## Arithmetic

Computers do not store numbers or letters, per se. They only store bit sequences. The bit sequences can be interpreted as representing integers or floating point numbers. Arithmetic is accomplished by the direct hardware implementation of arithmetic algorithms

## Converting Decimal to b-Bit Binary

- Let d be decimal number, less than $2^{\mathrm{b}}-1$



## Representation

- Terminology
- least significant bit (Isb): least magnitude bit.
- most significant bit (msb): greatest magnitude bit.
- unsigned integers: k bit sequence representing 0 to $2^{\mathrm{k}}$ - 1
- two's complement: number representation for signed integers.

| Unsigned Binary | Decimal |
| :--- | ---: |
| 0000000000000000 | 0 |
| 0000000000000001 | 1 |
| 0000000000000010 | 2 |

111111111111110165533
111111111111111065534
111111111111111165535

Unsigned gets a larger range at the expense of no negative representation

| 2s Complement Binary | Decimal |
| :--- | :---: |
| 0000000000000000 | 0 |
| 0000000000000001 | 1 |
| 0000000000000010 | 2 |

011111111111111132767 $1000000000000000-32768$ $1000000000000001-32767$ $1000000000000010-32766$
$1111111111111101 \quad-3$ since it has 1 more
$\begin{array}{ll}1111111111111110 & -2 \\ 1111111111111111 & -1\end{array}$ $1111111111111111 \quad-1$ than positive numbers. Why?

## Converting 2s Complement

A positive number is its unsigned equivalent, provided the msb is 0 .
Finding -x given $x$ : Complement all bits \& add 1
In $n$-bit 2 s complement, $x+-x=2^{n}$

| $0000000000001010_{2}$ | $10_{10}$ |
| ---: | ---: |
| Compl.: 1111111111110101 |  |
| Add 1: $+\frac{11111111111011_{2}}{1111110}$ | $-10_{10}$ |
| Compl.: 0000000000001001 | 1 |
| Add 1: $+\overline{0000000000001010_{2}}$ | $10_{10}$ |

## Subtraction

- The well-known rule that subtraction is equivalent to addition of a negated operand is important in computing

$$
a-b \equiv a+(-b)
$$

- In the arithmetic-logic unit of a computer, there is no subtraction circuitry per se, just negation



## Consequences of Signed Fields

A field of $b$ bits can represent $2^{b}$ configurations

- If the field is used for unsigned numbers ...

Range: 0 to $2^{\text {b }}-1$

- If the field is used for signed numbers ...

Range: $-2^{b-1}$ to $2^{b-1}-1$

- If the field is used so that some bits are always the same, then do not represent them

Range of byte addresses for instructions:
0 to $2^{b+2}$ since least significant bits are 00


## Representations

A bit sequence is neither signed nor unsigned, integer or floating point, character or pixel ... its just a bit sequence -- the key is how the bits are interpreted

- The interpretation nearly always matters, especially in comparisons:
- $10110<00110$ is true since $-10<6$ as 2 s complement
- $10110>00110$ is true since $22>6$ as unsigned
- MIPS has additional comparison operators:
- sltu \$8, \$9, \$10 \#set less than unsigned - sltiu \$8, \$9, 10 \#set less than immed. unsign

Why are there no unsigned variants for beq, bne, bgtz, blez, sh, sb, etc.?

## Facts of Finite Representation

When combining two numbers produces a result larger than can be represented in the available space, an overflow occurs


Overflow is not always bad, but it must be reportable Overflow is impossible when ...

- adding numbers with opposite signs because the result is numerically between the operands
- subtracting numbers with like signs, since the "addition" rule applies once B is negated


## Conditions Causing Overflow

- For add operations (add and subtract), the overflow can be detected by the sign of the operands and the result


| Operation Op A Op B Overflow |  |  |  |
| :---: | :---: | :---: | :---: |
| A+B | $\geq 0$ | $\geq 0$ | <0 |
| A+B | $<0$ | <0 | $\geq 0$ |
| A-B | $\geq 0$ | <0 | <0 |
| A-B | <0 | $\geq 0$ | $\geq 0$ |

Notice that the subtraction rule is simply the Addition rule applied when subtraction is negation of the second operand followed by addition

