Great Theoretical Ideas In Computer Science

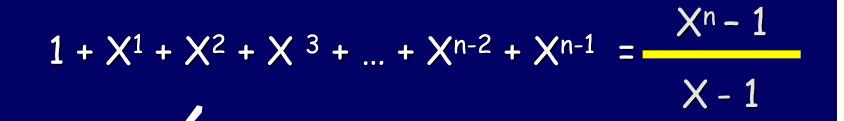
Steven Rudich

Lecture 3 Jan 20, 2004

CS 15-251 Spring 2004
Carnegie Mellon University

Unary, Binary, and Beyond





We are going to need this fundamental sum:

The Geometric Series

A Frequently Arising Calculation

$$(X-1)(1+X^1+X^2+X^3+...+X^{n-2}+X^{n-1})$$

=
$$X^1 + X^2 + X^3 + ...$$
 + $X^{n-1} + X^n$
- $1 - X^1 - X^2 - X^3 - ...$ - $X^{n-2} - X^{n-1}$

$$=$$
 $X^n - 1$

The Geometric Series

$$(X-1)(1+X^1+X^2+X^3+...+X^{n-1})=X^n-1$$

$$1 + X^{1} + X^{2} + X^{3} + ... + X^{n-2} + X^{n-1} = \frac{X^{n} - 1}{X - 1}$$

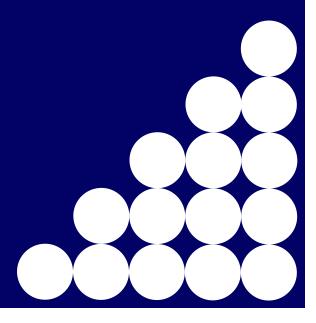
when X≠1



nth Triangular Number

$$\Delta_n = 1 + 2 + 3 + ... + n-1 + n$$

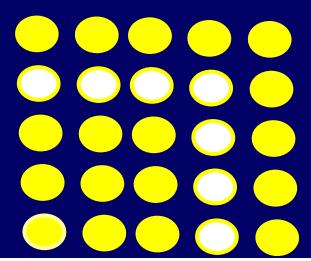
$$= n(n+1)/2$$



nth Square Number

$$n = 1 + 3 + ... + 2n-1$$

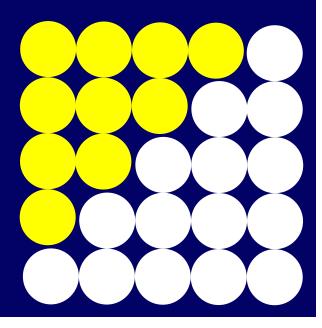
= Sum of first n odd numbers



nth Square Number

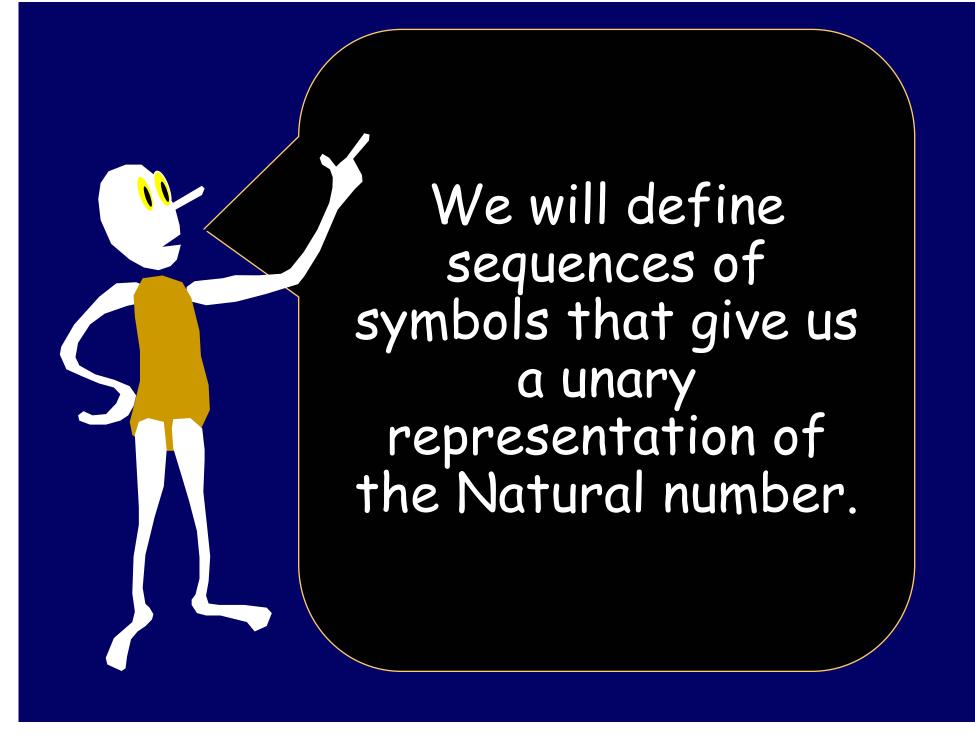
$$n = \Delta_n + \Delta_{n-1}$$

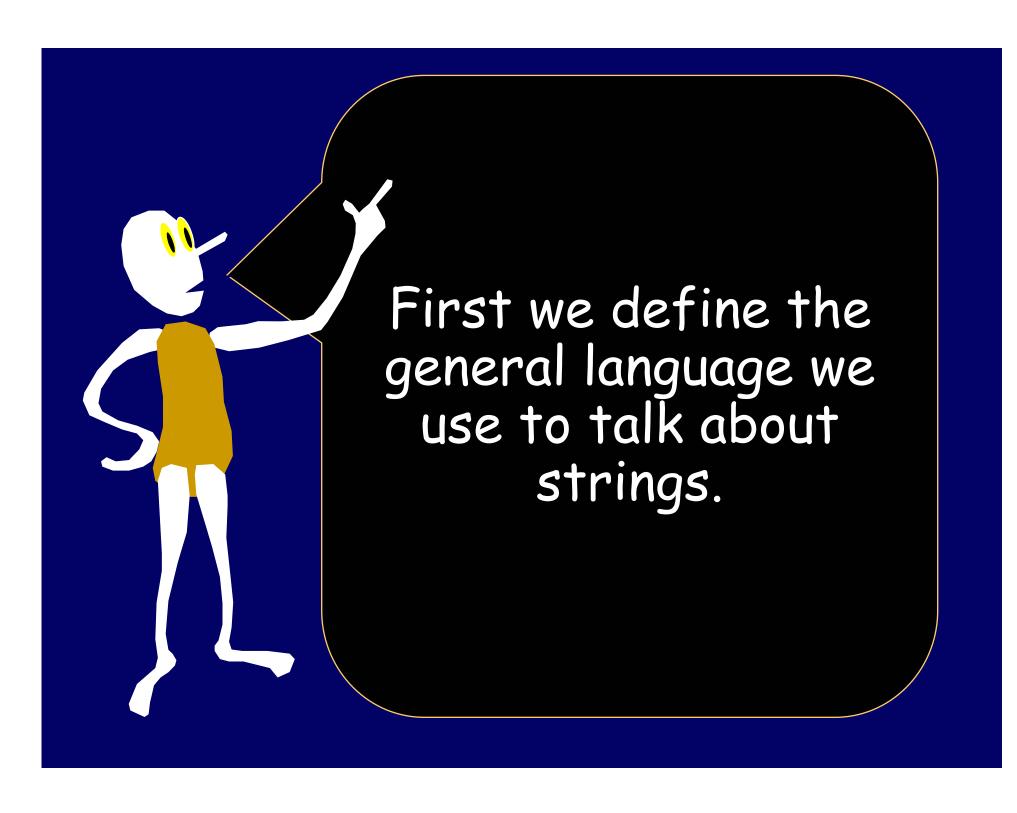
$$= n^2$$



$$(\Delta_n)^2 = (\Delta_{n-1})^2 + \square_n$$

$$(\Delta_n)^2 = \Box + \Box + \ldots + \Box_n$$





Strings Of Symbols.

We take the idea of symbol and sequence of symbols as primitive.

Let Σ be any fixed finite set of symbols. Σ is called an alphabet, or a set of symbols.

Examples:

 $\Sigma = \{0,1,2,3,4\}$

 $\Sigma = \{a,b,c,d,...,z\}$

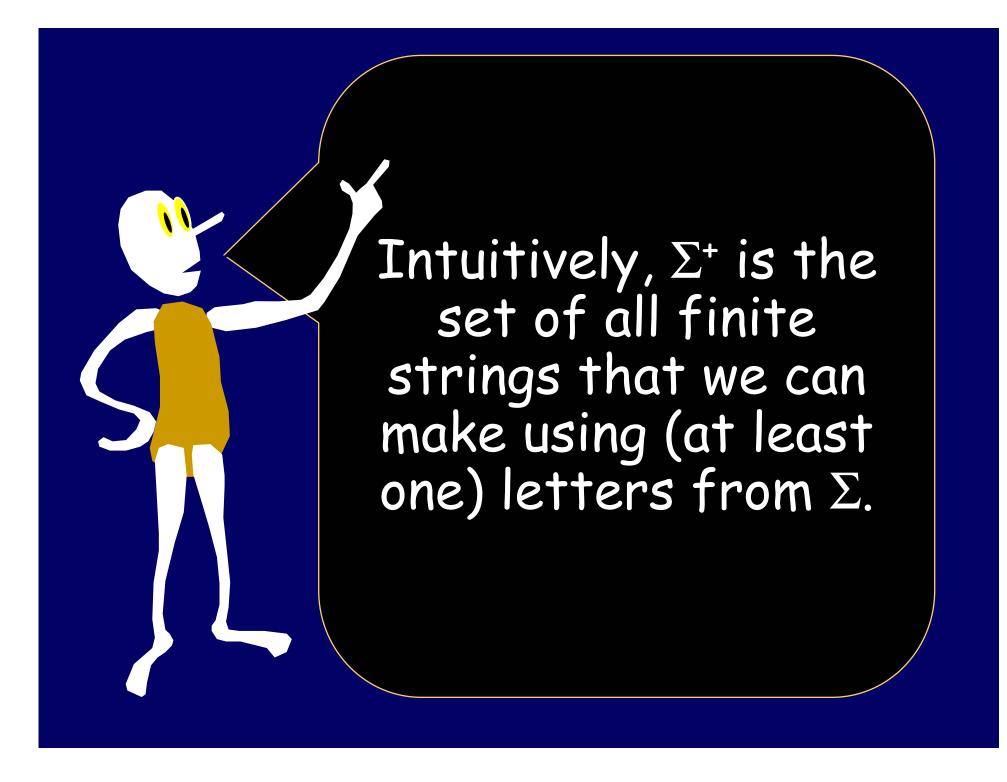
 Σ = all typewriter symbols.

Strings over the alphabet Σ .

A string is a sequence of symbols from Σ . Let s and t be strings. Let st denote the concatenation of s and t, i.e., the string obtained by the string s followed by the string t.

Define Σ^+ by the following inductive rules:

$$\mathbf{x}{\in}\Sigma\Rightarrow\mathbf{x}{\in}\Sigma^{\scriptscriptstyle\mathsf{+}}$$
 s,t $\in\Sigma^{\scriptscriptstyle\mathsf{+}}\Rightarrow$ st $\in\Sigma^{\scriptscriptstyle\mathsf{+}}$

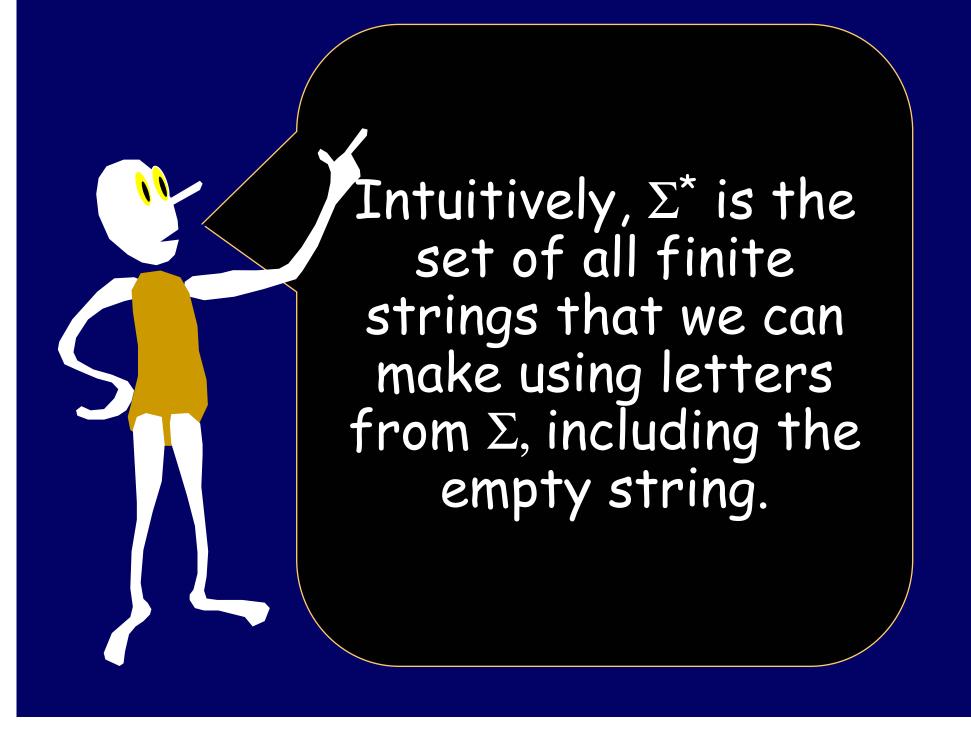


\sum^*

Define ϵ be the empty string. I.e., $X\epsilon Y=XY$ for all strings X and Y. ϵ is also called the string of length 0.

Define $\Sigma^0 = \{ \epsilon \}$

Define $\Sigma^* = \Sigma^+ \cup \{\epsilon\}$



The Natural Numbers

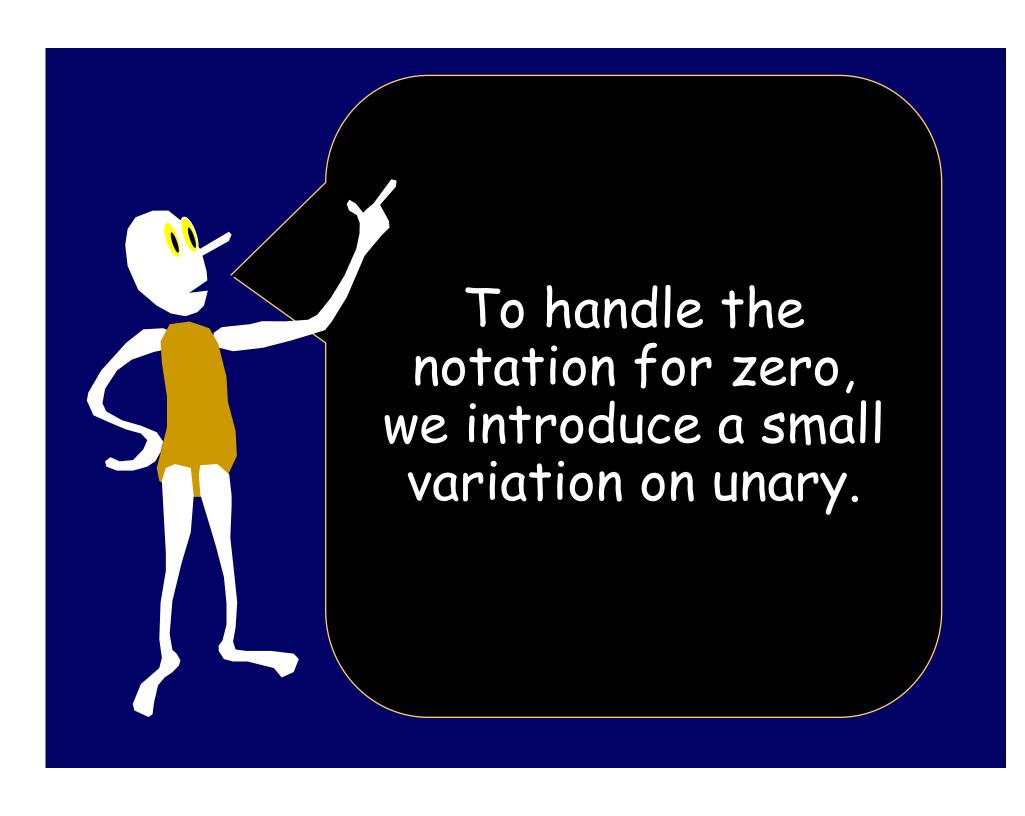
$$N = \{ 0, 1, 2, 3, \ldots \}$$

Notice that we include 0 as a Natural number.

"God Made The Natural Numbers. Everything Else Is The Work Of Man." Kronecker

$$N = \{ 0, 1, 2, 3, \ldots \}$$

Last Time: Unary Notation



Peano Unary (PU)

0123456

Giuseppe Peano [1889]

Each number is a sequence of symbols in {5,0}+

0 0

1 50

2 550

3 \$550

4 \$5550

6 555550

Peano Unary Representation of Natural Number

 $N = \{ 0, 50, 550, 5550, \ldots \}$

O is a natural number called zero.

Set notation: $0 \in N$

If X is a natural number, then SX is a <u>natural</u> <u>number</u> called <u>successor of X</u>.

Set notation: $X \in \mathbb{N} \Rightarrow SX \in \mathbb{N}$

Inductive Definition of +

$$N = \{ 0, 50, 550, 5550, \ldots \}$$

Inductive definition of addition (+):

X,
$$Y \in \mathbb{N} \Rightarrow$$

X "+" 0 = X
X "+" SY = S(X"+"Y)

Proof:

$$X, Y \in \mathbb{N} \Rightarrow$$

$$X "+" 0 = X$$

$$X "+" SY = S(X"+"Y)$$

Inductive Definition of *

 $N = \{ 0, 50, 550, 5550, \ldots \}$ Inductive definition of times (*):

$$X, Y \in \mathbb{N} \Rightarrow$$

$$X "*" 0 = 0$$

$$X "*" SY = (X"*"Y) + X$$

Inductive Definition of ^

$$N = \{ 0, 50, 550, 5550, \ldots \}$$

Inductive definition of raised to the (^):

X,
$$Y \in \mathbb{N} \Rightarrow$$

X "^" $0 = 1$ [or $X^0 = 1$]
X "^" $SY = (X''^"Y) * X$ [or $X^{SY} = X^Y * X$]

$N = \{ 0, 50, 550, 5550, \ldots \}$

Defining < for N:

```
\forall x,y \in \mathbb{N}
"x > y" is TRUE \Leftrightarrow "y < x" is FALSE
"x > y" is TRUE \Leftrightarrow "y > x" is FALSE
```

"x+1 > 0" is TRUE "x+1 > y+1" is TRUE \Rightarrow "x > y" is TRUE

$$N = \{ 0, 1, 2, 3, \ldots \}$$

Defining partial minus for N:

$$\forall x,y \in \mathbb{N}$$
 $x-0 = x$
 $x>y \Rightarrow$
 $(x+1) - (y+1) = x-y$

a = [a DIV b]*b + [a MOD b]

Defining DIV and MOD for N:

$$\forall a,b \in \mathbb{N}$$

 $a < b \Rightarrow$
 $a \text{ DIV } b = 0$
 $a \ge b > 0 \Rightarrow$
 $a \text{ DTV } b = 1 + (a - b)$

The maximum number of times b goes into a without going over.

a DIV
$$b = 1 + (a-b)$$
 DIV b

$$a MOD b = a - [b*(a DIV b)]$$

The remainder when a is divided by b.

Defining DIV and MOD for N:

```
\forall a,b \in \mathbb{N}

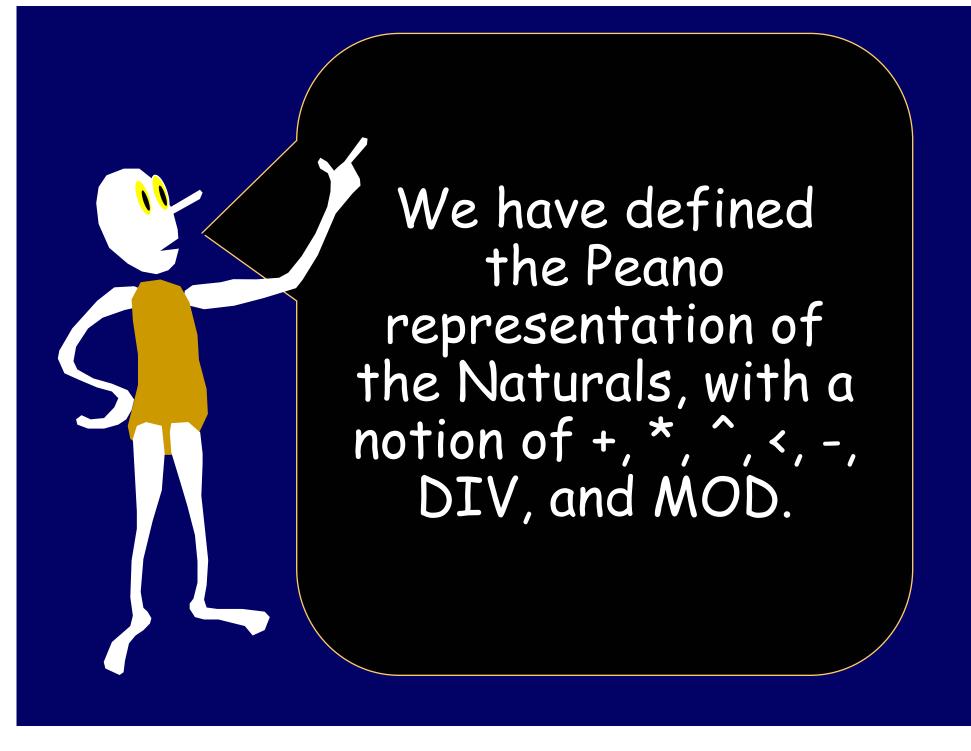
a < b \Rightarrow

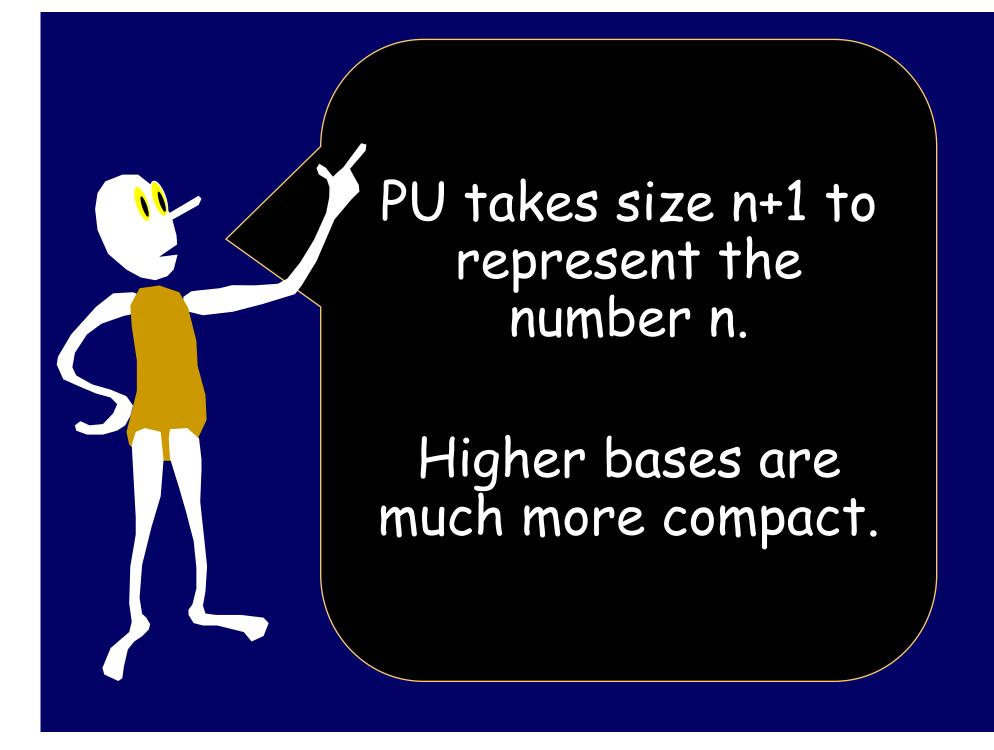
a \text{ DIV } b = 0

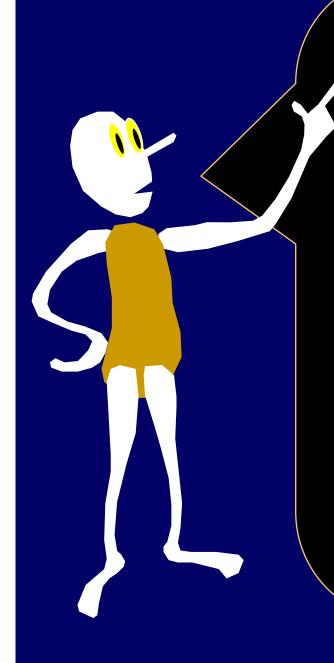
a \ge b > 0 \Rightarrow

a \text{ DIV } b = 1 + (a-b) \text{ DIV } b

a \text{ MOD } b = a - [b*(a \text{ DIV } b)]
```

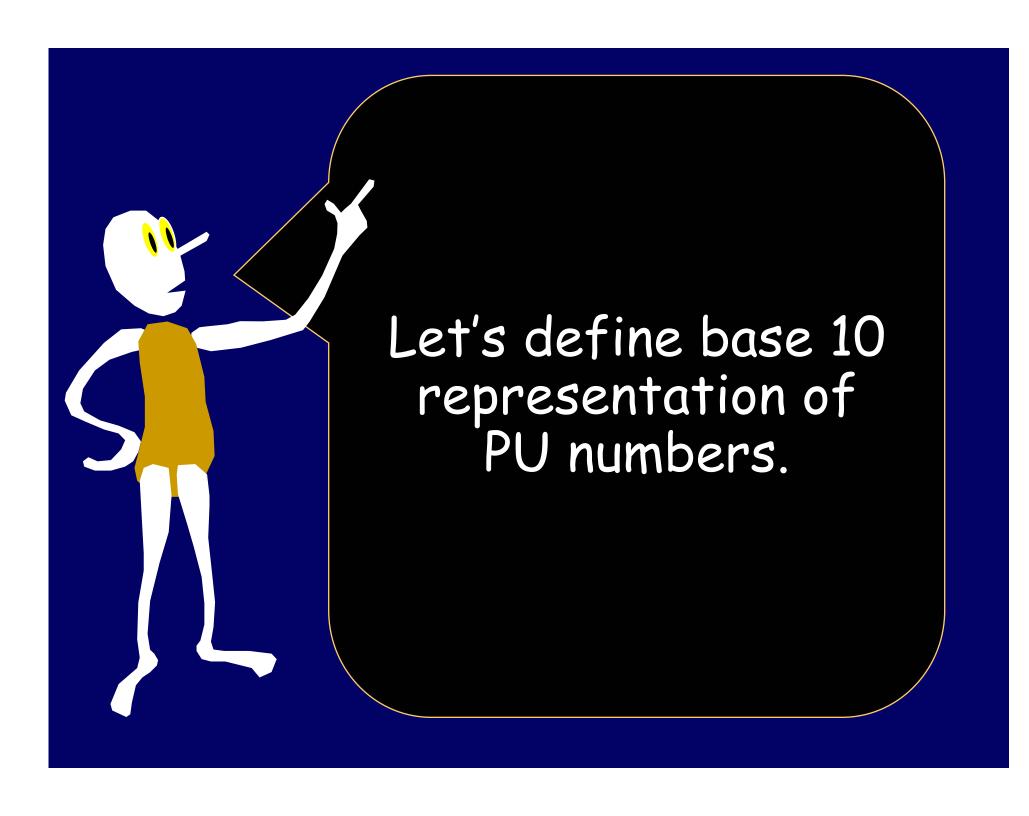


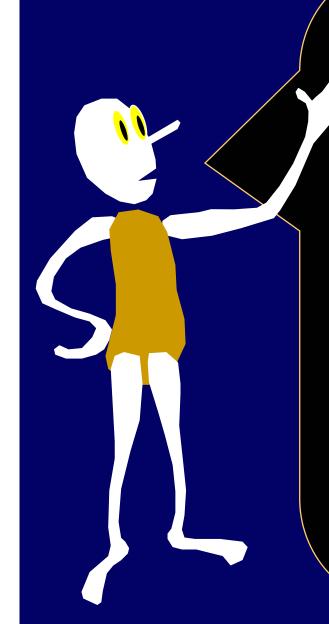




1000000 in Peano Unary takes one million one symbols to write.

1000000 only takes 7 symbols to write in base 10.





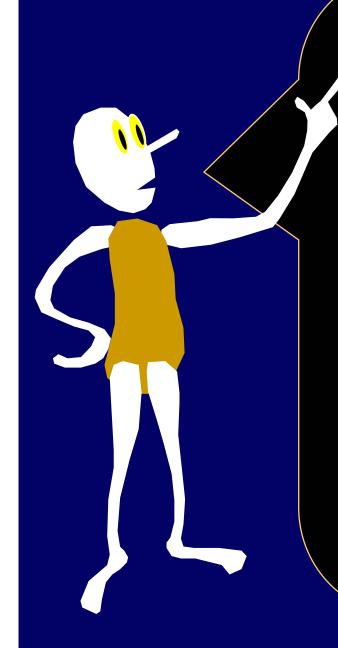
Let $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ be our symbol alphabet.

Any string in Σ^+ will be called a decimal number.



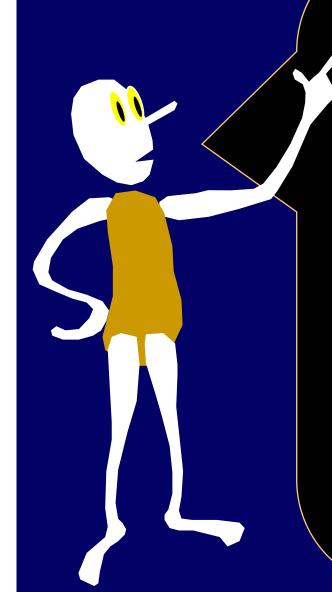
length(ε) = 0

 $X=aY, a \in \Sigma, Y \in \Sigma^* \Rightarrow$ length(X) = S(length(Y))



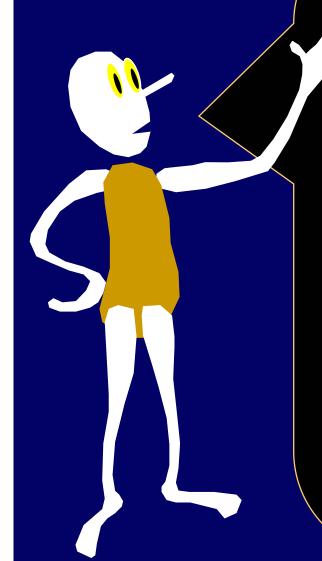
Let X be decimal. Let n (unary) = length(X).

For each unary in, we want to be able to talk about the ith symbol of X.



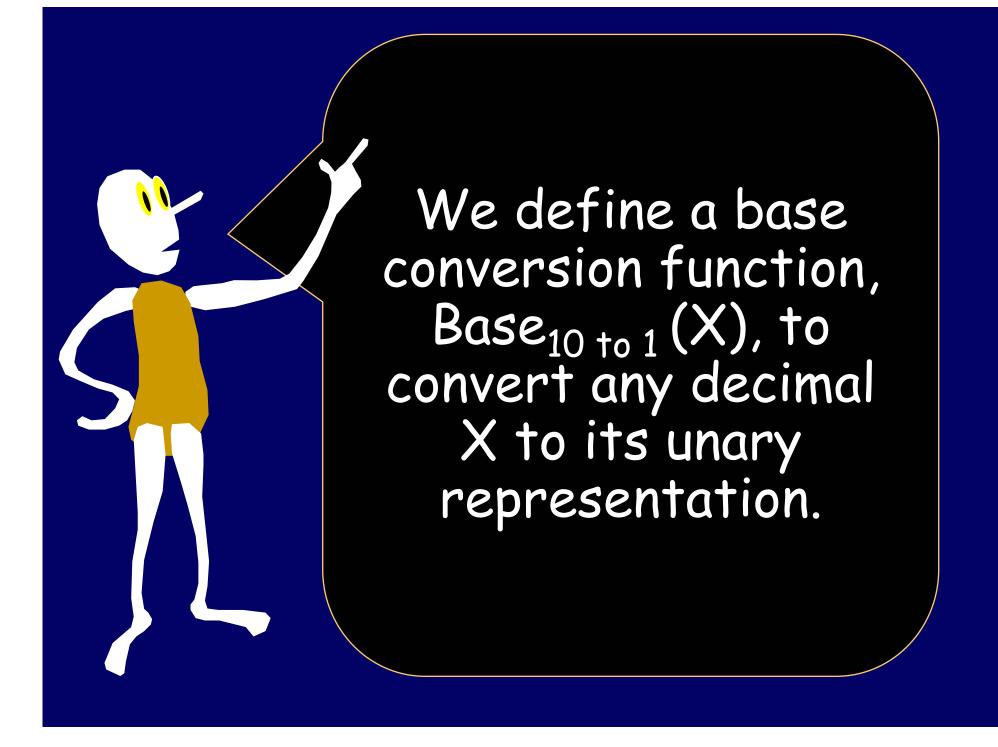
The ith symbol of X. Defining rule:

X = PaS where $P,S \in \Sigma^*$ and $a \in \Sigma$. i = length(S). \Rightarrow a is the i^{th} symbol of X.



For any string X, we can define its length $n \in PU$. For all $i \in PU$ s.t. i < n, we can define the i^{th} symbol a_i .

 $X = a_{n-1} a_{n-2} ... a_0$



Initial Cases, length(X)=1

Suppose n=length(X)>50

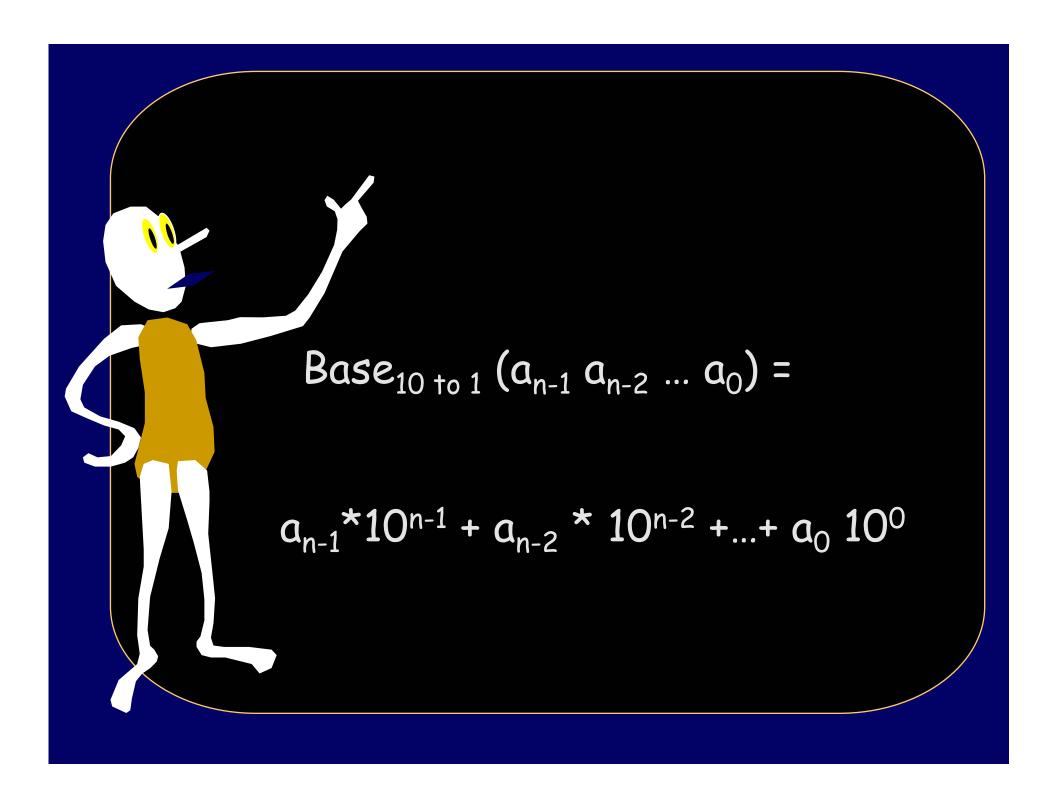
For all i<n, let a_i be the ith symbol of X. Hence, $X = a_{n-1} a_{n-2} ... a_0$.

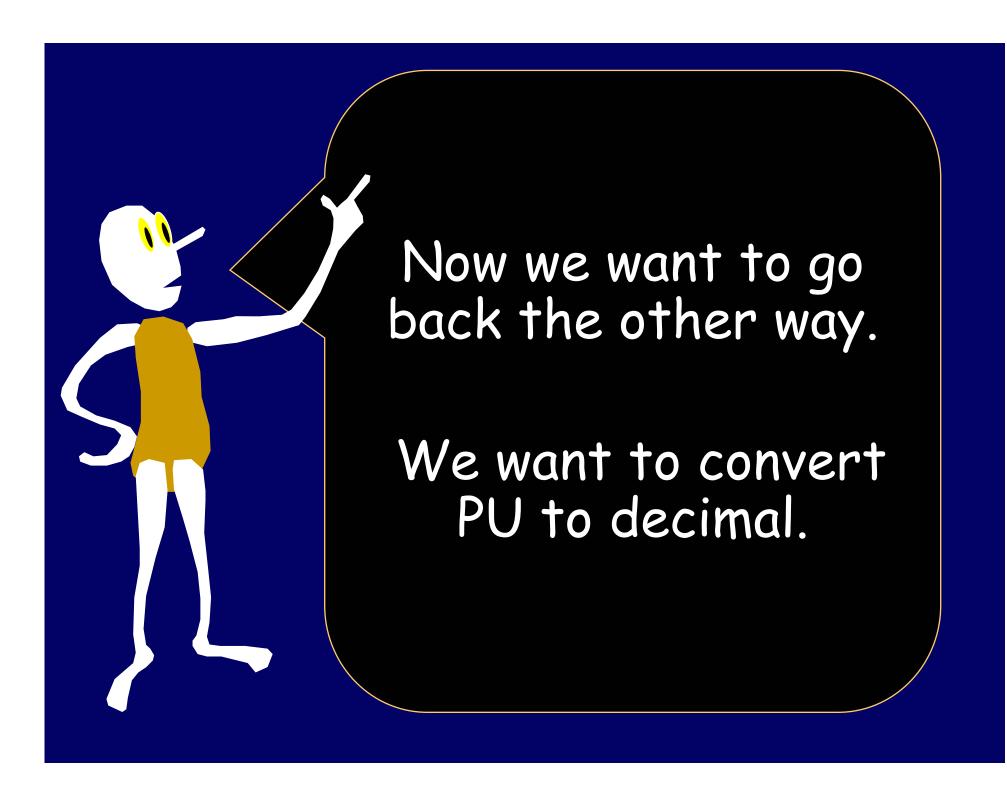
Define Base_{10 to 1}(X) =
$$\Sigma_{i < n}$$
 Base_{10 to 1}(α_i)*10ⁱ

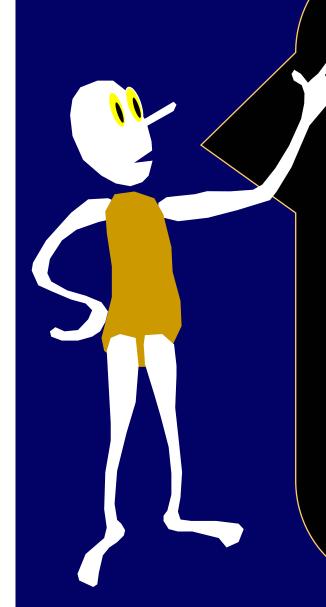
where +, *, and ^ are defined over PU

Example X= 238

Base_{10 to 1}
$$(238) = 2*100 + 3*10 + 8$$

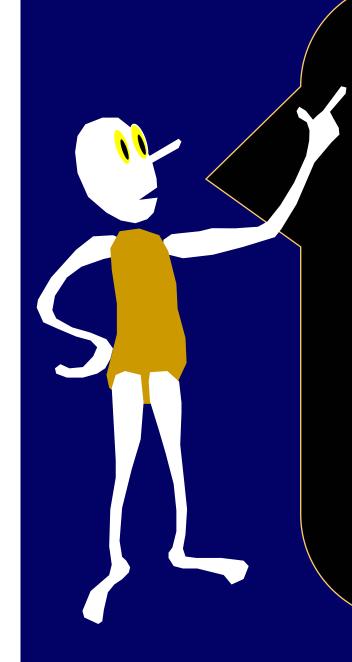






No Leading Zero:

Let NLZ be the set of decimal numbers with no leading 0 (leftmost symbol ≠ 0), or the decimal number 0.



We define Base_{1 to 10} from PU to NLZ

It will turn out to be the inverse function to Base_{10 to 1}.

One digit cases.

••••

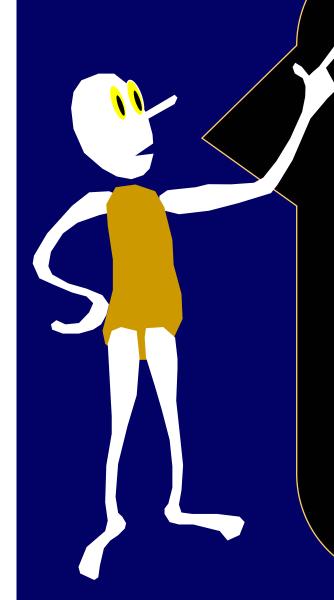
Base_{1 to 10}(SSSSSSSSS) = "9"

Suppose X > 9

Let n be the smallest unary number such that $10^n > X$, $10^{n-1} \le X$.

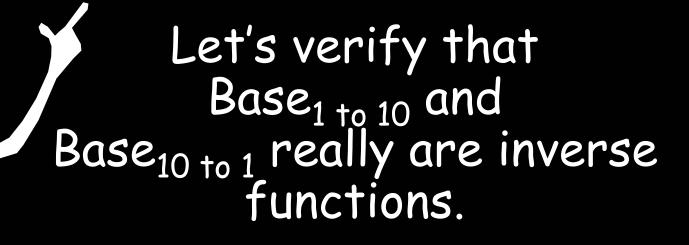
Let $d = X DIV 10^{n-1}$ [Notice that $1 \le d \le 9$] Let $Y = X MOD 10^{n-1}$

Define $Base_{1 to 10}(X) \in NLZ$ to be $Base_{1 to 10}$ (d) $Base_{1 to 10}(Y)$



For each n∈ PU define its decimal representation to be Base_{1 to 10} (n).

Base_{1 to 10} goes from PU to NLZ.



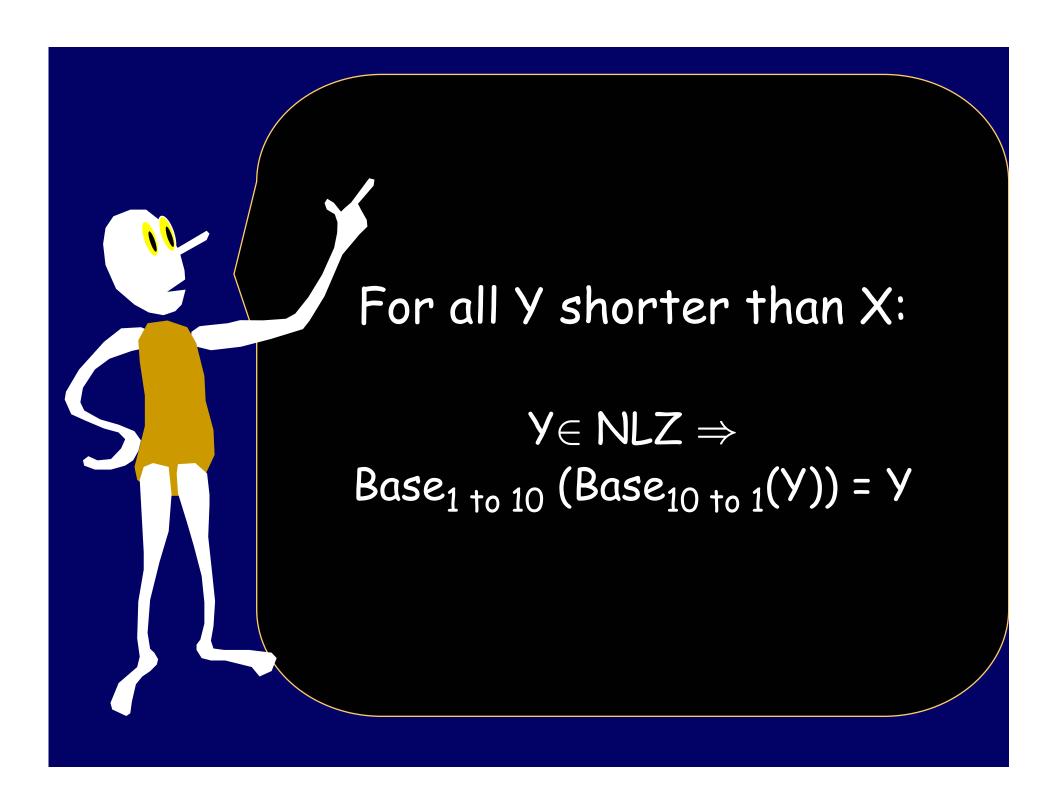
We need to show $X \in NLZ \Rightarrow$ Base_{1 to 10} (Base_{10 to 1} (X)) = X.

Clear when X is a single digit.

```
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (0)) = 0
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (1)) = 1
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (2)) = 2
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (3)) = 3
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (4)) = 4
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (5)) = 5
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (6)) = 6
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (7)) = 7
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (8)) = 8
Base<sub>1 to 10</sub> (Base<sub>10 to 1</sub> (9)) = 9
```



Base_{1 to 10} (Base_{10 to 1}(Z)) = Z



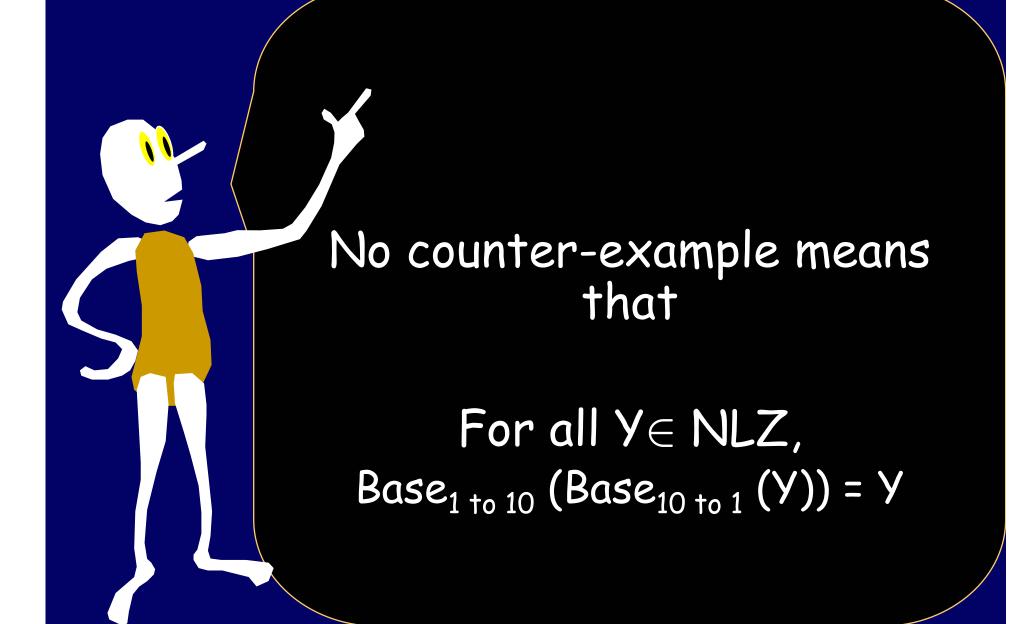
$X \in NLZ$, n=length(Z) > 1

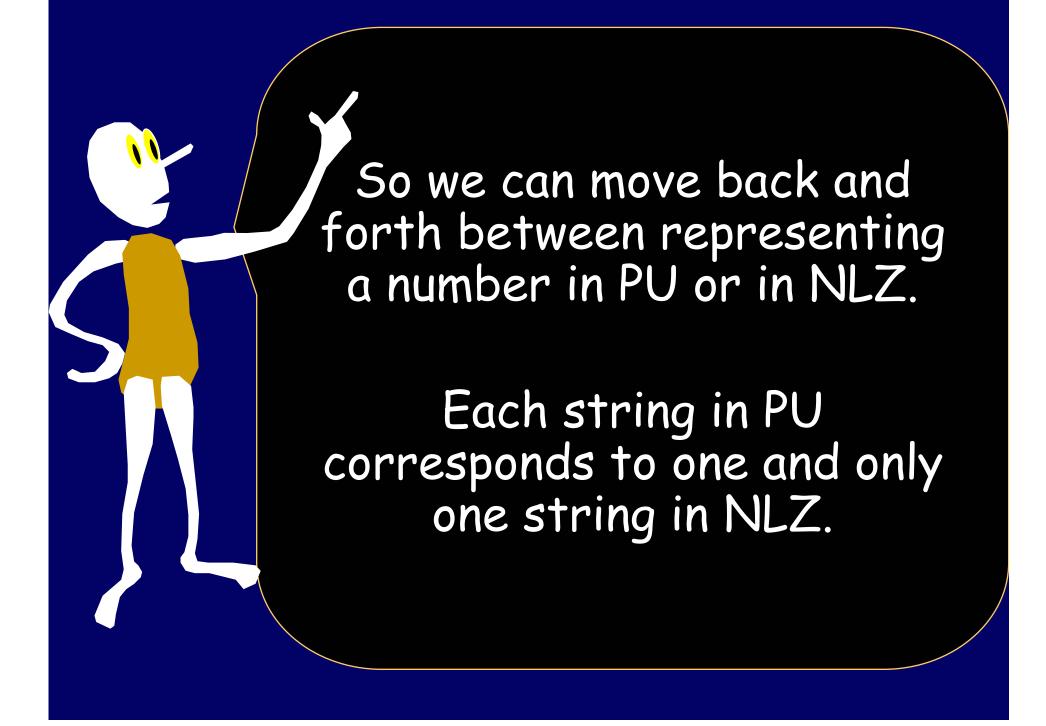
```
Suppose X=a_{n-1} a_{n-2} ... a_0

Base<sub>10 to 1</sub>(X) = Y =

\Sigma_{i < n} Base<sub>10 to 1</sub>(a_i)*10<sup>i</sup>
```

Base_{1 to 10} (Y) = ? n is smallest PU s.t. $10^{n-1} \le Y < 10^n$ Calculate d = Y DIV 10^{n-1} , Z= Y MOD 10^{n-1} d= a_{n-1} , Z= $\sum_{i < n-1}$ Base_{10 to 1}(a_i)* 10^i Output a_{n-1} Base_{1 to 10} (Z) [Z shorter than X] = a_{n-1} a_{n-2} ... a_1 a_0 Contradiction of X being a counter-example.





Base X Notation

Let Σ be an alphabet of size X. A base X digit is any element of Σ .

Let $S = a_{n-1}, a_{n-2}, ..., a_0$ be a sequence of base X digits.

Let Base_{X to 1}(S) = $a_{n-1} X^{n-1} + ... a_2 X^2 + a_1 X + a_0$

S is called the base X representation of the number $Base_{X to 1}(S)$.

$$S = a_{n-1}, a_{n-2}, ..., a_0$$

represents the number:
 $a_{n-1} X^{n-1} + a_{n-2} X^{n-2} + ... + a_0 X^0$

Base 2 [Binary Notation] 101 represents $1(2)^2 + 0(2^1) + 1(2^0)$

Base 7 015 represents $0(7)^2 + 1(7^1) + 5(7^0)$

$$S = a_{n-1}, a_{n-2}, ..., a_0$$

represents the number:
 $a_{n-1} X^{n-1} + a_{n-2} X^{n-2} + ... + a_0 X^0$

Bases In Different Cultures

Sumerian-Babylonian: 10, 60, 360

Egyptians: 3, 7, 10, 60

Africans: 5, 10

French: 10, 20

English: 10, 12,20



Biggest n "digit" number in base X would be:

$$(X-1)X^{n-1} + (X-1)X^{n-2} + ... + (X-1)X^{0}$$

Base 2 111 represents $1(2)^2 + 1(2^1) + 1(2^0)$

Base 7
666 represents $6(7)^2 + 6(7^1) + 6(7^0)$

Biggest n "digit" number in base X would be:

$$(X-1)X^{n-1} + (X-1)X^{n-2} + ... + (X-1)X^{0}$$

Base 2 111 represents $1(2)^2 + 1(2^1) + 1(2^0)$ 111 + 1 = 1000 represents 2^3 Thus, 111 represents $2^3 - 1$

Base 7
666 represents 6 $(7)^2$ + 6 (7^1) + 6 (7^0) 666 + 1 = 1000 represents 7^3 Thus, 666 represents 7^3 - 1

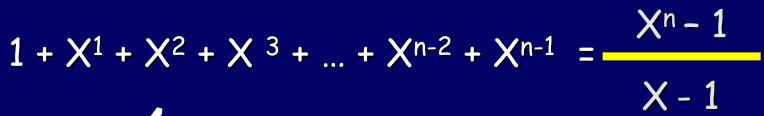
Biggest n "digit" number in base X would be:

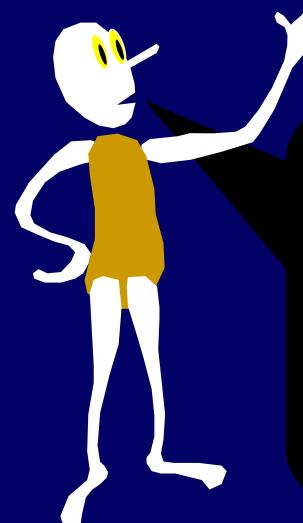
$$S = (X-1)X^{n-1} + (X-1)X^{n-2} + ... + (X-1)X^0$$
 \(Add 1 to get: X^n + 0 X^{n-1} + ... + 0 X^0

Thus,
$$S = X^n - 1$$

Base 2 111 represents $1(2)^2 + 1(2^1) + 1(2^0)$ 111 + 1 = 1000 represents 2^3 Thus, 111 represents $2^3 - 1$

Base 7 666 represents 6 $(7)^2$ + 6 (7^1) + 6 (7^0) 666 + 1 = 1000 represents 7^3 Thus, 666 represents 7^3 - 1





Recall the GEOMETRIC SERIES.

$$(X-1)X^{n-1} + (X-1)X^{n-2} + ... + (X-1)X^{0}$$

= $X^{n} - 1$

Proof:

Factoring out (X-1), we obtain:

$$(X-1)[X^{n-1} + X^{n-2} + + X + 1] =$$

$$(X-1)[(X^n-1)/(X-1)] =$$

$$X^n - 1$$

The highest n digit number in base X.

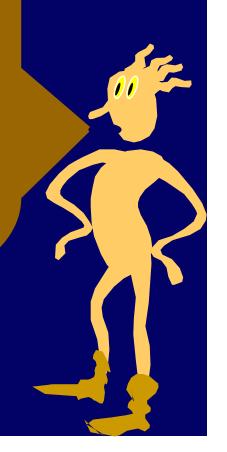
$$(X-1)(1+X^1+X^2+X^3+...+X^{n-1})=X^n-1$$

Base X. Let S be the sequence of n digits each of which is X-1. S is the largest number of length n that can be represented in base X.

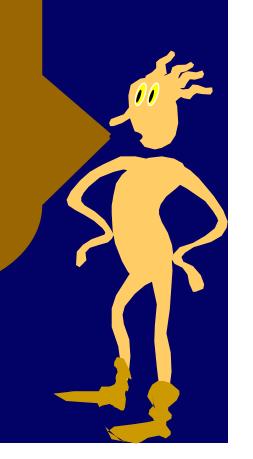
$$S = X^n - 1$$

Each of the numbers from 0 to $X^{n}-1$ is uniquely represented by an n-digit number in base X.

We could prove this using our previous method, but let's do it another way.



Our proof introduce an unfamiliar kind of base representation.



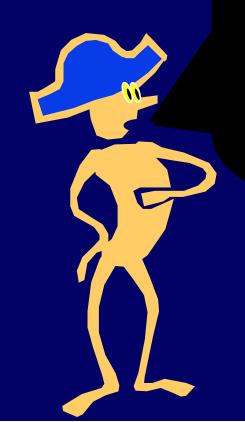
Plus/Minus Base X

An plus/minus base X digit is any integer -X < a < X

Let $S = a_{n-1}, a_{n-2}, ..., a_0$ be a sequence of plus/minus base X digits. S is said to be the plus/minus base X representation of the number:

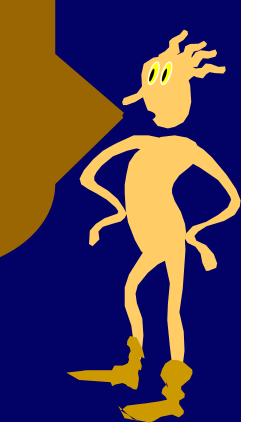
$$a_{n-1} X^{n-1} + a_{n-2} X^{n-2} + ... + a_0 X^0$$

Does each n-digit number in plus/minus base X represent a different integer?

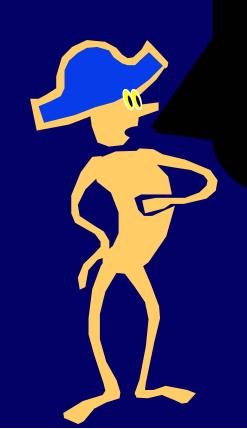


No.
Consider plus/minus binary.

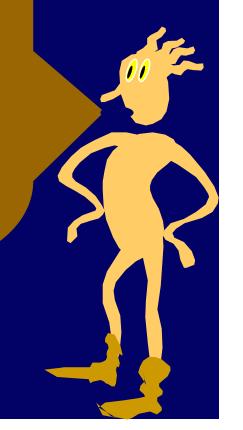
$$5 = 0 \ 1 \ 0 \ 1 = 1 \ 0 \ -1 \ -1$$



Humm.. So what is it good for?



Not every number has a unique plus/minus binary representation, but 0 does.



O has a unique n-digit plus/minus base X representation as all 0's.

Suppose $a_{n-1} X^{n-1} + a_{n-2} X^{n-2} + ... + a_0 X^0 = S = 0$, where there is some highest k such that $a_k \neq 0$. Wl.o.g. assume $a_k > 0$.

Because $X^k > (X-1)(1 + X^1 + X^2 + X^3 + ... + X^{k-1})$ no sequence of digits from a_{k-1} to a_0 can represent a number as small as $-X^k$

Hence $S \neq 0$. Contradiction.

Each of the numbers from 0 to Xⁿ-1 is uniquely represented by an n-digit number in base X.

We already know that n-digits will represent something between 0 and $X^n - 1$.

Suppose two distinct sequences represent the same number:

$$a_{n-1} X^{n-1} + a_{n-2} X^{n-2} + ... + a_0 X^0 = b_{n-1} X^{n-1} + b_{n-2} X^{n-2} + ... + b_0 X^0$$

The difference of the two would be an plus/minus base X representation of 0, but it would have a non-zero digit. Contradiction.

Each of the numbers from 0 to Xⁿ-1 is uniquely represented by an n-digit number in base X.

n digits represent up to $X^n - 1$ n-1 digits represents up to $X^{n-1} - 1$

Let k be a number: Xⁿ⁻¹ ≤ k ≤ Xⁿ - 1 So k can be represented with n digits.

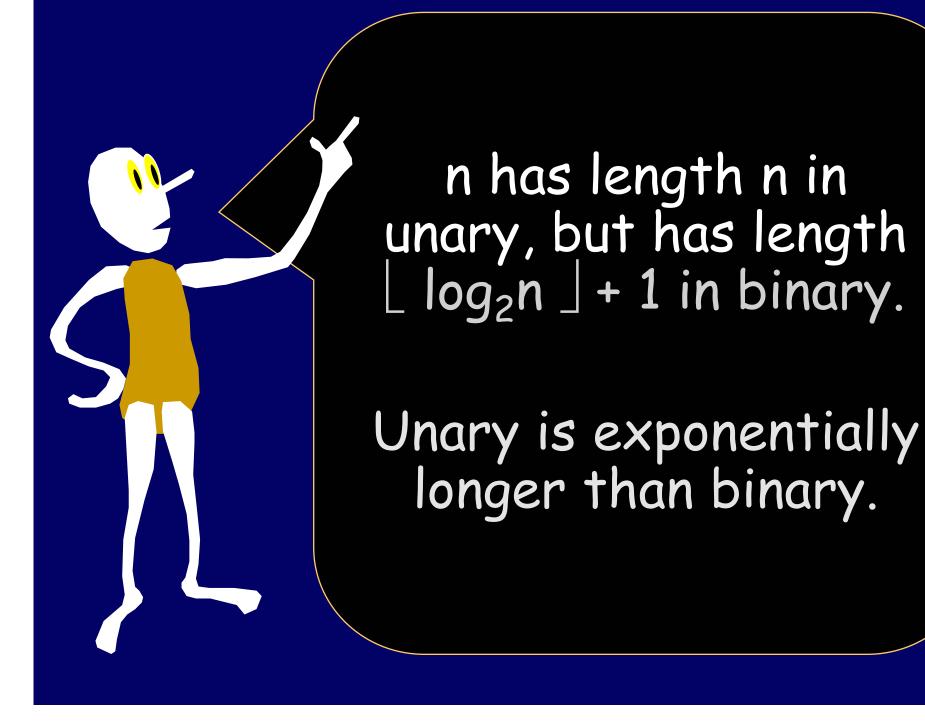
For all $k: \lfloor \log_{\times} k \rfloor = n - 1$

So k uses $\lfloor \log_{\times} k \rfloor + 1$ digits.

Fundamental Theorem For Base X:

Each of the numbers from 0 to Xⁿ-1 is uniquely represented by an n-digit number in base X.

k uses $\lfloor \log_x k \rfloor + 1$ digits in base X.



Egyptian Multiplication



The Egyptians used decimal numbers but multiplied and divided in binary

Egyptian Multiplication a times b by repeated doubling

b has some n-bit representation: b_n..b₀

Starting with a, repeatedly double largest so far to obtain: a, 2a, 4a,, 2^n a

Sum together all 2^k a where $b_k = 1$

Egyptian Multiplication 15 times 5 by repeated doubling

5 has some 3-bit representation: 101

Starting with 15, repeatedly double largest so far to obtain: 15, 30, 60

Sum together all $2^{k}(15)$ where $b_{k} = 1$: 15 + 60 = 75

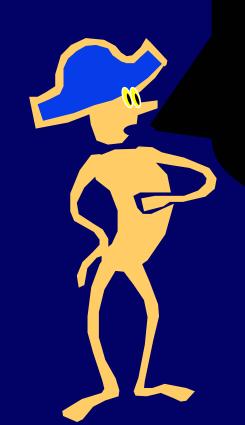
Why does that work?

$$b = b_0 2^0 + b_1 2^1 + b_2 2^2 + ... + b_n 2^n$$

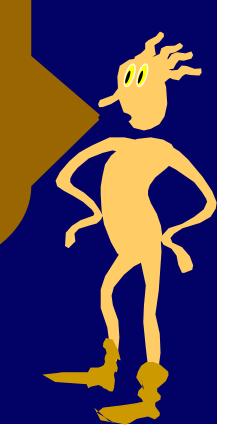
 $ab = b_0 2^0 a + b_1 2^1 a + b_2 2^2 a + ... + b_n 2^n a$

If b_k is 1 then 2^k a is in the sum. Otherwise that term will be 0.

Wait! How did the Egyptians do the part where they converted b to binary?



They used repeated halving to do base conversion. Consider ...



Egyptian Base Conversion

```
Output stream will print right to left.
Input X.
Repeat until X=0
{
    If X is even then Output O;
        Otherwise {X:=X-1; Output 1}

    X:=X/2
}
```

Egyptian Base Conversion

```
Output stream will print right to left.
Input X.
Repeat until X=0
{
    If X is even then Output O;
        Otherwise Output 1

    X:= \[ X/2 \]
}
```

```
Repeat until X=0

{    If X is even then Output O; Otherwise Output 1;    X:= \[ X/2 \]
```

01

```
Repeat until X=0
{
    If X is even then Output 0;
    Otherwise Output 1;
    X:= \[ X/2 \]
}
```

01

```
Repeat until X=0
{
    If X is even then Output 0;
    Otherwise Output 1;
    X:= \[ X/2 \]
}
```

101

```
Repeat until X=0
{
    If X is even then Output 0;
    Otherwise Output 1;
    X:= \[ X/2 \]
}
```

101

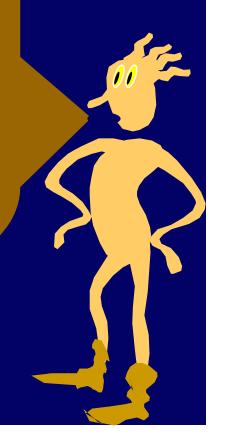
```
Repeat until X=0
{
    If X is even then Output 0;
    Otherwise Output 1;
    X:= \[ X/2 \]
}
```

And Keep Going until 0

```
Repeat until X=0

{    If X is even then Output 0;    Otherwise Output 1;        X:= \[ X/2 \] }
```

Sometimes the Egyptian combined the base conversion by halving and the multiplication by doubling into one algorithm



Rhind Papyrus (1650 BC) 70*13

70140280560

13 * 7063 * 3501 * 910

Rhind Papyrus (1650 BC) 70*13

Binary for 13 is
$$1101 = 2^3 + 2^2 + 2^0$$

 $70*13 = 70*2^3 + 70*2^2 + 70*2^0$

Rhind Papyrus (1650 BC)

173468136

184 48 14

Rhind Papyrus (1650 BC)

 17
 1

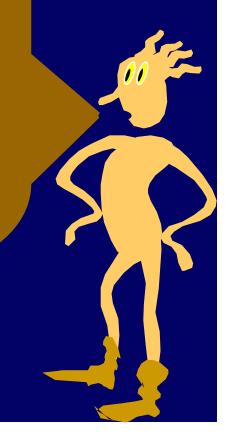
 34
 2

 68
 4

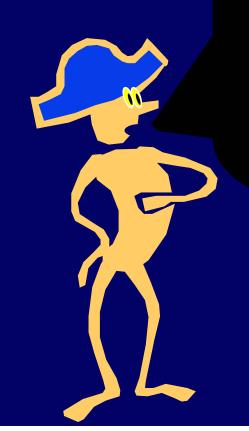
 136
 8

184 48 14

184 = 17*8 + 17*2 + 14184/17 = 10 with remainder 14 This method is called "Egyptian Multiplication/Division" or "Russian Peasant Multiplication/Division".



Wow. Those Russian peasants were pretty smart.



Standard Binary Multiplication = Egyptian Multiplication

Egyptian Base 3

We have defined

Base 3: Each digit can be 0, 1, or 2

Plus/Minus Base 3 uses -2, -1, 0, 1, 2

Here is a new one: Egyptian Base 3 uses -1, 0, 1

Example: 1-1-1=9-3-1=5

Unique Representation Theorem for Egyptian Base 3

No integer has 2 distinct, n-digit, Egyptian base-3 representations. We can represent all integers from $-(3^n-1)/2$ to $(3^n-1)/2$

Proof; If so, their difference would be a non-trivial plus/minus base 3 representation of 0. Contradiction. Highest number = $1111...1 = (3^n-1)/2$ Lowest number = $-1-1-1-1...-1 = -(3^n-1)/2$

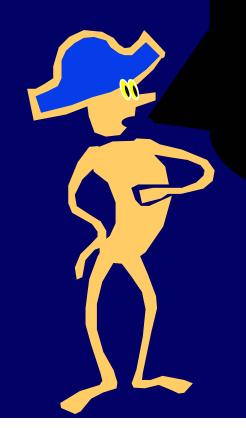
Unique Representation Theorem for Egyptian Base 3

No integer has 2 distinct, n-digit, Egyptian base-3 representations. We can represent all integers from $-(3^n-1)/2$ to $(3^n-1)/2$

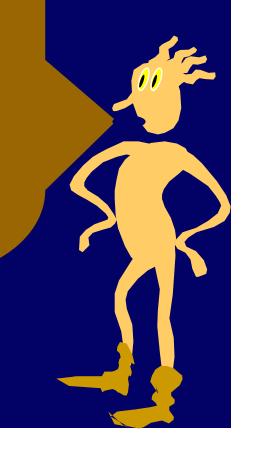
 3^n , n-digit, base 3 representations of the numbers from 0 to $3^n - 1$

Subtract 111111...111 = $(3^n - 1)/2$ from each to get an Egyptian base 3 representation of all the numbers from $-(3^{n}-1)/2$ to $(3^{n}-1)/2$.

How could this be Egyptian? Historically, negative numbers first appear in the writings of the Hindu mathematician Brahmagupta (628 AD).









One weight for each power of 3. Left = "negative". Right = "positive".

References

The Book Of Numbers, by J. Conway and R. Guy

History of Mathematics, Histories of Problems, by The Inter-IREM Commission