

TOKYO MACHINE LEARNING GYM

OCT 21 2017 HACKATHON OVERVIEW

PLAN

- introduction
- hackathon idea
 - small datasets
- little tutorial
 - chainer jumpstart
 - simple neural net architectures for images
- hacking!

SOME LOGISTICS

- 📲: MatsuWaveG / kawamura529
- 🛁: on 2F
- 🍕: shall we collect~ 500-1000¥ and order pizzas 4 lunch?
- **please refrain from entering other rooms**
- can leave any time

ROUND OF SELF-INTRODUCTION?

YOUR HOST: ALEKSANDR DROZD

- researcher at Tokyo Tech
- working on convergence of HPC and AI
 - disclaimer: I'm not a deep learning guru
 - although my particular interest is in NLP

SOME CONTACTS

-  <http://blackbird.pw>
-  <http://nightwind.in>

MY AI STORY

childish fascination with AGI



Good old fashioned NLP



u-turn to core HPC research



using HPC for AI ^_^

WHAT'S HPC BTW

AND WHAT DOES IT HAVE TO DO WITH AI?

TSUBAME 2.5 SUPERCOMPUTER



"A supercomputer is a computer that performs at or near the currently highest operational rate for computers. "

HIGH PERFORMANCE COMPUTING

”the use of super computers and parallel processing techniques for solving complex computational problems”. - technopedia.com

HIGH PERFORMANCE COMPUTING

"practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business." -insidehpc.com

SUPERCOMPUTERS ARE USED FOR

- classical supercomputing applications:
 - climate simulation
 - material science
 - drug discovery
- data science/ big data processing/ machine learning

HPC = PARALLEL

WHERE PERFORMANCE IS COMMING FROM

- still improvements in [sequential] hardware
- smarter algorithms
- smarter implementations
- parallel execution

SOME RANKINGS: TOP 500 LIST

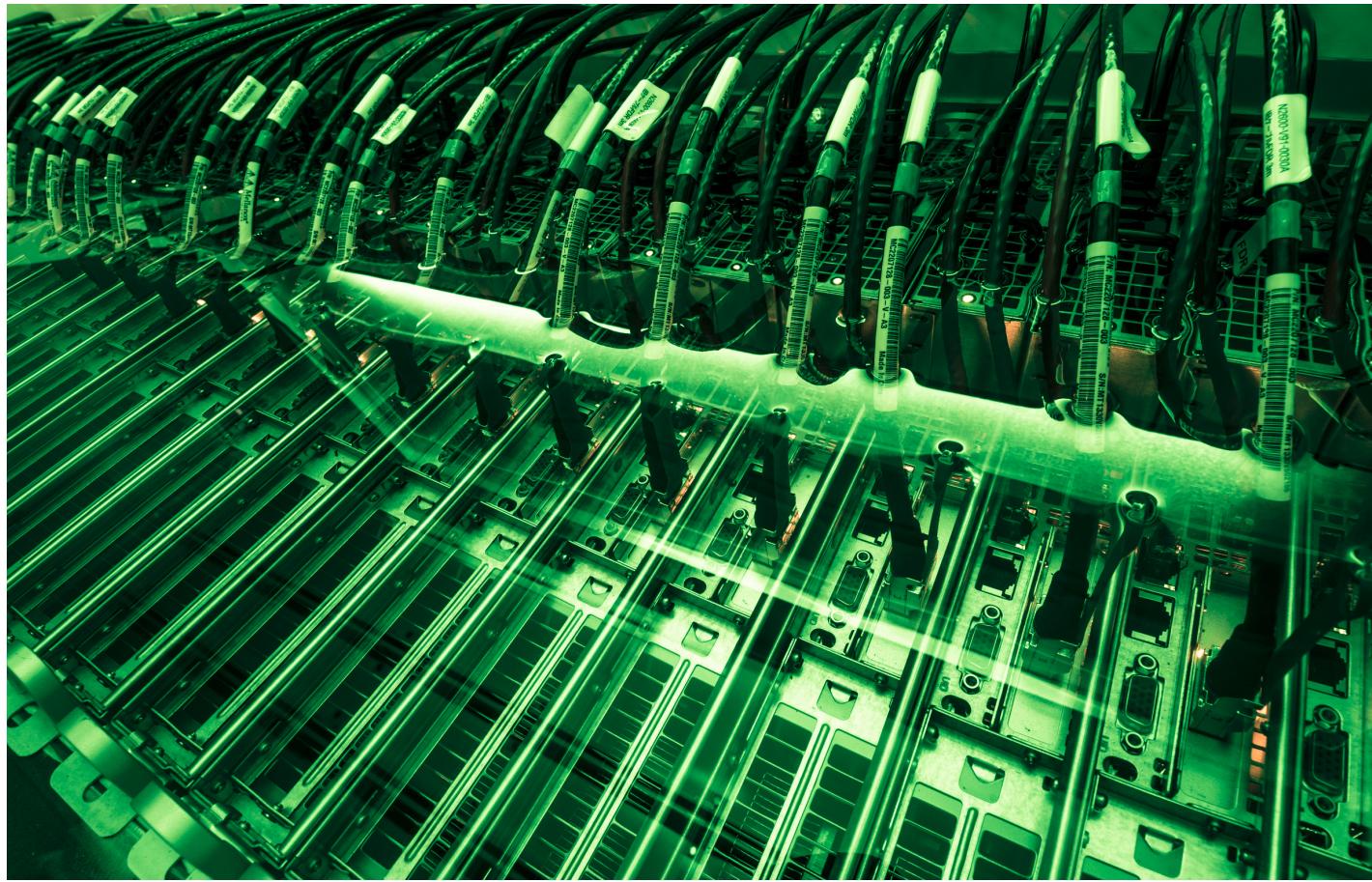
- twice a year since 1993
- high-performance LINPACK benchmark
- <http://top500.org/>
- also check out "green 500" list

SOME RANKINGS: GRAPH 500 LIST

[myth-busting]

- Breadth-First Search in Kronecker graphs
- measure Edge traversals per second (TEPS)
- No compute, all communication
- check out <http://www.graph500.org>
- not a single "cloud" platform there

SOME MORE SUPERCOMPUTER PORN



SO WHY HPC FOR AI?

.... switch to Andrew Ng's slides

"AI" SUPERCOMPUTERS

ubame 3
20 Racks
540 dual-socket nodes
2160 Nvidia P100 Pascal GPUs
Intel Omni-Path
47.2 AI (single precision)
PetaFLOPS



THIS IS NOT TO SAY BIG MODELS/DATASETS ARE BAD

QUITE THE OPPOSITE!

repeating one more time:

- **This hackathon is not about applying DL to small datasets!**
- If you solving some real-life problem and your data is small - go for kernel methods!

IMAGENET

IMAGENET LARGE SCALE VISUAL RECOGNITION CHALLENGE (ILSVRC)

- 1.2 M train images (256x256px)
- 100 K test images
- 1000 object classes

IMAGENET



IMAGENET RANT

AKA IMAGENET IS OVERFIT

In many senses:

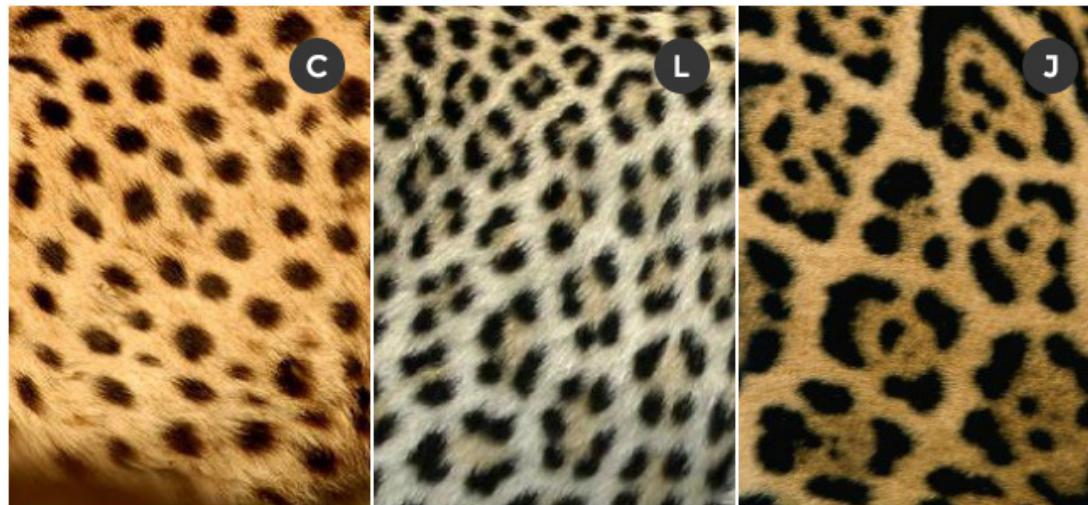
- The problem is rather dumb (no spatial reasoning required)
- Success is overhyped
- Imagenet winning models not necessarily rock in other image-processing tasks
- Basically consider as years ago solved problem by advanced AI research
- Vendors still using it to show-off certain hardware

**I T'S LIKE SAYING THAT CALCULATOR "SURPASSED HUMAN INTELLIGENCE"
BECAUSE IT CAN MULTIPLY LARGER NUMBERS %)**

WHAT'S THE DEAL ANYWAY?



MEMORIZING, NOT THINKING



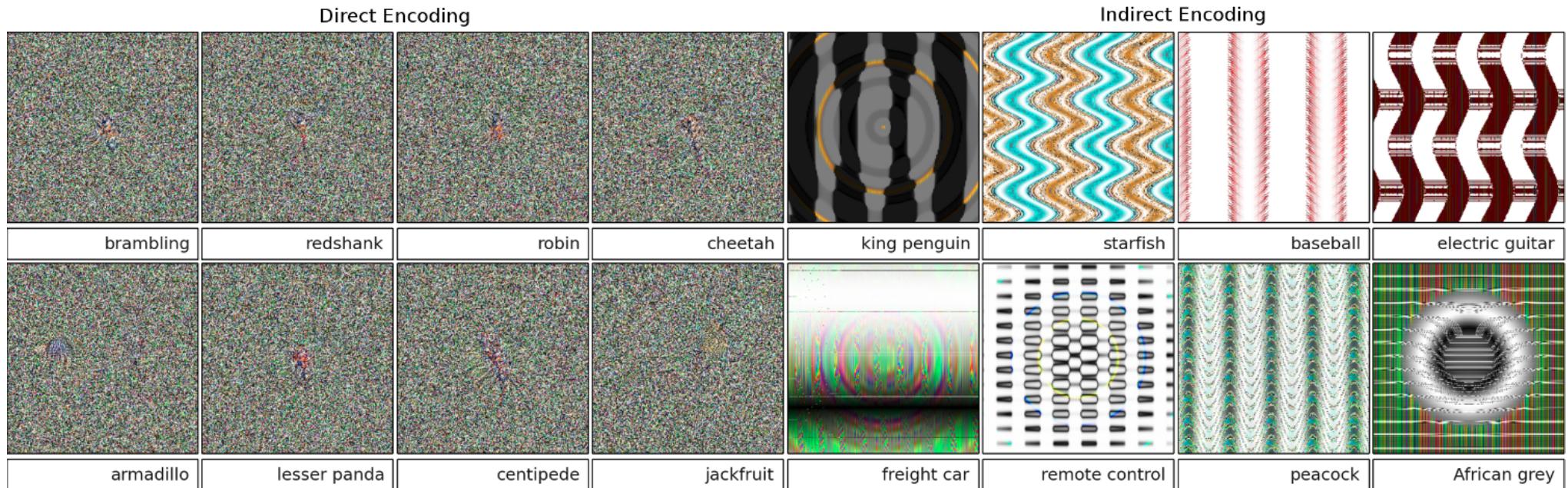
ULTIMATE SOFA TEST

let's take our best trained Imagenet-winning model....

WHAT'S THIS? JAGUAR???



ALSO ADVERSARIAL ATTACKS



see "Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images" by Nguyen et al.

THE QUESTION IS

do we need a dataset to be that big to develop good understanding?

MNIST

(MODIFIED NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY DATABASE

label = 5



label = 0



label = 4



label = 1



label = 9



label = 2



label = 1



label = 3



label = 1



label = 4



label = 3



label = 5



label = 3



label = 6



label = 1



label = 7



label = 2



label = 8



label = 6



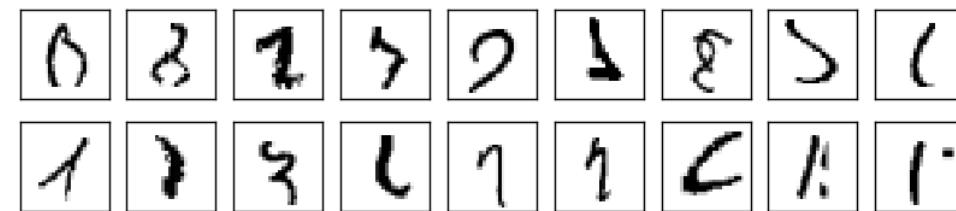
label = 9



MNIST

- handwritten digits
- 60,000 training images
- 10,000 testing images
- current lowest error 0.23 percent

NOT MUCH POINT IN FURTHER IMPROVEMENT



MNIST IS GOOD

- to get started with deep learning
- learn about, say, 2D convolutions

FINALLY...

THE IDEA FOR THIS [SERIES OF] HACKATHON[S]

HACKATHON IDEA

- find/create/modify small datasets
- which allow to illustrate different deep learning techniques
- and practice those techniques

TODAY'S FOCUS: IMAGES

- basic classification
- spatial invariance by global pooling
- image generation
 - with autoencoders
 - with GANS
- object counting

FUTURE TOPICS (YOU CAN START NOW, IF CONFIDENT)

More imaging:

- visual attention
- segmentation
- etc

also different modalities, multi-modal learning, reinforcement learning and more fun stuff.

TRYING TO MAKE EVEN SIMLER DATASET

ITERATION ONE



... maybe not the best idea

SCALE A BIT, MOVE OBJECTS AROUND



... STARTING TO LOOK LIKE SOMETHING

MAKING THE TASK MORE COMPLEX

[SAME CAN BE DONE FOR MNIST ACTUALLY]

train:



test:



OBJECT COUNTING

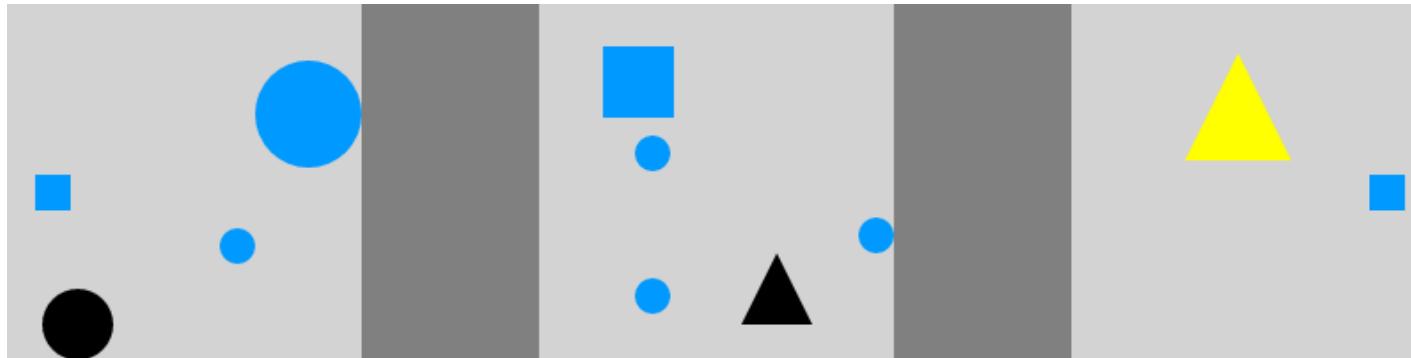


HOW FAR CAN YOU GO WITH THIS APPROACH?

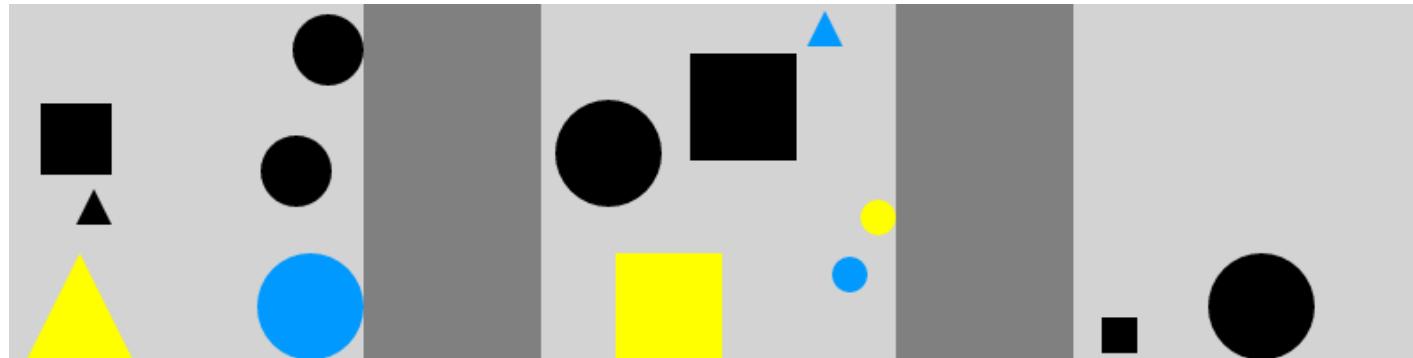
[some examples]

A CORPUS OF NATURAL LANGUAGE FOR VISUAL REASONING

SUHR ET AL, CORNELL + FACEBOOK AI



THERE IS EXACTLY ONE BLACK TRIANGLE NOT TOUCHING ANY EDGE: TRUE



THERE IS A BOX WITH MULTIPLE ITEMS AND ONLY ONE ITEM HAS A DIFFERENT COLOR: **FALSE**

NLVR:

- 92,244 pairs of natural language statements grounded in synthetic images.
- requires reasoning about sets of objects, quantities, comparisons, and spatial relations

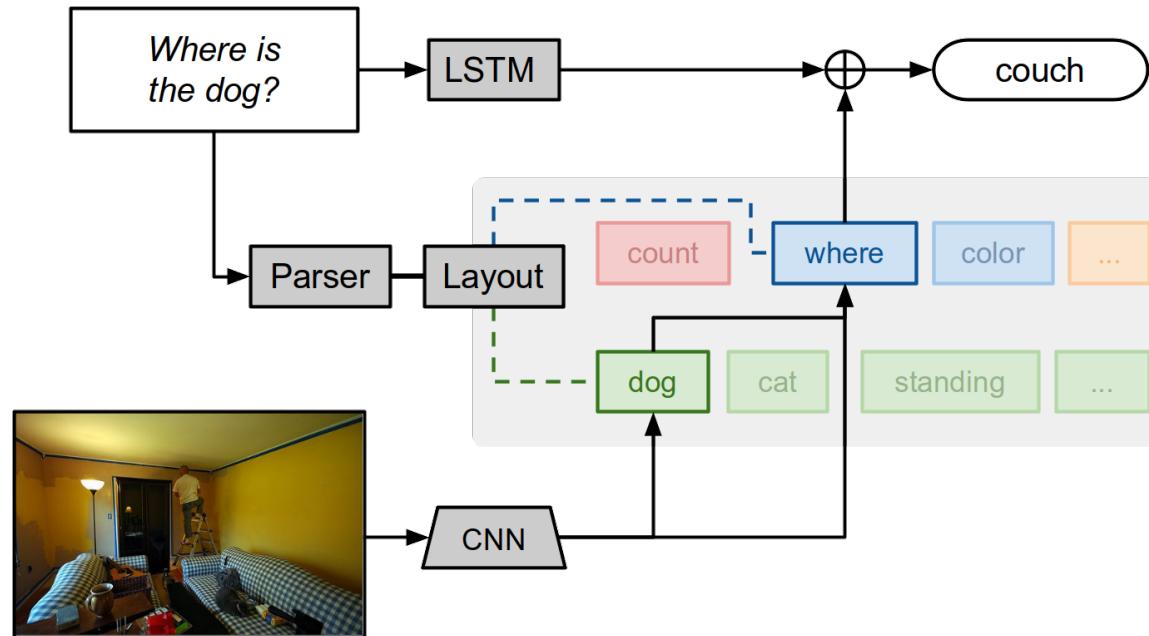
PREVIOUS WORK:

- VQA (Antol et al., 2015)
- CLEVR (Johnson et al., 2016)
- SHAPES (Andreas et al., 2016)

SOTA

The best model using images achieves an accuracy of 66.12%
using "Neural Module Networks" [Andreas et al, UC Berkeley 2017]

NEURAL MODULE NETWORKS



answer natural language questions about images using collections of jointly-trained neural “modules”, dynamically composed into deep networks based on linguistic structure.

RESOURCES AND LEADERBOARD

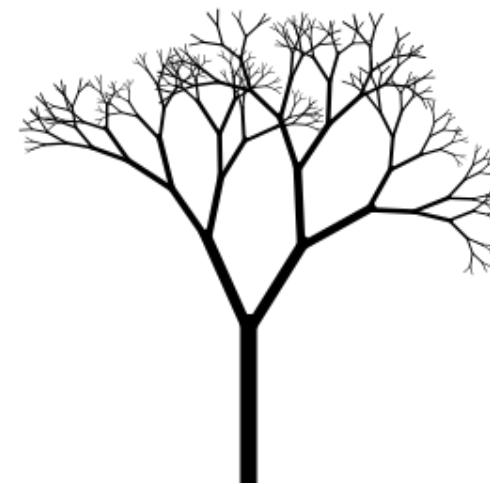
<http://lic.nlp.cornell.edu/nlvr>

HOW TO GENERATE SUCH EXAMPLES

[AKA MOMENT OF SHAME'S SELF-ADVERTIZING]

PYCONTEXTFREE

```
from contextfree.contextfree import *
@check_limits
def branch():
    line(0, 1)
    with translate(0, 0.9):
        with scale(0.7 + rnd(0.3)):
            with rotate(-0.4 + rnd(0.5)):
                branch()
            with rotate(0.4 + rnd(0.5)):
                branch()
init(canvas_size=(300, 300))
with translate(0, -1):
    with scale(0.6):
        branch()
display_ipython()
```



PYCONTEXTFREE

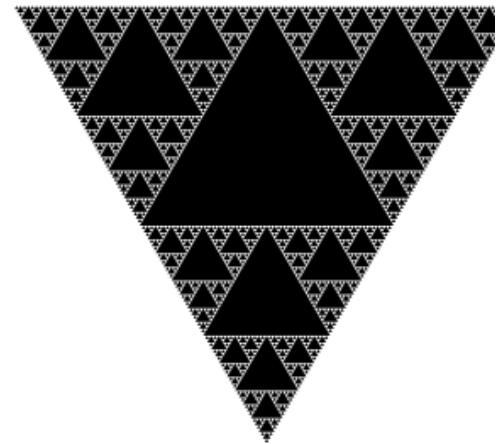
- Written by yours trully
- Inspired by CFDG
 - uses own language to define grammars
 - <https://www.contextfreeart.org>
- pip install contextfree
- <https://github.com/undertherain/pycontextfree>

MORE EXAMPLES

```
side = 1
height = math.sqrt(3) * side / 2

@check_limits
def serp():
    triangle(1)
    with scale(0.5):
        with translate(0, - height):
            serp()
        with translate(side, height):
            serp()
        with translate(-side, height):
            serp()

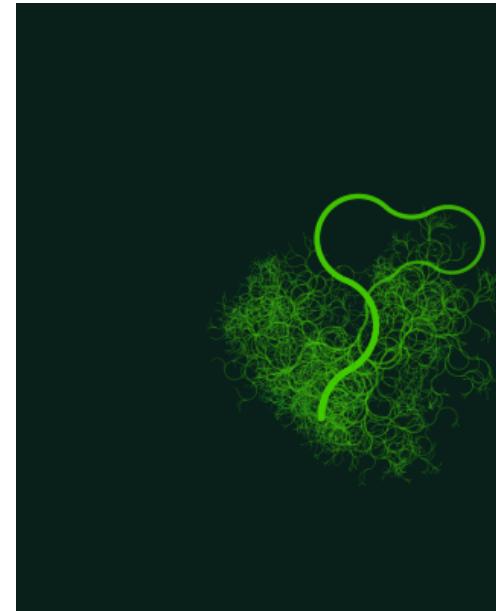
init(canvas_size=(300, 300))
serp()
```



MORE EXAMPLES

```
#check_limits
def branch():
    circle(1)
    with translate(0, 0.6):
        with color(alpha=0.997):
            with rotate(0.04):
                with scale(0.998):
                    if coinflip(50):
                        with flip_y():
                            branch()
                    else:
                        branch()
    if coinflip(50):
        with flip_y():
            with scale(0.9):
                branch()

init(canvas_size=(400, 400), face_color="#44cc00", background_color="black")
with scale(0.012):
    with translate(0, -30):
        branch()
```



CHAINER INTRODUCTION

switch to PFN slides

SOME RESOURCES

- datasets: <https://github.com/undertherain/dagen/>
- sample networks: <https://github.com/undertherain/nuts-and-bolts>

LET'S START HACKING!