
\begin{aligned}
\mathrm{MIPS} &= \frac{\text{CPU Clock Freq}}{\mathrm{CPI} \times 10^6} \\
&= \frac{10^9 \text{Cycles/s}}{2.6128 \text{Cycle/Million Inst} \times 10^6} \\
&\approx 383
\end{aligned}
$$

A.2 One Command Computer

Consider a computer which has only one command, subtract and branch if negative (SBrN D, S1, S2, Jump). Which does:

```
D = S1 - S2
if D < 0 goto Jump
```

Since there is only one command there is no need to include the opcode in the machine language instruction. The system is to have 1K of memory divided into 256 words of 4 bytes each. Since memory requires 1 bytes to specify the address of a memory location the instructions will have four fields of 1 byte each:

Destination	Source 1	Source 2	Jump Address

1. Design a CPU that implements this.

   Sol:

   See Figure 1

2. Alter your design to make it a four stage pipeline with forwarding.

   Sol:

   See Figure 2. Note that the control to the forwarding MUXs can come from tag bits on the RAM (first idea) or comparators on the destination (better solution).

3. Design the control for the CPU (hardwired or microcoded)

   Sol:

   In this case, most of the control signals are already handled. All that remains undone is the load commands to the program counter and instruction register, and the read and write commands to memory. The ifetch loop has only four states so the resulting logic table is:

S_1S_0	S_1S_0	Rpc	Rs1/Rs2	Wd
00	01	1	0	0
01	10	0	1	0
10	11	0	0	0
11	00	0	0	1

   Next $S_1 = S_1' \cdot S_0 + S_1 \cdot S_0'$

   Next $S_0 = S_0'$

   $Rpc = S_1' \cdot S_0'$

   $Rs1Rs2 = S_1' \cdot S_0$

   $Wd = S_1 \cdot S_0$

4. Show the tag bits (with their size), data field, and address of a 2-way associative write-back cache that uses NLLRU for this machine that has 8 locations. How many total bits must be stored?

   Sol:

   Main Memory has 2^8 locations (n=8)

   Cache has 2^3 locations (m=3)

   Associativity is 2^1 (k=1)

   Each location in cache has a total of 42 bits

   (a) Address tag bits: n-(m-k) = 8-(3-1) = 6
   (b) Valid tag bit: 1
   (c) Dirty tag bit: 1
   (d) NLLRU tag bits: 2 (the associativity)
   (e) Data bits: 32

   The entire cache has $8 \times 42 = 336$ bits

5. Show the cache accesses and calculate the hit ratio for the following memory values, assuming execution begins at location 0 and terminates when location 5 is reached. If a location is not specified below, the contents are not important. All values are in hex.

Address	D	S1	S2	J
00	87	88	80	01
01	86	80	8B	02
02	87	87	86	03
03	01	01	83	04
04	82	82	81	01

Address	Data			
80	00	00	00	00
81	FF	FF	FF	FF
82	FF	FF	FF	FE
83	00	00	01	00

   Sol:

   (I have grouped my cache table so the associated portions of cache are on the same row.)

   Initial condition (hits=0, misses=0)

NLLRU	V	D	Address	Loc	Data	NLLRU	V	D	Address	Loc	Data
00	0	0	000000	000	0x00000000	00	0	0	000000	100	0x00000000
00	0	0	000000	001	0x00000000	00	0	0	000000	101	0x00000000
00	0	0	000000	010	0x00000000	00	0	0	000000	110	0x00000000
00	0	0	000000	011	0x00000000	00	0	0	000000	111	0x00000000

command=0x87888001 (hits=0, misses=3) (read in 0 then 88, then 80 overwrote 0)

NLLRU	V	D	Address	Loc	Data	NLLRU	V	D	Address	Loc	Data
01	1	0	100000	000	0x00000000	00	1	0	100010	100	0x????????
00	0	0	000000	001	0x00000000	00	0	0	000000	101	0x00000000
00	0	0	000000	010	0x00000000	00	0	0	000000	110	0x00000000
01	1	1	100001	011	0x????????	00	0	0	000000	111	0x00000000

command=0x86808B02 (hits=1, misses=5)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
01	1	0	100000	000	0x00000000	00	1	0	100010	100	0x????????
01	1	0	000000	001	0x86808B02	00	0	0	000000	101	0x00000000
01	1	1	100001	010	0x????????	00	0	0	000000	110	0x00000000
00	1	1	100001	011	0x????????	10	1	0	100010	111	0x????????

command=0x87878603 (hits=3, misses=6)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
01	1	0	100000	000	0x00000000	00	1	0	100010	100	0x????????
01	1	0	000000	001	0x86808B02	00	0	0	000000	101	0x00000000
01	1	1	100001	010	0x????????	00	1	0	000000	110	0x87878603
01	1	1	100001	011	0x????????	00	1	0	100010	111	0x????????

command=0x01018304 (hits=4, misses=8)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
01	1	0	100000	000	0x00000000	00	1	0	100010	100	0x????????
01	1	1	000000	001	0x86808A02	00	0	0	000000	101	0x00000000
01	1	1	100001	010	0x????????	00	1	0	000000	110	0x87878603
01	1	1	100000	011	0x00000100	00	1	0	000000	111	0x01018304

command=0x82828101 (hits=4, misses=11) (NOTE: 82 is 0xFFFFFFFF at end of command then jumps to 01)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
00	1	0	100000	000	0x00000000	10	1	0	000001	100	0x82828101
00	1	1	000000	001	0x86808A02	10	1	0	100000	101	0xFFFFFFFF
00	1	1	100001	010	0x????????	10	1	0	100000	110	0xFFFFFFFF
01	1	1	100000	011	0x00000100	00	1	0	000000	111	0x01018304

command=0x0x86808A02 (hits=6, misses=12)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
00	1	0	100000	000	0x00000000	10	1	0	000001	100	0x82828101
01	1	1	000000	001	0x86808A02	00	1	0	100000	101	0xFFFFFFFF
00	1	1	100001	010	0x????????	10	1	0	100010	110	0x????????
01	1	1	100000	011	0x00000100	00	1	0	000000	111	0x01018304

command=0x0x87878603 (hits=7, misses=14)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
00	1	0	100000	000	0x00000000	10	1	0	000001	100	0x82828101
01	1	1	000000	001	0x86808A02	00	1	0	100000	101	0xFFFFFFFF
01	1	0	100001	010	0x87878603	00	1	0	100010	110	0x????????
00	1	1	100000	011	0x00000100	10	1	1	100001	111	0x????????

command=0x0x01018304 (hits=8, misses=16)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
00	1	0	100000	000	0x00000000	10	1	0	000001	100	0x82828101
01	1	1	000000	001	0x86808902	00	1	0	100000	101	0xFFFFFFFF
01	1	0	100001	010	0x87878603	00	1	0	100010	110	0x????????
01	1	0	000000	011	0x01018304	10	1	0	100003	111	0x00000100

command=0x0x82828101 (hits=10, misses=17) (loaded 82 as 0xFFFFFFFF then -(-1) to get 0 which ends)

LRU	V	D	Address	Loc	Data	LRU	V	D	Address	Loc	Data
00	1	0	100000	000	0x00000000	10	1	0	000001	100	0x82828101
01	1	1	000000	001	0x86808902	00	1	0	100000	101	0xFFFFFFFF
00	1	0	100001	010	0x87878603	10	1	1	100000	110	0x00000000
01	1	0	000000	011	0x01018304	10	1	0	100003	111	0x00000100

6. Assuming the cache has an access time of 4ns and the memory has an access time of 60ns, calculate the effective access time of the memory.

   Sol:

   $hr = \frac{\text{hit}}{\text{hit+miss}} = \frac{10}{27} \approx .37$

   $mr = 1 - hr \approx 1 - .37 = .63$

   $T_{\text{eff}} = hr \times T_{\text{cache}} + mr \times T_{\text{RAM}} \approx .37 \times 4 + .63 \times 60 \approx 39ns$

A.3 Multiple Issue Machine

You have a 1.5 GHz computer which can issue 2 instructions per cycle and a dynamic branch predictor that reduces the branch penalty from 4 cycles to 1 cycle, 90% of the time. Branch instructions are 15% of all instructions, loads are 20%, and stores are 5%.

The cache is split into 4k instruction cache and 4k data cache. The cache takes 2 ns to access. The instruction cache has a block size of 2 words, has an associativity of 4, and a miss rate of 2%. The data cache has a block size of 4 words, an associativity of 2, is write-back, is not write-allocate, has a read miss rate of 5%, a write miss rate of 2%, and 10% of the blocks are dirty.

The RAM is 8MB takes 50ns to access and can burst write subsequent accesses at 10ns.

1. How many cycles on average is the branch penalty?

$$\begin{aligned} Penalty_{branch} &= f_{\text{pred. correct}} Cost_{\text{pred. correct}} + f_{\text{pred. error}} Cost_{\text{pred. error}} \\ &= .9 \times 1 + .1 \times 4 \\ &= 1.3 \end{aligned}$$

2. How long does an instruction read miss take?

   Two words have to be loaded on a miss, which takes 70ns.

3. How long does a data read miss take?

   Four words have to be loaded on a miss, which takes 90ns. Now 10% of the time we also have to write four words, which takes the same as a read thus we have: $(1 + .1) \times 90ns = 99ns$

4. How long does a data write miss take?

   On a write miss, four words have to be written, which takes 90ns.

5. What is the effective access time for instruction loads?

$$\begin{aligned} T_{Inst} &= 2ns + .02 \times 70ns \\ &= 3.4ns \end{aligned}$$

6. What is the effective access time for data reads?

$$\begin{aligned} T_{read} &= 2ns + .05 \times 99ns \\ &= 6.95ns \end{aligned}$$

7. What is the effective access time for data writes?

$$\begin{aligned} T_{write} &= 2ns + .02 \times 90ns \\ &= 3.8ns \end{aligned}$$

8. What is the CPI of this machine?

$$\begin{aligned} CPI &= \frac{(1 + f_{branch}Penalty_{branch} + Penalty_{inst} + f_{read} \times Penalty_{read} + f_{write} \times Penalty_{write})}{\text{\# inst per cycle}} \\ &= \frac{(1 + f_{branch}Penalty_{branch} + Clock_{rate}(T_{inst} + f_{read} \times T_{read} + f_{write} \times T_{write}))}{\text{\# inst per cycle}} \\ &= \frac{1 + .15 \times 1.3 + 1.5GHz(3.4ns + .2 \times 6.95ns + .05 \times 3.8ns)}{2} \\ &= 1.0830625 \end{aligned}$$

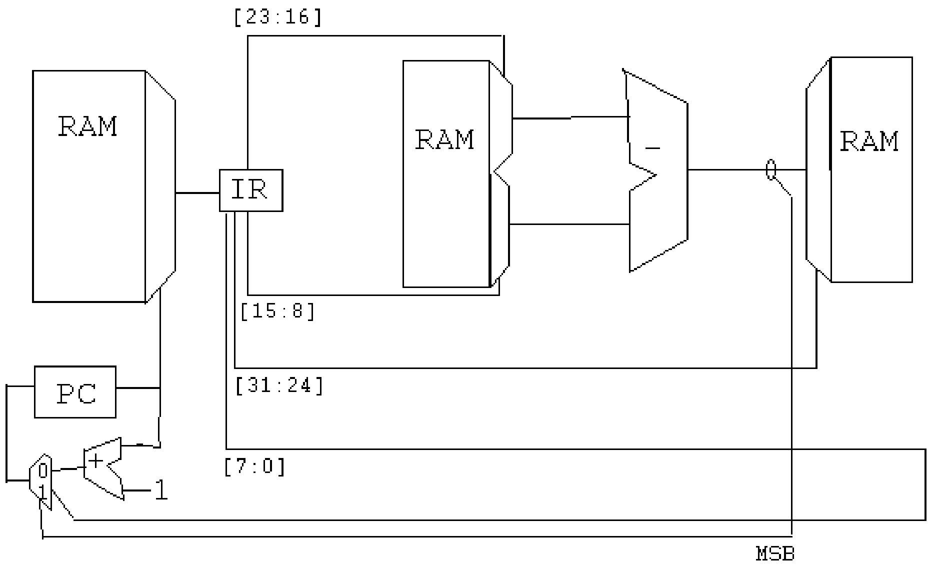


Figure 1

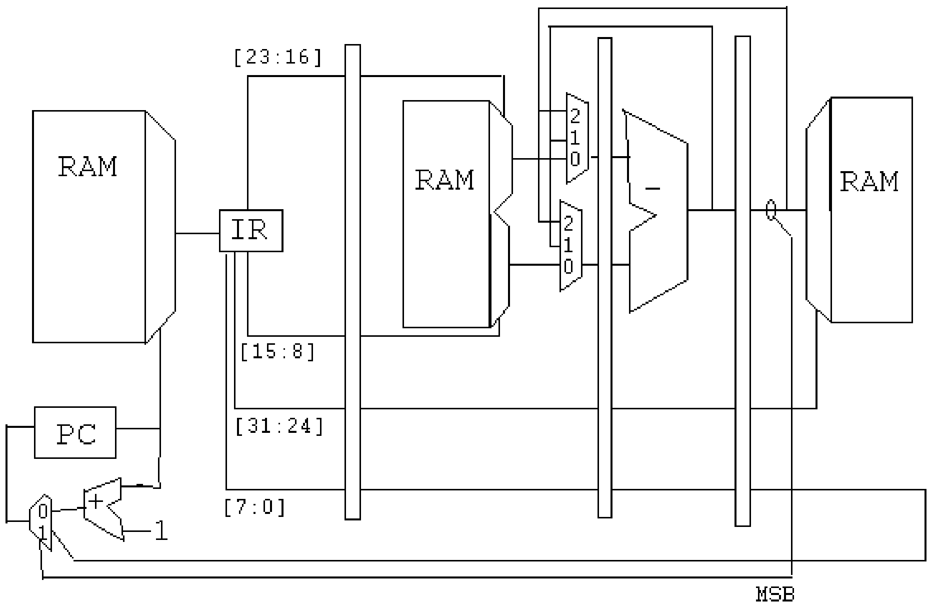


Figure 2

Figure A.1: One Command Computer

Appendix B

Encryption

B.1 Modular Arithmetic

B.1.1 Congruence

We say a is congruent to b modulus n when $a - b$ is divisible by n. In mathematical notation, we write $a \equiv b \pmod{n} \quad \Leftrightarrow \quad a - b = kn$ for some integer k. Several important properties of congruence are

1. $a \equiv a \pmod{n}$
2. $a \equiv b \pmod{n} \quad \Rightarrow \quad b \equiv a \pmod{n}$
3. $\{\{a \equiv b \pmod{n}\} \cdot \{b \equiv c \pmod{n}\}\} \quad \Rightarrow \quad a \equiv c \pmod{n}$

Example 27

$$\begin{aligned} 8 &\equiv 29 \pmod{7} \\ 8 - 29 &= -21 \\ &= (-3)7 \end{aligned}$$

$$\begin{aligned} 9 &\equiv -15 \pmod{6} \\ 9 - (-15) &= 24 \\ &= (4)6 \end{aligned}$$

B.1.2 Modulus

Invariably confusion happens with integer division, modulus, and remainder involving negative numbers. The problem arises in the basic definition. For a dividend, $a \in \mathbb{Z}$ and a divisor, $b \in \mathbb{Z}$, the quotient, q and remainder r must satisfy

1. $\{r, q\} \in \mathbb{Z}$,
2. $a = b * q + r$,
3. $|r| < |d|$.

The problem comes with the last requirement, because many choices can be made. The three most justifiable definitions are below[1]

1. Truncate division preserves the magnitudes of the quotient and remainder, independent of the signs of the dividend and divisor. This forces the remainder to have the same sign as the dividend.

2. Floor division forces the remainder to have the same sign as the divisor.

3. Euclidean division defines $r \geqslant 0$ and thus ensures $b \times q \leqslant a$.

Each is defensible.

Truncate

Remainder's definition is based off the definition of integer division. Integer division, a/b, is defined for positive a and b to be the number q such that

1. $b \times q \leqslant a$,

2. $b \times (q+1) \geqslant a$.

When negative numbers are allowed the following requirement is added

3 $(-a)/b = a/(-b) = -(a/b)$,

still for a and b positive. One could summarize this as:

$$c/d \quad = \quad \mathrm{sgn}(c)\mathrm{sgn}(d)(|c|/|d|)$$

Given we now have quotient or integer division defined we can then define remainder such that

$$\begin{aligned} a &= a/b + aremb \\ aremb &= a - a/b. \end{aligned}$$

Note that the sign of the remainder is the same as the

Example 28 *Consider the following:*

$$\begin{aligned} 5/2 &= 2 & 5rem2 &= 1 \\ (-5)/2 &= -2 & (-5)rem2 &= -1 \\ 5/(-2) &= -2 & 5rem(-2) &= 1 \\ (-5)/(-2) &= 2 & (-5)rem(-2) &= -1 \end{aligned}$$

B.1.3 Addition

$$\{\{a \equiv b \pmod{n}\} \cdot \{c \equiv d \pmod{n}\}\} \quad \Rightarrow \quad a + c \equiv b + d \pmod{n}$$

[1] other definitions exist such as ceiling division and rounding division, but they do not correspond to the what most people think of division for positive numbers. Note, from the requirements nothing says $5/2 = 3r - 1$ but this is hardly what most people would think of, and thus would probably not be programmed very well.

B.1.4 Additive Inverse

$$
\begin{aligned}
a+\bar{a} &\equiv 0 \pmod{n} \\
a+\bar{a} &= kn, \qquad k \in \mathbb{Z} \\
\bar{a} &= kn-a, \qquad k \in \mathbb{Z}
\end{aligned}
$$

Example 29 *Find the additive inverse(s) of 3 mod 7.*

$$
\begin{aligned}
\bar{a} &= kn-a, \qquad k \in \mathbb{Z} \\
&= 7k-3, \qquad k \in \mathbb{Z}
\end{aligned}
$$

k	$\bar{a}$	$(3+\bar{a}) \mod 7$
1	*4*	$(3+4) \mod 7 = 0$
2	*11*	$(3+11) \mod 7 = 0$
3	*18*	$(3+18) \mod 7 = 0$
4	*25*	$(3+25) \mod 7 = 0$
$\vdots$	$\vdots$	

B.1.5 Multiplication

$$\{\{a \equiv b \pmod{n}\} \cdot \{c \equiv d \pmod{n}\}\} \quad \Rightarrow \quad ac \equiv bd \pmod{n}$$

B.1.6 Multiplicative Inverse

$$
\begin{aligned}
a\bar{a} &\equiv 1 \pmod{n} \\
a\bar{a} &= 1+kn, \qquad k \in \mathbb{Z} \\
\bar{a} &= \frac{1+kn}{a}, \qquad k \in \mathbb{Z}
\end{aligned}
$$

Let $k_1 + ak_2 = k$ for k_1 and k_2 positive integers.

$$
\begin{aligned}
\bar{a} &= \frac{1+kn}{a}, \qquad k \in \mathbb{Z} \\
&= \frac{1+k_1n+ak_2n}{a}, \qquad k_1, k_2 \in \mathbb{Z}^+ \\
&= \frac{1+k_1n}{a} + k_2n, \qquad k_1, k_2 \in \mathbb{Z}^+
\end{aligned}
$$

We need a to divide $1+k_1n$, which means it divides with no remainder (aka divides evenly). Consider what would happen if $gcd(a,n) = a_1 > 1$, thus $a = a_1a_2$ and $n = a_1n_2$ for a_1, a_2, and n_2 positive integers. If a_1 is a factor of n then it is also a factor of k_1n If a_1 is a factor of k_1n then it cannot be a factor of k_1n+1 (it evenly divides k_1n and k_1n+k_1 but nothing in between).

Now assume $gcd(a,n) = 1$. For a to divide $1+k_1n$ implies $ak_3 = 1+k_1n$ for some positive integer k_3.

Example 30 *Find the multiplicative inverse(s) of 3 mod 7.*

$$
\begin{aligned}
\bar{a} &= \frac{1+kn}{a}, \qquad k \in \mathbb{Z} \\
&= \frac{1+7k}{3}, \qquad k \in \mathbb{Z}
\end{aligned}
$$

k	$\bar{a}$	$(3+\bar{a}) \mod 7$
1	$\frac{8}{3}$ no	
2	$\frac{15}{3} = 5$	$(3 \times 5) \mod 7 = 1$
3	$\frac{22}{3}$ no	
4	$\frac{29}{3}$ no	
5	$\frac{36}{3} = 12$	$(3 \times 12) \mod 7 = 1$
$\vdots$	$\vdots$	

B.2 Affine Encryption Program

Affine encryption is one of the simplest methods for doing encryption. Let P_i be the i^{th} character in the plain text message, and let C_i be the corresponding encoded character. Let there be n possible characters to encode, then the basic idea is to pick two numbers (a, b) to encode a message such that $\gcd(a, n) = 1$ (so a has an inverse). No requirement on b is needed if your modulus function has been encoded correctly. The encoded character can then be found by

$$a \times P_i + b \quad = \quad C_i \mod n.$$

Note that the " $\mod n$" at the end says the equation holds in $\mathbb{Z}_n$, the set of integers mod n with appropriately defined arithmetic.

To decrypt the message, the equation

$$\bar{a} \times (C_i + d) \quad = \quad P_i \mod n$$

is used. The term $\bar{a}$ is the inverse of a in $\mathbb{Z}_n$, which is found by solving

$$
\begin{aligned}
a \times \bar{a} &= 1 \mod n \\
&\textbf{or} \\
a \times \bar{a} &= m \times n + 1.
\end{aligned}
$$

Note that m is any whole number. The term d is the additive inverse of b in $\mathbb{Z}_n$, which is found by solving

$$d = n - (b \mod n).$$

We can summarize this by saying an affine cipher is an encryption technique that encodes using three integers: a, b, and n. If *plain* is the character to be encoded (with 'A'=0 and 'Z'=25) then $code = (a*plain+b) \mod n$. Decoding is also done using three integers: c, d, and n. If *code* is the character to be encoded (with 'A'=0 and 'Z'=25) then $plain = (c*(code+d)) \mod n$. The requirements on (a, b, c, d, n) are:

- $\gcd(a, n) = 1$
- $(ac) \mod n = 1$
- $(b + d) \mod n = 0$

Below is C code to implement a particular case of affine cyphers.

```
char affine_encode(char plain){
    // affine codes capital letter in plain using a=5, b=12 thus this is modulo 26
    int iCode, iPlain, a=3,b=0;

    // convert char to integer and shift so A=0
    iPlain=int(plain)-65;

    // do the encoding
    iCode = (a*iPlain+b)%26;

    // return the result as a char
    return char(iCode+65);
}

char affine_decode(char code){
    // affine decodes capital letter in plain using c=21, d=8 thus this is modulo 26
    int iCode, iPlain, c=9, d=0;

    // convert char to integer and shift so A=0
    iCode=int(code)-65;

    // do the decoding
    iPlain = (c*(iCode+d))%26;

    // return the result as a char
    return char(iPlain+65);
}
```

Appendix C

Projects for CSCI 313

C.1 Data Compression/Uncompression

Write in SPARC assembly a program that would use Huffman coding to compress an ASCII file and then uncompress the same file using Huffman coding in reverse.

The following table presents the relative frequencies of letters in the English language.

Letter	Freq.	Letter	Freq.	Letter	Freq.
A	0.0681	K	0.0037	U	0.0272
B	0.0123	L	0.0355	V	0.0095
C	0.0288	M	0.0257	W	0.0144
D	0.0406	N	0.0628	X	0.0025
E	0.1205	O	0.0671	Y	0.0146
F	0.0283	P	0.0210	Z	0.0004
G	0.0134	Q	0.0009	space	0.0600
H	0.0580	R	0.0514	.	0.0400
I	0.0577	S	0.0496	newline	0.0090
J	0.0018	T	0.0752		

You must first derive the decode tree using the above table and then create the translation table manually. The translation table can then be used to compress and decompress ASCII files.

C.2 Postfix Expression Evaluator

The project is to write in SPARC assembly a program that would evaluate a postfix expression. The postfix expression will contain the following arithmetic operators:

+ binary addition

- binary subtration

***** binary multiplication

/ binary division

? unary increment

! unary decrement

$\sim$ unary negation

The infix expression

$$5\,/\,2\,?\,+\,4 * 6 - 1 * 3$$

is equivalent to the postfix expression

$$5\,2\,?\,/\,4\,6 * +\,1\,3 * -.$$

The following is the algorithm for the postfix expression evaluator.

```
procedure EVAL (E)
    /*  Evaluate the postfix expression E. It is assumed
        that the last character in E is a NUL. A procedure
        NEXT-TOKEN is used to extract from E the next token.
        A token array STACK(1:n) is used as a stack.
    */

    top ← 0
    loop
        x ← NEXT-TOKEN(E)
            case
            :x = NUL: return
            :x is an operand: call PUSH(x,STACK)
            :else: remove the correct number of opeands
                for operator x from STACK, perform
                the operation and store the result,
                if any, onto the STACK
            end
    forever
end EVAL
```

Appendix D

Mini: ALU

D.1 Half Adder

Let's begin this section by considering a simple problem of how to design an adder for two bits. Call the bits "a" and "b". The sum will take two bits to hold, "carry" (c) and "sum" (s).

a	0	0	1	1
+b	+0	+1	+0	+1
cs	00	01	01	10

We can express this as a table.

a	b	c	s
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

From the table we can recognize that $c = a \cdot b$ and $s = a \oplus b$.

```
// name: half_adder
// desc: adds two single bits and outputs the the two bit answer [C,S]
// date:
// by  :
module half_adder(C,S,a,b); // you list all inputs and outputs, by convention outputs go first
  input  a, b;              // this tells the compiler which lines are inputs, outputs, and inouts
  output C, S;

  parameter delay=1;        // this creates a parameter that can be changed when it is
                            // instantiated, default value is 1

  and #delay carry(C,a,b);  // this instantiates a gate, sets its parameter to delay (time delay)
  xor #delay sum(S,a,b);    //  and passes the wires a,b as inputs to the gate and gets the
                            //  gate driven wires C or S as outputs
endmodule
```

D.2 Full Adders

We really want to have a way to add three bits, the two bits of the current digit and one bit carried from the previous sum.

$$\begin{array}{r} c_{prev} \\ \text{a} \\ \underline{+\text{b}} \\ \text{cs} \end{array}$$

As before we could make a table, but it is not necessary we can just add in pairs:

half-adder 1	half-adder 2
$\begin{array}{r} \text{a} \\ \underline{+\text{b}} \\ c_i s_i \end{array}$	$\begin{array}{r} c_{prev} \\ \underline{+s_i} \\ c_j s \end{array}$

and we **or** (as in the gate) the two intermediate carries together ($c = c_i + c_j$). Thus we can implement it as two half adders. Create a module "full_adder" and instantiate two half adders per the table to generate the sum and the intermediate carries, then combine the carries with an or gate to generate the carry out. Make sure you also include the parameter delay, or the next level up will not be able to change the levels below it.

D.3 Adder-Subtractor

In the previous section you made a module of a full adder. In this preparation you will make a four bit adder subtractor using your full adder module. We will add a feature called carry enable, which when set (equal 1) causes the adder-subtractor to act normally, but when unset (equal 0) stops the carry from being passed, and thus turn the adder-subtractor into an **xor** gate (from the half-adder).

1. Create a new module with two four bit inputs for the numbers and a four bit output for the result. Your module should also have a carry in and a carry out line.

```
// name: four_bit_adder
// desc: four bit ripple carry adder with carry enable,
//         if C_en then [C_out,Z] = x+y+C_in
//                 else Zi = Xi xor Yi
// date:
// by  :
module four_bit_adder(Z,C_out,x,y,C_in,c_en);
  input  C_in, c_en;
  input  [3:0] x,y;
  output C_out;
  output [3:0] Z;

  parameter delay=1;        // this creates a parameter that can be changed when it is
                            // instantiated, default value is 1

endmodule
```

2. Now create 7 wires to hold the intermediate carries between the full adders and the and gates that will connect them.

3. Make four instances of your full adder, being sure to pass it the delay parameter.

4. Create four **and** gates to do the carry enable logic. Be sure to give them the parameter "delay" so we can do timing later. The outputs of each And gate should be connected to the carry in of one of the full adders. One of the inputs of each and gate should be connected to C_en.

5. Finally, connect the carry outs of the first three adders to the **and** gate of the next bit. The first **and** gate gets C_in.

Test your full adder with the following module:

```
// name: test1
// desc: tests four bit ripple carry adder
// date:
// by  :
module test1();
  reg [3:0] a, b;
  reg c0, cen;
  wire [3:0] s;
  wire c4;

  // create instance of adder
  four_bit_adder #1 adder(c,c4,a,b,c0,cen);

  // set up the monitoring
  initial
  begin
    $display("A    B    C0   C4 S     Time");
    $monitor("%b   %b   %b   %b %b    %d", a,b,c0,c4,s,$time);
  end

  // run through a series of numbers
  initial
  begin
        a=4'b0000; b=4'b0000; c0=1'b0; cen=1'b1;
    #10 a=4'b0100; b=4'b0000; c0=1'b0; cen=1'b1;
    #10 a=4'b0100; b=4'b0011; c0=1'b0; cen=1'b1;
    #10 a=4'b0100; b=4'b0011; c0=1'b1; cen=1'b1;
    #10 a=4'b1100; b=4'b0011; c0=1'b1; cen=1'b1;
    #10 a=4'b1100; b=4'b0011; c0=1'b0; cen=1'b1;
    #10 a=4'b0100; b=4'b0000; c0=1'b0; cen=1'b0;
    #10 a=4'b0100; b=4'b0011; c0=1'b0; cen=1'b0;
    #10 a=4'b0100; b=4'b0011; c0=1'b1; cen=1'b0;
    #10 a=4'b1100; b=4'b0011; c0=1'b1; cen=1'b0;
    #10 a=4'b1100; b=4'b0011; c0=1'b0; cen=1'b0;
    #10 $finish;
  end

endmodule
```

Once your four bit adder is working, you need to make a four bit adder subtractor from it. See Figure 4-13 in Morris Mano, Digital Design, page 127 for a diagram of a ripple carry adder-subtractor. For simplicity we will not calculate overflow (V). Follow these steps:

1. Create a new module.

```
// name: four_bit_adder_subtractor
// desc: four bit ripple carry adder, [C_out,Z] = x+(-)y+C_in
// date:
// by  :
module four_bit_adder_subtractor(Z,C_out,x,y,sub,mode_arith);
  input  sub, mode_arith;
  input  [3:0] x,y;
  output C_out;
  output [3:0] Z;

  parameter delay=1;        // this creates a parameter that can be changed when it is
                            // instantiated, default value is 1

endmodule
```

2. Add a four bit wire called “w”, which will hold the output of the four **xor** gates in Figure 4-13. Don’t forget the delay parameter.

3. Create four **xor** gates whose inputs are the bits of “y” with ”sub” outputs are the bits of “w”. Don’t forget the delay parameter.

4. Make an instance of your adder subtractor and pass it “x”, “w”, “sub”, and “mode_arith”. Don’t forget the delay parameter.

5. Modify test1 to verify the design.

Appendix E

Mini: Register File

E.1 Register File

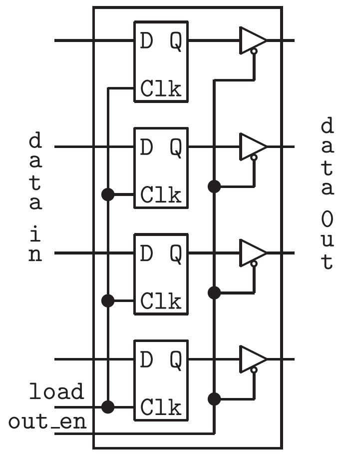
In this lab you will be making the register file (memory) for the Mini. In the preparation you will be designing the register file in Verilog. First read section 5-5 in the book (pages 190-197). The registers in the Mini each hold one nibble (half a byte, i.e.: four bits). The register file is made up of four registers. We will design our register file in four steps:

1. **create a D flip-flop**

   A D flip-flop must hold 1 bit of data, and it only changes its data when the clock changes. We want a positive edge triggered flip-flop. Enter the D flip-flop, "D_FF" from example 5-2 on page 192 of the book.

2. **make a four bit register with D flip-flops**

   Create a module to hold our four bit register. Just like the picture.




```
// name: Nibble_Reg
// desc: four bit register with output enable (low),
//        made from D flip-flops
// date:
// by  :
module Nibble_Reg(data_out,data_in,load,out_en);
  input  [3:0] data_in;
  input        load,out_en;
  output [3:0] data_out;

  // wires between flip-flops and tri-state gates
  wire   [3:0] dff_out;

  // instantiate tri-state gates to do output enable
  bufif0 tri3(data_out[3],dff_out[3],out_en);
  bufif0 tri2(data_out[2],dff_out[2],out_en);
  bufif0 tri1(data_out[1],dff_out[1],out_en);
  bufif0 tri0(data_out[0],dff_out[0],out_en);
```

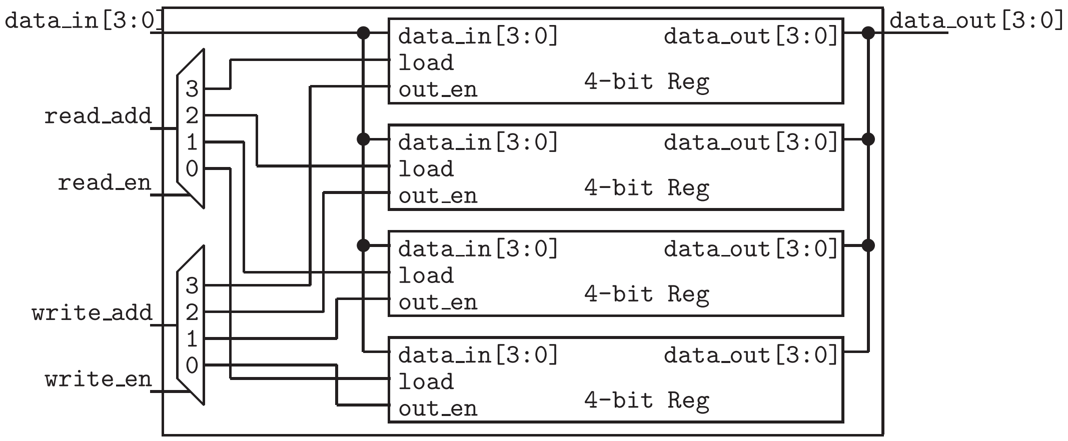
```
    //instantiate D flip-flops here
    D_FF Reg_Bit_3(dff_out[3],data_in[3],load);

       // you finish making instances

  endmodule
```

3. **create a 2 to 4 line decoder**

   We will need two decoders in the final step of our design so we will create them now. Enter the 2 to 4 line decoder, "decoder_df" from example 4-3 on page 153 of the book. To follow standard design practices we will make a few modifications.

   - Put "D" first in the port list. As a general rule, outputs are always first.
   - The ports "A" and "B" are actually the address bits so combine them into one new port "A" that has two bits. Note you will have to change the port list, input line, and the assignments.
   - Change the bit ordering of "D" from "[0:3]"(big endian) to "[3:0]"(little endian) to be consistent with the rest of the design

4. **build the register file from the registers**



   Create a module to hold our register file. Just like the picture

```
// name: Register_File
// desc: 4x4 register file
// date:
// by  :
module Register_File(data_out,data_in,read_add,read_en,write_add,write_en);
  input  [3:0] data_in;            // data to write
  input  [2:0] read_add,write_add; // read address and write address
  input        read_en,write_en;   // read and write enable
  output [3:0] data_out;           // data to read

  wire   [3:0] read_sel,write_sel;
```

```
  //instantiate registers here
  decoder_df Dec_Read(read_sel,read_en,read_add);
  decoder_df Dec_Write(write_sel,write_en,write_add);


  //instantiate registers here
  Nibble_Reg Reg_0(data_out,data_in,write_sel[0],read_sel[0]);

     // you finish making instances

endmodule
```

Appendix F

Mini: Timing

F.1 Timing

One of the main advantages of using a Hardware Description Language (HDL) like Verilog is the ability to simulate timing and performance of a circuit and work out any problems quickly before fabricating. In this lab we will be looking at the basic techniques of how this is done.

1. Use the VeriLogger Pro software that came with your book to do the following.
    (a) If you have not done so already, install Verilogger Pro.
    (b) Launch Verilogger Pro.
    (c) Under the "Project" tab, select "Add File(s)..." and add the files you created for Lab **??** and Lab **??**. They should appear in the "Project" window and show you all the modules that are defined in them.
    (d) Press the green play arrow. VeriLogger will automatically check your syntax, compile, and run if no errors are found. If it runs you will see your signals automatically plotted in the "Diagram" window.
2. Add gate delays by adding "parameter delay=0" to the top of each module with gates, which sets the default value to be zero (no delay). You can edit a module by double clicking its name in the "Project" window. We set a clock parameter because it allows us to easily change it later when we need. Parameters can even be changed when we instantiate them by placing a "#(value)" between the between the module name and the instance name when instantiating. At each gate declaration modify them so that you pass the time delay to them by adding a "#(delay)" before the gate name (see HDL Example 3.2 in the book). For example an xor gate would now look like "xor #(delay) x0(T[0], M, B[0]);". The delays are used by the simulator to see how long it takes for the signal to propagate through the circuit. We can graph the signals over time and thus see what is happening in any system we design. Make sure you modify all the following modules.
    - halfadder
    - fulladder
    - four_bit_adder_subtractor
    - four_bit_alu
3. Run the Verilog code. It should produce the same results since the delay is zero.

4. Modify the test module for the four bit alu so that the instantiation is now "four_bit_alu #(5) alu(s,c4,a,b,m,cen)" and run it. What happens and why?

5. Modify the delay a few times and see if you can predict what will happen each time. How long does it take to get the solution? How long is that in terms of gate delays? Can you express it as a formula?

F.2 Assembling

In this lab we will be timing a simple version of our cpu.

1. Create a module to contain our simple computer.

2. Add two four bit registers named "ACC" and "Op2".

3. Now make two registers to hold the signals "sub" and "mode".

4. Next create four wires called "result" and a single wire called carry.

5. Make an instance of the adder-subtractor and pass the registers and wires you created to it.

6. Just like you did for the test units create an initial unit and set the values of the registers to

   - ACC=0
   - Op2=5
   - sub=0
   - mode=1

   and setup a "$monitor" command to track the registers and wires.

7. Make a parameter called "wait" and set its value to the time you calculated in the preparation to get the solution.

8. Then make an always unit to control the flow of data in the computer. This essentially tells the accumulator to load the result of the alu.

```
always begin
    #(wait) ACC=result;
end
```

9. Run the computer. What does it do? Show the output to the instructor.

10. Set "wait" to twice its value. Does it still give the correct results? Why or why not?

11. Now set "wait" to half its initial value. Does it still work? Why or why not?

Nibble 1	Nibble 2	Nibble 3	Nibble 4	Instruction
0000				Add
0001				Sub
0010				Two Op Codes
	0000	S1	S2	Unsigned Multiplication, (U,V)<-S1 x S2
	0001	S1	S2	Signed Multiplication, (U,V)<-S1 x S2
	0010	S1	S2	Unsigned Division, U<- S1/S2, V<-S1 mod S2
	0011	D1	D2	Move D1 <- U, D2 <- V
	0100	D/S	ShiftAmt	Shift left logical by ShiftAmt
	0101	D/S	ShiftAmt	Shift left circulant by ShiftAmt
	0110	D/S	ShiftAmt	Shift right arithmetic by ShiftAmt
	0111	D	S	Not
0011	D	Imm		Set, D <- SE(Imm)
0100				And
0101				Or
0110				Xor
0111	D	S	Imm	Addi, D <- S + SE(Imm)
1000				
1001				
1010				Branching
	0leg	Address		branch conditionally, leg are flags for less, equal, or greater; PC <- PC + SE(Address)
	1000	S1	S2	Compare R0 <- S1-S2, set condition codes
	1100	R		Jump, PC <- PC+R, r15 <- PC+1
	1101	R		Jump, PC <- R
	1110	Code		Trap, call Trap(Code)
	1111			Return from Interrupt
1011	D	Imm		LEA, D <- PC+SE(Imm)
1100	D	S1	S2	Load Indexed, D <- m[S1+S2]
1101	D	S	Imm	Load Displaced D <- m[S + ZE(Imm)]
1110	S3	S1	S2	Store Indexed, m[S1+S2] <- S3
1111	S2	S1	Imm	Store Displaced m[S1 + ZE(Imm)] <- S2

Appendix G

7400 Series Part Numbers

Part	Description
00	4x Two input NAND
01	4x Two input NAND, Open collector
02	4x Two input NOR
03	4x Two input NAND, Open collector
04	6x Inverter (NOT)
05	6x Inverter (NOT), Open collector
06	6x Inverter (NOT), High voltage Open collector
07	6x Buffer (NO-OP), High voltage Open collector
08	4x Two input AND
09	4x Two inout AND, Open collector
10	3x Three input NAND
11	3x Three inout AND
12	3x Three input NAND, Open collector
13	2x Four input, Schmitt Trigger NAND
14	6x Inverter (NOT), Schmitt Trigger
15	3x Three input AND, Open collector
16	6x Inverter (NOT), High voltage Open collector
17N	6x Buffer (NO-OP), High voltage Open collector
19	6x Inverter (NOT), Schmitt Trigger
20	2x Four input NAND
21	2x Four input AND
22	2x Four input NAND, Open collector
23	2x Four input NOR with Strobe
25	2x Four input NOR with Strobe
26	4x Two input NAND, High voltage
27	3x Three input NOR
28	4x Two input NOR
30	Eight input NAND
31	6x DELAY (6nS to 48nS)
32	4x Two input OR
33	4x Two input NOR, Open collector
37	4x Two inout NAND
38	4x Two input NAND, Open collector
39	4x Two input NAND, Open collector
40	4x Two input NAND, Open collector

Part	Description
42	Four-to-Ten (BCD to Decimal) DECODER
45	Four-to-Ten (BCD to Decimal) DECODER, High current
46	BCD to Seven-Segment DECODER, Open Collector, lamp test and leading zero handling
47	BCD to Seven-Segment DECODER, Open Collector, lamp test and leading zero handling
48	BCD to Seven-Segment DECODER, lamp test and leading zero handling
49	BCD to Seven-Segment DECODER, Open collector
50	2x (Two input AND) NOR (Two input AND), expandable
51	(a AND b AND c) NOR (c AND e AND f) plus (g AND h) NOR (i AND j)
53	NOR of Four Two input ANDs, expandable
54	NOR of Four Two input ANDs
55	NOR of Two Four input ANDs
56P	3x Frequency divider, 5:1, 5:1, 10:1
57P	3x Frequency divider, 5:1, 6:1, 10:1
64	4-3-2-2 AND-OR-INVERT
65	4-3-2-2 AND-OR-INVERT
68	2x Four bit BCD decimal COUNTER
69	2x Four bit binary COUNTER
70	1x gated JK FF with preset and clear
72	1x gated JK FF with preset and clear
73	2x JK FF with clear
74A	2x D FF, edge triggered with preset and clear
75	4x D LATCH, gated
76A	2x JK FF with preset and clear
77	4x D LATCH, gated
78A	2x JK FF with preset and clear
83	Four bit binary ADDER
85	Four bit binary COMPARATOR
86	4x Two input XOR (exclusive or)
90	Four bit BCD decimal COUNTER
91	Eight bit SHIFT register
92	Four bit divide-by-twelve COUNTER
93	Four bit binary COUNTER
94	Four bit SHIFT register
95B	Four bit parallel access SHIFT register
96	Five bit SHIFT register
107A	2x JK FF with clear
109A	2x JK FF, edge triggered, with preset and clear
112A	2x JK FF, edge triggered, with preset and clear
114A	2x JK FF, edge triggered, with preset
116	2x Four bit LATCH with clear
121	Monostable Multivibrator
122	Retriggerable Monostable Multivibrator
123	Retriggerable Monostable Multivibrator
124	2x Clock Generator or Voltage Controlled Oscillator
125	4x Buffer (NO-OP), (low gate) Tri-state
126	4x Buffer (NO-OP), (high gate) Tri-state
130	Retriggerable Monostable Multivibrator
128	4x Two input NOR, Line driver
132	4x Two input NAND, Schmitt trigger
133	Thirteen input NAND
134	Twelve input NAND, Tri-state
135	4x Two input XOR (exclusive or)
136	4x Two input XOR (exclusive or), Open collector

Part	Description
137	3-8 DECODER (demultiplexer)
138	3-8 DECODER (demultiplexer)
139A	2x 2-4 DECODER (demultiplexer)
140	2x Four input NAND, 50 ohm Line Driver
143	Four bit counter and latch with 7-segment LED driver
145	BCD to Decimal decoder and LED driver
147	10-4 priority ENCODER
148	8-3 gated priority ENCODER
150	16-1 SELECTOR (multiplexer)
151	8-1 SELECTOR (multiplexer)
153	2x 4-1 SELECTOR (multiplexer)
154	4-16 DECODER (demultiplexer)
155A	2x 2-4 DECODER (demultiplexer)
156	2x 2-4 DECODER (demultiplexer)
157	4x 2-1 SELECTOR (multiplexer)
158	4x 2-1 SELECTOR (multiplexer)
159	4-16 DECODER (demultiplexer), Open collector
160A	Four bit synchronous BCD COUNTER with load and asynchronous clear
161A	Four bit synchronous binary COUNTER with load and asynchronous clear
162A	Four bit synchronous BCD COUNTER with load and synchronous clear
163A	Four bit synchronous binary COUNTER with load and synchronous clear
164	Eight bit parallel out SHIFT register
165	Eight bit parallel in SHIFT register
166A	Eight bit parallel in SHIFT register
169A	Four bit synchronous binary up+down COUNTER
170	4x4 Register file, Open collector
174	6x D LATCH with clear
175	4x D LATCH with clear and dual outputs
170	Four bit parallel in and out SHIFT register
180	Four bit parity checker
181	Four bit ALU
182	Look-ahead carry generator
183	2x One bit full ADDER
190	Four bit Synchronous up and down COUNTER
191	Four bit Synchronous up and down COUNTER
192	Four bit Synchronous up and down COUNTER
193	Four bit Synchronous up and down COUNTER
194	Four bit parallel in and out bidirectional SHIFT register
195	Four bit parallel in and out SHIFT register
198	Eight bit parallel in and out bidirectional SHIFT register
199	Eight bit parallel in and out bidirectional SHIFT register, JK serial input
221	2x Monostable multivibrator
240	8x Inverter (NOT), Tri-state
241	8x Buffer (NO-OP), Tri-state
242	4x Bidirectional, Tri-state inverting transceiver
243	4x Bidirectional, Tri-state transceiver
244	8x Buffer (NO-OP), Tri-state Line driver
245	8x Bidirectional Tri-state BUFFER
259	Eight bit addressable LATCH
260	2x Five input NOR
273	8x D FF with clear
279	4x SR LATCH
283	Four bit binary full ADDER
373	8x Transparent (gated) LATCH, Tri-state
374	8x Edge-triggered LATCH, Tri-state

Part	Description
629	Volatge controlled OSCILLATOR
688	Eight bit binary COMPARATOR

Bibliography