On Buffon Machines & Numbers

Philippe FLAJOLET,

INRIA-Rocquencourt, France

With Maryse PELLETIER & Michèle SORIA, LIP6, Paris

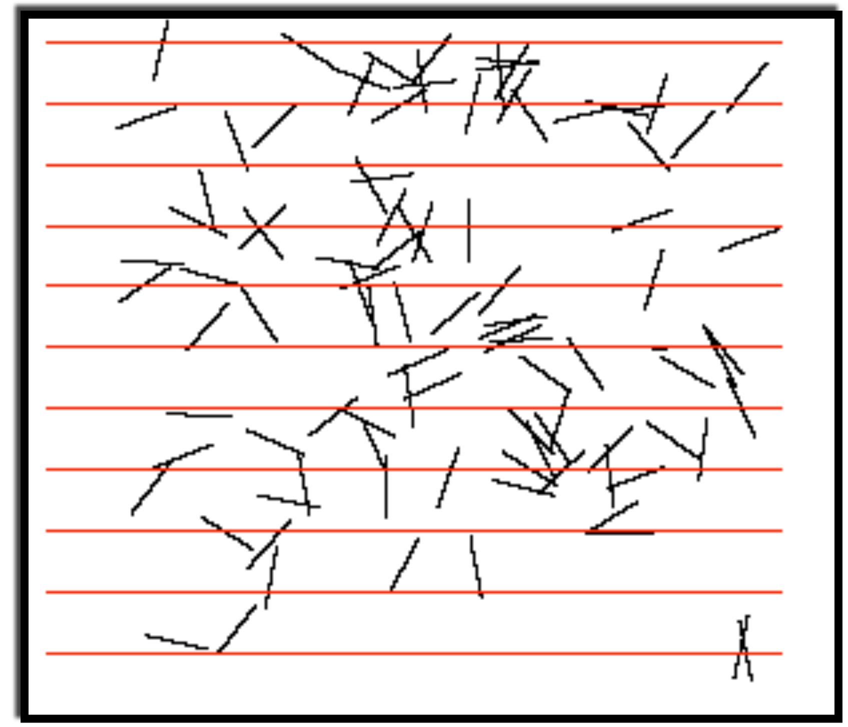
~~~ ALÉA, Luminy, March 2010 ~~~

Arxiv & http://algo.inria.fr/flajolet/
~~~

1733: Countess Buffon drops her knitting kit on the floor.

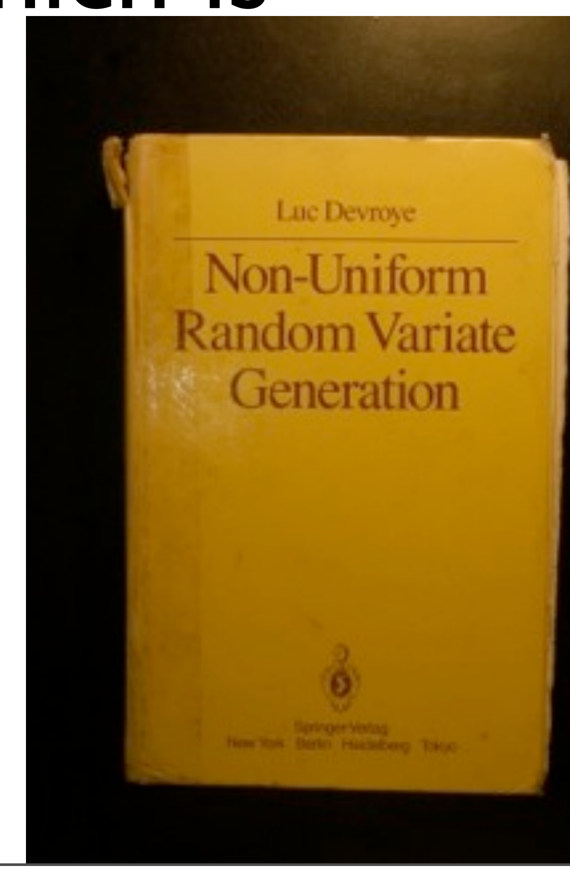
Count Buffon picks it up and notices that about 63% of the needles intersect a line on the floor.

Oh-Oh! 0.6366 is almost 2/pi (!)...



- A large body of literature on **real-number simulations**, starting with von Neumann, Ulam, Metropolis,...

- Luc Devroye's monumental synthesis, which is available on the web: @ http://cg.scs.carleton.ca/~luc/



What to do if you travel and don't want to carry floor planks and knitting needles?

Assume you have a coin!

Insist on PERFECT simulations!

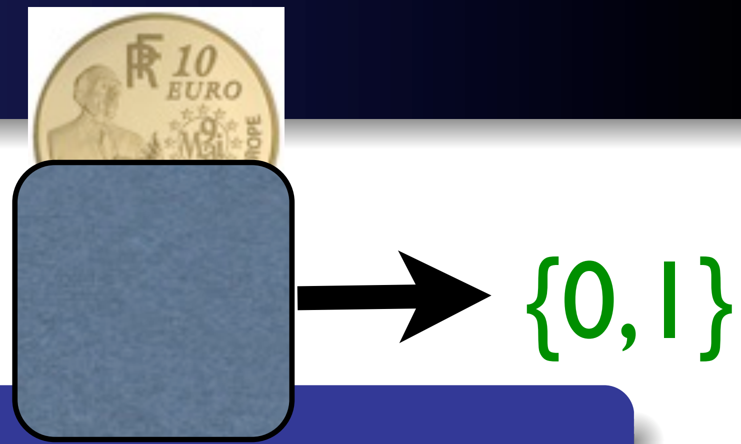
Themes:

- Computability theory: the power of probabilistic devices
- Simulation: how to be discrete & perfect?

- Boltzmann samplers for combinatorics
- Further connexions: special functions, analytic combinatorics, discrete processes, analysis of algorithms ....

1. The framework

Basic Buffon Machines



Definition

A **basic Buffon Machine** is an algorithm or program that can call an external procedure *"flip"* that provides a source of independent unbiased coin flips. Its output is in $\{0, 1\}$ (also, in $\mathbb{Z}_{>0}$) and stops. It is assumed to terminate with probability 1.

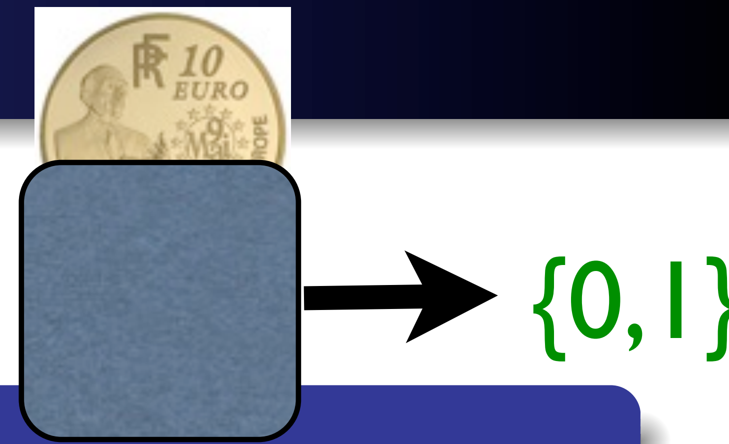
Also: interpret **1** as success $(\top)$; **0** as failure $(\bot)$.

- Can be viewed as device, such as a Turing machine, with an external tape, or **oracle** that is a random uniform $\{0, 1\}^\infty$ string.

- Read the first two symbols on the tape and output **1** if the tape starts **01**.... Succeeds with probability $\frac{1}{4}$.

- Scan tape until first **1** is encountered; output **1** if the position is even. Succeeds with probability $\frac{1}{4} + \frac{1}{16} + \cdots = \frac{1}{3}$.

Basic Buffon Machines

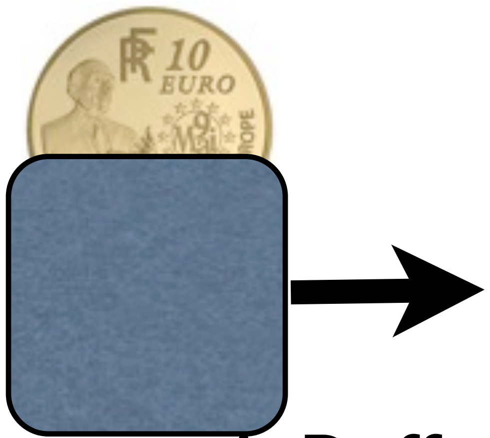


Definition

A **basic Buffon Machine** is an algorithm or program that can call an external procedure *"flip"* that provides a source of independent unbiased coin flips. Its output is in $\{0, 1\}$ (also, in $\mathbb{Z}_{>0}$) and stops. It is assumed to terminate with probability 1.

- Read the first two symbols on the tape and output **1** if the tape starts **01...**. Succeeds with probability $\frac{1}{4}$.
- Scan tape until first **1** is encountered; output **1** if the position is even. Succeeds with probability $\frac{1}{4} + \frac{1}{16} + \cdots = \mathbf{\frac{1}{3}}$.

Facts about Buffon Machines (1)



$\{0,1\}$

- A Buffon Machine, when used repeatedly, produces i.i.d random variables. Since these are in $\{0,1\}$, the BM produces a Bernoulli random variable with a certain probability p of success $(1;\top)$. That probability is a computable real $\in [0,1]$:

$$p = \sum \frac{A_n}{2^n},$$

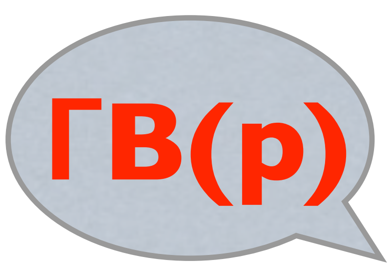
where A_n is the number of successful oracles of length n.

- **Buffon machines have no permanent memory => *they can only produce i.i.d random variables; typically, Bernoulli*.**

Facts about Buffon Machines (2)

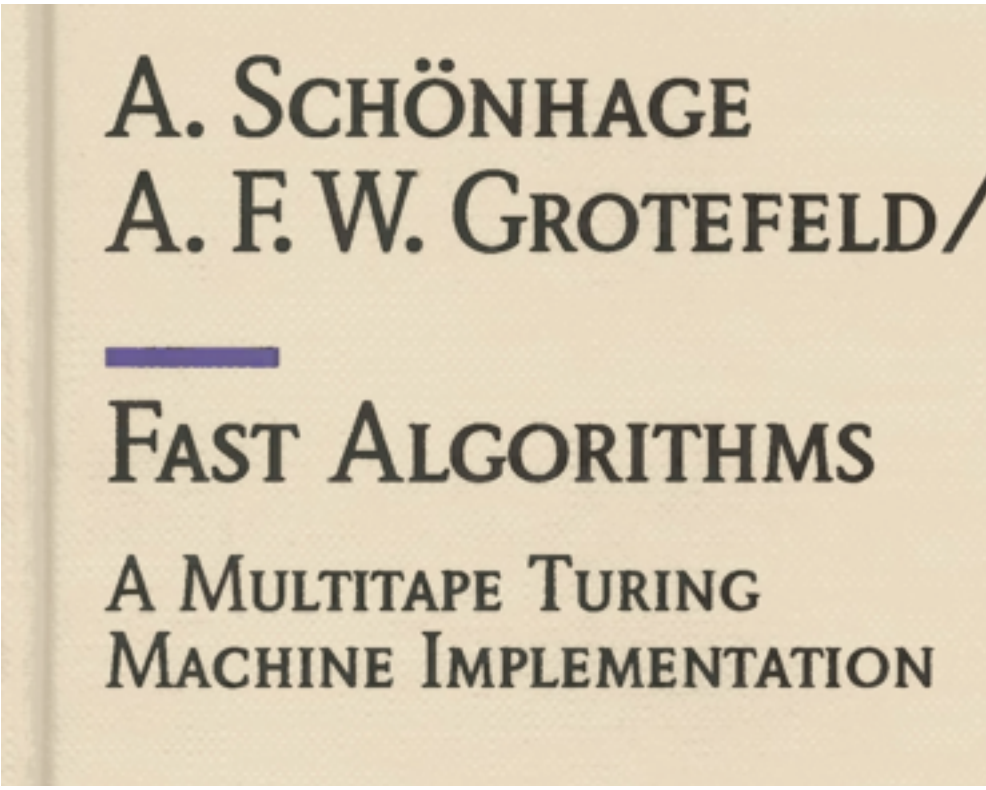
- Conversely, given any computable real $x \in [0,1]$, we can construct a simple Buffon Machine that has probability of success x. Simplified version:
  - Compute on demand as many bits of $x = (0.b_1 b_2 b_3 \cdots)_2$ as needed;
  - Compare with the oracle until a discrepant digit is found; output **1** if oracle loses (is smaller).

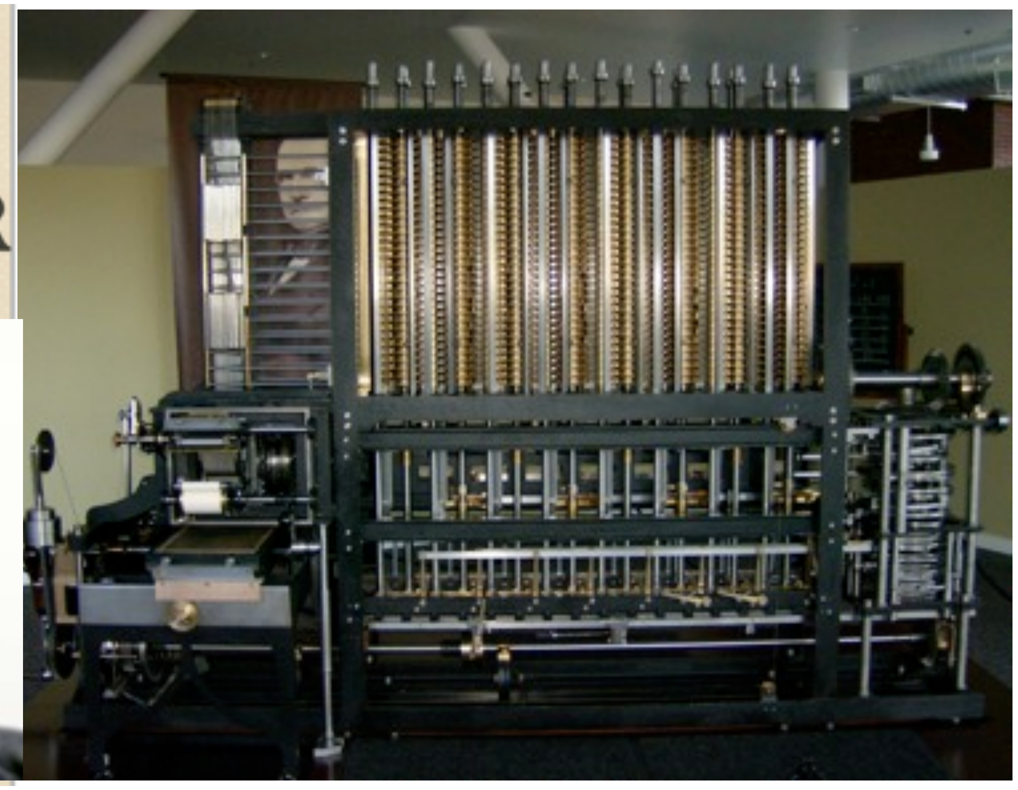
Optimization. To get a Bernoulli generator $\Gamma B(p)$, do:



$$\text{return}\ \ \mathrm{bit}_G(p), \qquad \text{where}\ G \in 1 + \mathrm{Geom}(\tfrac{1}{2}).$$

Generally, make use of a computable sequence of “framing” intervals.



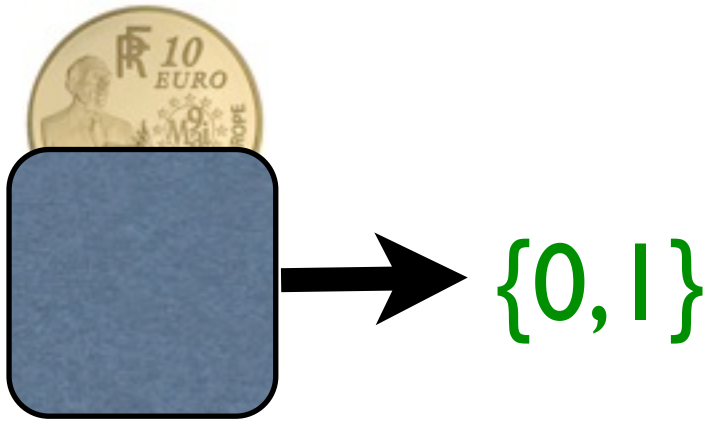


Problems with the universal construction of a $\Gamma B(p)$:

☠ • Requires arbitrary-precison routines, so that program size is HUGE

☢ • Does not qualify as "simple process"; e.g., is not human implementable.

Main purpose here is *algorithmic design.*



- Can you do such numbers as

$$\boxed{1/\sqrt{2}, \quad e^{-1}, \quad \log 2, \quad \frac{1}{\pi}, \quad \pi - 3, \quad \frac{1}{e-1}, \quad ???}$$

with only basic coin flips and no arithmetics.

- Simulation: expected # flips is finite.
  Strong simulation: + has exponential tails.

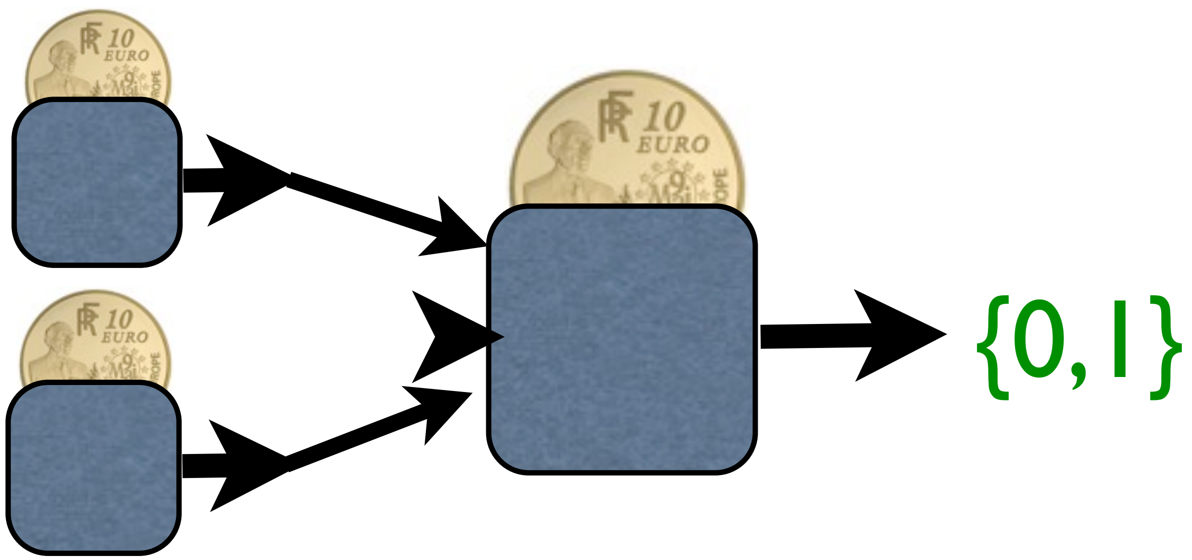
Composition of Buffon Machines

Definition (Extended Buffon Machines)

Extend the notion of (basic) Buffon machine, so that it can read from several input streams (of type $\{0, 1\}$). In particular, it may call at will other Buffon Machines.

This way, we can **compose** BMs.

- Read input_1; read input_2. Output 1 only if received 1 and 1. Computes a logical and ($\wedge$).
- If we plug in $\Gamma B(p_1)$ and $\Gamma B(p_2)$, we get $\Gamma B(p_1 p_2)$; this *without knowing* p_1, p_2 *explicitly*. Computes a product ($p_1 \times p_2$). !!!



Composition of Buffon Machines

Definition (Extended Buffon Machines)

Extend the notion of (basic) Buffon machine, so that it can read from several input streams (of type $\{0, 1\}$). In particular, it may call at will other Buffon Machines.

This way, we can **compose** BMs.

- Read input_1; read input_2. Output 1 only if received 1 and 1.
Computes a logical and ($\wedge$).

- If we plug in $\Gamma B(p_1)$ and $\Gamma B(p_2)$, we get $\Gamma B(p_1 p_2)$; this *without knowing* p_1, p_2 *explicitly.* Computes a product ($p_1 \times p_2$). !!!

Definition (Function computed by a BM)

An extended BM is said to compute $\phi(p)$ if, given as input a machine that is a $\Gamma B(p)$ [p unknown!], it produces a $\Gamma B(\phi(p))$.

Which of these can be computed?

$$p^2, \quad 2p - p^2, \quad \frac{p_1 + p_2}{2}, \quad 2p.$$

Meta-theorem

Theorem (Class of BM computable functions $\varphi(p)$)

You can do constructively, simply, and efficiently:

- *Closure under half sum, product, composition.*
- *All rational numbers $p \in \mathbb{Q}$; many polynomials and rational functions with rational coeffs.*
- *Positive $\mathbb{Q}$-algbebraic functions including $\sqrt{p}$.*
- *Exponentials; logarityhms; trig functions.*
- *Closure under integration; inverse trigs.*
- *Hypergeometrics of binomial type.*
- *$+$ Poisson and logarithmic-series generators.*

[Nacu–Peres-Mossel]. With suitable (but costly) arithmetics, can do all polynomials and rational functions that map $[0, 1]$ to $(0, 1)$. [Keane–O'Brien] Cannot do $2p$, without restrictions; do continuous functions by approximation.



- Builds on ideas of
  von Neumann, Knuth-Yao
- Encapsulates constructions by
  Wästlund, Nacu, Peres, Mossel
- Develops new constructions:
  VN-generator, integration; Poisson & logarithmic distributions.

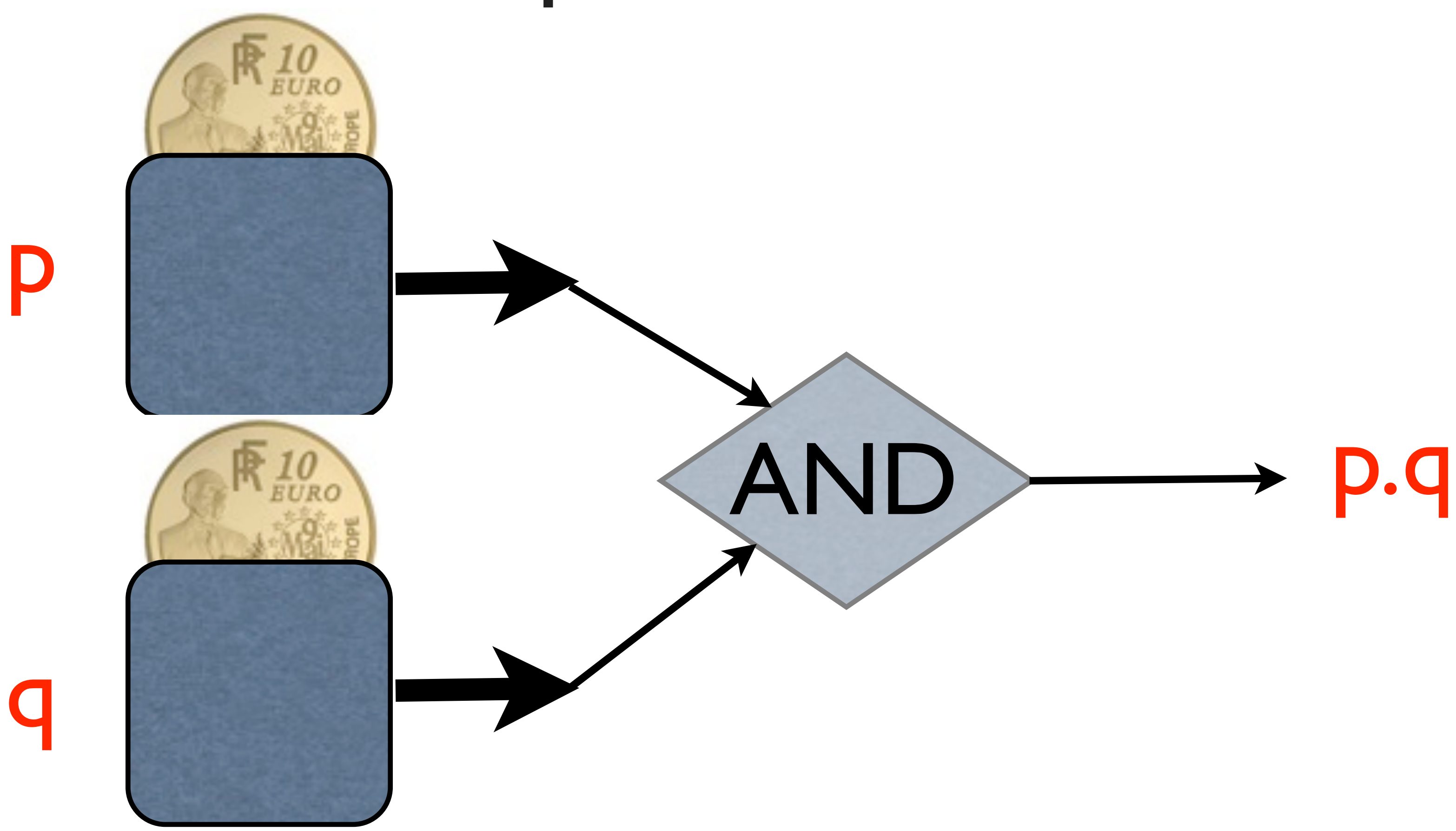
- We shall see nine ways to get π, some with 5 coin flips on average, with typically about a dozen lines of code...

2. Basic construction rules

- **Decision trees and loopless programs**

  Do Bernoulli of param. 3/8,5/8; *dyadic rationals*

  "Compute" *Boolean combinations*



p.q
1-p
(p+q)/2
p^2

Name	*realization*	*function*
Conjunction $(P \wedge Q)$	**if** $P() = 1$ **then** return$(Q())$ **else** return(0)	$p \wedge q = p \cdot q$
Disjunction $(P \vee Q)$	**if** $P() = 0$ **then** return$(Q())$ **else** return(1)	$p \vee q = p + q - pq$
Complementation $(\neg P)$	**if** $P() = 0$ **then** return(1) **else** return(0)	$1 - p$
Squaring	$(P \wedge P)$	p^2
Conditional $(P \rightarrow Q \vert R)$	**if** $R() = 1$ **then** return$(P())$ **else** return$(Q())$	$rp + (1 - r)q.$

• Finite graphs and Markov chains

- <u>*Can do all rational p:*</u>
  To do a ΓB(3/7), flip three times; in 3 cases, return(1); in 4 cases return(0); otherwise repeat.

- $\boxed{\textit{do a geometric } \Gamma G(p) \textit{ from a Bernoulli } \Gamma B(p)}$

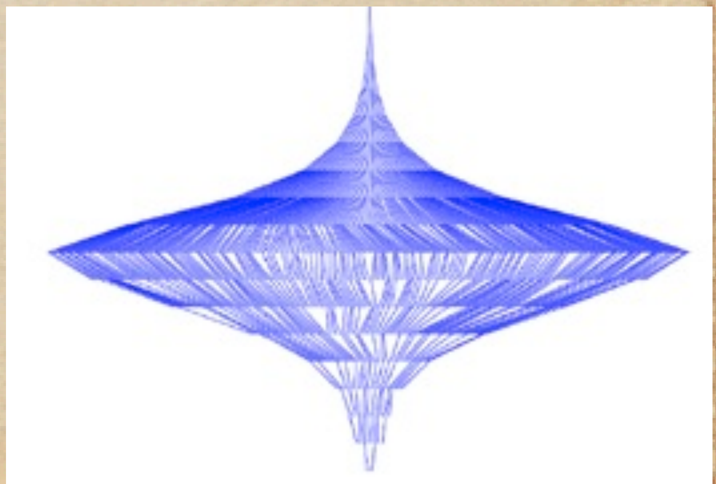
- From a ΓB(p); repeatedly try till **1** is observed. If number of trials is even, then return(**1**).
  <u>*Computes 1/(1+p)*</u> $= (1-p)[1+p^2+p^4+ \ldots]$

- Mossel, Nacu, Peres, Wästlund:

Theorem 1 ([21, 22, 27]). *(i) Any polynomial $f(x)$ with rational coefficients that maps $(0,1)$ into $(0,1)$ is strongly realizable by a finite graph. (ii) Any rational function $f(x)$ with rational coefficients that maps $(0,1)$ into $(0,1)$ is strongly realizable by a finite graph.*

- .... but it requires arbitrary-precision routines.

3. The von Neumann schema



Von Neumann Schema (I)

- Choose a class of permutations with P_n the number of those of size *n*.
- *Draw* N∈Geo(λ) *uniform Random Variables over* [0,1].
- *Succeed if the order type is good = in* P_n.

geometric

$\Gamma\mathrm{VN}[\mathcal{P}](\lambda) := \{$ **do** $\{$
$N := \Gamma\mathrm{G}(\lambda);$
let $\mathbf{U} := (U_1, \ldots, U_N)$ be a vector of $[0,1]$–uniform variables.
$\{$ *bits of the* U_j *are produced on a call-by-need basis to determine* σ *and* τ $\}$
set $\tau := \mathrm{trie}(\mathbf{U})$; **let** $\sigma := \mathrm{type}(\mathbf{U})$;
if $\sigma \in \mathcal{P}_N$ **then** return$(N)\}\}$.

Von Neumann Schema (2)

- Choose a class of permutations with P_n the number of those of size n. Draw N=Geom(lambda).

- Probability of success with $N{=}n$ is

$$\frac{(1-\lambda)P_n\lambda^n/n!}{(1-\lambda)\sum_n P_n\lambda^n/n!} = \frac{1}{P(\lambda)}\frac{P_n\lambda_n}{n!}n.$$

- Thus, get *Poisson and logarithmic distributions*

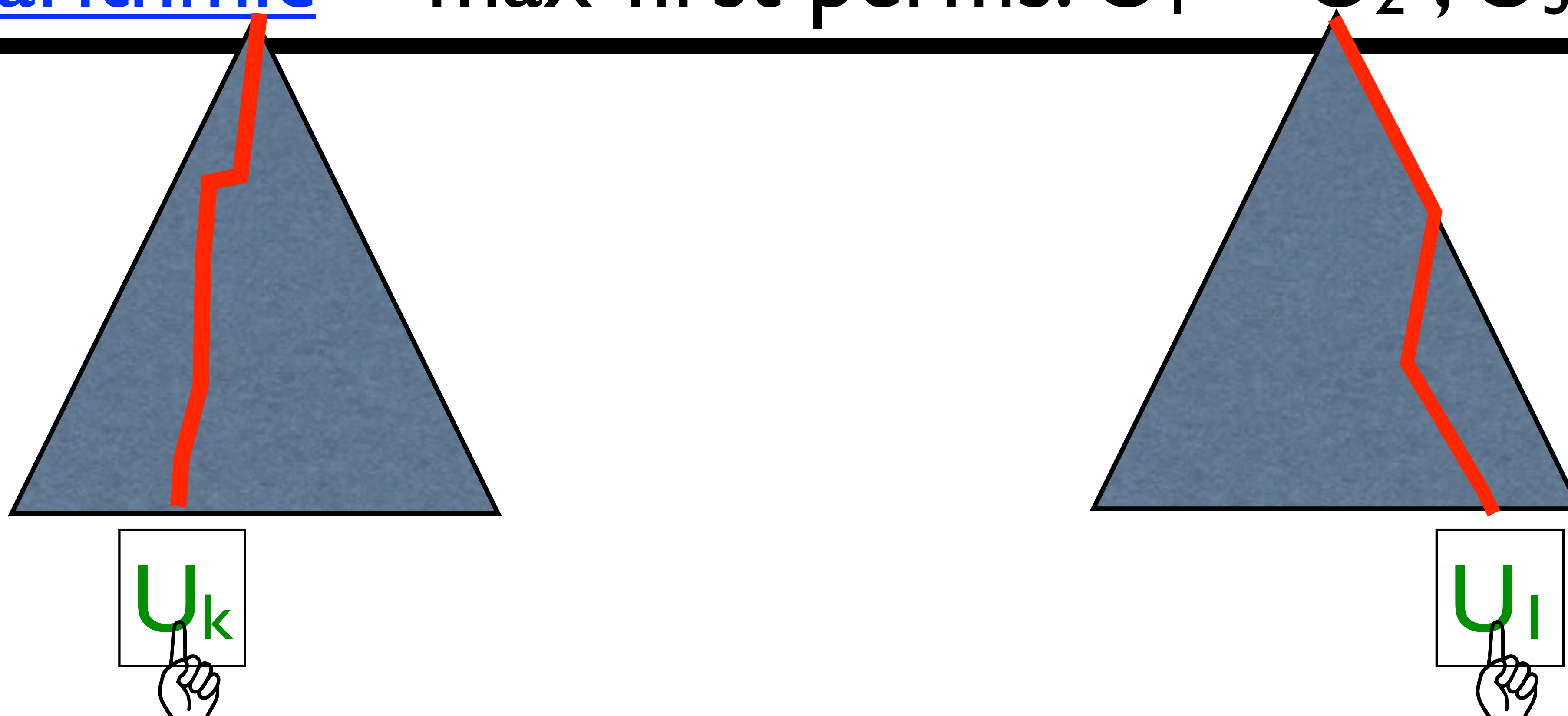
permutations $(\mathcal{P})$:	all $(\mathcal{Q})$	sorted $(\mathcal{R})$	cyclic $(\mathcal{S})$
distribution:	$(1-\lambda)\lambda^n$ geometric	$e^{-\lambda}\frac{\lambda^n}{n!}$ Poisson	$\frac{1}{L}\frac{\lambda^n}{n}$, $\quad L := \log(1-\lambda)^{-1}$ logarithmic.

Von Neumann Schema (3)

- Using a *digital tree* (aka *trie*), we only need a single string register to recognize perm classes for Poisson and logarithmic distribs!

- Poisson = sorted perms: $U_1 < U_2 < U_3$

- Logarithmic = max-first perms: $U_1 > U_2, U_3$



cf **Leader election**: Prodinger; Fill, Mahmoud, Szpankowski, Janson,...

- For **VN schema**, *path-length of tries* determines # coin flips.

PGF:

$$h_n(q) = \frac{1}{1-q^n 2^{1-n}} \sum_{k=1}^{n-1} \frac{1}{2^n} \binom{n}{k} h_k(q) h_{n-k}(q).$$

Proposition 1. *(i) Given a class $\mathcal{P}$ of permutations and a parameter $\lambda \in (0,1)$, the von Neumann schema $\Gamma\mathrm{VN}[\mathcal{P}](\lambda)$ produces* exactly *a discrete random variable with probability distribution*

$$\mathbb{P}(N=n) = \frac{1}{P(\lambda)} \frac{P_n \lambda^n}{n!}.$$

(ii) The number K of iterations has expectation $1/s$, where $s = (1-\lambda)P(\lambda)$, and its distribution is $1 + \mathrm{Geo}(s)$.

(iii) The number C of flips consumed by the algorithm (not counting[2] *the ones in $\Gamma\mathrm{G}(\lambda)$) is a random variable with probability generating function*

$$\mathbb{E}(q^C) = \frac{H^+(\lambda, q)}{1 - H^-(\lambda, q)}. \tag{10}$$

where H^+, H^- are determined by (9):

$$H^+(z,q) = (1-z)\sum_{n=0}^{\infty} \frac{P_n}{n!} h_n(q) z^n, \qquad H^-(z,q) = (1-z)\sum_{n=0}^{\infty} \left(1 - \frac{P_n}{n!}\right) h_n(q) z^n.$$

The distribution has exponential tails.

Theorem 2. *The Poisson and logarithmic distributions of parameter $\lambda \in (0,1)$ have a strong simulation by a Buffon machine that only uses a single string register.*

- ☛ ***Poisson***: Declare success (1) if N=0; failure o.w. Get exp(-λ), etc.
- ☛ **Check P:** Do only one run; return(1) if success. E.g, for Poisson, gives (1-λ)exp(λ)
- ☛ Use alternating (zigzag) perms & get trigs!

Theorem 3. *The following functions admit a strong simulation:*

$$e^{-x},\ e^{x-1},\ (1-x)e^{x},\ xe^{1-x},$$
$$\frac{x}{\log(1-x)^{-1}},\ \frac{1-x}{\log(1/x)},\ (1-x)\log\frac{1}{1-x},\ x\log(1/x),$$
$$\frac{1}{\cos(x)},\ x\cot(x),\ (1-x)\cos(x), (1-x)\tan(x).$$

- Polylogarithms, Bessel,...: do *r* experiments

$$\mathrm{Li}_r(z) := \sum_{n=1}^{\infty} \frac{z^n}{n^r},$$

$$\mathrm{Li}_2(1/2) = \frac{\pi^2}{12} - \frac{1}{2}\log^2 2, \qquad \mathrm{Li}_3(1/2) = \frac{1}{6}\log^3 2 - \frac{\pi^2}{12}\log 2 + \frac{7}{8}\zeta(3).$$

Get log(2), then $\pi^2/24$, in less than 10 flips on average

4. Square roots, algebraic & hypergeometric functions

- Generate N∈Geo(λ) and succeed if we get a balanced score from 2N flips.
- The probability of success:

$$s(\lambda) := \sum_{n=0}^{\infty}(1-\lambda)\lambda^n \varpi_n = \sqrt{1-\lambda}$$

$$\varpi_n = \frac{1}{2^{2n}}\binom{2n}{n}$$

Theorem 4. *The square-root construction of Equation* (11) *provides an exact Bernoulli generator of parameter* $\sqrt{1-\lambda}$, *given a* $\Gamma\mathrm{B}(\lambda)$. *The mean number of coin flips required, not counting the ones involved in the calls to* $\Gamma\mathrm{B}(\lambda)$, *is* $\frac{2\lambda}{1-\lambda}$. *Hence the function* $\sqrt{1-x}$ *is strongly realizable.*

Theorem 5 ([21]). *To each bistoch grammar* G *and non-terminal* S, *there corresponds a construction (Figure 3), which can be implemented by a deterministic pushdown automaton and calls to a* $\Gamma\mathrm{B}(\lambda)$ *and is of type* $\Gamma\mathrm{B}(\lambda) \longrightarrow \Gamma\mathrm{B}\left(S\left(\frac{\lambda}{2}\right)\right)$, *where* $S(z)$ *is the algebraic function canonically associated with the grammar* G *and non-terminal* S.

- Get hypergeometrics of binomial type.

Ramanujan:

$$\frac{1}{\pi} = \sum_{n=0}^{\infty} \binom{2n}{n}^3 \frac{6n+1}{2^{8n+4}};$$

procedure Rama(); $\{$*returns the value 1 with probability* $1/\pi\}$
S1. let $S := X_1 + X_2$, where $X_1, X_2 \in \mathrm{Geom}(\frac{1}{4})$;
S2. with probability $\frac{5}{9}$ do $S := S + 1$;
S3. for $j = 1, 2, 3$ do
S4. draw a sequence of $2S$ coin flippings;
if (# Heads − # Tails) $\neq 0$ then return(0);
S5. return(1).

<11 coin flips on average

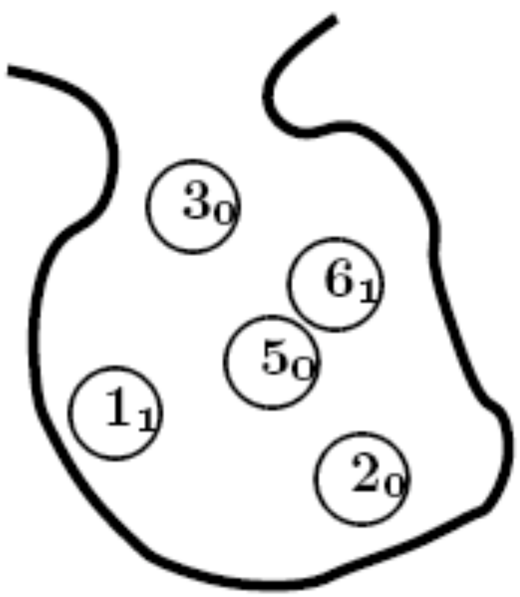
5. A Buffon integrator

- In a construction of a ΓB(φ(λ)) from a ΓB(λ), we substitute a ΓB(Uλ), with U uniform. Get an *integrator*:

$$\Phi(\lambda) = \frac{1}{\lambda}\int_0^\lambda \phi(w)\,dw.$$

- We can do *a product ΓB(Uλ)=ΓB(U).ΓB(λ) by an AND* (∧), while *emulating a uniform U with a "bag"*:

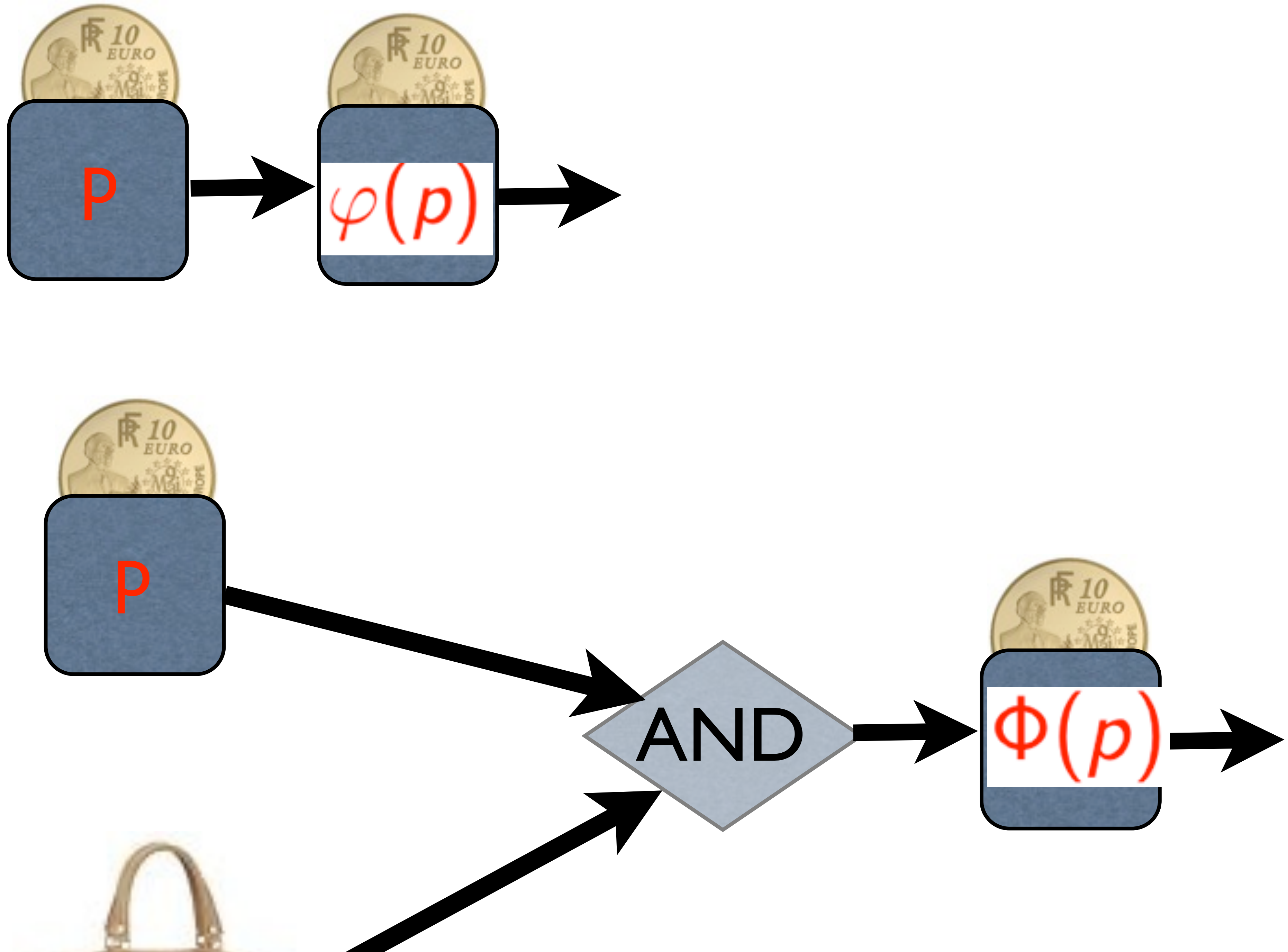
$U =$



$$= \begin{array}{rc|c|l} 9 & : & ? & \\ 8 & : & ? & \\ 7 & : & ? & \Longleftarrow J \\ 6 & : & 1 & \\ 5 & : & 0 & \\ 4 & : & ? & \\ 3 & : & 0 & \\ 2 & : & 0 & \\ 1 & : & 1 & \end{array}$$

```
Ghalf:=proc() local K;
# a geometric RV of param. 1/2
      K:=-1; do K:=K+1; if flip()=0
             then return(K) fi; od;

bag:=proc(U)  local J;
           J:=1+Ghalf();
           if type(U[J],name)
              then U[J]:=flip(); fi;
           return(U[J]); end;
```



$$\Phi(p) := \frac{1}{p} \int_0^p \varphi(w)\, dw$$

Theorem 6. *Any construction* $\mathbf{C}$ *that produces a* $\Gamma\mathrm{B}(\phi(\lambda))$ *from a* $\Gamma\mathrm{B}(\lambda)$ *can be transformed into a construction of a* $\Gamma\mathrm{B}(\Phi(\lambda))$*, where* $\Phi(\lambda) = \frac{1}{\lambda}\int_0^\lambda \phi(w)\,dw$*, by addition of a geometric bag. In particular, if* $\phi(\lambda)$ *is realizable, then its integral taken starting from* 0 *is also realizable. If in addition* $\phi(\lambda)$ *is analytic at 0, then its integral is strongly realizable.*

- Chain: $p \rightarrow p^2 \rightarrow 1/(1+p^2) \xrightarrow{\int} \arctan(x)$

Theorem 7. *The following functions are strongly realizable* $(0 < x < 1)$*:*

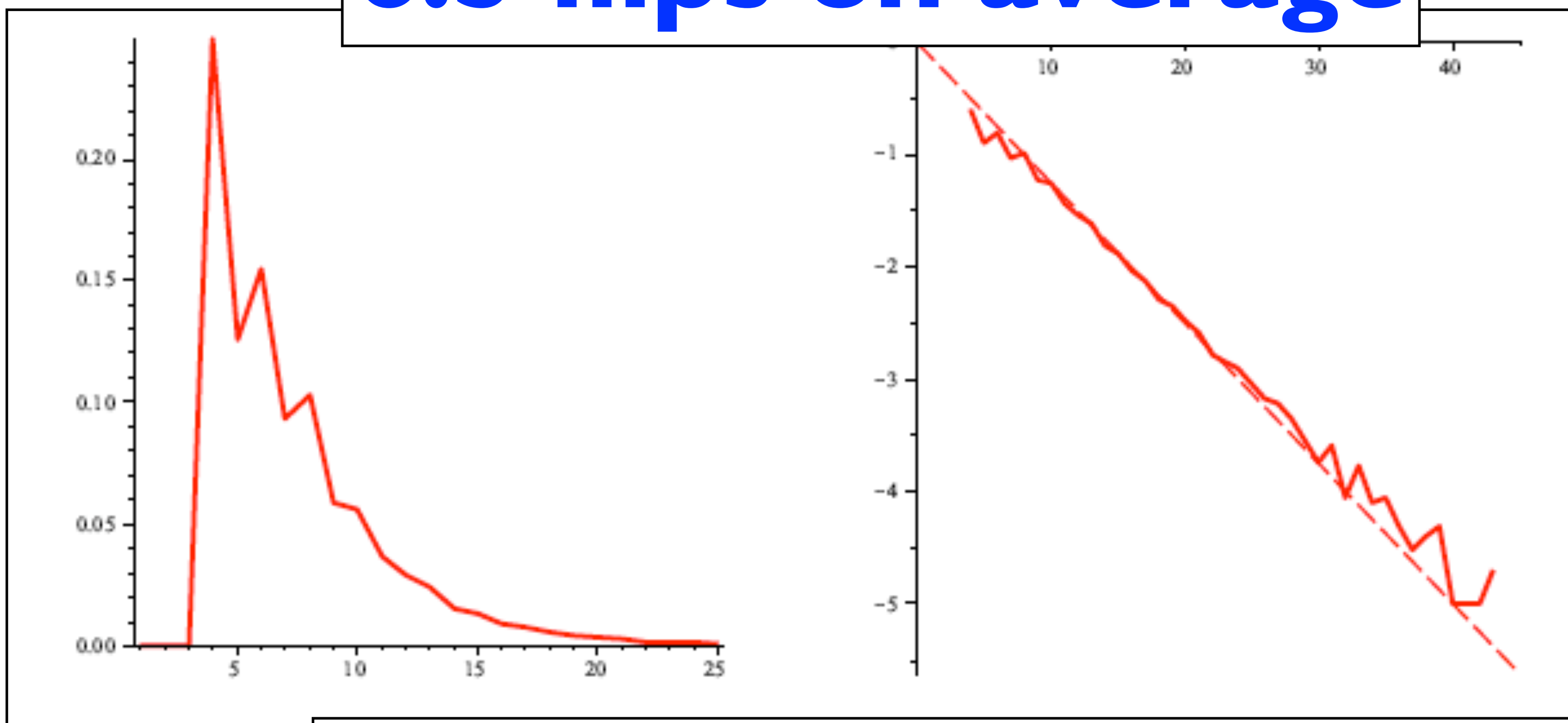
$$\log(1+x),\ \arctan(x),\ \frac{1}{2}\arcsin(x),\ \int_0^x e^{-w^2/2}\,dw.$$

- Madhava-Gregory-Leibniz: arctan(1)=π/4

```
MGL:=proc()     do
        if bag(U)=0 then return(1) fi;  if bag(U)=0 then return(1) fi;
        if bag(U)=0 then return(0) fi;  if bag(U)=0 then return(0) fi; od;  end.
```

- Machin machine: arctan(1/2)+arctan(1/3)=π/4.

6.5 flips on average



Distribution of costs (plain & log.)

6. Experiments

MAPLE: an interpreter ~ 60 lines

```
> Z4:=expn(compl(ave(flip,ave(int1(int1(int1(even(prod(z, prod
  (x, y))), x, ONE), y, ONE), z, ONE),compl(sqrt0(int0(ave(logp
  (flip), sqrt0(prod(flip, int0(prod(Y, even(int0(even(prod(X,
  X)), X, expn(prod(flip,Z))))), Y, expn(prod(flip,flip)))))),
  Z, flip))))))):
> test(Z4,10000);
mean_number_of_flips = 103.1645000
```

$$0.6313000000$$

```
> val(Z4);
```

$$e^{-\frac{1}{2}+\frac{3}{16}\zeta(3)-\frac{1}{4}\sqrt{2}\sqrt{\int_0^{\frac{1}{2}}\left(\frac{1}{2}\ln(2)+\frac{1}{4}\sqrt{\frac{e^{-\frac{1}{4}}}{1+\frac{\arctan\left(e^{-\frac{1}{2}Z}\right)}{e^{-\frac{1}{2}Z}}}}\right)\mathrm{d}Z}}$$

```
> evalf(val(Z4));
```

$$0.6356033009$$

- Implements all earlier constructions: *it works*!
- Results for π-related constants:

$\mathrm{Li}_2(1/2)$	Rama	arcsin $[1; \frac{1}{\sqrt{2}}; \frac{1}{2}]$			arctan $[1/2 + 1/3; 1]$		$\zeta(4)$	$\zeta(2)$
$\frac{\pi^2}{24}$	$\frac{1}{\pi}$	$\frac{\pi}{4}$	$\frac{\pi}{8}$	$\frac{\pi}{12}$	$\frac{\pi}{4}$	$\frac{\pi}{8}$	$\frac{7\pi^4}{720}$	$\frac{\pi^2}{12}$
7.9	10.8	76.5 (∞)	16.2	4.9	4.5	26.7 (∞)	6.2	7.2.

Method; constant; mean # flips