

# Inferring Graphics Programs from Images

Anonymous Author(s)

Affiliation

Address

email

## Abstract

1

## 2 1 Introduction

3 How could an agent go from noisy, high-dimensional perceptual input to a symbolic, abstract object,  
 4 like a computer program? Here we consider this problem within a graphics program synthesis domain.  
 5 We develop an approach for converting natural images, such as hand drawings, into executable source  
 6 code for drawing the original image. The graphics programs in our domain draw simple figures like  
 7 those found in machine learning papers (see Fig.1). [The use of ‘graphics programs / visual programs’  
 8 in the paper title, title of this section, and the body of this section feels too broad. ‘Graphics program’  
 9 could conjur a lot of different ideas (esp. 3D graphics); don’t want to set the reader up to expect one  
 10 thing and then be disappointed that what you’ve done isn’t that. You bring up diagram-drawing later  
 11 in the intro; I think it should be made clear sooner (and certainly mentioned explicitly in the abstract,  
 12 when you get around to writing that).]

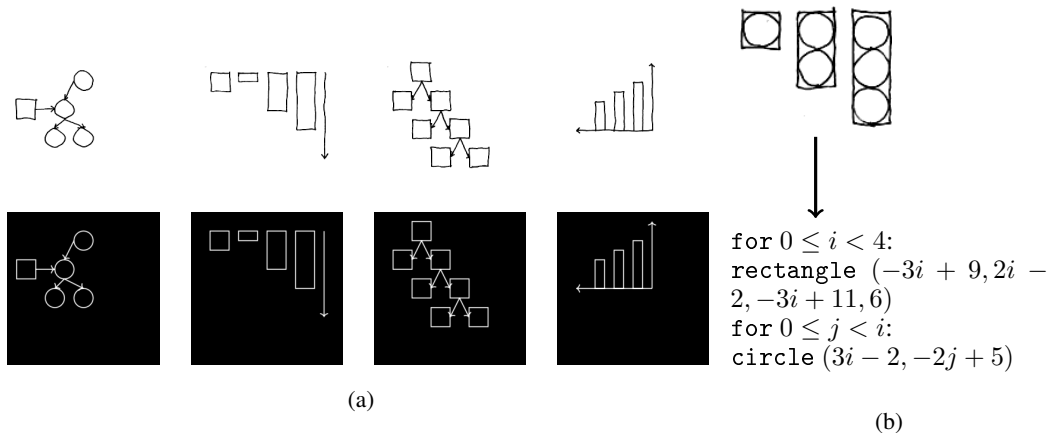


Figure 1: (a): Model learns to convert hand drawings (top) into L\textsuperscript{A}T\textsubscript{E}X (bottom). (b) Synthesizes high-level *graphics program* from hand drawing.

13 High dimensional perceptual input may seem ill matched to the abstract semantics of a programming  
 14 language. But programs with constructs like recursion or iteration produce a simpler *execution trace*  
 15 of primitive actions; for our domain the primitive actions are drawing commands. Our hypothesis is  
 16 that the execution trace of the program is better aligned with the perceptual input, and that the trace  
 17 can act as a kind of bridge between perception and programs. We test this hypothesis by developing  
 18 a model that learns to map from an image to the execution trace of the graphics program that drew  
 19 it. With the execution trace in hand, we can bring to bear techniques from the program synthesis

community to recover the latent graphics program. [This is an *excellent* explanation! I would add maybe one more sentence / citation to elaborate on what you mean by ‘techniques from the program synthesis community’, as this will be an unfamiliar concept to some readers, I imagine.]

We develop a hybrid architecture for inferring graphics programs. Our approach uses a deep neural network infer an execution trace from an image; this network recovers primitive drawing operations such as lines, circles, or arrows [and their parameters?] . For added robustness, we use the deep network as a proposal distribution for a stochastic search over execution traces. Finally, we use techniques in the program synthesis community to recover the program from its trace. [This paragraph is all about making things a bit more specific, so you really need more specifics about program synth here.]

Each of these three components – the deep network, the stochastic search, the program synthesizer – confers its own advantages. From the deep network, we get a very fast system that can recover plausible execution traces in about a minute [A minute seems slow to me, for deep net inference. Are you talking about training time, here, or...?]. From the stochastic search we get added robustness; essentially, the stochastic search can correct mistakes made by the deep network’s proposals. From the program synthesizer, we get abstraction: our system recovers coordinate transformations, for loops, and subroutines, which are useful for downstream tasks [and can help correct some mistakes of the earlier stages?]. [I wonder if this would work even better as a bulleted list...]

## 2 Related work

attend infer repeat: [1]. Crucial distinction is that they focus on learning the generative model jointly with the inference network. Advantages of our system is that we learn symbolic programs, and that we do it from hand sketches rather than synthetic renderings.

ngpm: [2]. We build on the idea of a guide program, extending it to scenes composed of objects, and then show how to learn programs from the objects we discover.

Sketch-n-Sketch: [3]. Semiautomated synthesis presented in a nice user interface. Complementary to our work: you could pass a sketch to our system and then pass the program to sketch-n-sketch

Converting hand drawings into procedural models using deep networks: [4, 5].

## 3 Neural architecture for inferring image parses

We developed a deep network architecture for efficiently inferring a execution trace,  $T$ , from an image,  $I$ . Our model constructs the trace one drawing command at a time. When predicting the next drawing command, the network takes as input the target image  $I$  as well as the rendered output of previous drawing commands. Intuitively, the network looks at the image it wants to explain, as well as what it has already drawn. It then decides either to stop drawing or proposes another drawing command to add to the execution trace; if it decides to continue drawing, the predicted primitive is rendered to its “canvas” and the process repeats.

Figure 2 illustrates this architecture. We first pass the target image and a rendering of the trace so far to a convolutional network. Given the features extracted by the convolutional network, a multilayer perceptron then predicts a distribution over the next drawing command to add to the trace. We predict the drawing command token-by-token, and condition each token both on the image features and on the previously generated tokens. For example, the network first decides to emit the `circle` token conditioned on the image features, then it emits the  $x$  coordinate of the circle conditioned on the image features and the `circle` token, and finally it predicts the  $y$  coordinate of the circle conditioned on the image features, the `circle` token, and the  $x$  coordinate. [There are some more details that are important to provide about this architecture, though possibly in an Appendix: the functional form(s) of the probability distributions over tokens, the network layer sizes, which MLPs share parameters, etc.]

[Planning to move the description of SMC / beam search up here, too?]

<code>circle(<math>x, y</math>)</code>	Circle at $(x, y)$
<code>rectangle(<math>x_1, y_1, x_2, y_2</math>)</code>	Rectangle with corners at $(x_1, y_1)$ & $(x_2, y_2)$
<code>LINE(<math>x_1, y_1, x_2, y_2</math>, arrow <math>\in \{0, 1\}</math>, dashed <math>\in \{0, 1\}</math>)</code>	Line from $(x_1, y_1)$ to $(x_2, y_2)$ , optionally with an arrow and/or dashed
<code>STOP</code>	Finishes execution trace inference

Table 1: The deep network in (2) predicts drawing commands, shown above.

Program	$\rightarrow$ Command; $\dots$ ; Command
Command	$\rightarrow$ circle(Expression, Expression)
Command	$\rightarrow$ rectangle(Expression, Expression, Expression, Expression)
Command	$\rightarrow$ LINE(Expression, Expression, Expression, Expression, Boolean, Boolean)
Command	$\rightarrow$ for( $0 \leq \text{Var} < \text{Expression}$ ) { Program }
Command	$\rightarrow$ REFLECT(Axis) { Program }
Expression	$\rightarrow$ Z * Var + Z
Var	$\rightarrow$ A free (unused) variable
Z	$\rightarrow$ an integer
Axis	$\rightarrow$ X = Z
Axis	$\rightarrow$ Y = Z

Table 2: Grammar over graphics programs. We allow loops (for), vertical/horizontal reflections (REFLECT), and affine transformations ( $Z * \text{Var} + Z$ ).

67 The distribution over the next drawing command factorizes:

$$\mathbb{P}_\theta[t_1 t_2 \dots t_K | I, T] = \prod_{k=1}^K \mathbb{P}_\theta[t_k | f_\theta(I, \text{render}(T)), \{t_j\}_{j=1}^{k-1}] \quad (1)$$

68 where  $t_1 t_2 \dots t_K$  are the tokens in the drawing command,  $I$  is the target image,  $T$  is an execution trace,  
69  $\theta$  are the parameters of the neural network, and  $f_\theta(\cdot, \cdot)$  is the image feature extractor (convolutional  
70 network). The distribution over execution traces factorizes as:

$$\mathbb{P}_\theta[T | I] = \prod_{n=1}^{|T|} \mathbb{P}_\theta[T_n | I, T_{1:(n-1)}] \times \mathbb{P}_\theta[\text{STOP} | I, T] \quad (2)$$

71 where  $|T|$  is the length of execution trace  $T$ , and the STOP token is emitted by the network to signal  
72 that the execution trace explains the image.

73 We train the network by sampling execution traces  $T$  and target images  $I$  for randomly generated  
74 scenes, and maximizing (2) wrt  $\theta$  by gradient ascent. Training does not require backpropagation across  
75 the entire sequence of drawing commands: drawing to the canvas ‘blocks’ the gradients, effectively  
76 offloading memory to an external visual store. In a sense, this model is like an autoregressive variant  
77 of AIR [1] without attention.

78 [I like that you make this connection, but it could be made more precisely. Specifically, (1) the  
79 architecture isn’t *really* recurrent (it uses no hidden state cells), so it’d be good to use a different term  
80 or drop this part of the point: (2) training of recurrent nets is also typically fully-supervised (Most  
81 RNNs lack latent variables per timestep)—if you’re thinking about AIR specifically, maybe just say  
82 that, and (3) it’s like an autoregressive AIR *without attention*.] [Something related to this that’s also  
83 cool to point out: training this model doesn’t require backpropagation across the entire sequence of  
84 drawing commands (drawing to the canvas ‘blocks’ the gradients, effectively offloading memory to  
85 an external (visual) store, so in principle it might be scalable to much longer sequences.]

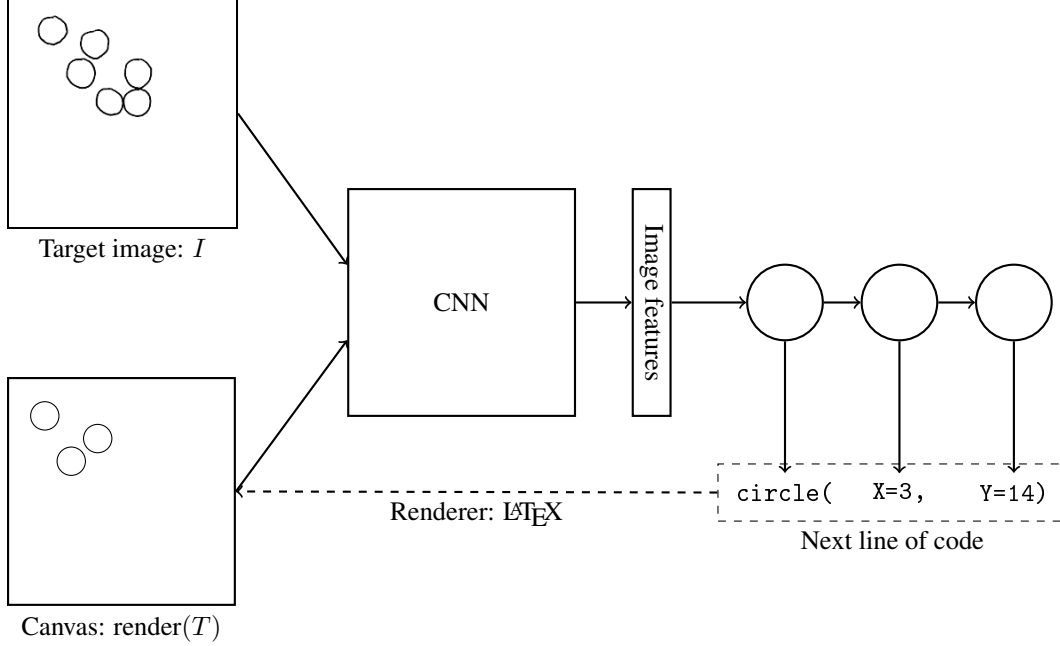


Figure 2: Our neural architecture for inferring the execution trace of a graphics program from its output. [Thoughts on improving this figure: (1) Convnet diagrams typically show the sequence of layers, if possible (space might not permit it here, but those thin arrows just aren't doing it for me). (2) Are the target image / canvas convolved down independently, or jointly (i.e. starting as a 2-channel image)? That's an important detail that's not clear with the current figure/explanation. (3) The three circles downstream from 'Image Features' are supposed to be MLPs, I assume(?), but it took me a little while to parse that. Having some visual way of clearly separating network operations from data (color, perhaps) would go a long way.]

## 86 4 Generalizing to hand drawings

## 87 5 Synthesizing graphics programs from execution traces

## 88 6 Neural networks for guiding SMC

89 Let  $L(\cdot) : \text{image}^2 \rightarrow \mathcal{R}$  be our likelihood function: it takes two images, an observed target image  
 90 and a hypothesized program output, and gives the likelihood of the observed image conditioned on  
 91 the program output. We want to sample from:

$$\mathbb{P}[p|x] \propto L(x|\text{render}(p))\mathbb{P}[p] \quad (3)$$

92 where  $\mathbb{P}[p]$  is the prior probability of program  $p$ , and  $x$  is the observed image.

93 Let  $p$  be a program with  $L$  lines, which we will write as  $p = (p_1, p_2, \dots, p_L)$ . Assume the prior  
 94 factors into:

$$\mathbb{P}[p] \propto \prod_{l \leq L} \mathbb{P}[p_l] \quad (4)$$

95 Define the distribution  $q_L(\cdot)$ , which happens to be proportional to the above posterior:

$$q_L(p_1, p_2, \dots, p_{L-1}, p_L) \propto q_{L-1}(p_1, p_2, \dots, p_{L-1}) \times \frac{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}, p_L))}{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}))} \times \mathbb{P}[p_L] \quad (5)$$

96 Now suppose we have some samples from  $q_{L-1}(\cdot)$ , and that we then sample a  $p_L$  from a distribu-  
 97 tion proportional to  $\frac{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}, p_L))}{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}))} \times \mathbb{P}[p_L]$ . The resulting programs  $p$  are distributed  
 98 according to  $q_L$ , and so are also distributed according to  $\mathbb{P}[p|x]$ .

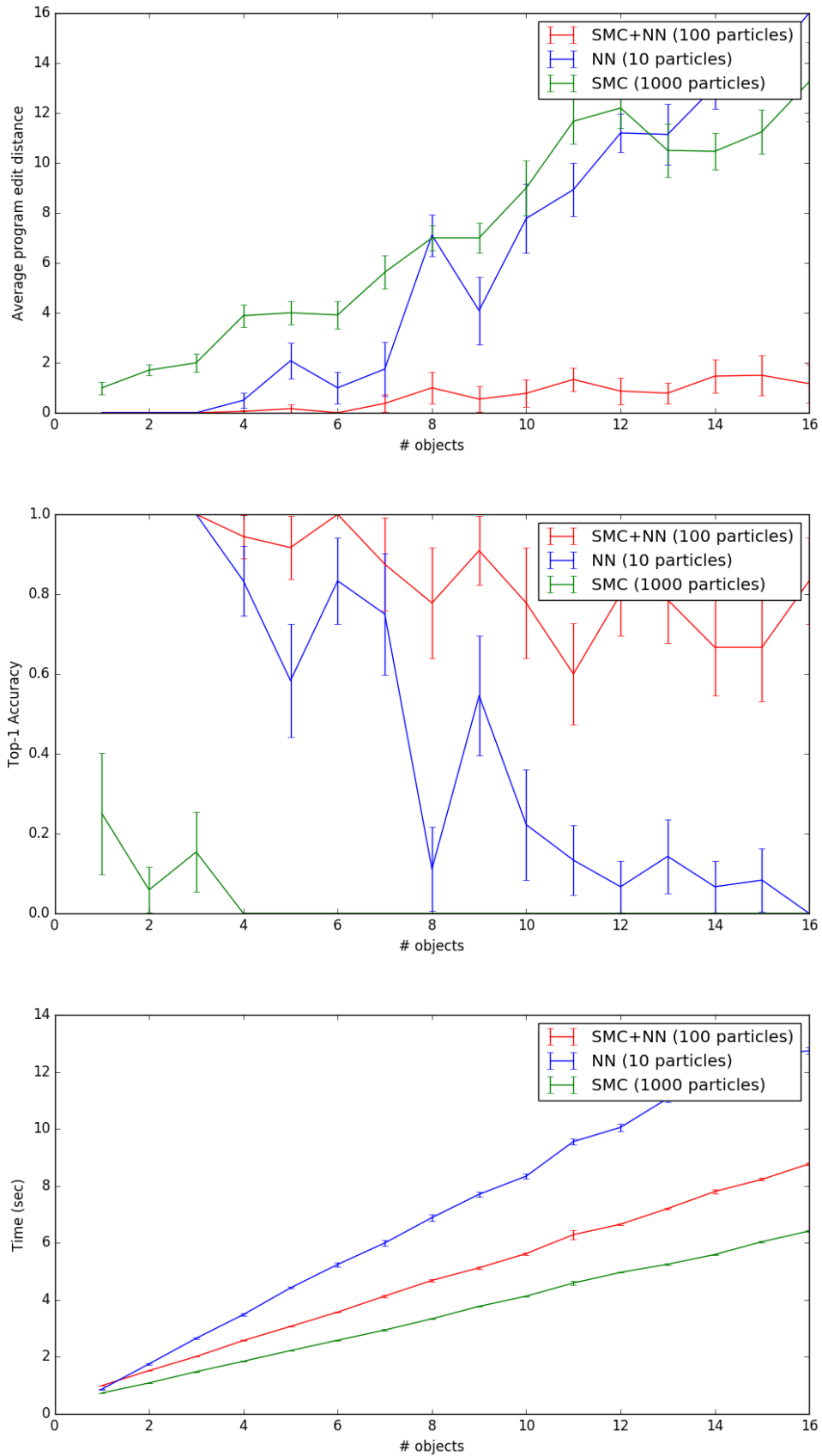


Figure 3: Using the model to parse latex output. The model is trained on diagrams with up to 8 objects. As shown above it generalizes to scenes with many more objects. Neither the stochastic search nor the neural network are sufficient on their own.

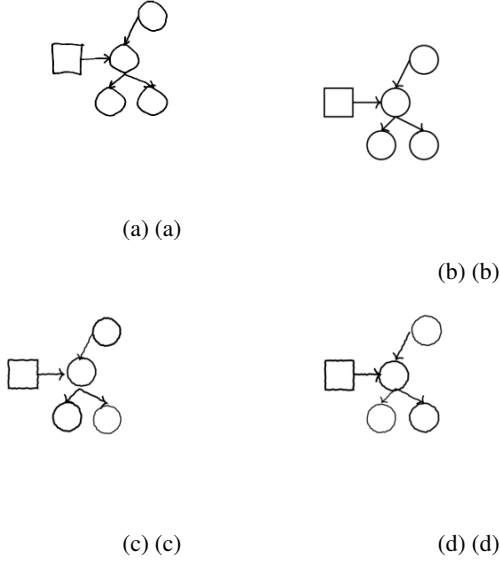


Figure 4: (a): a hand drawing. (b): Rendering of the parse our model infers for (a). We can generalize to hand drawings like these because we train the model on images corrupted by a noise process designed to resemble the kind of noise introduced by hand drawings - see (c) & (d) for noisy renderings of (b).

99 How do we sample  $p_L$  from a distribution proportional to  $\frac{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}, p_L))}{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}))} \times \mathbb{P}[p_L]$ ? We  
 100 have a neural network that takes as input the target image  $x$  and the program so far, and produces  
 101 a distribution over next lines of code ( $p_L$ ). We write  $\text{NN}(p_L|p_1, \dots, p_{L-1}; x)$  for the distribution  
 102 output by the neural network. So we can sample from NN and then weight the samples by:

$$w(p_L) = \frac{\mathbb{P}[p_L]}{\text{NN}(p_L|p_1, \dots, p_{L-1}; x)} \times \frac{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}, p_L))}{L(x|\text{render}(p_1, p_2, \dots, p_{L-1}))} \quad (6)$$

103 Then we can resample from these now weighted samples to get a new population of particles (here  
 104 programs are particles), where each program now has  $L$  lines instead of  $L - 1$ .

105 This procedure can be seen as a particle filter, where each successive latent variable is another line of  
 106 code, and the emission probabilities are successive ratios of likelihoods under  $L(\cdot|\cdot)$ .

107 **Comments for Dan.** Right now I'm not actually sampling from the neural network - instead, I  
 108 enumerate the top few hundred lines of code suggested by the network, and then weight them by their  
 109 likelihoods. So actually the form of NN is:

$$\text{NN}(p_L|p_1, \dots, p_{L-1}; x) \propto \begin{cases} 1, & \text{if } p_L \in \text{top hundred neural network proposals} \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

110 Do you think this is a problem? The neural network puts almost all of its mass on a few guesses. In  
 111 order to get the correct line of code I sometimes need to get something like the 50th top guess, so I  
 112 don't want to literally just sample from the distribution suggested by the neural network.

---

**Algorithm 1** Neurally guided SMC

---

**Input:** Neural network NN, beam size  $N$ , maximum length  $L$ , target image  $x$

**Output:** Samples of the program trace

Set  $B_0 = \{\text{empty program}\}$

**for**  $1 \leq l \leq L$  **do**

**for**  $1 \leq n \leq N$  **do**

$p_n \sim \text{Uniform}(B_{l-1})$

$p'_n \sim \text{NN}(\text{render}(p), x)$

    Define  $r_n = p'_n \cdot p_n$

    Set  $\tilde{w}(r_n) = \frac{L(x|r_n)}{L(x|p_n)} \times \frac{\mathbb{P}[p'_n]}{\mathbb{P}[p'_n = \text{NN}(\text{render}(p), x)]}$

**end for**

  Define  $w(p) = \frac{\tilde{w}(p)}{\sum_{p'} \tilde{w}(p')}$

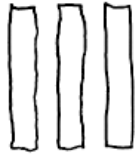
  Set  $B_l$  to be  $N$  samples from  $r_n$  distributed according to  $w(\cdot)$

**end for**

**return**  $\{p : p \in B_{l \leq L}, p \text{ is finished}\}$

---

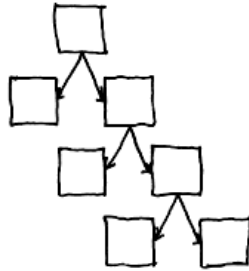
## 113 7 Preliminary Synthesis results



114

```
Rectangle(0,0,1,6)
Rectangle(2,0,3,6)
Rectangle(4,0,5,6)
```

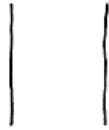
115



116

```
Rectangle(2,9,4,11)
for (3)
  Line(-2*i + 7,3*i + 3,-2*i + 6,3*i + 1,arrow = True,solid
  Line(2*i + 3,-3*i + 9,2*i + 4,-3*i + 7,arrow = True,solid
  Rectangle(2*i,-3*i + 6,2*i + 2,-3*i + 8)
  Rectangle(-2*i + 8,3*i,-2*i + 10,3*i + 2)
```

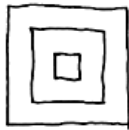
117



118

```
Line(0,0,0,5,arrow = False,solid = True)
```

119



120

```
Rectangle(0,0,5,5)  
Rectangle(1,1,4,4)  
Rectangle(2,2,3,3)
```

121



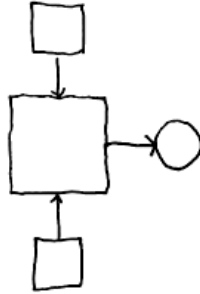
122

```
Circle(1,1)
```

123



124



```
Rectangle(0,4,4,8)
  reflect(y = 12)
  Circle(7,6)
  Line(2,2,2,4,arrow = True,solid = True)
  Line(4,6,6,6,arrow = True,solid = True)
  Rectangle(1,0,3,2)
```

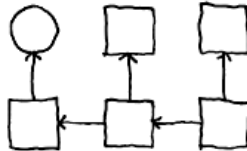
125



126

```
Rectangle(2,2,5,3)
Rectangle(0,0,3,1)
Rectangle(4,4,7,5)
```

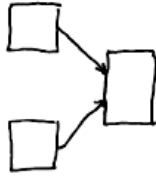
127



128

```
Circle(1,5)
for (2)
  Line(-5*i + 9,-1*i + 2,-7*i + 9,-3*i + 4,arrow = True,solid = True)
  Rectangle(-4*i + 8,4*i,-4*i + 10,4*i + 2)
  reflect(x = 6)
  Line(7*i + 1,-1*i + 2,5*i + 1,-3*i + 4,arrow = True,solid = True)
  Rectangle(8*i,4*i,8*i + 2,4*i + 2)
```

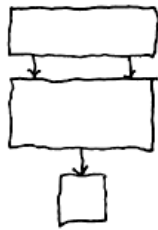
129



130

```
Rectangle(4,2,6,5)
    reflect(y = 7)
    Line(2,6,4,4,arrow = True,solid = True)
    Rectangle(0,5,2,7)
```

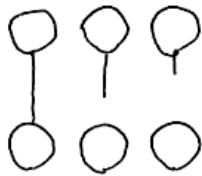
131



132

```
Line(3,3,3,2,arrow = True,solid = True)
Rectangle(0,7,6,9)
Rectangle(2,0,4,2)
Rectangle(0,3,6,6)
    Line(1,7,1,6,arrow = True,solid = True)
```

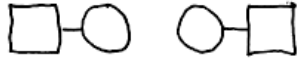
133



134

```
for (3)
    Circle(-3*i + 7,1)
    Circle(3*i + 1,6)
    Line(3*i + 1,2,3*i + 1,5,arrow = False,solid = True)
```

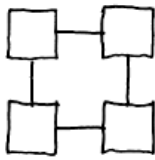
135



136

```
reflect(x = 12)
Circle(4,1)
Line(9,1,10,1,arrow = False,solid = True)
Rectangle(10,0,12,2)
```

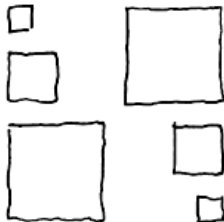
137



138

```
reflect(x = 6)
Line(5,2,5,4,arrow = False,solid = True)
    reflect(y = 6)
        Line(2,5,4,5,arrow = False,solid = True)
            Rectangle(0,4,2,6)
```

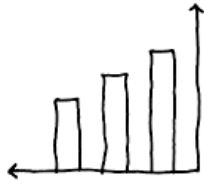
139



140

```
for (2)
    Rectangle(0,-3*i + 8,1,-2*i + 9)
    Rectangle(-3*i + 8,5*i,9,8*i + 1)
    Rectangle(-7*i + 7,-2*i + 2,-5*i + 9,4)
```

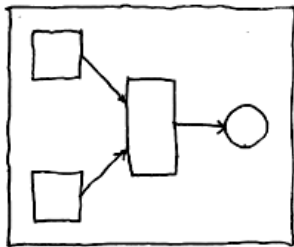
141



142

```
Rectangle(4,0,5,4)
for (2)
    Line(8,0,-8*i + 8,-7*i + 7,arrow = True,solid = True)
    Rectangle(-4*i + 6,0,-4*i + 7,-2*i + 5)
```

143



144

```
Circle(10,5)
Line(7,5,9,5,arrow = True,solid = True)
Rectangle(5,3,7,7)
Rectangle(0,0,12,10)
    reflect(y = 10)
    Line(3,2,5,4,arrow = True,solid = True)
    Rectangle(1,1,3,3)
```

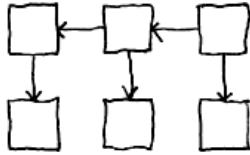
145



146

```
Line(0,0,0,2,arrow = False,solid = True)
Line(0,2,2,2,arrow = False,solid = True)
```

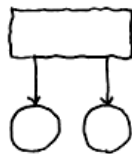
147



148

```
Line(8,5,6,5,arrow = True,solid = True)
Line(4,5,2,5,arrow = True,solid = True)
for (3)
    Line(-4*i + 9,4,-4*i + 9,2,arrow = True,solid = True)
    Rectangle(4*i,0,4*i + 2,2)
    Rectangle(-4*i + 8,4,-4*i + 10,6)
```

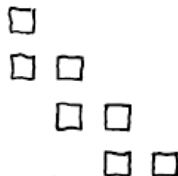
149



150

```
Rectangle(0,4,5,6)
reflect(x = 5)
Circle(4,1)
Line(1,4,1,2,arrow = True,solid = True)
```

151



152

```
Rectangle(0,6,1,7)
for (3)
    Rectangle(-2*i + 6,2*i,-2*i + 7,2*i + 1)
    Rectangle(-2*i + 4,2*i,-2*i + 5,2*i + 1)
```

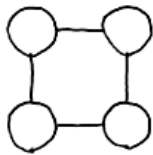
153



154

```
for (3)
  Circle(4*i + 1,1)
  Rectangle(4*i,0,4*i + 2,2)
```

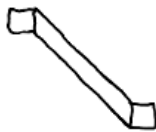
155



156

```
reflect(y = 6)
Line(2,1,4,1,arrow = False,solid = True)
  reflect(x = 6)
  Circle(1,1)
  Line(1,2,1,4,arrow = False,solid = True)
```

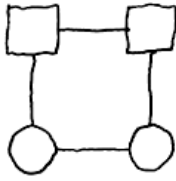
157



158

```
Line(1,5,5,1,arrow = False,solid = True)
Line(1,4,5,0,arrow = False,solid = True)
Rectangle(5,0,6,1)
Rectangle(0,4,1,5)
```

159



160

```
for (2)
  Circle(-5*i + 6,1)
  Line(-4*i + 6,-1*i + 2,-1*i + 6,-4*i + 5,arrow = False,solid
  Line(-1*i + 2,-4*i + 6,-4*i + 5,-1*i + 6,arrow = False,solid
  Rectangle(5*i,5,5*i + 2,7)
```

161



162

```
Line(0,0,0,5,arrow = False,solid = False)
Line(4,1,4,5,arrow = False,solid = False)
Line(4,0,4,1,arrow = False,solid = False)
```

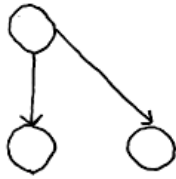
163



164

```
Rectangle(0,0,3,4)
```

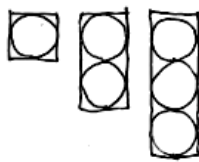
165



166

```
Circle(1,1)
for (2)
  Circle(-5*i + 6,5*i + 1)
  Line(1,5,5*i + 1,2,arrow = True,solid = True)
```

167



168

```
Circle(7,3)
for (3)
  Circle(-3*i + 7,5)
  Circle(-3*i + 7,2*i + 1)
  Rectangle(-3*i + 6,2*i,-3*i + 8,6)
```

169



170

```
Circle(1,8)
for (2)
  Line(4*i + 1,-5*i + 7,2*i + 3,5,arrow = False,solid = True)
  Rectangle(-4*i + 4,5*i,6,7*i + 2)
```

171





172

```
Circle(7,1)
Circle(4,1)
Circle(1,1)
```

173



174

```
Line(1,2,1,5,arrow = False,solid = True)
for (2)
  Line(2,-4*i + 4,-4*i + 6,4,arrow = False,solid = True)
  Line(0,-2*i + 6,-2*i + 2,6,arrow = False,solid = True)
  Line(1,5,2*i + 2,5,arrow = False,solid = True)
```

175

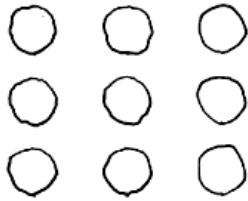


176

```
Rectangle(5,0,8,3)
Rectangle(2,1,4,3)
Rectangle(0,2,1,3)
```

177

178



179

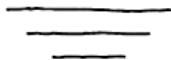
```
for (3)
  Circle(-4*i + 9,4)
  Circle(-4*i + 9,1)
  Circle(4*i + 1,7)
```

180



181

```
Circle(1,5)
Line(1,4,1,2,arrow = True,solid = True)
Rectangle(0,0,2,2)
```



182

```
Line(0,2,7,2,arrow = False,solid = True)
Line(1,1,6,1,arrow = False,solid = True)
Line(2,0,5,0,arrow = False,solid = True)
```

183

184



```
Circle(1,9)
Line(1,8,1,6,arrow = True,solid = True)
Line(1,3,1,2,arrow = True,solid = True)
Line(1,3,1,4,arrow = False,solid = True)
    reflect(y = 6)
    Circle(1,1)
```

185



186

```
Line(1,2,3,2,arrow = False,solid = True)
Line(2,1,4,1,arrow = False,solid = False)
Line(3,0,5,0,arrow = False,solid = True)
Line(0,3,2,3,arrow = False,solid = False)
```

187



188

```
Line(4,4,2,2,arrow = True,solid = True)
Rectangle(3,4,5,6)
Rectangle(0,0,2,2)
```

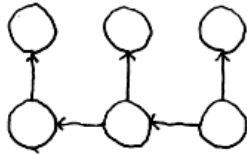
189

!

190

```
Line(0,0,0,4,arrow = False,solid = True)
```

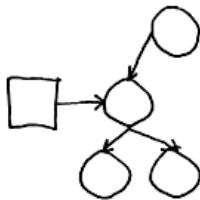
191



192

```
Line(4,1,2,1,arrow = False,solid = True)
for (2)
  Line(-4*i + 5,2,-4*i + 5,4,arrow = True,solid = True)
  for (3)
    Circle(-4*j + 9,4*i + 1)
    Line(8,1,3*i + 6,3*i + 1,arrow = True,solid = False)
```

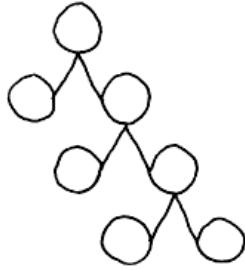
193



194

```
Rectangle(0,3,2,5)
for (2)
  Circle(3*i + 4,1)
  Circle(-2*i + 7,-3*i + 7)
  Line(5,3,-3*i + 7,2,arrow = True,solid = True)
  Line(4*i + 2,3*i + 4,4,4,arrow = True,solid = True)
```

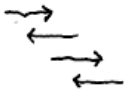
195



196

```
Circle(9,1)
for (3)
  Circle(2*i + 3,-3*i + 10)
  Circle(-2*i + 5,3*i + 1)
  Line(2*i + 2,-3*i + 7,2*i + 3,-3*i + 9,arrow = False,solid
  Line(-2*i + 7,3*i + 3,-2*i + 8,3*i + 1,arrow = False,solid
```

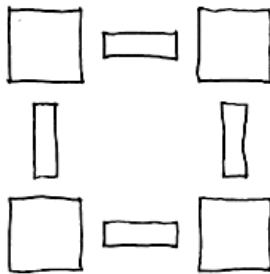
197



198

```
Line(5,0,3,0,arrow = True,solid = True)
Line(0,3,2,3,arrow = True,solid = True)
Line(3,2,1,2,arrow = True,solid = True)
Line(2,1,4,1,arrow = True,solid = True)
```

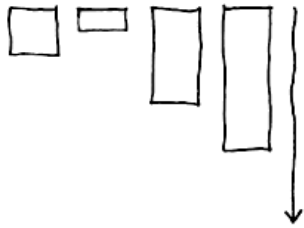
199



200

```
reflect(y = 11)
Rectangle(4,9,7,10)
  reflect(x = 11)
    Rectangle(8,0,11,3)
    Rectangle(1,4,2,7)
```

201



202

```
Line(12,9,12,0,arrow = True,solid = True)
for (2)
  Rectangle(-3*i + 6,3*i + 5,-3*i + 8,9)
  Rectangle(9*i,-4*i + 7,9*i + 2,9)
```

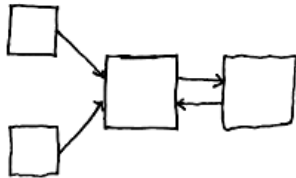
203

-----

204

```
Line(4,0,0,0,arrow = False,solid = False)
```

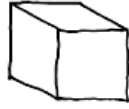
205



206

```
Line(8,3,9,3,arrow = False,solid = True)
for (2)
  Line(-6*i + 8,3*i + 3,-3*i + 7,3,arrow = True,solid = True)
  Line(5*i + 2,3*i + 1,5*i + 4,3,arrow = True,solid = True)
  Rectangle(9*i,2*i,10*i + 2,3*i + 2)
  Rectangle(-4*i + 4,3*i + 2,-5*i + 7,2*i + 5)
```

207



208

```

Line(0,1,0,4,arrow = False,solid = True)
Rectangle(2,0,5,3)
for (2)
    Line(0,3*i + 1,2,3*i,arrow = False,solid = True)
    Line(-3*i + 3,4,-2*i + 5,3,arrow = False,solid = True)

```

## 209 References

- 210 [1] SM Eslami, N Heess, and T Weber. Attend, infer, repeat: Fast scene understanding with generative models.  
 211 arxiv preprint arxiv:..., 2016. URL <http://arxiv.org/abs/1603.08575>.
- 212 [2] Daniel Ritchie, Anna Thomas, Pat Hanrahan, and Noah Goodman. Neurally-guided procedural models:  
 213 Amortized inference for procedural graphics programs using neural networks. In *Advances In Neural*  
 214 *Information Processing Systems*, pages 622–630, 2016.
- 215 [3] Brian Hempel and Ravi Chugh. Semi-automated svg programming via direct manipulation. In *Proceedings*  
 216 *of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pages 379–390, New  
 217 York, NY, USA, 2016. ACM.
- 218 [4] Haibin Huang, Evangelos Kalogerakis, Ersin Yumer, and Radomir Mech. Shape synthesis from sketches via  
 219 procedural models and convolutional networks. *IEEE transactions on visualization and computer graphics*,  
 220 2017.
- 221 [5] Gen Nishida, Ignacio Garcia-Dorado, Daniel G. Aliaga, Bedrich Benes, and Adrien Bousseau. Interactive  
 222 sketching of urban procedural models. *ACM Trans. Graph.*, 35(4), 2016.