# From rules to compression 

## to thinking

From Instructions to AI

## HUGE CONCEPT 1

All computers do is math (we did this with binary)

## HUGE CONCEPT 2

All computers do is follow a very precise list of instructions that one or more people wrote.

## Understanding Power

10 PRINT "John is AWESOME";
20 GOTO 10

## Teaching the robots to escape

1) If there's a door in arms-reach, exit - you're done, else
2) If you can, take one step forward then goto 1), else
3) Rotate to the left until there's not a wall in front of you then goto 1)
(this will get you out of any "regular" empty room)

## An almost random bit on recursion

- In computers, it's actually okay to define something with itself. PSUEDOCODE!

```
Define function="EscapeFromRoom"{
1) If there's a door in arms-reach, exit - you're done, else
2) If you can, take one step forward then EscapeFromRoom, else
3) Rotate to the left until there's not a wall in front of you then EscapeFromRoom \}
```

(this will get you out of any "regular" empty room)

Go to the store; if they have 2\% lactose free chocolate milk, then get me a carton.

## Misusing Power

go to the store;
if [[ they have 2\% lactose free chocolate milk]]
then

> get me a carton.

The Magic Genie Recursion, trees, and "crowdsourcing"
\{0) Start with "Is it Batman"?\}

1) Ask my (yes/no) questions down the tree
2) If win, "yay"
3) If lose, add/replace new last question to one for which my guess was wrong and her guess was right (optionally, try to be general or"half-y"?)

- repeat until genius


## The Magic Genie (can be used for evil too...)

- What about instead of
"Is your person a DC character?"
you ask real questions about real people?
(more on this later, but this demonstrates why surveillance is easy and anonymity is hard.)


## Computers don't do "magic"

## Not even A.I.

They just take data and mess around with it. e.g our "Magic Genie..

## Remember:

- We're presuming that:
-     - the local computer is fast
- and
- File storage is costly
-     - the network is slow
=> thus, small filesizes are better.


## File Compression (e.g. ZIP)

- Lossless (or "Perfect") file compression.

To make smaller, so as to be able to store more, or send faster.

But also, to reproduce PERFECTLY.
(it's not magic)

## Consider:

- "Penelope and Robert Jones Smith went to the car and grabbed the bat and the ball and the chair and the Doritos and Penelope's coat and Robert Jones Smith's favorite suit and the directions to the park. Penelope told Robert Jones Smith that they and the other people were going to have a wonderful and fun and lovely day. Robert Jones Smith told Penelope that he agreed. Also, that her name started with P."
(413)


## First step:

- Turn all the ands into \&.
- Penelope \& Robert Jones Smith went to the car \& grabbed the bat \& the ball \& the chair \& the Doritos \& Penelope's coat \& Robert Jones Smith's favorite suit \& the directions to the park. Penelope told Robert Jones Smith that they \& the other people were going to have a wonderful \& fun \& lovely day. Robert Jones Smith told Penelope that he agreed. Also, that her name started with $P$.
\& $\rightarrow$ and
(394)


## But this doesn't have to read like English...

Penelope \& Robert Jones Smith went to \# car \& grabbed \# bat \& \# ball \& \# chair \& \# Doritos \& Penelope's coat \& Robert Jones Smith's favorite suit \& \# directions to \# park. Penelope told Robert Jones Smith that \#y \& \# o\#r people were going to have a wonderful \& fun \& lovely day. Robert Jones Smith told Penelope that he agreed. Also, that her name started with P.
\&->and
\# $\rightarrow$ the
(379)

## Robert has a long name...

Penelope \& RJS went to \# car \& grabbed \# bat \& \# ball \& \# chair \& \# Doritos \& Penelope's coat \& RJS's favorite suit \& \# directions to \# park. Penelope told RJS that \#y \& \# o\#r people were going to have a wonderful \& fun \& lovely day. RJS told Penelope that he agreed. Also, that her name started with P.
\&->and
\#->the
RJS->Robert Jones Smith
(344)

## Penelope too, but wait...

- "Also, that her name started with P"


## Penelope too, but wait...

- "Also, that her name started with P"
"Also, that her name started with Penelope"


## So, then...

- P \& RJS went to \# car \& grabbed \# bat \& \# ball \& \# chair \& \# Doritos \& P's coat \& RJS's favorite suit \& \# directions to \# park. P told RJS that \#y \& \# o\#r people were going to have a wonderful \& fun \& lovely day. RJS told $P$ that he agreed. Also, that her name started with PP.
\&->and
- \#->the
- RJS->Robert Jones Smith
- $\mathrm{P} \rightarrow$ Penelope \PP $\rightarrow P$
(337)


## But what if the next file is?

"My favorite letter? P
P all day
Man I love the $P$
Gonna say it a bunch
PPPPPPPPPPPPPPPPPP P"
110

## But what if the next file is?

"My favorite letter? PP
PP all day
Man I love \# PP
PP \& PP \& PP
Gonna say it a bunch
PP PP PP PP PP PP PP PP PP PP PP PP PP PP PP PP PP PP PP PP"
(140)

## Again, there is no magic

There is NO LOSSLESS compression that makes EVERY example smaller;

All lossless compression has at least one "file" that will get BIGGER.

## Pictures on the web:

- Basically, two formats:

Raster and Vector.
Most of the time you're looking at RASTER, which is basically a grid*
*kinda

## Pictures on the web:



## Lossless (RAW) Raster

©

R $93 \%$
G $93 \%$
B $93 \%$

G $35 \%$

| R | $90 \%$ |
| :--- | ---: |
| G | $90 \%$ |
| B | $0 \%$ |

## But this math ain't mathin

1920×1080 = 2 million
2 million X 8 colors $=2 \mathrm{MB}$ (megabytes) ?
$3480 \times 2160=8$ million
only 8 colors would be 8 megabytes
but we don't use 8 colors
WE USE 16 MILLION?
\#ABCDEF

## Vector Graphics

- Use math (specifically, geometry) to tell the computer how to draw the lines and do the colors - then render it somehow



## Pictures on the web:



## Pictures on the web:



## Lossless (RAW) Raster

©

R $93 \%$
G $93 \%$
B $93 \%$

G $35 \%$

| R | $90 \%$ |
| :--- | ---: |
| G | $90 \%$ |
| B | $0 \%$ |

## Lossless image compression

- Instead of
" $0,0=$ white $0,5=$ white
- $0,1=$ white $0,6=$ white
- $0,2=$ white $0,7=$ white
- $0,3=$ white $0,8=$ white
- $0,4=$ white $0,9=$ white".. .


## Just say..

- 0,0 through $0,10=$ white.


## This also works for audio...

1 sec. = silence

- 2 sec. = silence
- 3 sec. $=$ silence....etc
just say
$1-10$ sec. $=$ silence


## ..and video

"at 0,0 for 10 seconds $=$ white"

## But without getting into detail

- This math still ain't mathin

Let's take advantage of the fact that we are analogue:

## Lossy Compression

EVEN better, for us humans.
We can't see or hear tiny differences, so this is very good for e.g. pictures/video/sound.

## Lossy compression = Good for analogue/multimedia

Tiny filesizes; takes advantage of the limitations in human perception.

We cannot percieve individual pixels (or samples) when surrounded by other meaningful pixels/samples. They blur, frequently in predictable patterns.

## "Goldilocks and the 3 bears" (with lossy compression)

- Some nosey girl went up in some bears' house, ate their food, fell asleep and freaked out when the bears came back.
- "if 0,0 is white, 0,1 is likely to be white, or light gray" "if 0,1 is light gray, 0,2 is likely to be slightly darker gray..."

Okay, so just ( $x+1$ ) the darkness every pixel to the left.
$y=x+1$

- What happens if you compress something losslessly over and over?

What happens if you compress something lossi-ly over and over?


## Can also "fix" images "denoising"



## Smart people terms

How to store knowledge:
I'm going to put the number PI on the next slide.
the whole thing
Seriously

## Smart people terms

"The ratio of a circle's circumference to its diameter"

BOOM.

## Kolmogorov Complexity

Basically "what do you need to completely describe the thing."

Stored as a numeral, Pi, is infinitely 'complex'
But it's not "Kolmogorov Complex" at all.

## Lets try something

1. Get every image on the internet that has words.
2. Deliberately add noise to them and tell the computer to watch for math patterns


## Now do that backwards

3. Take words, e.g. "womans face," then "denoise" according to what you learned by mathematically analyzing

## (this is Al art)



## Okay, AI

## First, lets try words.

## This is actually not hard...

## You could literally just do random words..but that's not great...

## Fill in the blank.

## Seven ___?

## Fill in the blank.

## Seven ___?

## Seas? <br> Continents? <br> Nation Army?

Sure, pick one, sometimes randomly.

## Fill in the blank.

## Four Score and Seven ___?

## Again, not hard.

The computer just picks the one (or one of the ones) that usually follows FROM THE DATA IT SEES.

## By the way

## Remember "the" and \& et al from before?

The text Al's mostly don't use words, but "tokens"

## Not too different from our magic genie

Just a series of choices/options.

## Al

## Lots of moved goalposts <br> - chess <br> - basic questions TURING COMPLETENESS

## Big Picture Ideas

## Some have moved from "Al" "AGI"

"Artificial General Intelligence"

## Big Picture Ideas

"The Singularity"
The point at which Al intelligence surpasses ours, and therefore there's literally no point in trying to think about after that...



# Literally just my opinion, but: meh. 

After YEARS of doing
"todo/ideas/2nd brain stuff"
I've landed on some conclusions:

# We can think; the computers can just "play back" what we've done. 

## That's ALL.



## Looks cool. Appealing.

## but IMHO, our brains are WAY better at this part. Kinda pointless.

The computers ARE good at "repeating verbatim," tho

