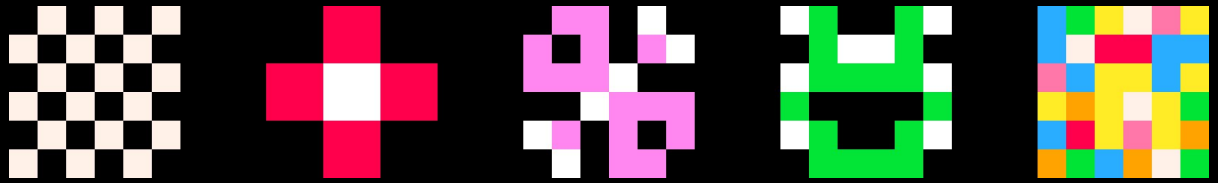
Why Oatmeal is Cheap

A Fundamental Theorem for Procedural Generators

$$|G^*| \geq K^*(G) - p(G) \geq 0$$



by Younès Rabii
@pyrofoux

Younès Rabii
Indie gamedev
PhD Student at QMUL
IGGI programme

Michael Cook
Senior Lecturer
at King's College London

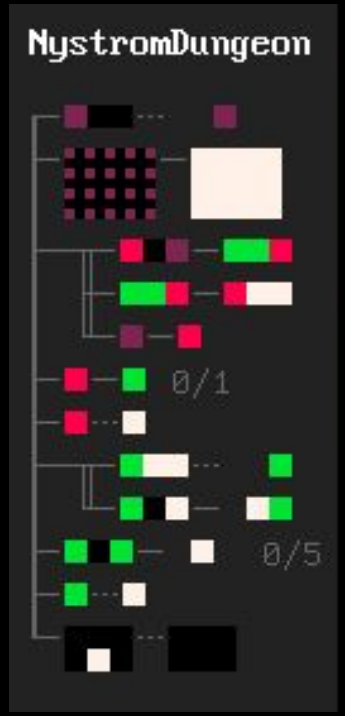
Queen Mary University of London

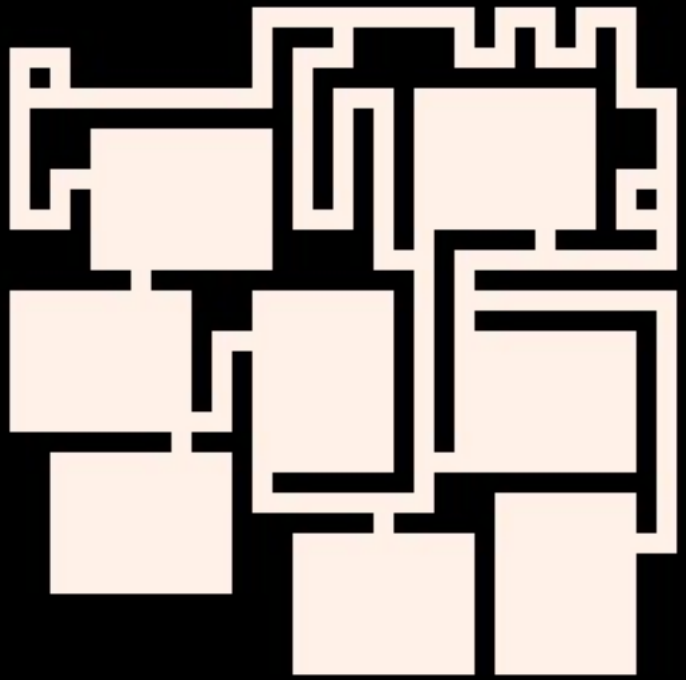
KING'S College LONDON

Procedural Content Generation

and Oatmeal

Procedures that make something





MarkovJunior

BAKE THE PERFECT COOKIE

What's the difference between crisp and cakey chocolate chip cookies? Just a slight modification to the amount of butter and sugar.

THIN & CRISP	SOFT & CHEWY	LIGHT & CAKEY
2½ sticks unsalted butter, softened	2 sticks unsalted butter, softened	1¾ sticks unsalted butter, softened
1¼ cups granulated sugar	½ cup granulated sugar	¾ cup granulated sugar
¾ cup brown sugar	1 cup brown sugar	¼ cup brown sugar

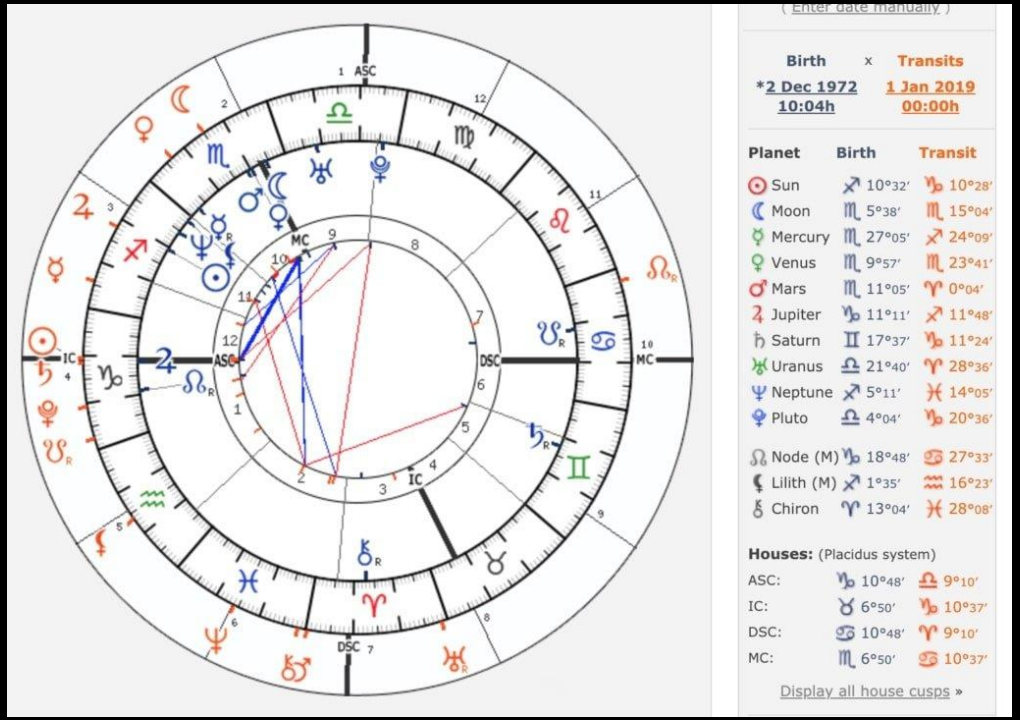
PLUS ...

- 2¼ cups all-purpose flour
- ½ teaspoon baking soda
- 1 teaspoon coarse salt
- 2 teaspoons pure vanilla extract
- 2 large eggs, room temperature
- 2 cups (about 12 ounces) chocolate chips

1. Preheat oven to 350 degrees. In a bowl, whisk together flour and baking soda. In another bowl, beat butter with both sugars until fluffy. Beat in salt, vanilla and eggs until well combined. Mix in flour mixture until just combined. Stir in chips.
2. Drop tablespoon-size balls of dough 2 inches apart onto parchment-lined baking sheets.
3. Bake until cookies are golden around edges but still soft in center, 8 to 10 minutes. Transfer to a wire rack; let cool.

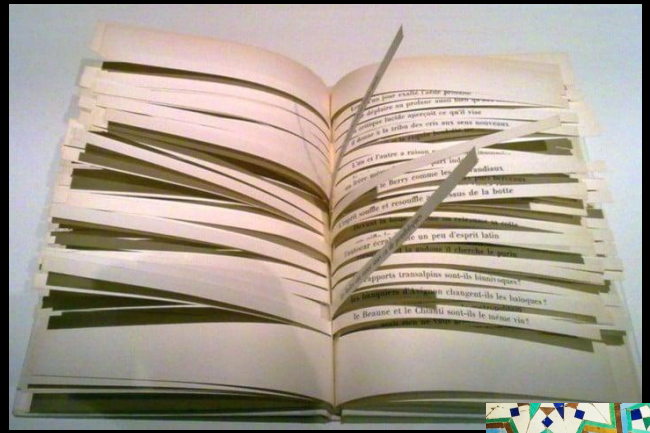
Karen's Cooking Korner

Cooking recipes

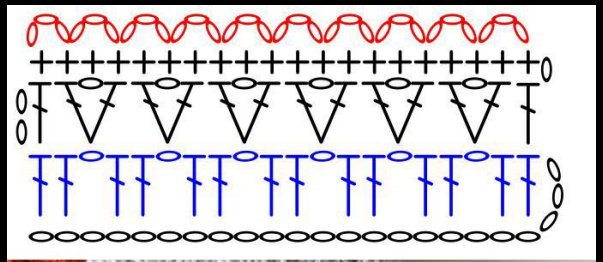


Astrology fortunes

Cent mille milliards de poèmes– Raymond Queneau



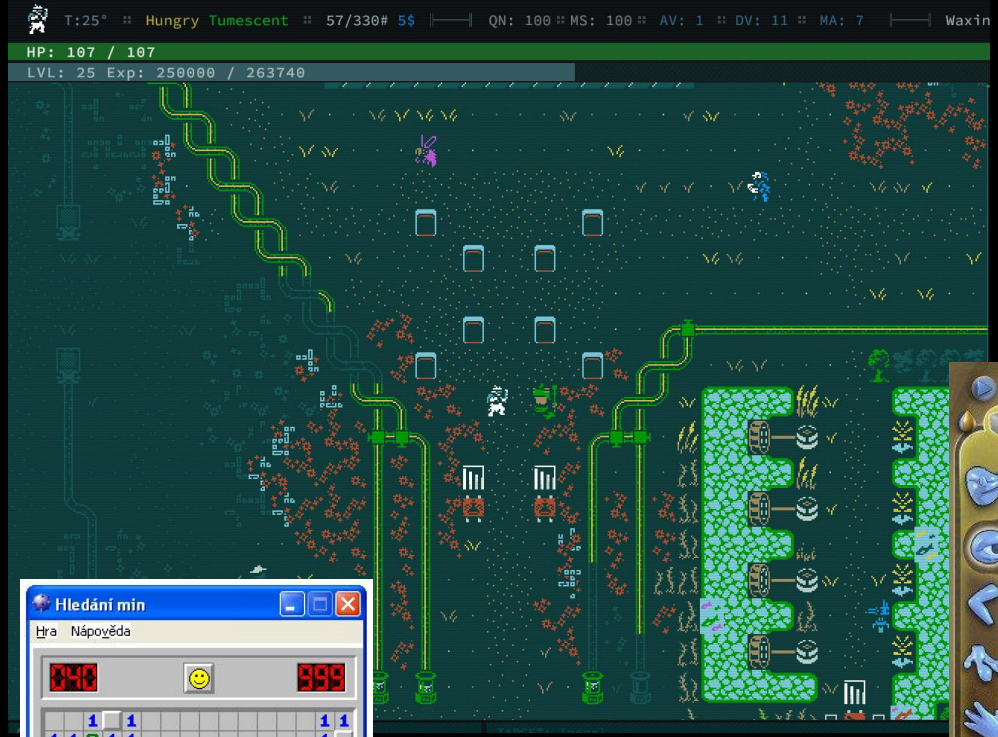
Poetry



Crochet

Mosaïcs

Moroccan Zellij

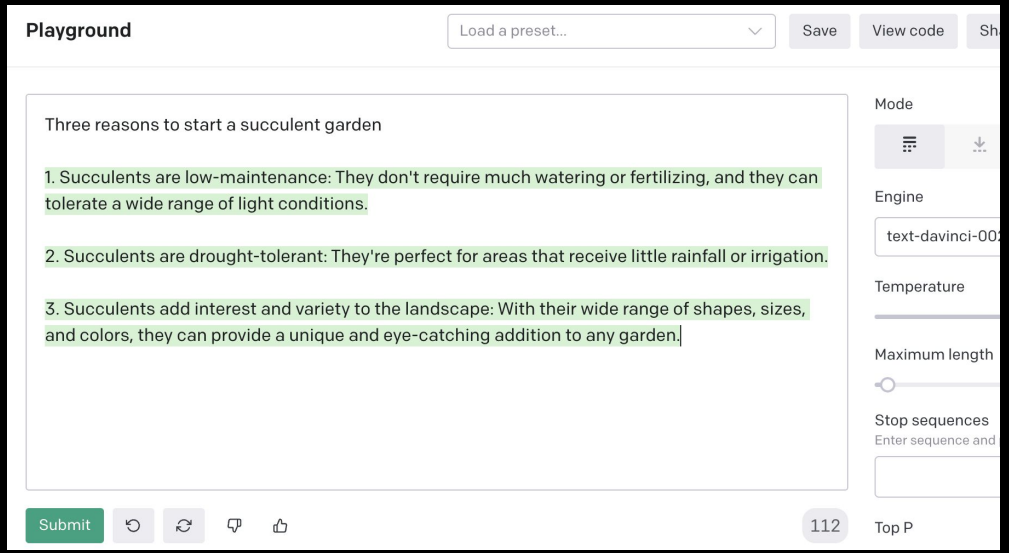

T:25° :: Hungry Tumescent :: 57/330# 5$ QN: 100 MS: 100 AV: 1 DV: 11 MA: 7 Waxin
HP: 107 / 107
LVL: 25 Exp: 250000 / 263740

Strength
Speed
Stealth
Herbivore
Carnivore
150

Hledání min
Hra Nápověda

Games

StableDiffusion

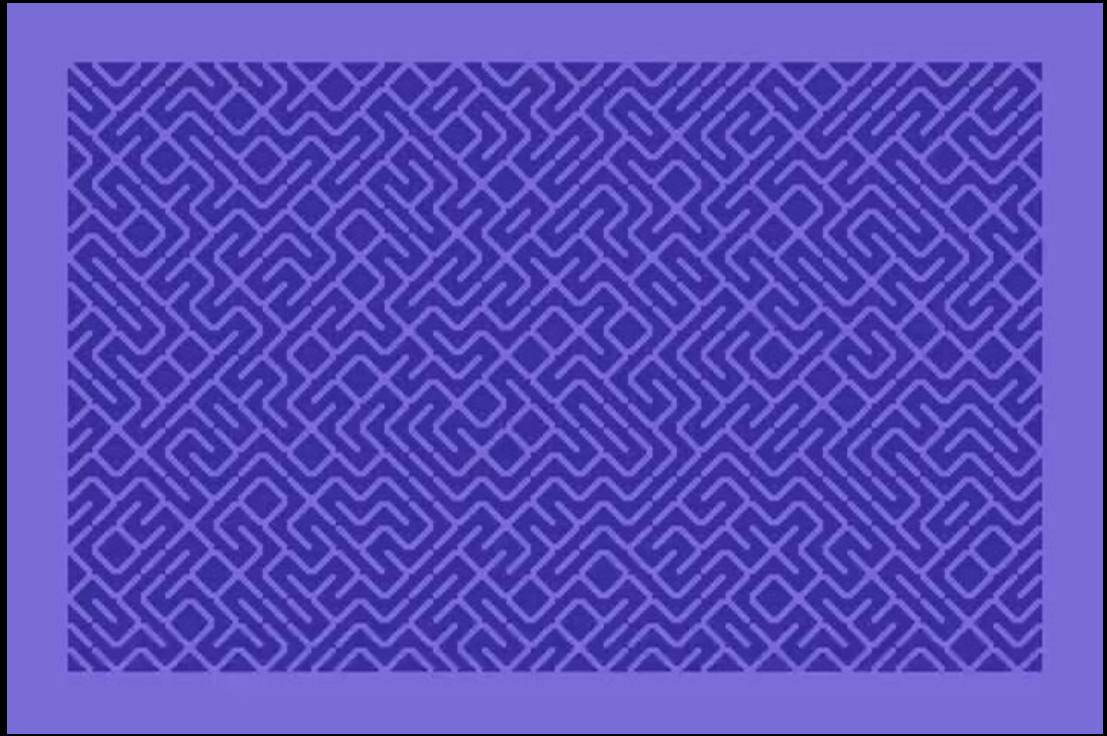


GPT-3

SIMPLE GENERATORS

```
10 PRINT CHR$(205.5+RND(1)); : GOTO 10
```

INFINITE OUTPUT



“There’s a trade-off between the breadth of content you can make in one of these tools, and the depth, the quality of that content.”

Chaim Gingold - GDC 2007

Orteil
@Orteil42

Follow

thanks to procedural generation, I can produce twice the content in double the time

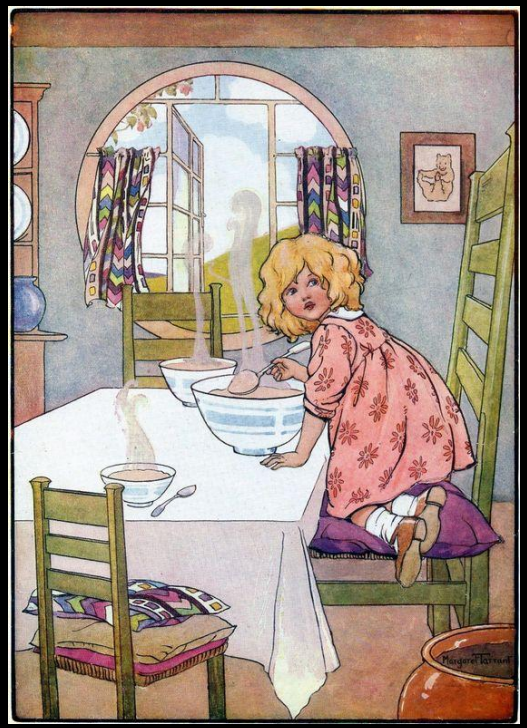
1:10 PM - 25 Nov 2016

The 10.000 bowls of Oatmeal problem

These bowls of oatmeal are all technically unique…

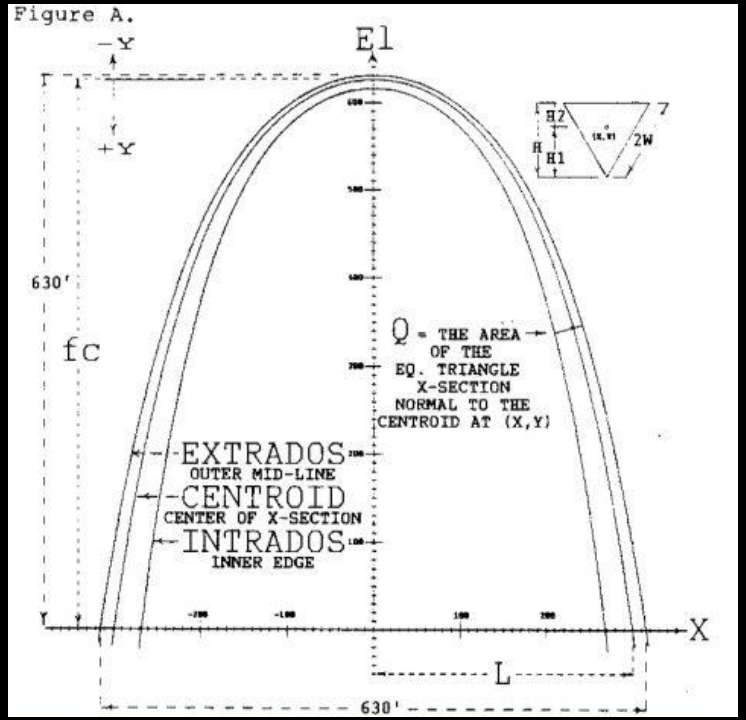
…but they taste exactly the same.

Margaret W Tarrant – Illustration of Goldilocks



(sometime Oatmeal is fine)

Fundamental theory as a bridge



PCG is extensively studied empirically, but rarely on a fundamental mathematical level

I. 3 mathematical properties of generators
→ 3 intuitive concepts used by designers

II. a fundamental theorem made of 3 inequalities
→ insight about the challenges of designing generators

Definitions

What we're studying: Ideal Generators

an **ideal** generator G has these properties:

- ★ **Terminability** : It always outputs something
- ★ **Fixed Input Size**: It accepts input of a specific length
- ★ **Injectivity** : Different inputs give different outputs

Inputs & Outputs are binary strings

`0110010` → `011001001101010`

We can take a non-ideal generator and turn into one that is ideal.

1. Length

```
--set palette
pal({1,8,3,9,14,4})
cls()
function _update()
  local x = rnd(128)
  local y = rnd(128)
  local e = t()
  for c=-2,1 do for r=-1,1 do
  local a=64+c*54+r%2*27
  local b=64+r*48
  line()
  -- draw hexagons
  for i=0,1,1/6 do
    line(a+sin(i)*30,b+cos(i)*30,7)
  end
  -- draw planets, bit by bit
  if(sqrt((x-a)^2+(y-b)^2) < 15+sin(c/7)*4+sin(r/2)*4)
    pset(x,y,2+c+r*2+sin(x/81+e)*2+cos(y/(14*c)))
  end
  end
end
```

(206 bytes)

|G| is the length of G's source code

```
pal({1,8,3,9,14,4})cls()s=sin::_::
x=rnd(128)y=rnd(128)e=t() for c=-2,1
do for r=-1,1do a=64+c*54+r%2*27b=64
+r*48line()for i=0,1,1/6do line(a+s(i)
*30,b+cos(i)*30,7)end if sqrt((x-a)^2+
(y-b)^2)<15+s(c/7)*4+s(r/2)*4) pset(x,y
,2+c+r*2+s(x/81+e)*2+cos(y/(14*c)))end
end goto _--
```

(140 bytes)

written as shortly as possible

→ Encoded Knowledge

Knowledge…

- Learn how to compose artefacts
- Specify a precise procedure to make them

… you have to Encode

- Implement the procedure
- Test it
- Debug it
- Document it
- Optimise it

It’s costly.

2. Size of Possibility Space

Tea Garden's sprite generator

$P(G)$ is the number of unique artefacts in G's possibility space.

$$P(G) = 100$$

$p(G)$ is the $\log_2$ of that number.

$$p(G) = \log_2(100) \approx 6.64$$

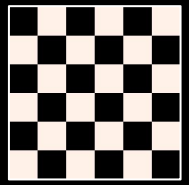
→ Scale

Borderlands 3

Animal Crossing: New Horizons

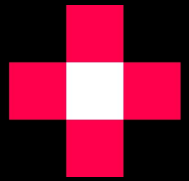
Picrew by @ ASTROLAVAS

3. Kolmogorov Complexity (simplified)



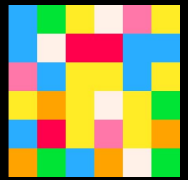
"alternate black and **white**"

4 words



"a 9x9 red cross overlaid with a 3x3 **white square** in the center"

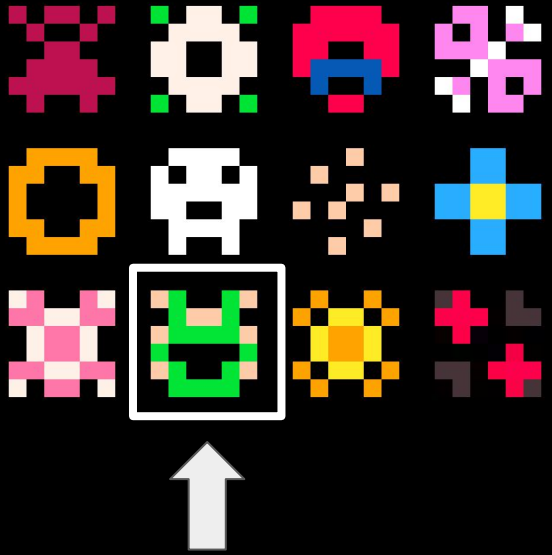
13 words



"**blue** then **green** then **yellow** then **white** then **pink** then **yellow** then **blue** then **white** then **two red** then **two blue** then…"

72 words

3. Kolmogorov Complexity (simplified)



Most complex artefact

$K(A)$ is the complexity of an artefact A

$K^*(G)$ is the complexity of the most complex artefact of G

$\bar{K}(G)$ is the complexity of the most complex artefact of G

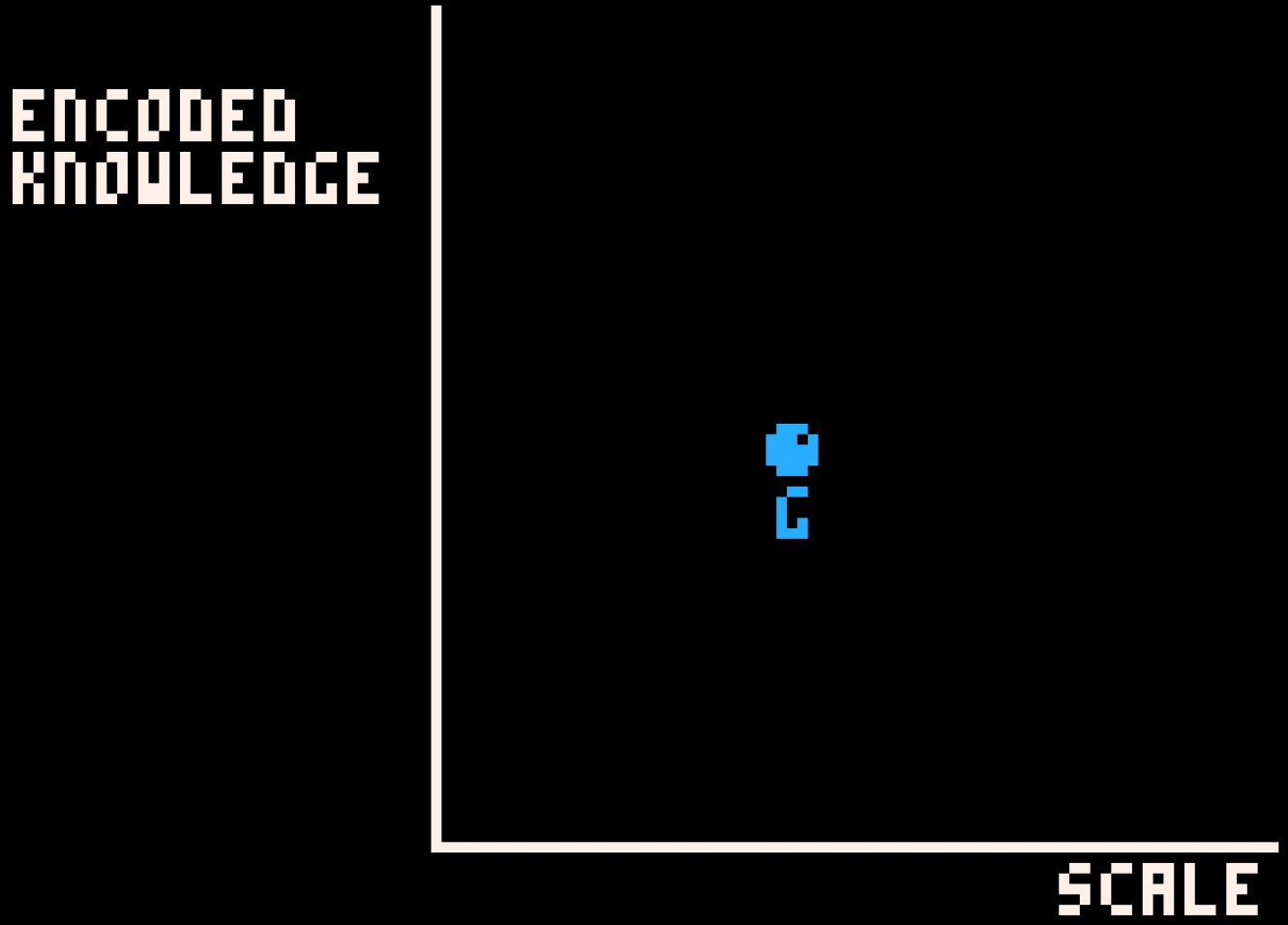
→ Pattern Density

Low K-Complexity

- Repetitive
- Patterns are easy to spot

High K-Complexity

- Noisy
- No spottable patterns


ENCODED
KNOWLEDGE
SCALE

Inequality 1 / 3

$$K^*(G) \geq p(G)$$

$A_0, A_1, A_2 \ldots A_{\square}$ are the artefacts in G's possibility space

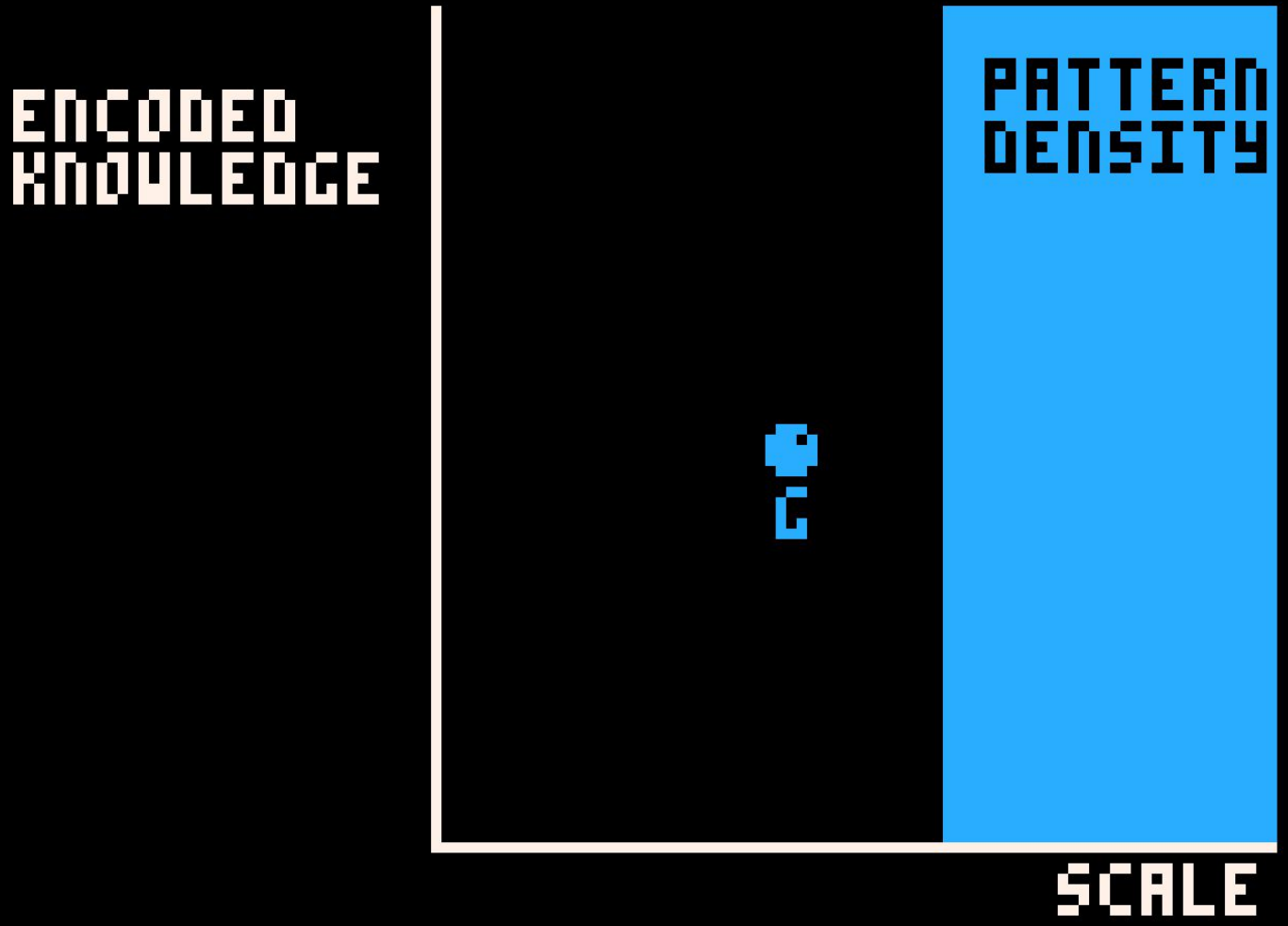
$P_0, P_1, P_2 \ldots P_{\square}$ are the shortest programs that output $A_0, A_1, A_2 \ldots A_{\square}$

P_0	0
P_1	1
P_2	00
P_3	01
…	…
$P_{\square}$	10110

$$|P_{\square}| \geq \ln_2(n)$$

$$K(P_{\square}) \geq \ln_2 P(G)$$

$$K^*(G) \geq p(G)$$


ENCODED
KNOWLEDGE
PATTERN
DENSITY
SCALE

Inequality 2 / 3

$$|G| + p(G) \geq K^*(G)$$

A^* is the most complex artefact in G's possibility space

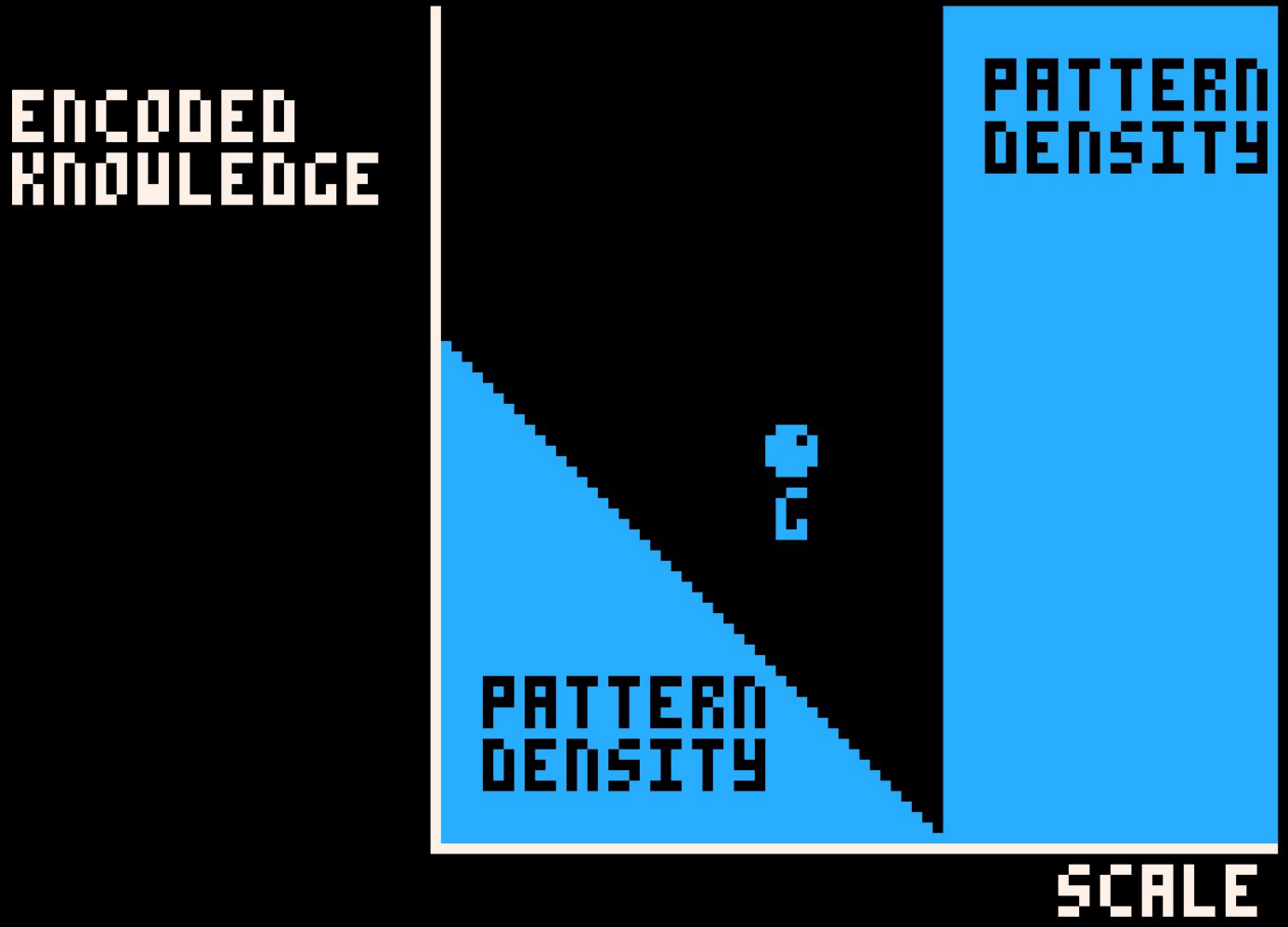
$$|\text{any procedure that outputs } A^*| \geq K(A^*)$$

$$|G(A^* \text{ seed})| \geq K(A^*)$$

$$|G| + |A^* \text{ seed}| \geq K(A^*)$$

$$|G| + p(G) \geq K(A^*)$$

$$|G| + p(G) \geq K^*(G)$$


ENCODED
KNOWLEDGE
PATTERN
DENSITY
PATTERN
DENSITY
SCALE

Inequality 3 / 3

$$[???] \geq |G|$$

$$C_1 + P(G) \cdot C_0 + \sum |Pi|$$

$$C_1 + P(G) \cdot C_0 + P(G) \cdot \overline{K}(G)$$

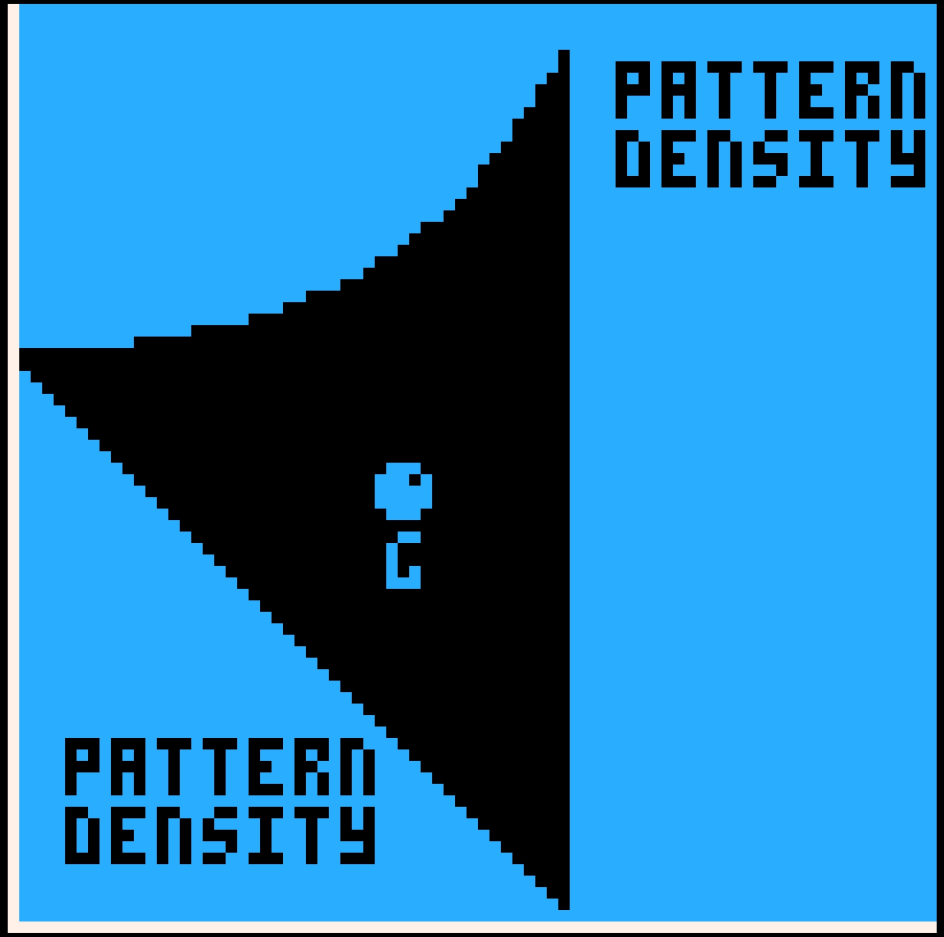
$$C_1 + P(G) \cdot (C_0 + \overline{K}(G)) \geq |G|$$

A_0, A_1, A_2 ... $A_□$ are the artefacts in G's possibility space

P_0, P_1, P_2 ... $P_□$ are the shortest programs that output A_0, A_1, A_2 ... $A_□$

```
def G(seed):
    p = [P0, P1, P2, ... P□]
    return p[seed]
```

ENCODED KNOWLEDGE



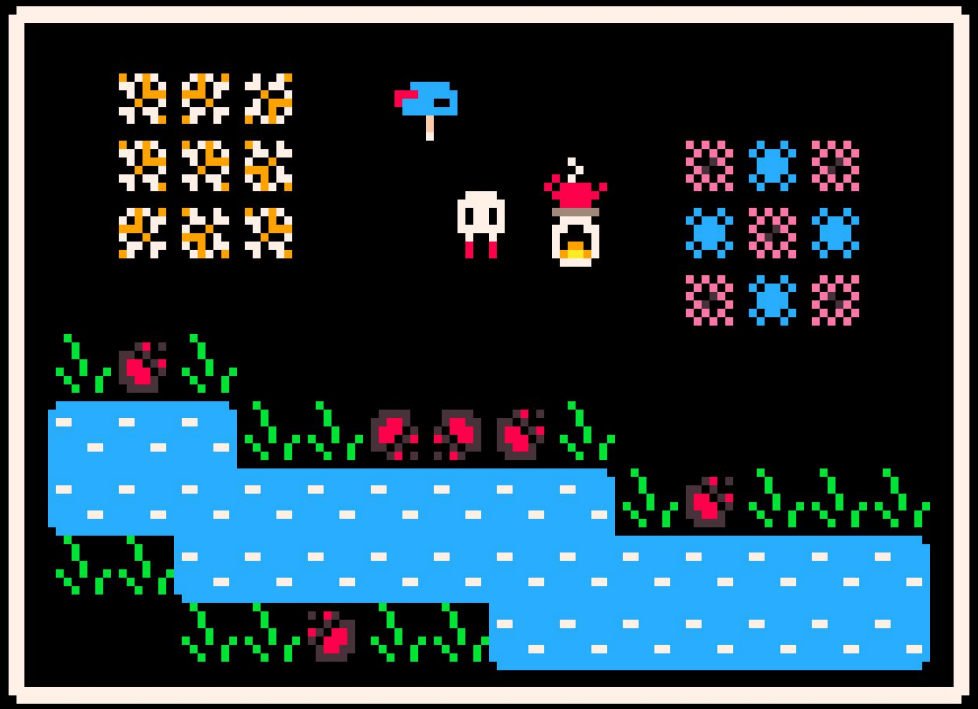
Trade-offs

When designing generators

Minecraft’s village update

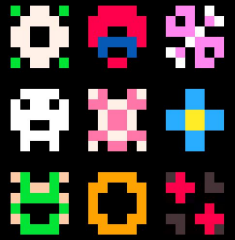
Goals

- Add a new **pattern**
- Can’t change **scale**



Tea Garden's sprite generator

12x12



6x6

Borderlands 3

Goals

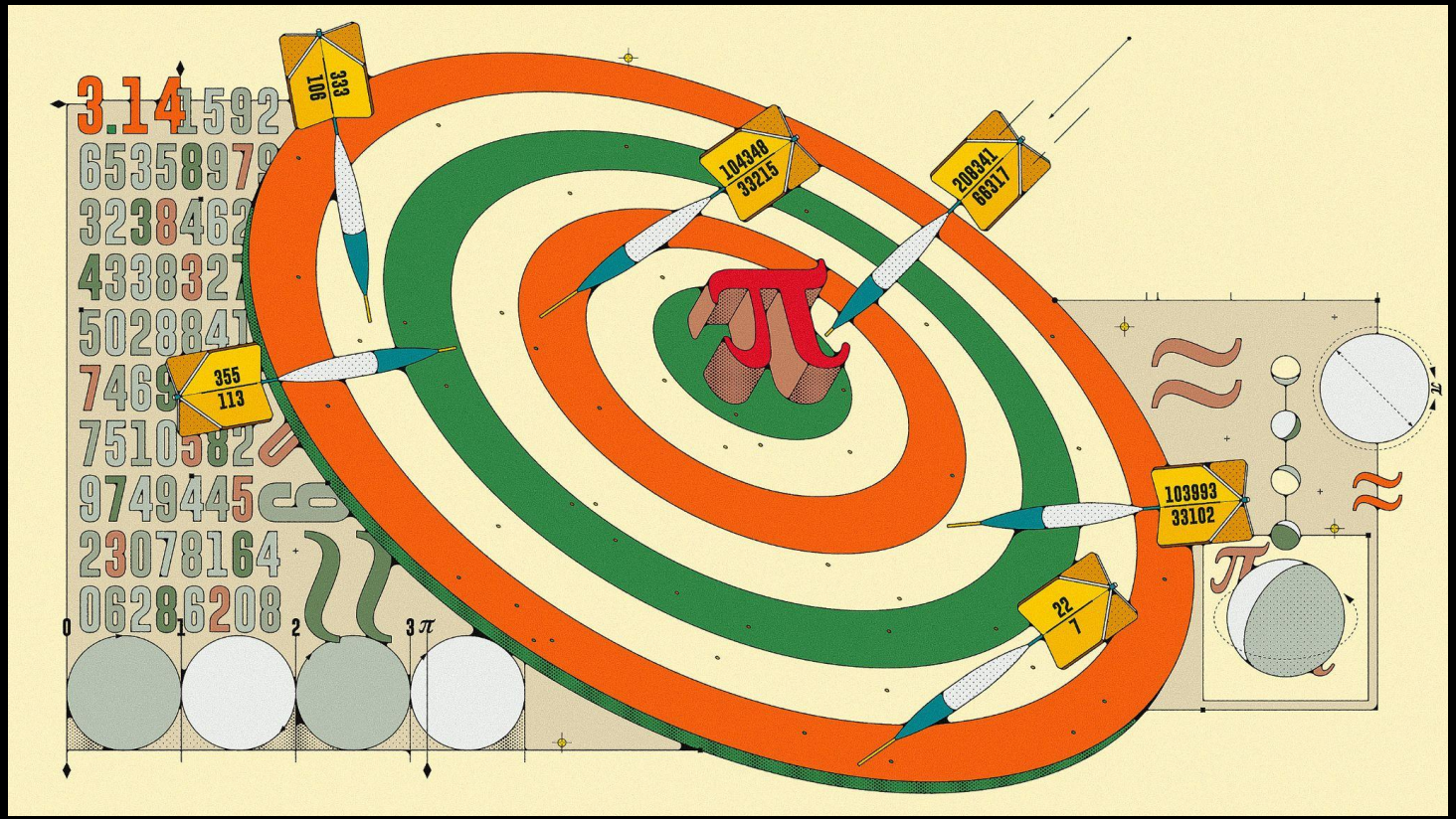
- Maximize scale
- Minimize costs

Forbes

FORBES > INNOVATION > GAMES

It Takes 5 Billion Years To Visit Each Planet In 'No Man's Sky'

Approximations



How many different pictures can Stable Diffusion generate?

How many different answers can ChatGPT generate?

How many different articles can Wikipedia generate?

saint louis

WIKIPEDIA
The Free Encyclopedia

Create account Log in

St. Louis

113 languages

Article Talk Read View source View history Tools

From Wikipedia, the free encyclopedia

Coordinates: 38°37′38″N 90°11′52″W

This article is about the city in Missouri, United States. For other uses, see St. Louis (disambiguation).

St. Louis (/seɪnt ˈluːɪs, sənt ˈluːɪs/)[10] is the second-largest city in Missouri, United States. It is located near the confluence of the Mississippi and the Missouri rivers. In 2020, the city proper had a population of 301,578,[8] while its bi-state metropolitan area, which extends into Illinois, had an estimated population of over 2.8 million. It is the largest metropolitan area in Missouri and the second largest in Illinois.

Before European settlement, the area had been occupied for thousands of years by various Native American cultures. From roughly 900 to 1500 CE, it was a regional center of Mississippian culture, based in Cahokia east of the river, and extending across the continent along the Mississippi and its tributaries.

St. Louis

Independent city

Downtown St. Louis and the Old Courthouse (St. Louis) and Gateway Arch

P(G) = Number of articles

K(G) = Average compressed size of an article

|G| = Size of Wikipedia's source code, compressed

$$C_1 + P(G) \cdot (C_0 + K^*(G)) \geq |G|$$

$$P(G) \cdot (K^*(G)) \geq |G|$$

$$P(G) \cdot (K^*(G)) \approx |G|$$

Estimation: 1.9e7 pages * 2.5 Kb = 47 Gb

Size of Wikipedia's archive : 16 Gb

P(G) = Number of different pictures

K(G) = Average compressed size of a picture

|G| = Size of StableDiffusion source code, compressed

$$C_1 + P(G) \cdot (C_0 + K^*(G)) \geq |G|$$

$$P(G) \cdot (K^*(G)) \geq |G|$$

$$P(G) \cdot (K^*(G)) \approx |G|$$

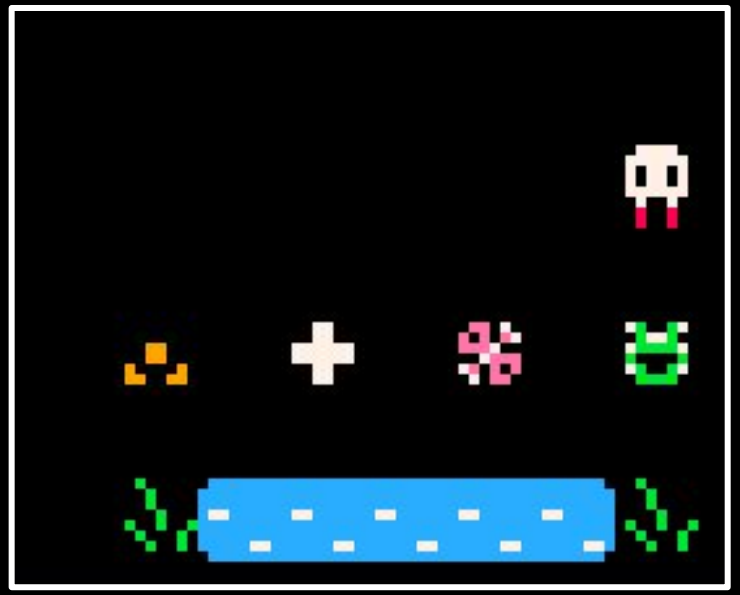
Estimation: 10 Gb / 1.26 Mb = 7900 different pictures

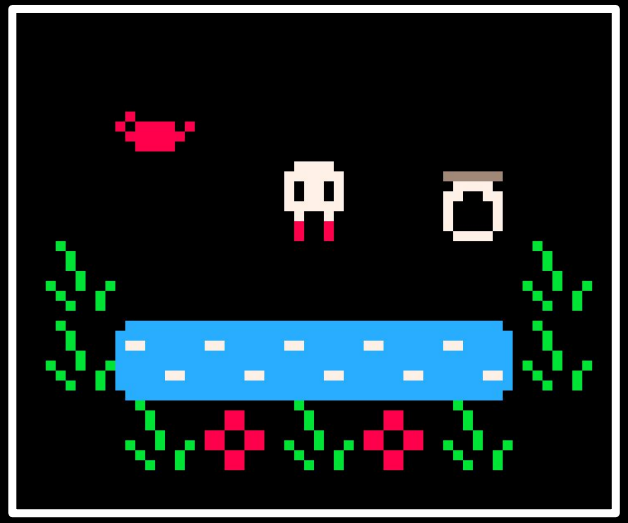
Thank you!

Paper + Demo
pyrofoux.github.io/why-oatmeal-is-cheap

@pyrofoux

yrabii.eggs@gmail.com





THANK
YOU