INTRODUCTION TO COMPUTER SCIENCE · WEEK 2 · STATION I

Size

Bits, orders of magnitude, state spaces — and why testing shows presence, not absence.

WHERE WE LEFT OFF

One line, in a formal language.

```
$ ./selfie
./selfie { -c { source } | -o binary | [ -s | -S ] assembly | -l binary } [ ( -m | -d | -r | -y ) 0-4096
```

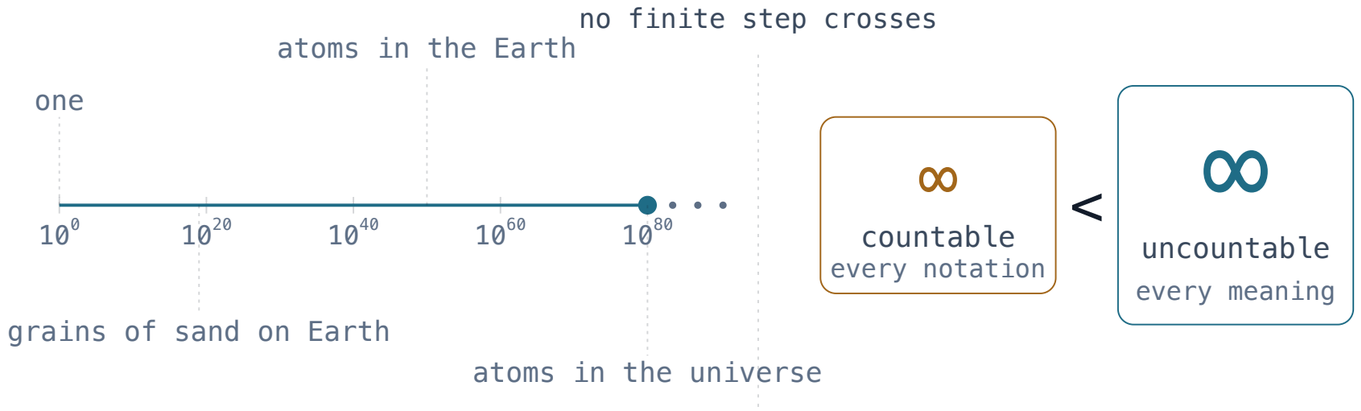
This is selfie's synopsis: a regular expression that says exactly how selfie may be invoked. Braces mean *any number of*, brackets mean *optional*, the bar means *or*. We learn to read it properly in week 4.

For today: `-c selfie.c` compiles selfie's own source. Watch what it prints about itself.

```
$ ./selfie -c selfie.c
./selfie: selfie compiling selfie.c to 64-bit RISC-U with 64-bit starc
./selfie: 365784 characters read in 12394 lines and 1741 comments
./selfie: with 219888(60.11%) characters in 51329 actual symbols
./selfie: 491 global variables, 661 procedures, 512 string literals
./selfie: 188392 bytes generated with 43492 instructions and 14424 bytes of data
```

THE MAP

One line carries the whole class: small, vast, then two sizes of endless.



log scale · then off the end of every scale

Left. From one thing to every atom there is — the whole physical world in eighty steps, each one ten times the last.

The wall. No number of steps crosses it. Infinity is not the far end of the line; it is what the line never reaches.

Right. Past the wall, endlessness comes in sizes: one counts every notation, the other every meaning — and is strictly bigger.

ANCHORS

A million, a billion, a trillion. Three words one syllable apart — three different worlds.

10^6 s

A MILLION SECONDS · 11½ DAYS

10^9 s

A BILLION SECONDS · 32 YEARS

10^{12} s

A TRILLION SECONDS · 32,000 YEARS

A million seconds ago you had not yet packed for today. A billion seconds ago you were a child, or not yet born. A trillion seconds ago somebody was painting the walls of the Chauvet cave.

THE HABIT

Never accept a number you cannot picture. Find the anchor — a fortnight, a career, an ice age — or admit you are repeating the number, not reading it.

UNITS

Kilo, mega, giga, tera: the same three zeros, again and again.

PREFIX	FACTOR	BYTES · HOW MUCH IT HOLDS	HERTZ · HOW OFTEN IT HAPPENS
kilo	10^3	KB — a long paragraph	kHz — the top of human hearing, 20 kHz
mega	10^6	MB — a novel	MHz — the first IBM PC, 4.77 MHz, 1981
giga	10^9	GB — a thousand novels	GHz — this laptop; Wi-Fi at 2.4
tera	10^{12}	TB — a million novels	THz — infrared light

A byte is eight bits — one keystroke. A hertz is one beat per second. The left column counts things, the right counts events, and the prefixes do not care which.

HOLD A NANOSECOND

At 1 GHz a beat lasts a billionth of a second, and light gets 30 cm in that time. Grace Hopper handed those 30 cm out as lengths of wire, so a room could feel why a signal cannot cross a desk and return inside one beat.

TWO KILOS

Kilo means a thousand. In memory it means 1,024. Both are in daily use.

	BASE 10	BASE 2	GAP
kilo	KB = 10^3	KiB = 2^{10} = 1,024	+2.4%
mega	MB = 10^6	MiB = 2^{20}	+4.9%
giga	GB = 10^9	GiB = 2^{30}	+7.4%
tera	TB = 10^{12}	TiB = 2^{40}	+10.0%

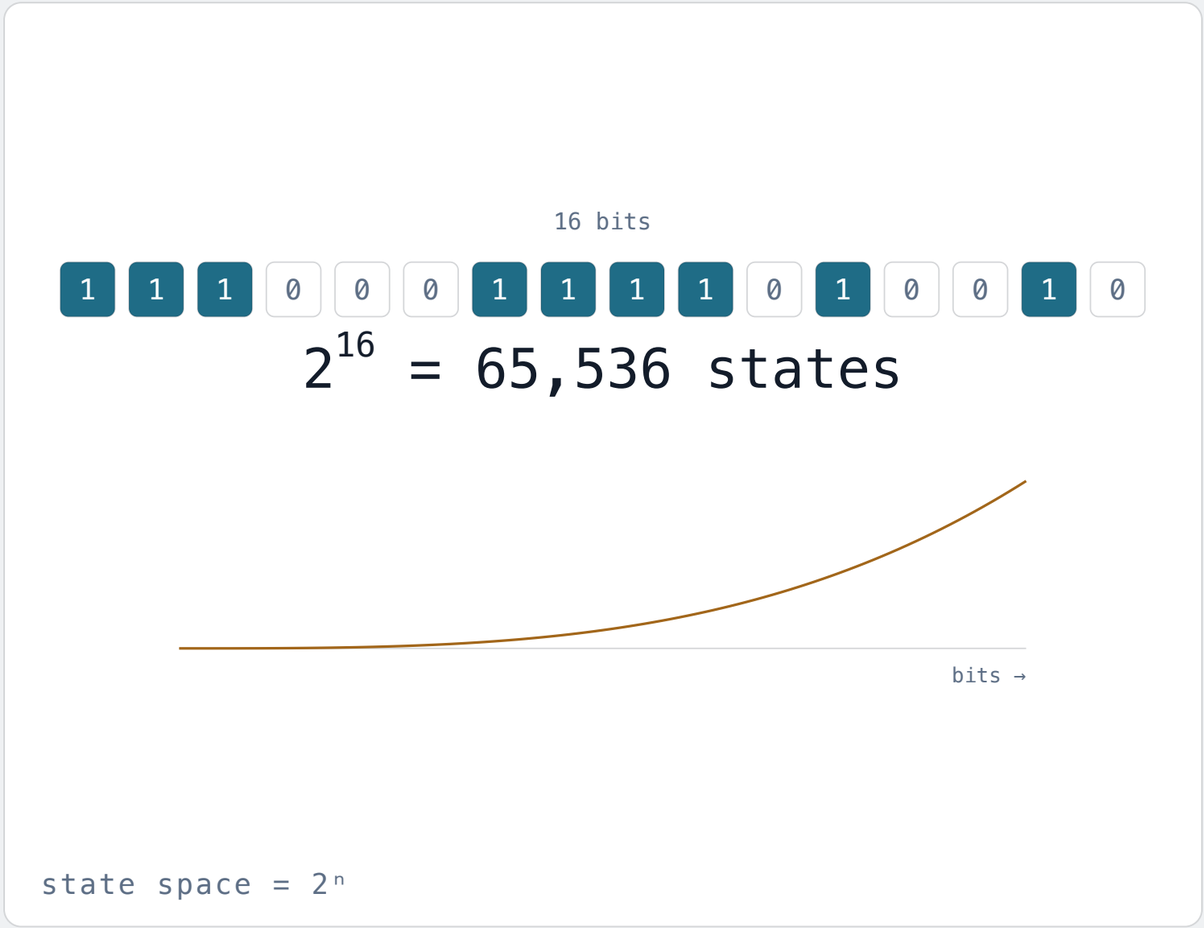
Memory is addressed with bits, so its natural steps are powers of two — and 2^{10} lands so close to 1,000 that the prefix was simply borrowed. Memory comes in powers of two; disks and networks in powers of ten. A "1 TB" drive holds 10^{12} bytes and shows up as 931 GiB.

NOTICE WHAT JUST HAPPENED

One notation, two meanings, and nothing in the notation to tell you which you were handed. Keep hold of that. The rest of the semester is about that gap.

ONE BIT

A bit is one distinction: this, not that.



Add a bit and you don't *add* states. You double them.

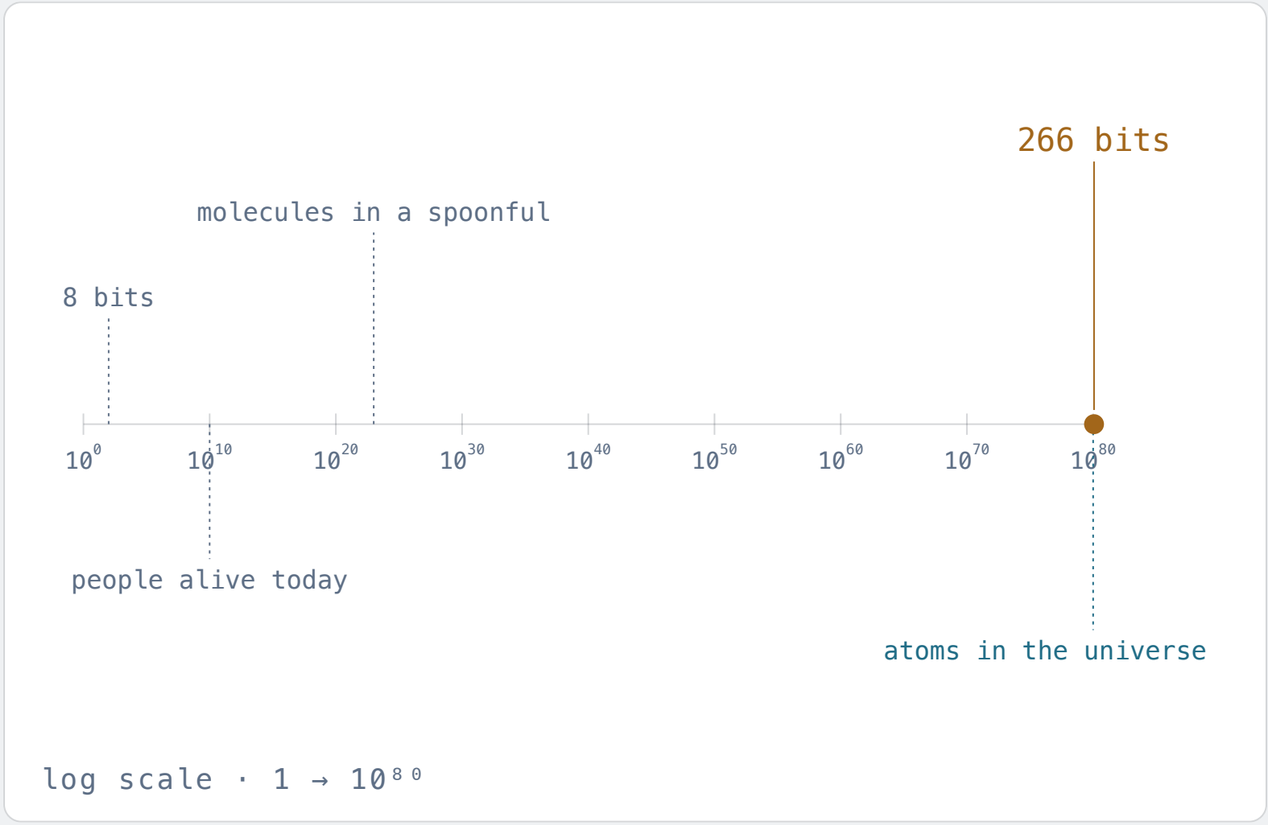
Ten bits: a thousand states. Twenty: a million. Thirty: a billion.

Doubling is the most underestimated operation in human reasoning — and the entire engine of computing.

How many bits does it take to have more states than the universe has atoms?

SCALE CHECK

34 bytes of memory has more states than the observable universe has atoms.



10^{80}

ATOMS IN THE OBSERVABLE UNIVERSE

2^{266}

STATES OF 266 BITS ≈ 1.2×10^{80}

37 bytes to beat every photon, about 10^{89}.

YOUR POCKET · AND THE SPECIMEN

A phone with 8 GB of memory has $2^{68,719,476,736}$ states.

Written out in decimal, that number has about 20.7 billion digits. At 3,000 digits a page: 6.9 million pages. Bound into 500-page books: 13,800 volumes — some 400 metres of shelf.

That shelf does not hold the phone's states. It holds the *number* of them, written down once.

Every app you will ever install, every photo you will ever take, and every bug you will ever hit is one point in this space.

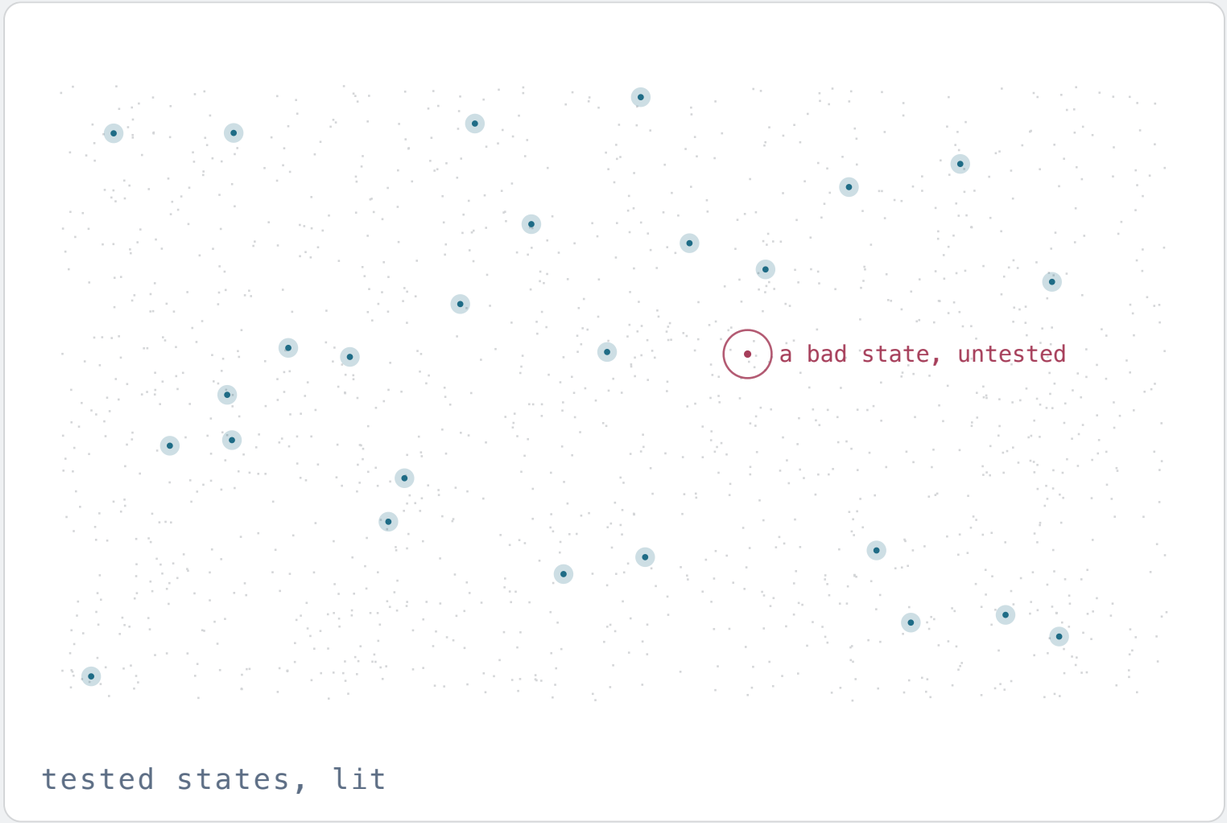
SELFIE'S MACHINE

RISC-U has 4 GB of byte-addressed memory: 2^{35} bits, so $2^{34,359,738,368}$ states — a number with 10.3 billion decimal digits. Selfie, the program that runs on it, is 12,394 lines. One person can read the program to the end. Nobody can inspect the machine.

```
2^34359738368 = 23873117371686870504051737186753437
90316897974753805346507569675911986411509205033986 9
94105429253718329002209917514457363727964159248766 7
34550493444566589626653439050806025736458878735872 0
56453636861539278309337170472140168557499483813718 7
63620556826923243128918037860241764100458175521435 5
```

THE CONSEQUENCE

In a space that big, good states and bad states look alike.



Test a billion states per second, starting at the Big Bang, and by now you would have checked about 2^{88} of them. Out of $2^{68,719,476,736}$.

The fraction is not small. It is indistinguishable from zero.

> *Testing shows the presence, not the absence of bugs.*
>
> EDSGER W. DIJKSTRA, 1969

REFRAME · 1 OF 5

The vastness is not the problem. It is the inventory.

WHAT IT FORBIDS

Certainty by inspection. You will never check your system exhaustively — not with more testers, not with faster machines, not ever.

WHAT IT OPENS

Every symphony not yet written, every protein not yet folded, every proof not yet found, every program not yet imagined — all of them are already in there, waiting to be addressed. The same enormity that hides your bugs is the reason novelty is inexhaustible.

First appearance of the device that every limit in this class gets: what it forbids, and what it opens. A space small enough to audit would be a space too small to be interesting.

INTERLUDE · LITERATURE

Borges built this space in 1941 and called it a library.

The Library of Babel holds every possible book: every truth, every refutation — and overwhelmingly, gibberish. Its inhabitants go mad: containing every truth is worthless without a way of finding one.

Modern echo: a large language model already "contains" astonishing amounts of text. Containment was never the hard part.

BORGES · 1941

410 pages, 40 lines, 80 characters, 25 symbols. The shelves hold $25^{1,312,000}$ books $\approx 10^{1,834,097}$ volumes.

THE SHIFT

From storage to search: from having the answer somewhere to knowing it when you see it.

LIFE · FROM THE BOOK

A machine that counts through all its states would, eventually, write the Mona Lisa. It would not notice.

Write a program that treats all of memory as one integer and increments it. It enumerates every state the machine can be in. It will have turned to dust before getting anywhere — but suppose it ran.

Once in a while, an enormous while, the state would encode Beethoven's Fifth, or the human genome. Without any intelligence whatsoever. And the machine could not *select* those states over the others.

Evolution is that program with a selector attached: enumerate DNA by mutation, select by reproductive success. A physical process bound by physics, modelled by a computer bound by logic.

QUANTITIES MATTER

Lots of time, or mutate fast. Lots of space, or be tiny. Lots of energy, or be efficient. A few more zeroes in a number make all the difference — in life and in computing.

BEFORE NEXT WEEK

Recommended exercises.

1. Read the Size chapter up to *Numbers*, and Borges's *Library of Babel*.
2. Rerun `./selfie -c selfie.c`. Where on the map do 365,784 characters sit? And 43,492 instructions? And the machine's state count?
3. How many bits do you need for a thousand states? A million? Every person alive? Every atom in the Earth (about 10^{50})?
4. A "512 GB" phone: how many GiB does it report, and why?
5. Find three anchors of your own for a million, a billion, and a trillion — of anything.

NEXT WEEK

Everything is bits: numbers, negative numbers, overflow, characters, text, files, images, code. The same 85 in five different notations, and what the machine makes of each.