

John Lafferty
Lecture 12 Oct 5, 2005

CS 15-251

Fall 2006

Carnegie Mellon University

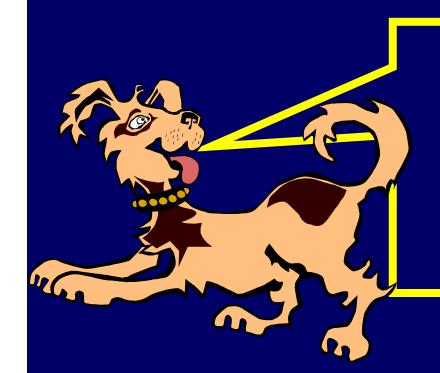
Ancient Wisdom: Primes, Continued Fractions, The Golden Ratio, and Euclid's GCD

$$\frac{3+\sqrt{13}}{2} = 3 + \frac{1}{3+\frac$$



1. Recap, and finishing up probability

2. Something completely different...



What might be is surely possible!

Goal: show exists object of value at least v. Proof strategy:

- · Define distribution D over objects.
- Define RV: X(object) = value of object.
- Show $E[X] \ge v$. Conclude it must be possible to have $X \ge v$.





Pigeonhole principle: Given n boxes and m > n objects, at least one box must contain more than one object.



Letterbox principle: If the average number of letters per box is a, then some box will have at least a letters. (Similarly, some box has at most a.)



Independent Sets

An *independent set* in a graph is a set of vertices with no edges between them.

All of the vertices in such a set can be given the same color, so the size of the largest independent set i(X) gives a bound on the number of colors required c(G):

$$c(G) i(X) >= n$$

(A coloring divides up the graph into independence sets, and each one is no bigger than i(X) in size.)

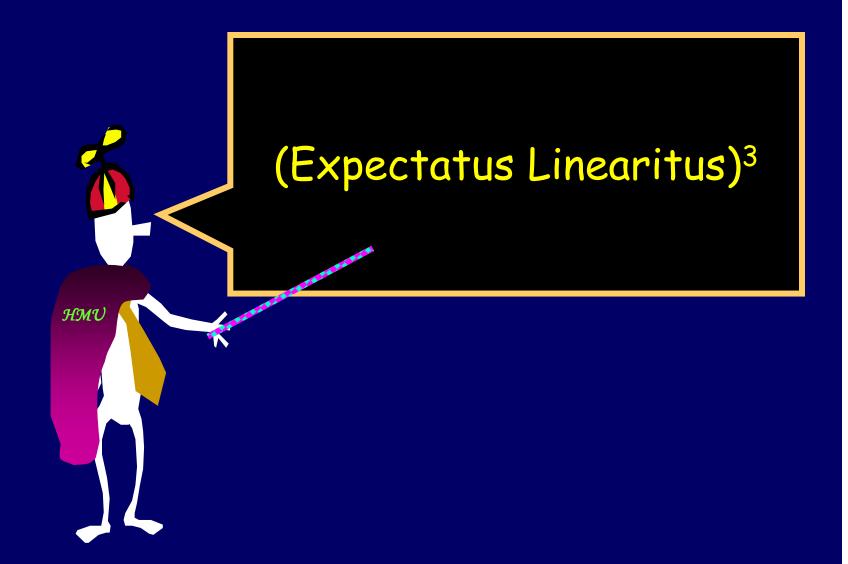
Theorem: If a graph G has n vertices and m edges, then it has an independent set with at least n²/4m vertices.

Let d = 2m/n be the average degree.
Randomly take away vertices and edges:

- 1. Delete each vertex of G (together with its incident edges) with probability 1-1/d
- 2. For each remaining edge remove it and one of its vertices.

The remaining vertices form an independent set. How big is it expected to be?





Theorem: If a graph G has n vertices and m edges, then it has an independent set with at least n²/2m vertices.

Let X be the number of *vertices* that survive the first step:

$$E[X] = n/d.$$

Let Y be the number of *edges* that survive the first step:

$$E[Y] = m(1/d)^2 = nd/2 (1/d)^2 = n/2d$$
.

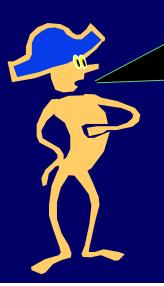
The second step removes all the remaining edges and at most Y vertices. So size of final set of vertices is at least X-Y and

$$E[X-Y] = n/d - n/2d = n/2d = n^2/4m$$



An easy question

What is
$$\sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i$$
? A: 2.



But it never actually gets to 2. Is that a problem?



But it never actually gets to 2. Is that a problem?

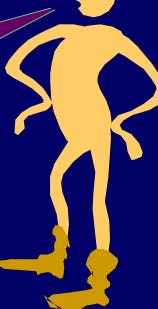




No, by $\sum_{i=0}^{\infty} f(i)$, we really mean $\lim_{n\to\infty} \sum_{i=0}^{n} f(i)$.

[if this is undefined, so is the sum]

In this case, the partial sum is $2-(\frac{1}{2})^n$ which goes to 2.





A related question

Suppose I flip a coin of bias p, stopping when I first get heads.

What's the chance that I:

- ·Flip exactly once?
 - Ans: p
- ·Flip exactly two times?
 - Ans: (1-p)p
- ·Flip exactly k times?
 - Ans: $(1-p)^{k-1}p$
- ·Eventually stop?

Ans: 1. (assuming p>0)





A related question

Pr(flip once) + Pr(flip 2 times) + Pr(flip 3 times) + ... = 1:

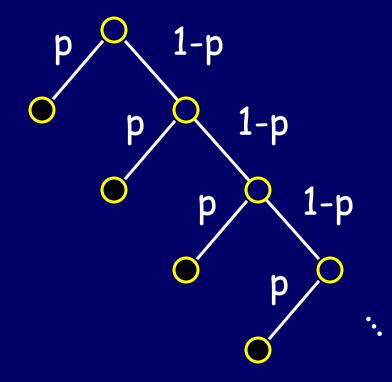


Or, using q = 1-p,

$$\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}$$

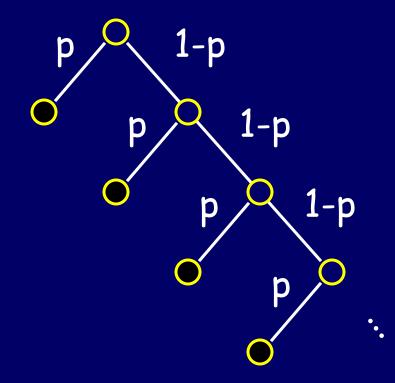


<u>Pictorial view</u>



Sample space S = leaves in this tree. Pr(x) = product of edges on path to x. If p>0, prob of not halting by time n goes to 0 as $n \to \infty$.

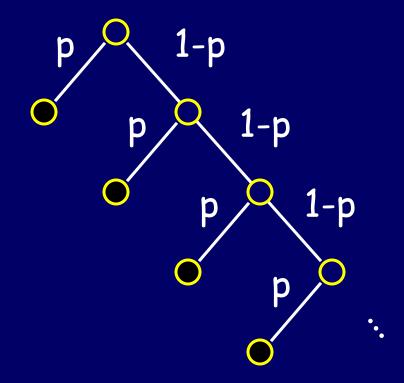
Use to reason about expectations too



Pr(x|A)=product of edges on path from A to x. $E[X] = \sum_{x} Pr(x)X(x)$. $E[X|A] = \sum_{x \in A} Pr(x|A)X(x)$. I.e., it is as if we started the game at A.

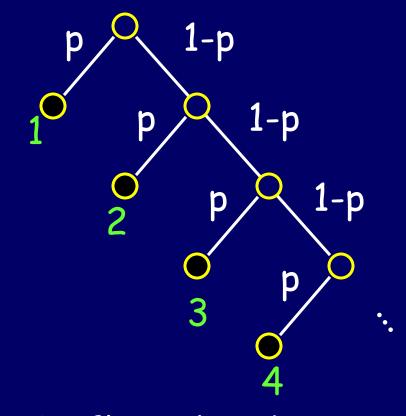


Use to reason about expectations too



Flip bias-p coin until heads. What is expected number of flips?

Use to reason about expectations too



Let X = # flips.

Let $A = \text{event that } 1^{\text{st}} \text{ flip is heads.}$

 $E[X] = E[X|A]Pr(A) + E[X|\neg A]Pr(\neg A)$ = 1*p + (1 + E[X])*(1-p).

Solving: pE[X] = p + (1-p), so E[X] = 1/p.

Infinite Probability spaces

Notice we are using infinite probability spaces here, but we really only defined things for <u>finite</u> spaces so far.

Infinite probability spaces can sometimes be weird. Luckily, in CS we will almost always be looking at spaces that can be viewed as choice trees where

Pr(haven't halted by time t) \rightarrow 0 as t $\rightarrow\infty$.

General picture

Let 5 be a sample space we can view as leaves

of a choice tree.

Let $S_n = \{ \text{leaves at depth} \leq n \}.$

For event A, let $A_n = A \cap S_n$.

If $\lim_{n\to\infty} \Pr(S_n)=1$, can define:

$$Pr(A)=\lim_{n\to\infty}Pr(A_n).$$



Setting that doesn't fit our model

Flip coin until #heads > 2*#tails.

There's a reasonable chance this will never stop...

You go into a casino with \$k, and at each time step you bet \$1 on a fair game. Leave when you are broke or have \$n.



Question 1: what is your expected amount of money at time t?

Let X_t be a R.V. for the amount of money at time t.

You go into a casino with \$k, and at each time step you bet \$1 on a fair game. Leave when you are broke or have \$n.

Question 1: what is your expected amount of money at time t?

 $X_t = k + \delta_1 + \delta_2 + ... + \delta_{t_i}$ where δ_i is a RV for the change in your money at time i.

 $E[\delta_i] = 0$, since $E[\delta_i|A] = 0$ for all situations A at time i.

So,
$$E[X_{t}] = k$$
.



You go into a casino with \$k, and at each time step you bet \$1 on a fair game. Leave when you are broke or have \$n.

Question 2: what is the probability you leave with \$n?



You go into a casino with \$k, and at each time step you bet \$1 on a fair game. Leave when you are broke or have \$n.

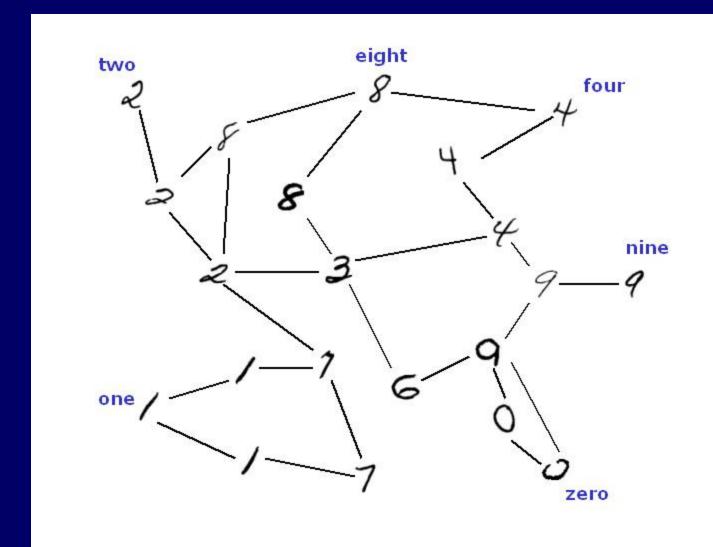
Question 2: what is the probability you leave with \$n?

One way to analyze:

- $E[X_+] = k$.
- $E[X_{+}] = E[X_{+}|X_{+}=0]*Pr(X_{+}=0) + E[X_{+}|X_{+}=n]*Pr(X_{+}=n) + E[X_{+}|neither]*Pr(neither).$
- So, $E[X_t] = 0 + n*Pr(X_t=n) + something*Pr(neither).$
- As $t \to \infty$, Pr(neither) $\to 0$. Also 0 < something < n.

So,
$$\lim_{t\to\infty} \Pr(X_t=n) = k/n$$
.
So, $\Pr(\text{leave with } \$n) = k/n$.

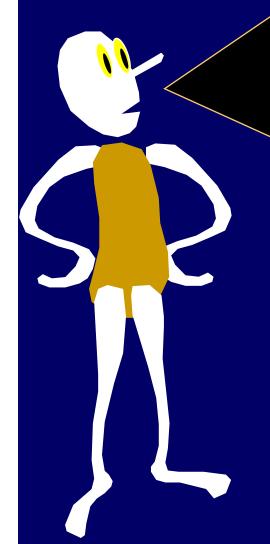






And now, for something completely different....





Definition: A number > 1 is prime if it has no other factors, besides 1 and itself.

Each number can be factored into primes in a unique way. [Euclid]



Theorem: Each natural has a unique factorization into primes written in non-decreasing order.

Definition: A number > 1 is prime if it has no other factors, besides 1 and itself.

Primes: 2, 3, 5, 7, 11, 13, 17, ...

Factorizations:



Multiplication might just be a "one-way" function

Multiplication is fast to compute Reverse multiplication is apparently slow

We have a feasible method to multiply 1000 bit numbers [Egyptian multiplication]

Factoring the product of two random 1000 bit primes has no known feasible approach.



Grade School GCD algorithm

GCD(A,B) is the greatest common divisor, i.e., the largest number that goes evenly into both A and B.

What is the GCD of 12 and 18? $12 = 2^2 * 3$ $18 = 2*3^2$

Common factors: 21 and 31

Answer: 6



How to find GCD(A,B)?

A Naïve method:

Factor A into prime powers.

Factor B into prime powers.

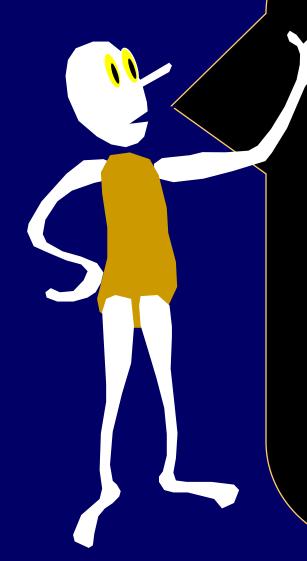
Create GCD by multiplying together each common prime raised to the highest power that goes into both A and B.

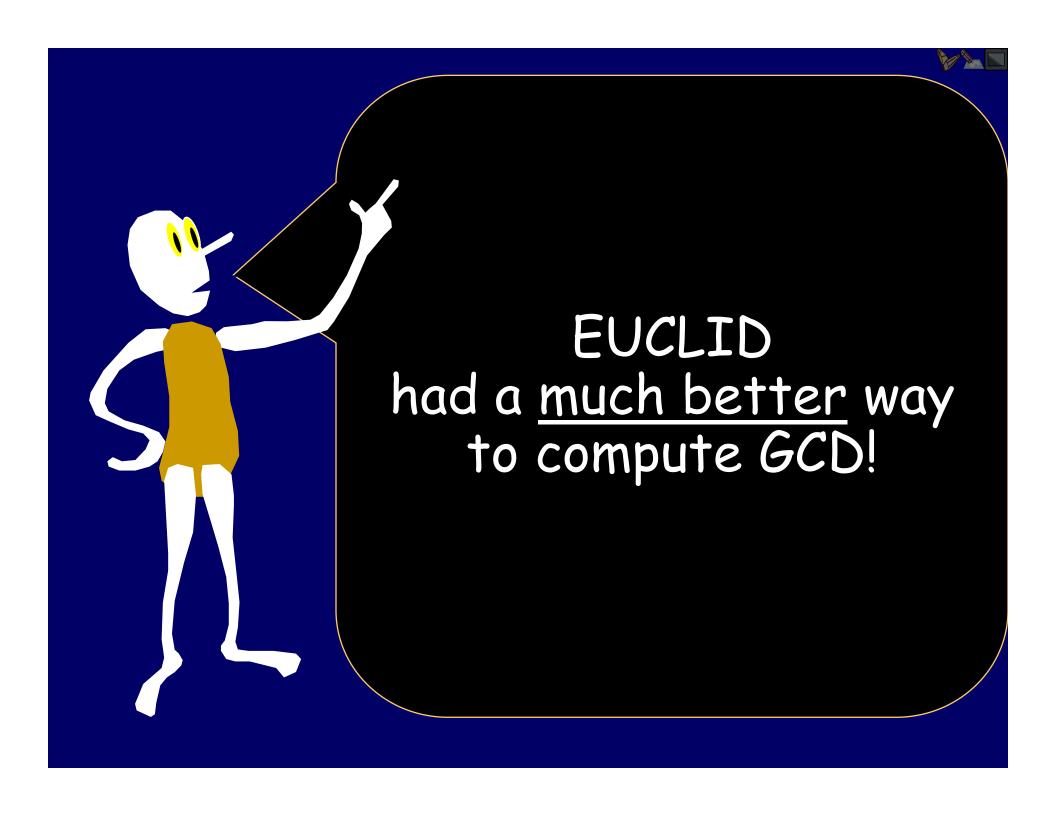




This requires factoring A and B.

No one knows a particularly fast way to factor numbers in general.







Ancient Recursion: Euclid's GCD algorithm

Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)



A small example

```
Euclid(A,B)
If B=0 then return A
  else return Euclid(B, A mod B)
```

```
Note: GCD(67, 29) = 1

Euclid(67,29) 67 mod 29 = 9
Euclid(29,9) 29 mod 9 = 2
Euclid(9,2) 9 mod 2 = 1
Euclid(2,1) 2 mod 1 = 0
Euclid(1,0) outputs 1
```



Important questions to ask

Is the algorithm correct?

Does the algorithm stop?

How many steps does the algorithm run for?



But is it correct?

Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

Claim: $GCD(A,B) = GCD(B, A \mod B)$



But is it correct?

```
Euclid(A,B)
If B=0 then return A
  else return Euclid(B, A mod B)
```

Claim: GCD(A,B) = GCD(B, A mod B) value of GCD is an invariant!



But is it correct?

Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

Claim: $GCD(A,B) = GCD(B, A \mod B)$

 $d|A \text{ and } d|B \Leftrightarrow d|(A - kB)$

The set of common divisors of A, B equals the set of common divisors of B, A-kB.



Does the algorithm stop?

Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

Claim: After first step, $A \ge B \ge 0$



Does the algorithm stop?

```
Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)
```

```
Claim: A mod B < \frac{1}{2} A

Proof:

If B = \frac{1}{2} A then A mod B = 0

If B < \frac{1}{2} A then any X Mod B < B < \frac{1}{2} A

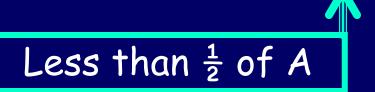
If B > \frac{1}{2} A then A mod B = A - B < \frac{1}{2} A
```



Does the algorithm stop?

Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

GCD(A,B) calls GCD(B, A mod B)





Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

GCD(A,B) calls $GCD(B, \langle \frac{1}{2}A)$



Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

GCD(A,B) calls $GCD(B, \langle \frac{1}{2}A)$

which calls $GCD(\langle \frac{1}{2}A, B \mod \langle \frac{1}{2}A \rangle)$





Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

Every two recursive calls, the input numbers drop by half.



Euclid(A,B)
If B=0 then return A
else return Euclid(B, A mod B)

Theorem: If two input numbers have an n bit binary representation, Euclid's Algorithm will not take more than 2n calls to terminate.

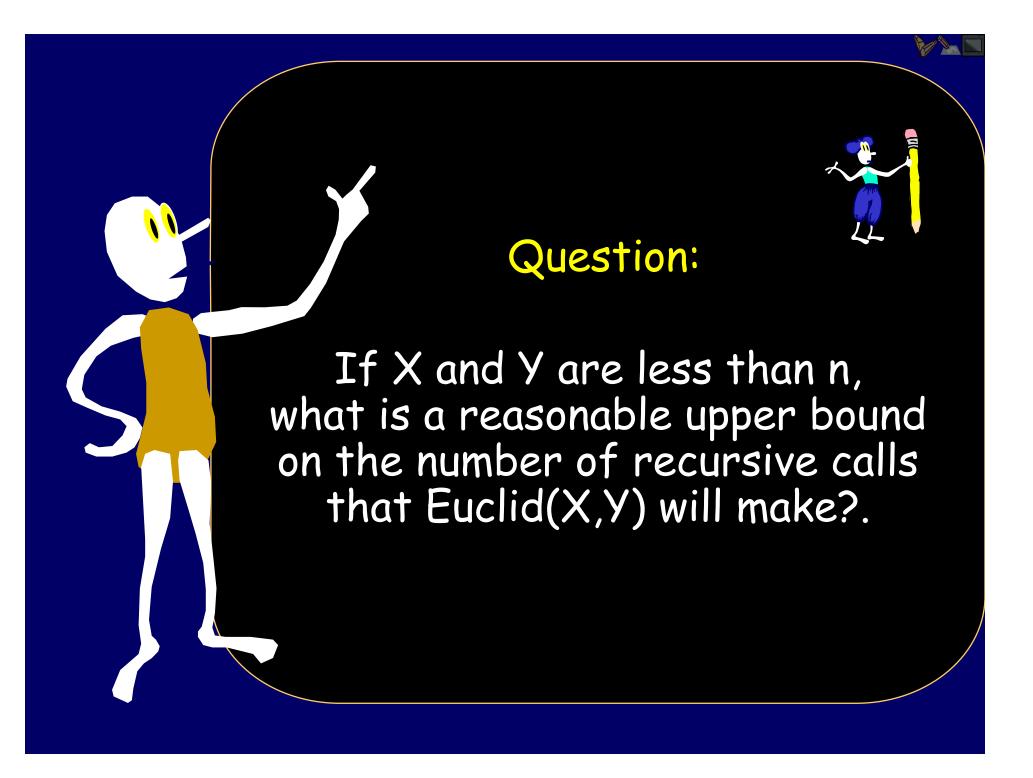


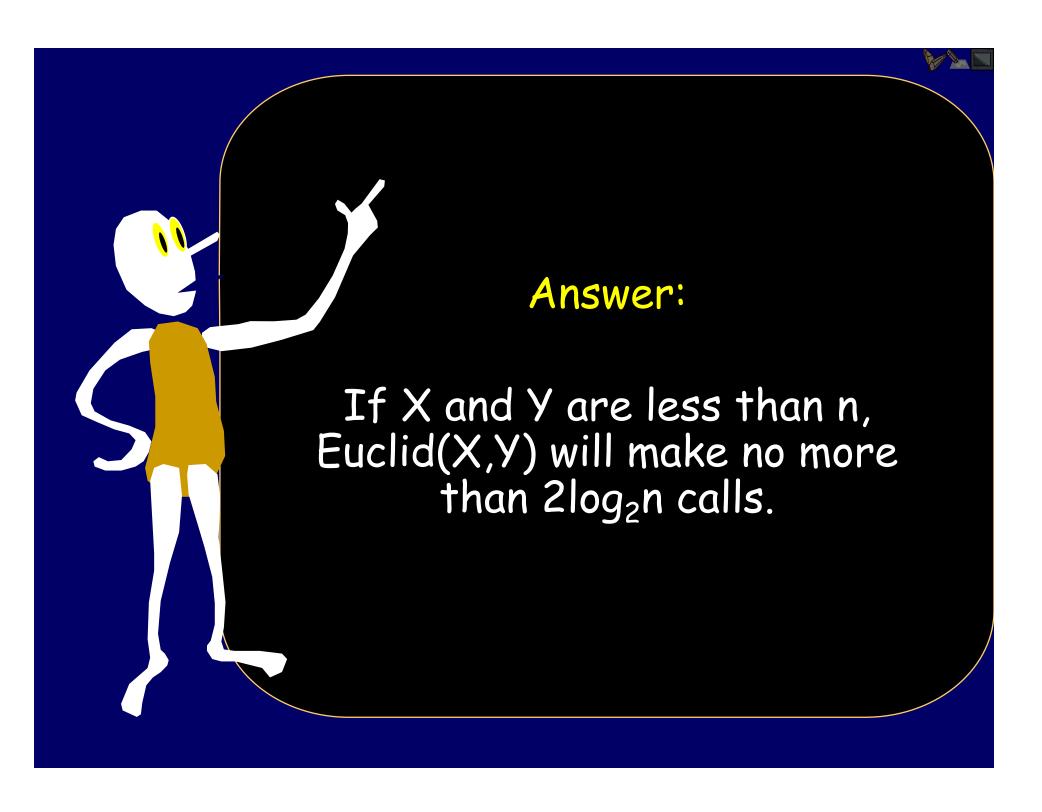
Important questions to ask

Is the algorithm correct?

Does the algorithm stop?

How many steps does the algorithm run for?







Euclid(A,B)

If B=0 then return A
else return Euclid(B, A mod B)

```
Euclid(67,29)
Euclid(29,9)
Euclid(9,2)
Euclid(2,1)
Euclid(1,0) outputs 1
```



Let <r,s> denote the number r*67 + s*29. Calculate all intermediate values in this representation.

Euclid(67,29)	9=<1,0> - 2*<0,1>	9 =<1,-2>
Euclid(29,9)	2= <0,1> - 3 *<1,-2>	2=<-3,7>
Euclid(9,2)	1=<1,-2> - 4*<-3,7>	1=<13,-30>
Euclid(2,1)	0=<-3,7> - 2*<13,-30>	0=<-29,67>

Euclid(1,0) outputs
$$1 = 13*67 - 30*29$$



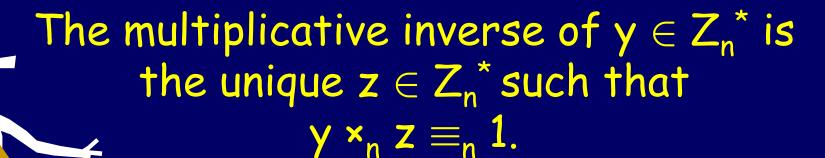
Euclid's Extended GCD algorithm

Input: X,Y

Output: r,s,d such that rX+sY = d = GCD(X,Y)

```
67=<1,0> 29=<0,1>
Euclid(67,29) 9=67 - 2*29 9=<1,-2>
Euclid(29,9) 2=29 - 3*9 2=<-3,7>
Euclid(9,2) 1=9 - 4*2 1=<13,-30>
Euclid(2,1) 0=2 - 2*1 0=<-29,67>
```

Euclid(1,0) outputs 1 = 13*67 - 30*29



The unique inverse of a must exist because the y row contains a permutation of the elements and hence contains a unique 1.

Z₅*

×	1	Z	3	4
1	1	2	3	4
2	2	4	1	3
У	3	1	4	2
4	4	3	2	1



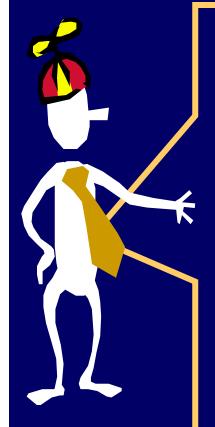
The multiplicative inverse of $y \in Z_n^*$ is the unique $z \in Z_n^*$ such that $y \times_n z \equiv_n 1.$

To quickly compute the inverse of y:

Run ExtendedEuclid(x,n).

returns a, b, and d such that ay+bn = dBut d = GCD(y,n) = 1, so ay + bn = 1

Hence $ay = 1 \pmod{n}$ Thus, a is the multiplicative inverse of y.



The RSA story

Pick 2 distinct, random 1000 bit primes, p and q.

Multiply them to get n = (p*q)Multiply (p-1) and (q-1) to compute $\phi(n)$ Randomly pick an e s.t. GCD(e,n) = 1. Publish n and e

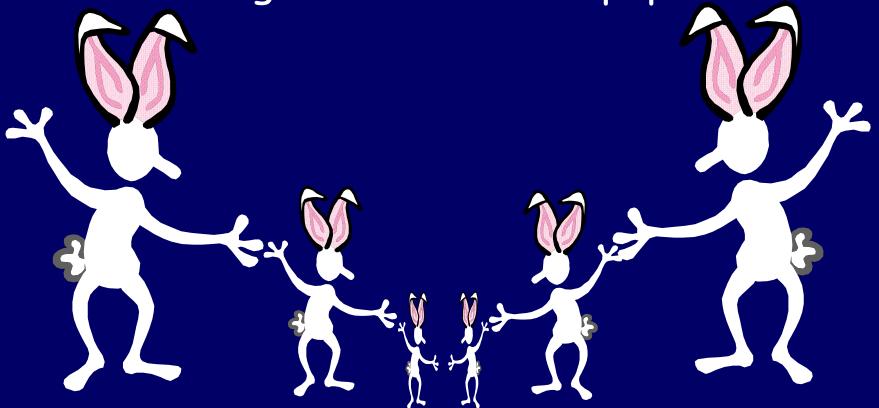
Compute multiplicative inverse of e mod $\phi(n)$ to get a <u>secret</u> number d.

 $(m^e)^d = m^{ed} = m^1 \pmod{n}$



Leonardo Fibonacci

In 1202, Fibonacci proposed a problem about the growth of rabbit populations.







Inductive Definition or Recurrence Relation for the Fibonacci Numbers

Stage 0, Initial Condition, or Base Case: Fib(0) = 0; Fib (1) = 1

Inductive Rule

For n>1, Fib(n) = Fib(n-1) + Fib(n-2)

n	0	1	2	3	4	5	6	7
Fib(n)	0	1	1	2	3	5	8	13

A (Simple) Continued Fraction Is Any Expression Of The Form:

$$a + \frac{1}{b + \frac{1}{c + \frac{1}{d + \frac{1}{e + \frac{1}{f + \frac{1}{i + \frac{1}{j + \dots}}}}}}}$$

where a, b, c, ... are whole numbers.



A Continued Fraction can have a finite or infinite number of terms.

$$a + \frac{1}{b + \frac{1}{c + \frac{1}{d + \frac{1}{e + \frac{1}{f + \frac{1}{i + \frac{1}{j + \dots}}}}}}}$$

We also denote this fraction by [a,b,c,d,e,f,...]



A Finite Continued Fraction

$$2 + \frac{1}{3 + \frac{1}{4 + \frac{1}{2}}}$$

Denoted by [2,3,4,2,0,0,0,...]



An Infinite Continued Fraction



Recursively Defined Form For CF

CF = whole number, or

= whole number
$$+\frac{1}{CF}$$



$$\frac{5}{3} = 1 + \frac{1}{1 + \frac{1}{2}}$$



$$\frac{5}{3} = 1 + \frac{1}{1 + \frac{1}{1}}$$

$$1 + \frac{1}{1 + \frac{1}{1}}$$

$$= [1,1,1,1,0,0,0,...]$$



$$? = 1 + \frac{1}{1 + \frac{$$



$$\frac{8}{5} = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}}$$

$$1 + \frac{1}{1 + \frac{1}{1}}$$

= [1,1,1,1,1,0,0,0,...]



$$\frac{13}{8} = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}}$$

$$1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}}$$

$$= [1,1,1,1,1,0,0,0,0,...]$$



A Pattern?

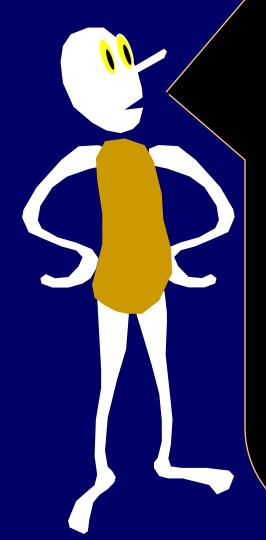
Let
$$r_1 = [1,0,0,0,...] = 1$$

 $r_2 = [1,1,0,0,0,...] = 2/1$
 $r_3 = [1,1,1,0,0,0...] = 3/2$
 $r_4 = [1,1,1,1,0,0,0...] = 5/3$
and so on.

Theorem:

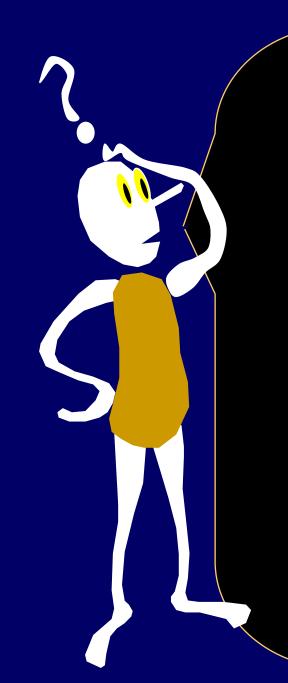
$$r_n = Fib(n+1)/Fib(n)$$





Proposition:
Any finite continued fraction evaluates to a rational.

Theorem (proof later) Any rational has a finite continued fraction representation.



Hmm.
Finite CFs = Rationals.

Then what do infinite continued fractions represent?



An infinite continued fraction

$$\sqrt{2} = 1 + \frac{1}{2 + \dots}}}}}}}$$



Quadratic Equations

$$X^2 - 3x - 1 = 0$$

$$X = \frac{3 + \sqrt{13}}{2}$$

$$X^2 = 3X + 1$$

 $X = 3 + 1/X$

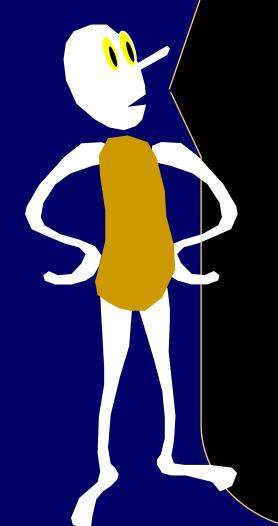
$$X = 3 + 1/X = 3 + 1/[3 + 1/X] = ...$$



A Periodic CF

$$\frac{3+\sqrt{13}}{2} = 3 + \frac{1}{3+\frac{1}{3+\frac{1}{3+\frac{1}{3+\frac{1}{3+\frac{1}{3+\frac{1}{3+\dots}}}}}}}$$





Theorem:

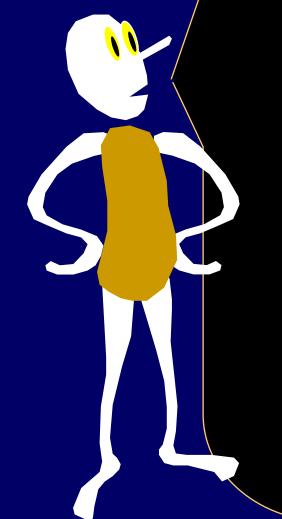
Any solution to a quadratic equation has a periodic continued fraction.

Converse:

Any periodic continued fraction is the solution of a quadratic equation.

(try to prove this!)





So they express more than just the rationals...

What about those non-recurring infinite continued fractions?



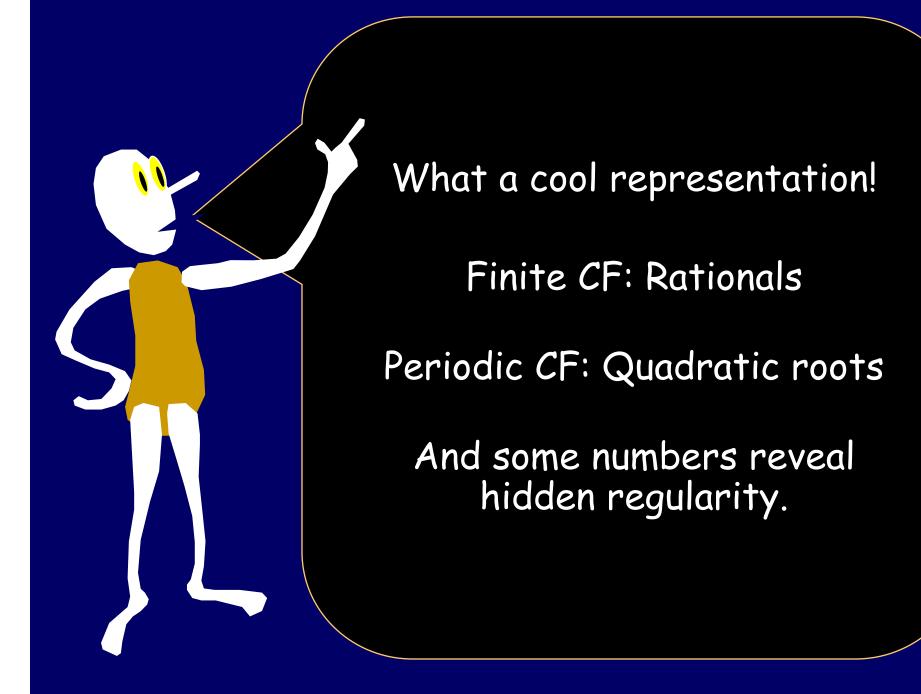
Non-periodic CFs

$$e-1=1+\frac{1}{1+\frac{1}{2+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\dots}}}}}}}$$



What is the pattern?

$$\pi = 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1 + \frac$$





More good news: Convergents

Let $\alpha = [a_1, a_2, a_3, ...]$ be a CF.

Define: $C_1 = [a_1, 0, 0, 0, 0, 0...]$

 $C_2 = [a_1, a_2, 0, 0, 0, ...]$

 $C_3 = [a_1, a_2, a_3, 0, 0, ...]$ and so on.

 C_k is called the k-th convergent of α

 α is the limit of the sequence C_1 , C_2 , C_3 ,...



Best Approximator Theorem

A rational p/q is the <u>best approximator</u> to a real α if no rational number of denominator smaller than q comes closer to α .

BEST APPROXIMATOR THEOREM:

Given any CF representation of α , each convergent of the CF is a best approximator for α !



Best Approximators of π

$$C_1 = 3$$

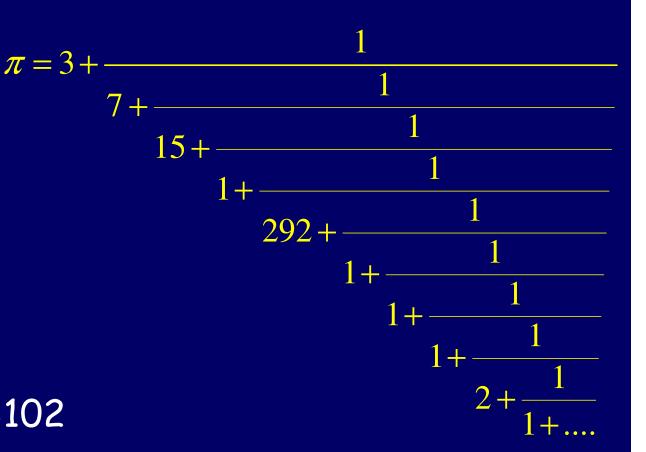
$$C_2 = 22/7$$

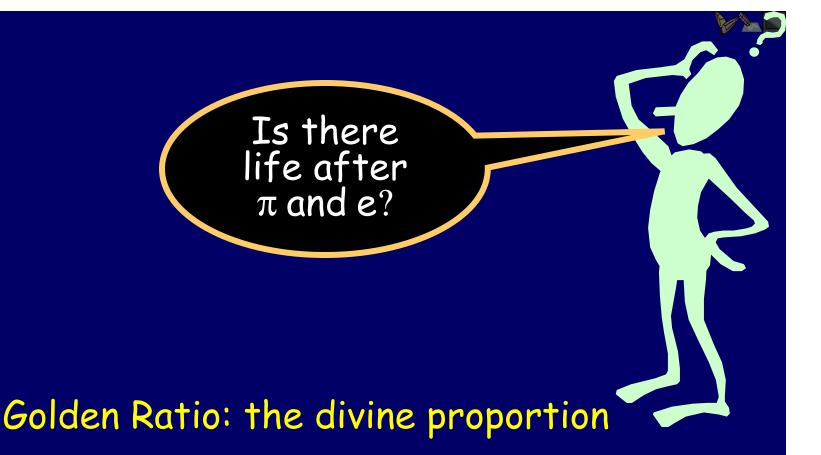
$$C_3 = 333/106$$

$$C_4 = 355/113$$

$$C_5 = 103993/33102$$

$$C_6 = 104348/33215$$



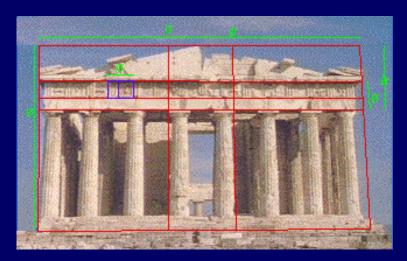


φ = 1.6180339887498948482045...

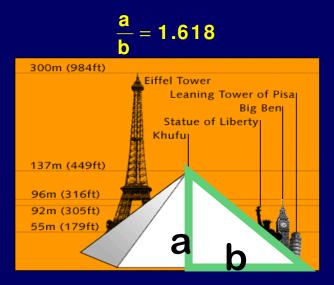
"Phi" is named after the Greek sculptor Phidias



Golden ratio supposed to arise in...



Parthenon, Athens (400 B.C.)



The great pyramid at Gizeh

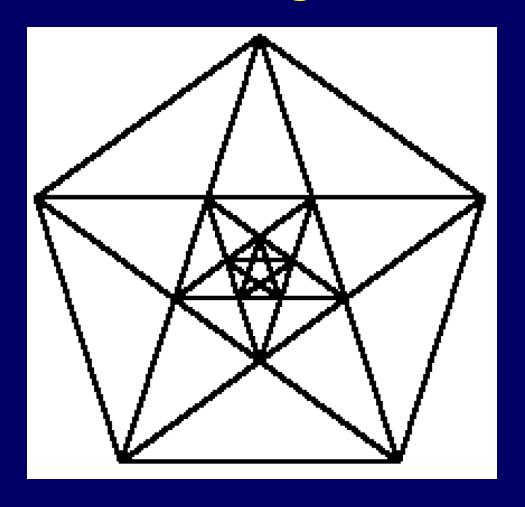


Ratio of a person's height to the height of his/her navel

Mostly circumstantial evidence...



Pentagon





Definition of ϕ (Euclid)

Ratio obtained when you divide a line segment into two unequal parts such that the ratio of the whole to the larger part is the same as the ratio of the larger to the smaller.

$$\phi = \frac{AC}{AB} = \frac{AB}{BC}$$

$$\phi^2 = \frac{AC}{BC}$$

$$\phi^2 - \phi = \frac{AC}{BC} - \frac{AB}{BC} = \frac{BC}{BC} = 1$$

$$\phi^2 - \phi - 1 = 0$$



Expanding Recursively

$$\phi = 1 + \frac{1}{\phi}$$



Expanding Recursively

$$\phi = 1 + \frac{1}{\phi}$$

$$= 1 + \frac{1}{1 + \frac{1}{\phi}}$$



Expanding Recursively

$$\phi = 1 + \frac{1}{\phi}$$

$$= 1 + \frac{1}{1 + \frac{1}{\phi}}$$

$$= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{\phi}}}$$

$$= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{\phi}}}$$



Continued Fraction Representation

$$\phi = 1 + \frac{1}{1 + \dots}}}}}}}$$



Continued Fraction Representation

$$\frac{1+\sqrt{5}}{2} = 1 + \frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\dots}}}}}}}$$



Remember?

We already saw the convergents of this CF
[1,1,1,1,1,1,1,1,1,1,1,1]
are of the form
Fib(n+1)/Fib(n)

Hence:
$$\lim_{n\to\infty} \frac{F_n}{F_{n-1}} = \phi = \frac{1+\sqrt{5}}{2}$$



1,1,2,3,5,8,13,21,34,55,....

```
2/1 = 2

3/2 = 1.5

5/3 = 1.666...

8/5 = 1.6

13/8 = 1.625

21/13 = 1.6153846...

34/21 = 1.61904...
```

φ = 1.6180339887498948482045



Continued fraction representation of a standard fraction

$$\frac{67}{29} = 2 + \frac{1}{3 + \frac{1}{4 + \frac{1}{2}}}$$



$$\frac{67}{29} = 2 + \frac{1}{\frac{29}{9}} = 2 + \frac{1}{3 + \frac{2}{9}} 2 + \frac{1}{3 + \frac{1}{\frac{1}{2}}}$$

e.g.,
$$67/29 = 2$$
 with remainder $9/29 = 2 + 1/(29/9)$



A Representational Correspondence

$$\frac{67}{29} = 2 + \frac{1}{\frac{29}{9}} = 2 + \frac{1}{3 + \frac{2}{9}} 2 + \frac{1}{3 + \frac{1}{\frac{1}{2}}}$$

```
Euclid(67,29) 67 div 29 = 2

Euclid(29,9) 29 div 9 = 3

Euclid(9,2) 9 div 2 = 4

Euclid(2,1) 2 div 1 = 2

Euclid(1,0)
```



Euclid's GCD = Continued Fractions

$$\frac{A}{B} = \left\lfloor \frac{A}{B} \right\rfloor + \frac{1}{B}$$

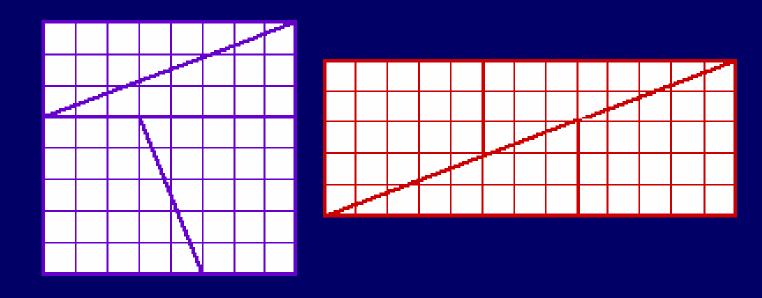
$$A \mod B$$

Euclid(A,B) = Euclid($B, A \mod B$) Stop when B=0

Theorem: All fractions have finite continuous fraction expansions



Fibonacci Magic Trick





REFERENCES

Continued Fractions, C. D. Olds

The Art Of Computer Programming, Vol 2, by Donald Knuth

"Misconceptions About the Golden Ratio", George Markowsky, College Mathematics Journal, Jan 92.

The Golden Ratio: The Story of PHI, the World's Most Astonishing Number, by Mario Livio

Fibonacci Numbers and the Golden Section, Ron Knott's excellent website





GCD

Euclid's algorithm Extended Euclid's algorithm Given X,Y, outputs r,s and GCD(X,Y)such that rX + sY = GCD(X,Y)Use it to find X^{-1} (for X in Z_n^*)

Continued Fractions

Finite CFs = rationals
Periodic CFs = roots of quadratics
Convergents
e.g. convergents of $[1,1,1,...] = F_n/F_{n-1}$

Study Bee

Golden Ratio ϕ

Solution to quadratic $x^2 - x - 1 = 0$. $\phi = [1,1,1,1,...]$