Learning Task-General Representations with Generative Neuro-Symbolic Modeling

Reuben Feinman & Brenden Lake
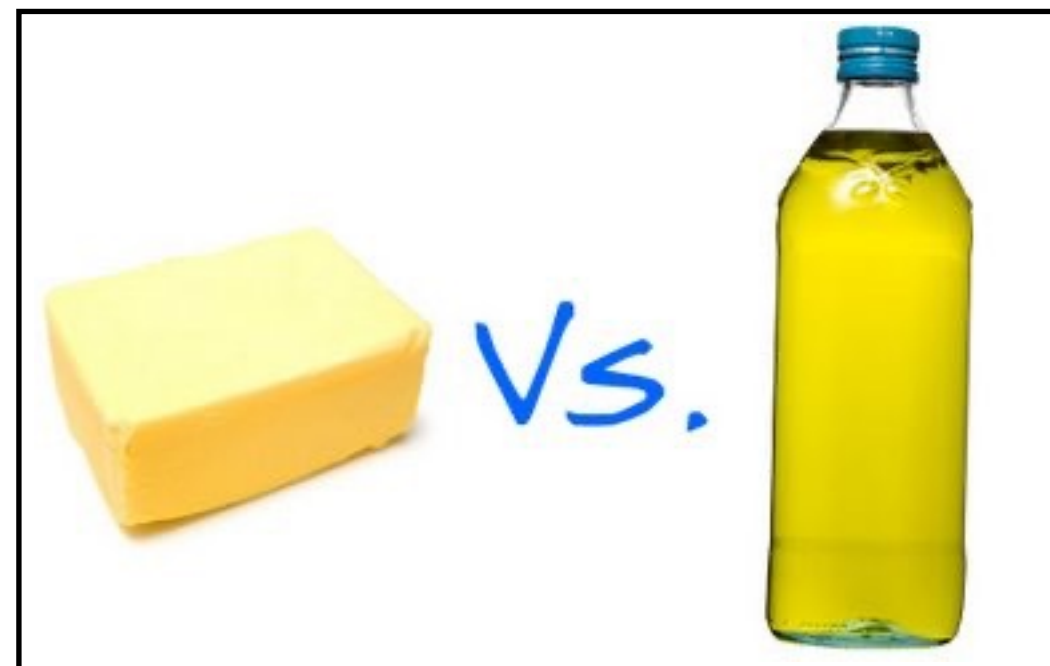
Structure and statistics in conceptual representations

Ingredients (parts)

Correlations

substitutions   co-occurences
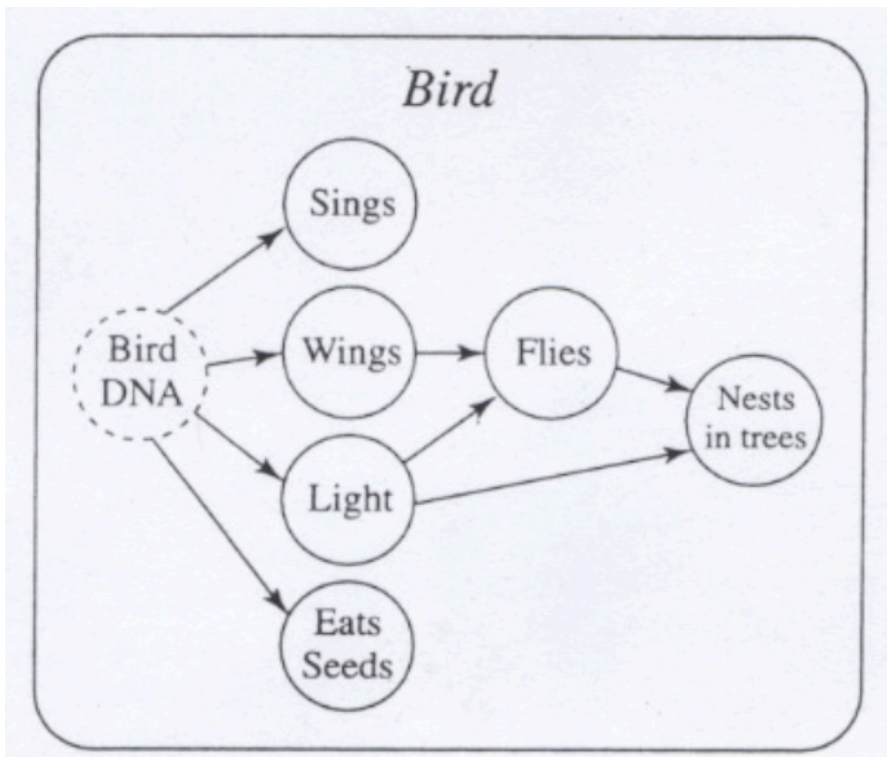
Causal processes

Modeling structure and statistics
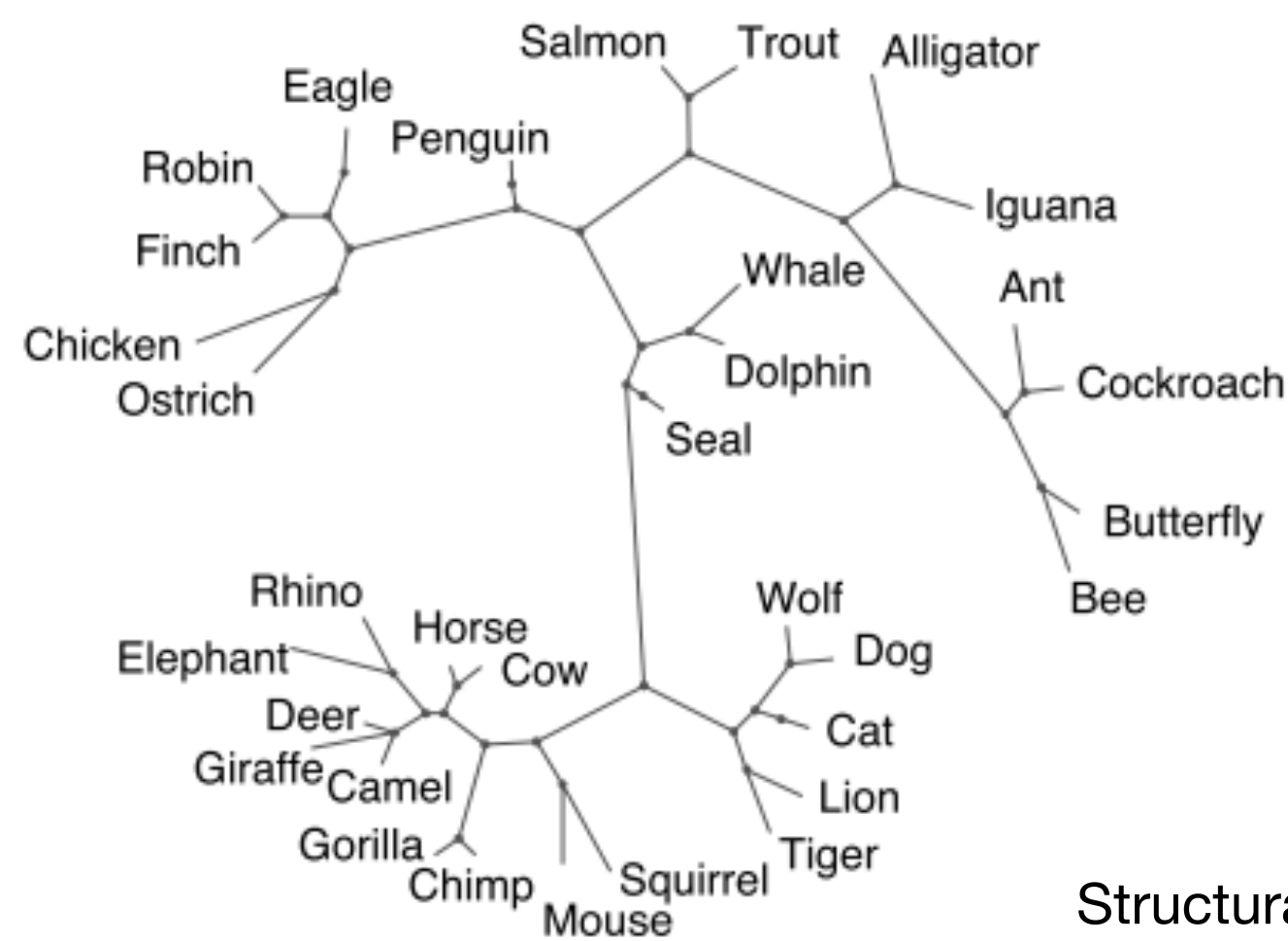
Structured knowledge
(symbolic models)

Emergent "statistical" knowledge
(neural networks)

Causal-Model Theory
(Rehder, 2007)

Bayesian Program Learning
(Lake et al., 2015)

ALCOVE
(Kruschke, 1992)

Semantic Cognition Network
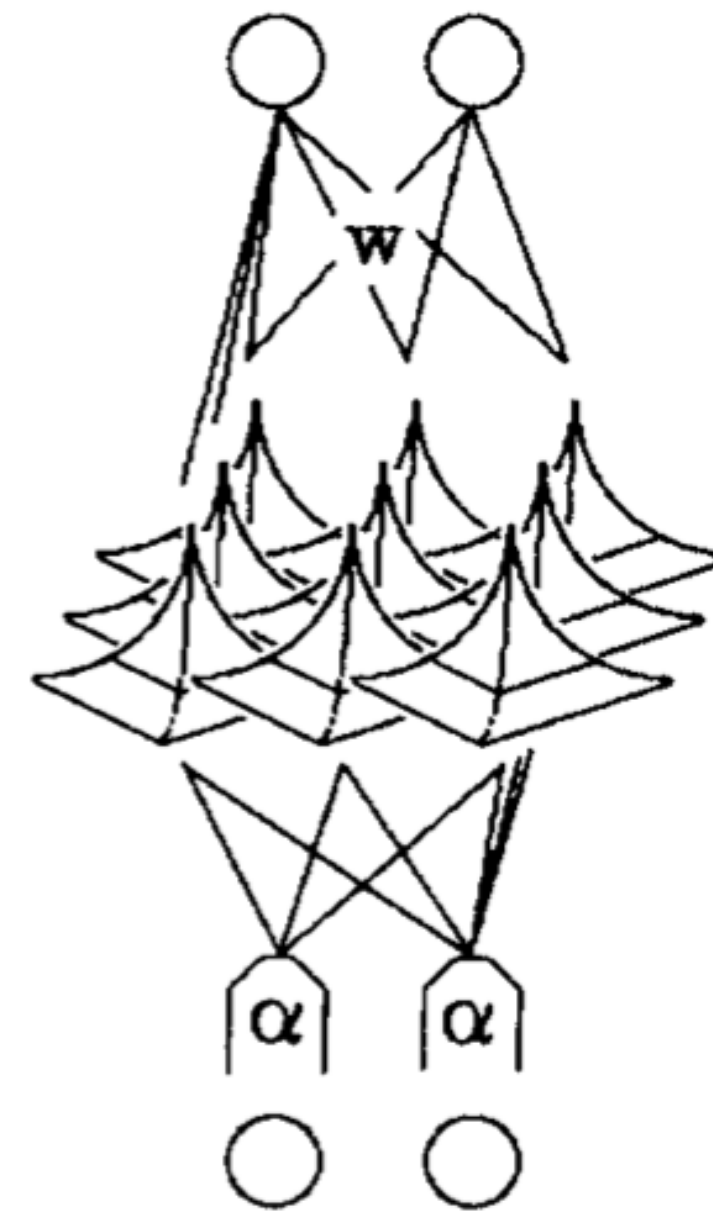(Rogers & McClelland, 2003)

procedure GENERATETYPE
$\kappa \leftarrow P(\kappa)$ ▷ Sample number of parts
for $i = 1 \dots \kappa$ **do**
$\quad n_i \leftarrow P(n_i | \kappa)$ ▷ Sample number of sub-parts
$\quad$ **for** $j = 1 \dots n_i$ **do**
$\quad\quad s_{ij} \leftarrow P(s_{ij} | s_{i(j-1)})$ ▷ Sample sub-part sequence
$\quad$ **end for**
$\quad R_i \leftarrow P(R_i | S_1, \dots, S_{i-1})$ ▷ Sample relation
end for
$\psi \leftarrow \{\kappa, R, S\}$
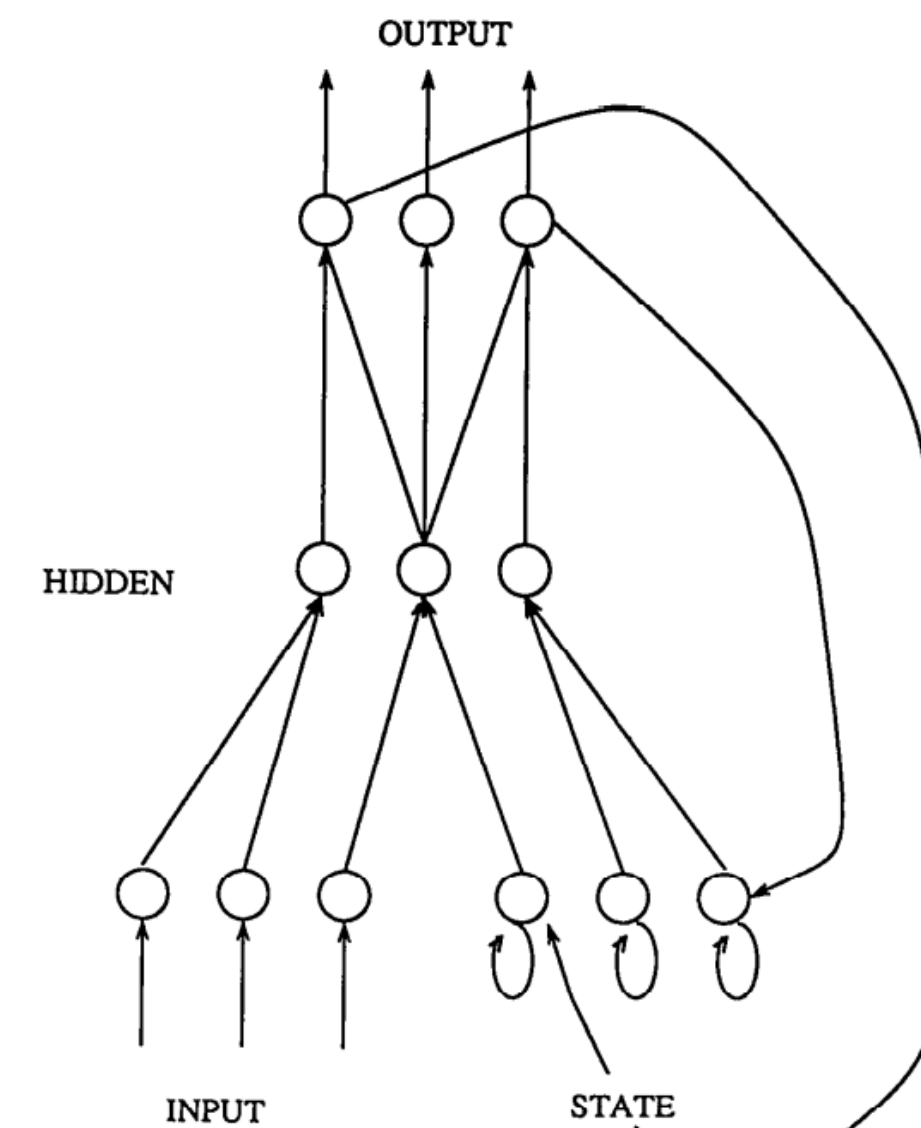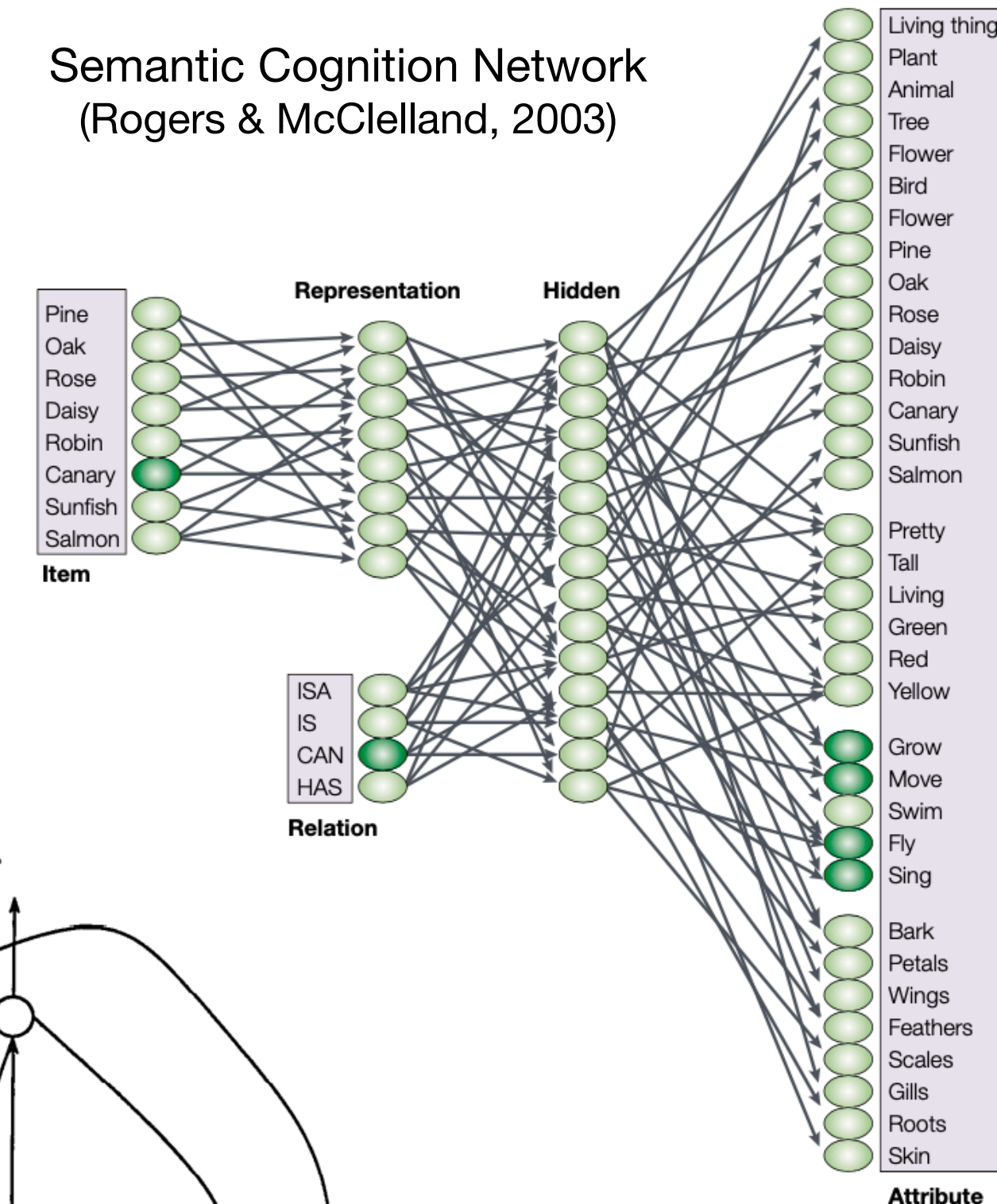**return** @GENERATETOKEN(ψ) ▷ Return program

Structural Forms
(Kemp & Tenenbaum, 2008)

Finding Structure in Time
(Elman, 1990)

Proposal:
Generative Neuro-Symbolic (GNS) Modeling

- Goal: model the *compositional* and *causal* structure in how concepts are formed, while simultaneously modeling nonparametric statistical relationships

- Proposal: probabilistic programs with neural network sub-routines

procedure GENERATECONCEPT
  $M \leftarrow 0$ ▷ Initialize memory state
  while $True$ **do**
    $x_i, r_i \sim p(x, r \mid M)$ ▷ Sample part and relation from neural net
    $M \leftarrow f_{render}(x_i, r_i, M)$ ▷ Render part to memory (differentiable)
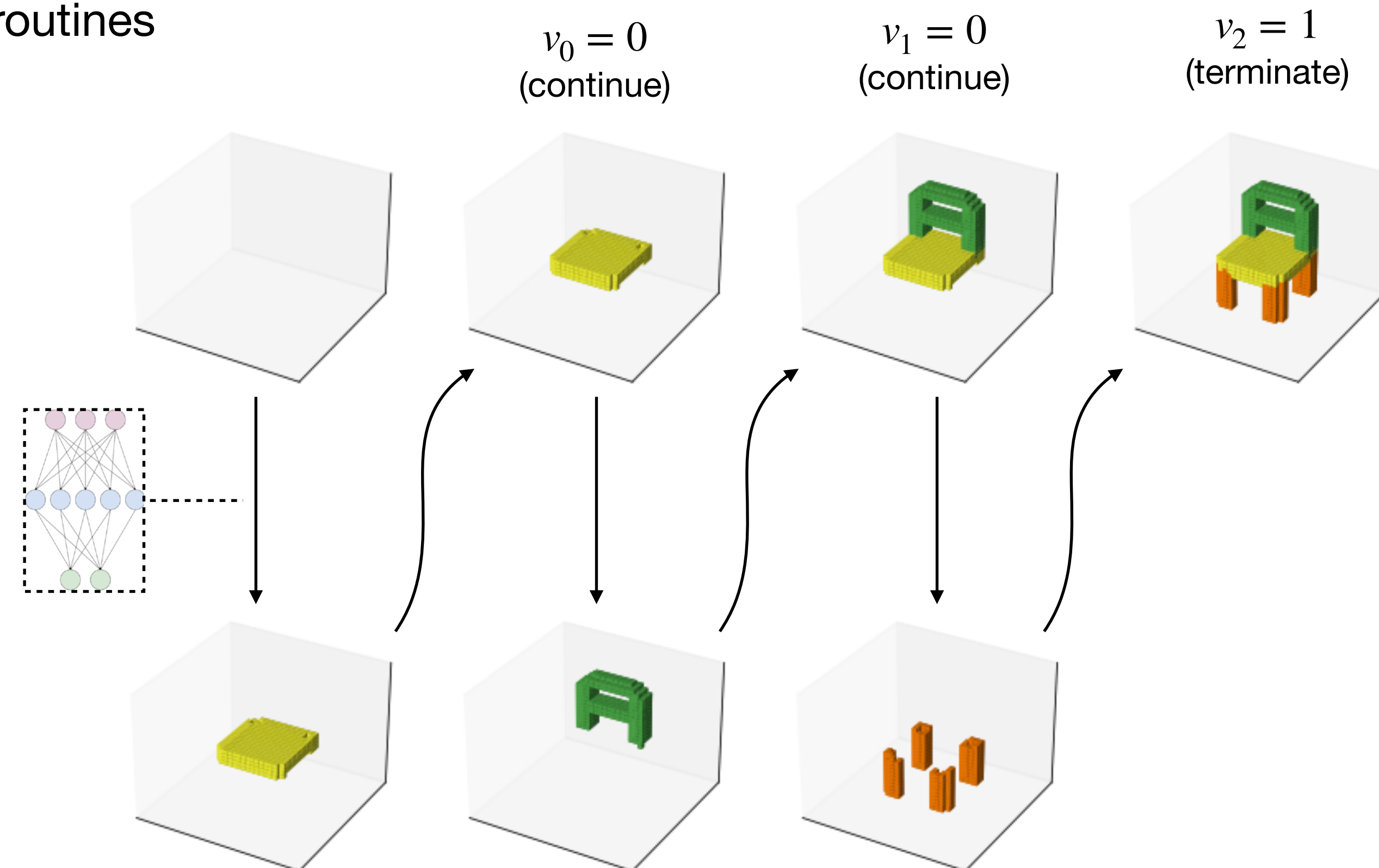 $v_i \sim p(v \mid M)$ ▷ Sample termination indicator
    if v_i **then**
      break
  return $\{X, R\}$ ▷ Return concept type
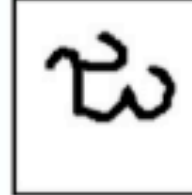
Omniglot concept learning challenge

(Lake et al., 2015)

GNS model of character concepts
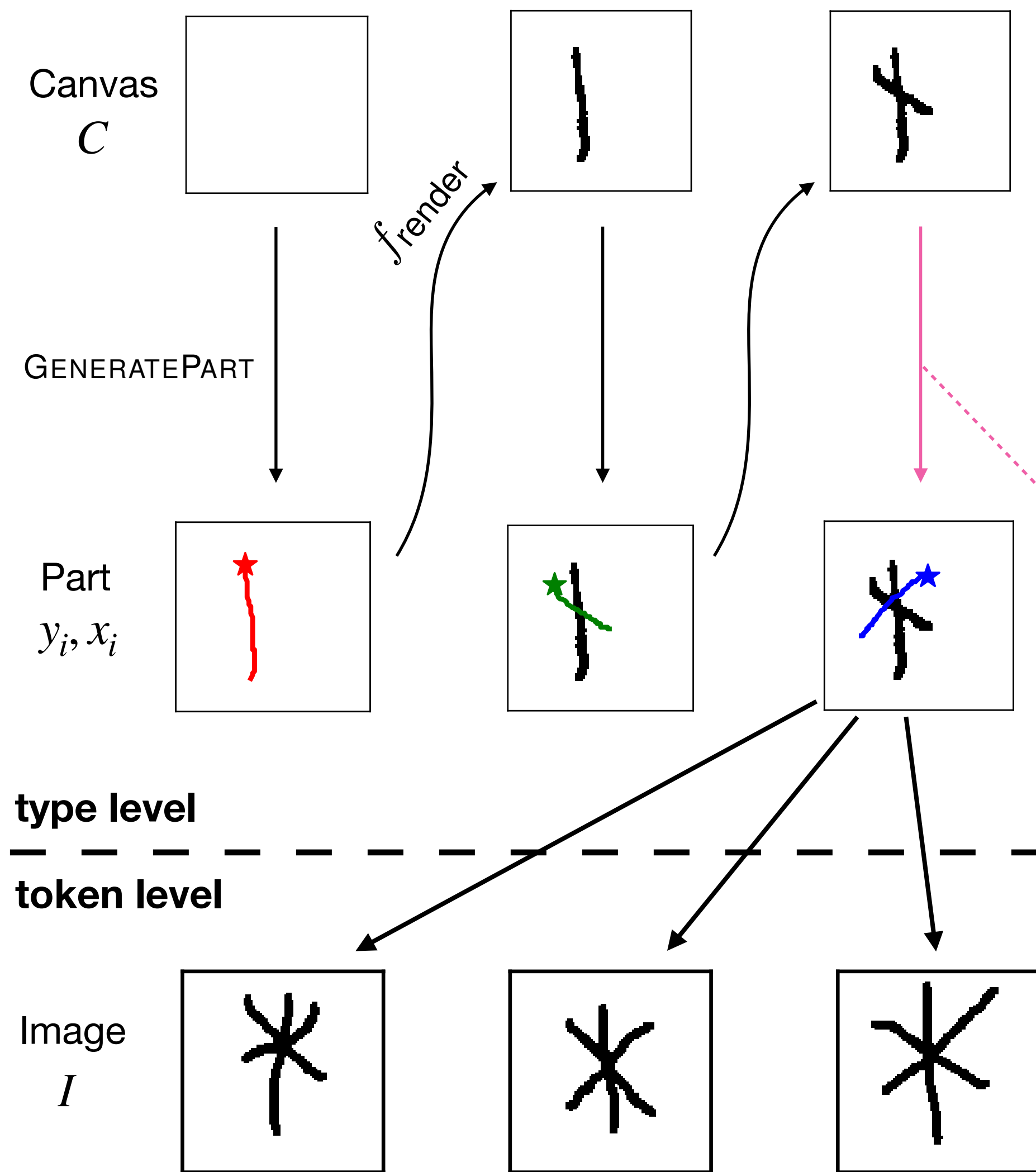
procedure GENERATETYPE
    $C \leftarrow 0$ ▷ Initialize blank image canvas
    while $true$ **do**
        $[y_i, x_i] \leftarrow$ GENERATEPART(C) ▷ Sample part location & parameters
        $C \leftarrow f_{\mathrm{render}}(y_i, x_i, C)$ ▷ Render part to image canvas
        $v_i \sim p(v \mid C)$ ▷ Sample termination indicator
        if v_i **then**
            break   ▷ Terminate sample
    $\psi \leftarrow \{\kappa, y_{1:\kappa}, x_{1:\kappa}\}$
    return ψ ▷ Return concept type

GENERATEPART(C)

location model $p(y \mid C)$ stroke model $p(x \mid y, C)$

Generating new concepts (unconstrained)

procedure GENERATETYPE

1. Evaluations on held-out concepts

Test loss (negative log-likelihood)

Neuro-Sym	**19.51**
H-LSTM	20.16
Baseline	19.66

Replicates across different train/test splits

2. Sampling new concepts

Humans

GNS model

fully-symbolic model (BPL)
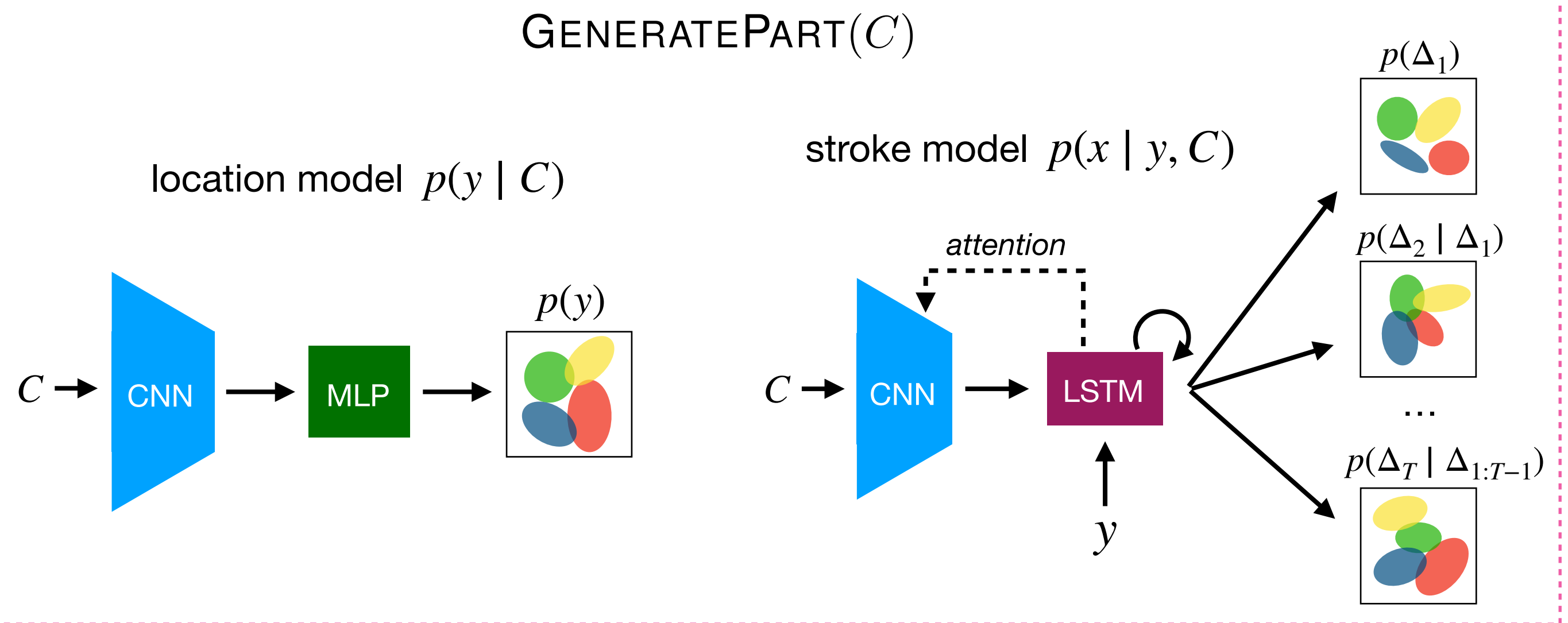
One-Shot Classification

Training item

probabilistic inference

Re-fitting

Test items

fit score: 25.3

fit score: 12.1
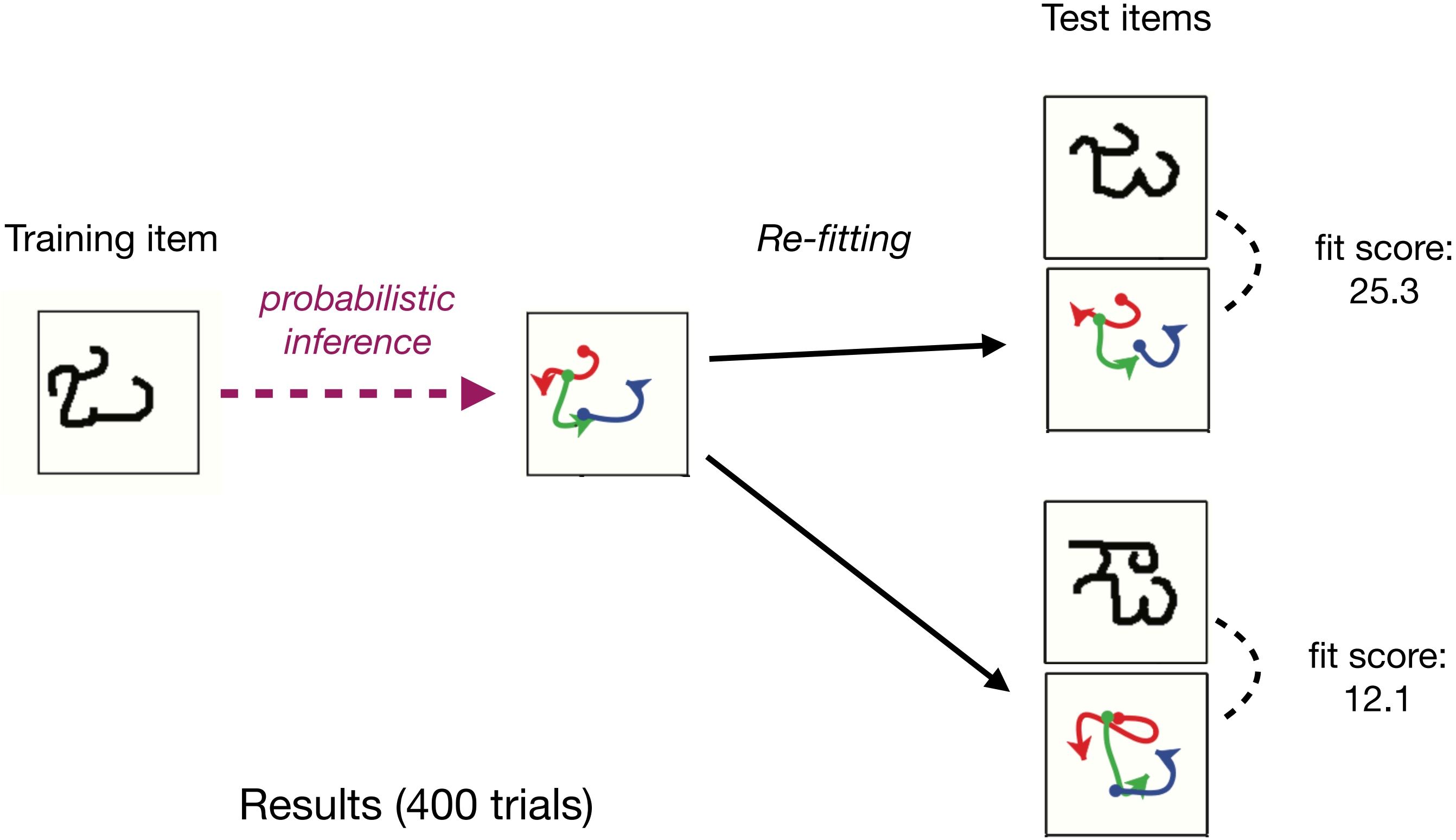
Results (400 trials)

	Accuracy
GNS	94.3%
Humans	95.5%

Parsing
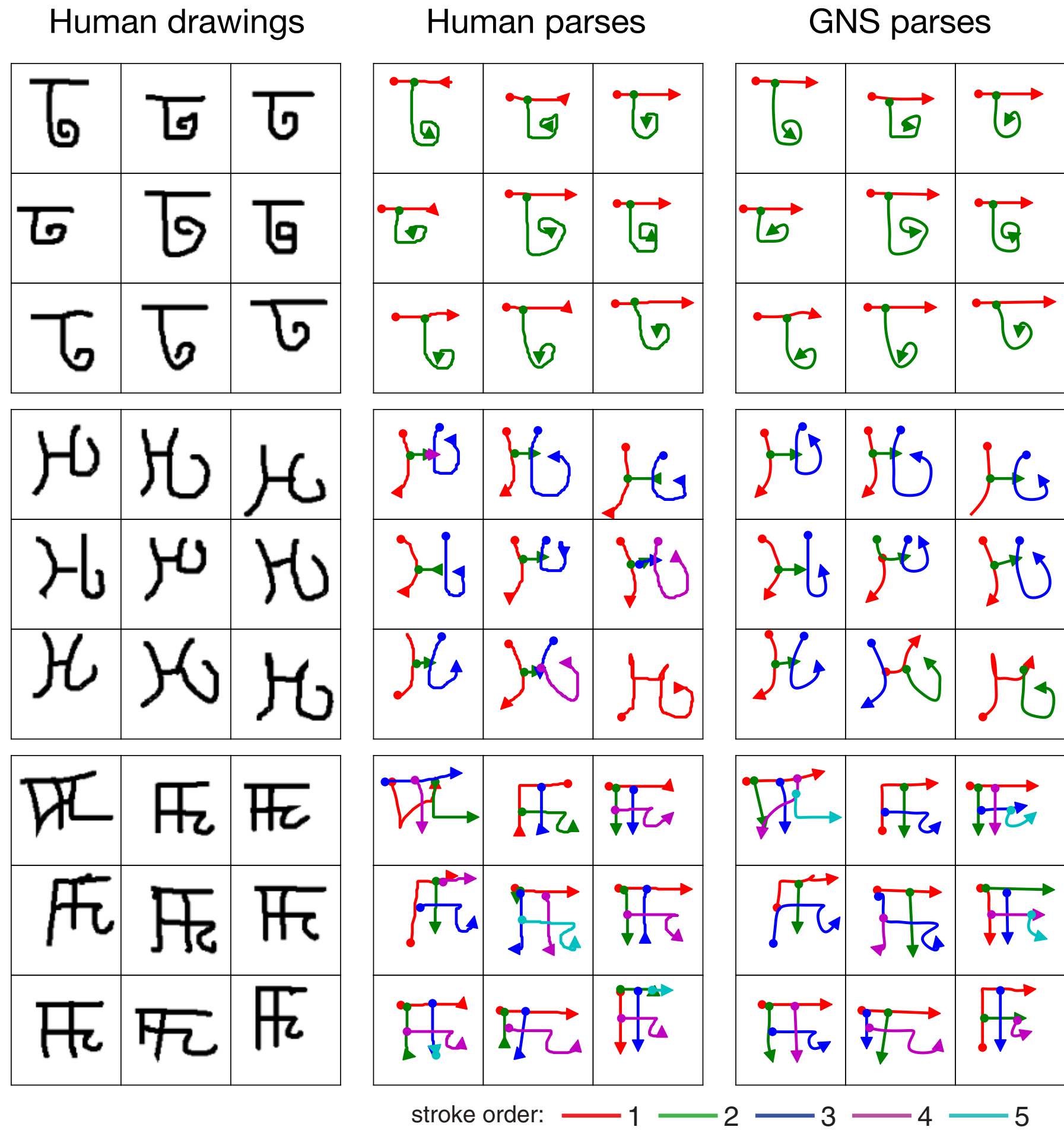
Generating new exemplars

end for
$A^{(m)} \leftarrow P(A^{(m)})$ ▷ Sample affine transform
$I^{(m)} \leftarrow P(I^{(m)}|T^{(m)}, A^{(m)})$ ▷ Sample image
return $I^{(m)}$

characters. (**A**) New types are generated by choosing primitive actions (color coded) from a library (i), ... (iii), and combining parts with relations to define simple programs (iv). New tokens are generated by running ... as raw data (vi). (**B**) Pseudocode for generating new types ψ and new token images $I^{(m)}$ for $m = 1, ..., M$. The ... ce and start location into a trajectory.

B Human drawings   Human parses   Machine parses

Human drawings

-695   -7...

Target

GNS parses

GNS

es. Parts are distinguis...
...ginning of a stroke an...
...w shows the five best...
...heir log-probability sc...
...ts. For classification, each...
...eft in image triplets),...
...n its image reconstruc...
...rior predictive probabi...
...much higher classifica...
...the best programs indu...
...wings of three charac...
...s (middle) and best m...

Human

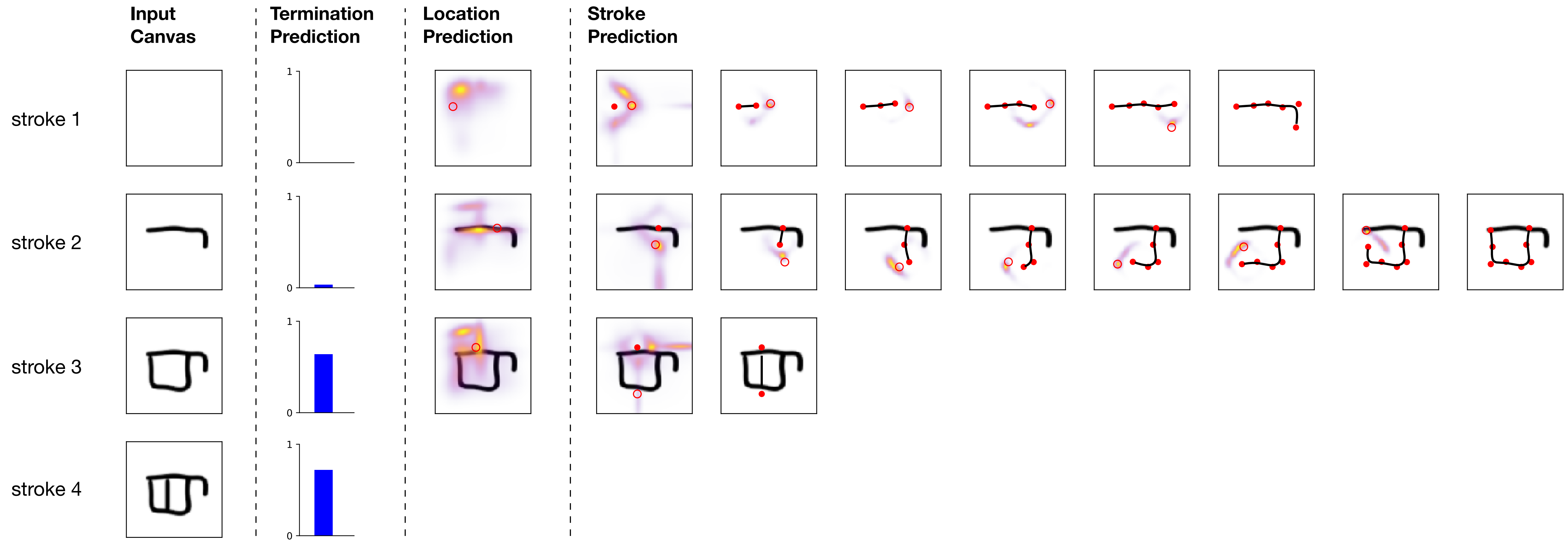stroke order: — 1 — 2 — 3 — 4 — 5 — 4 — 5
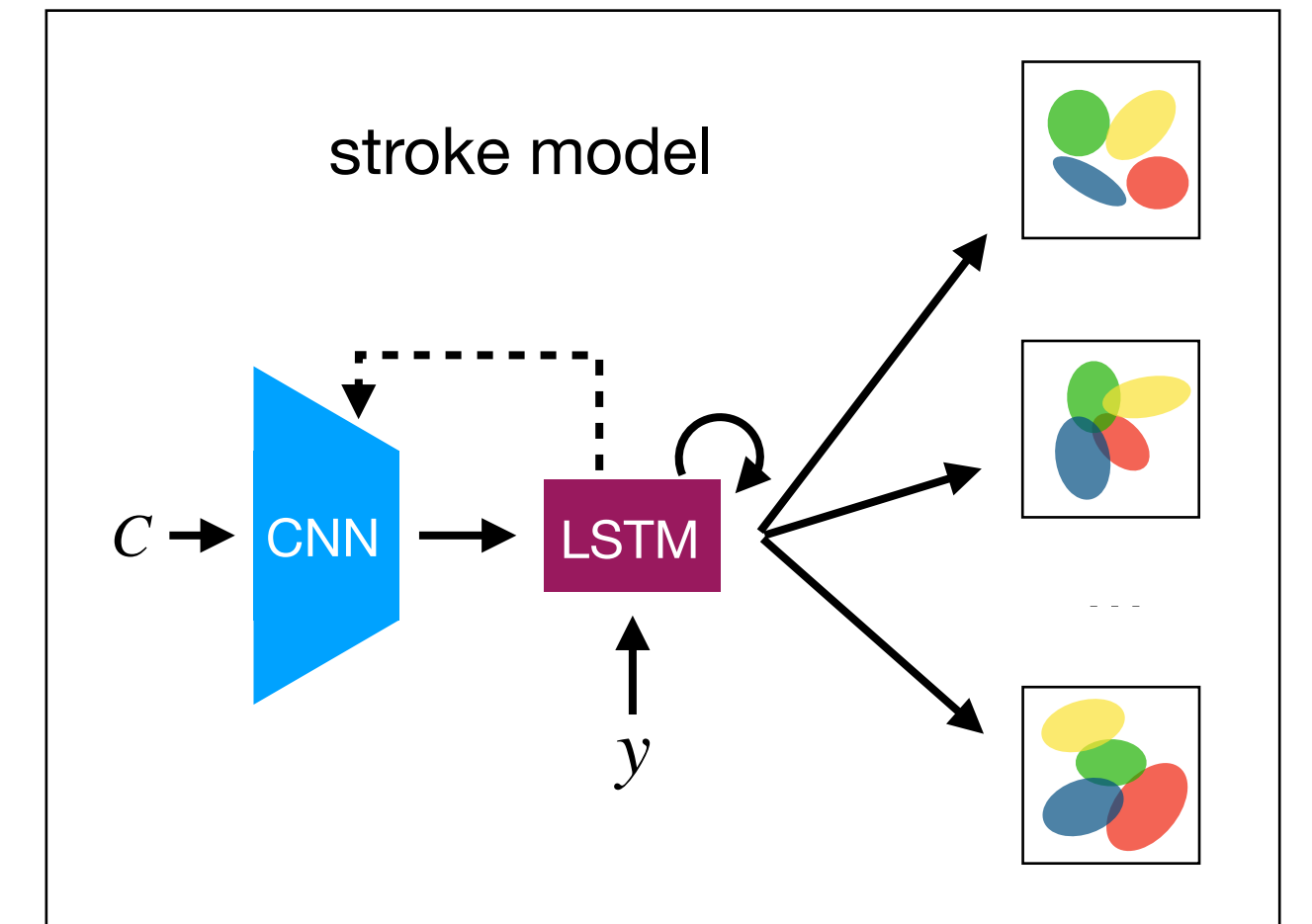
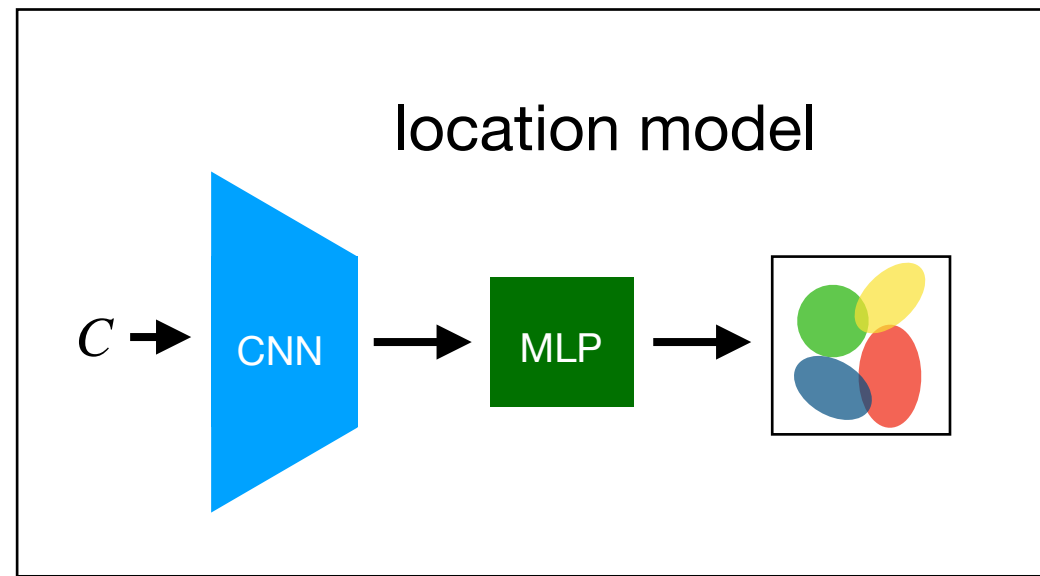More examples in Feinman & Lake (2020)

Thanks
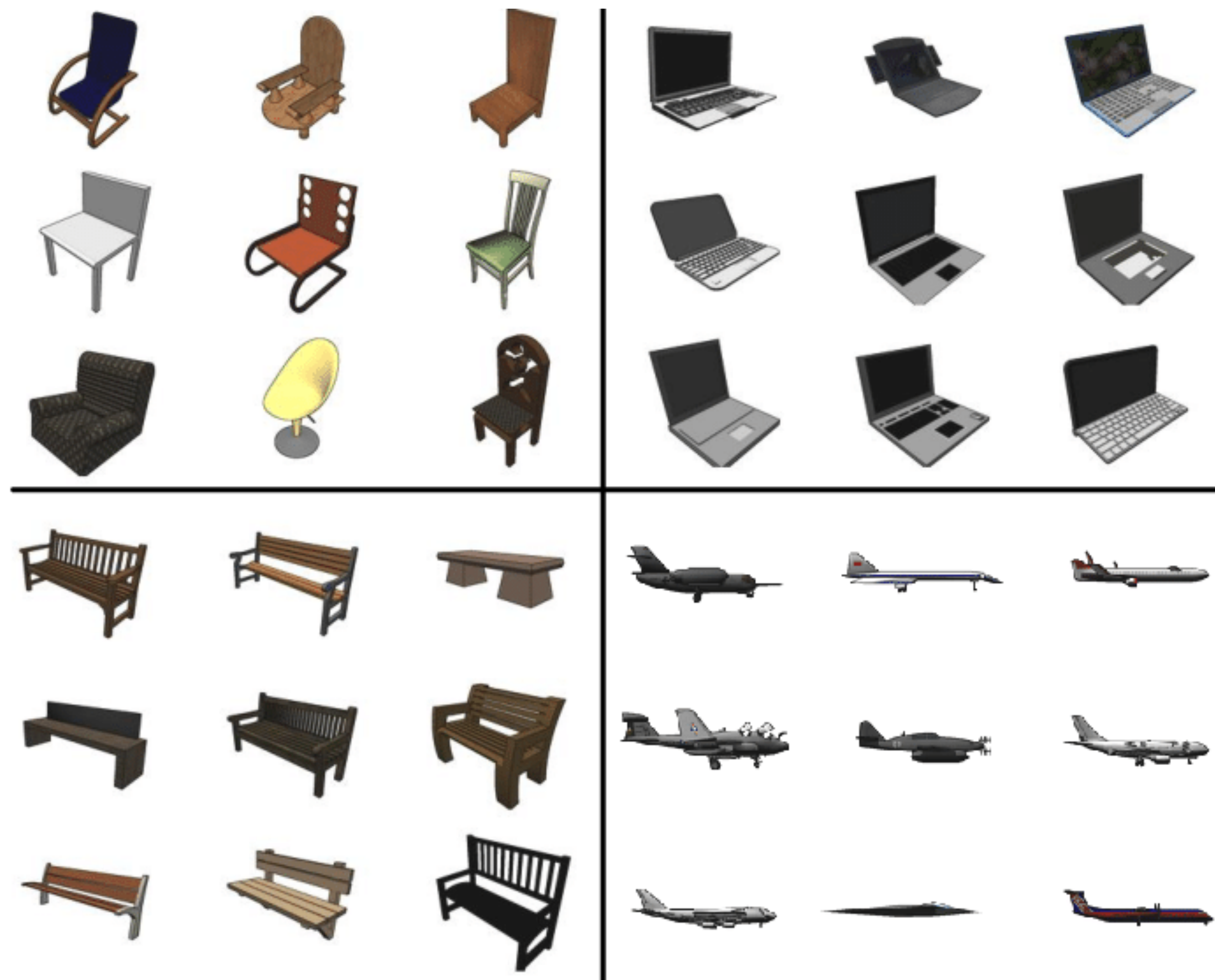
For more information, come to my poster!

Extras

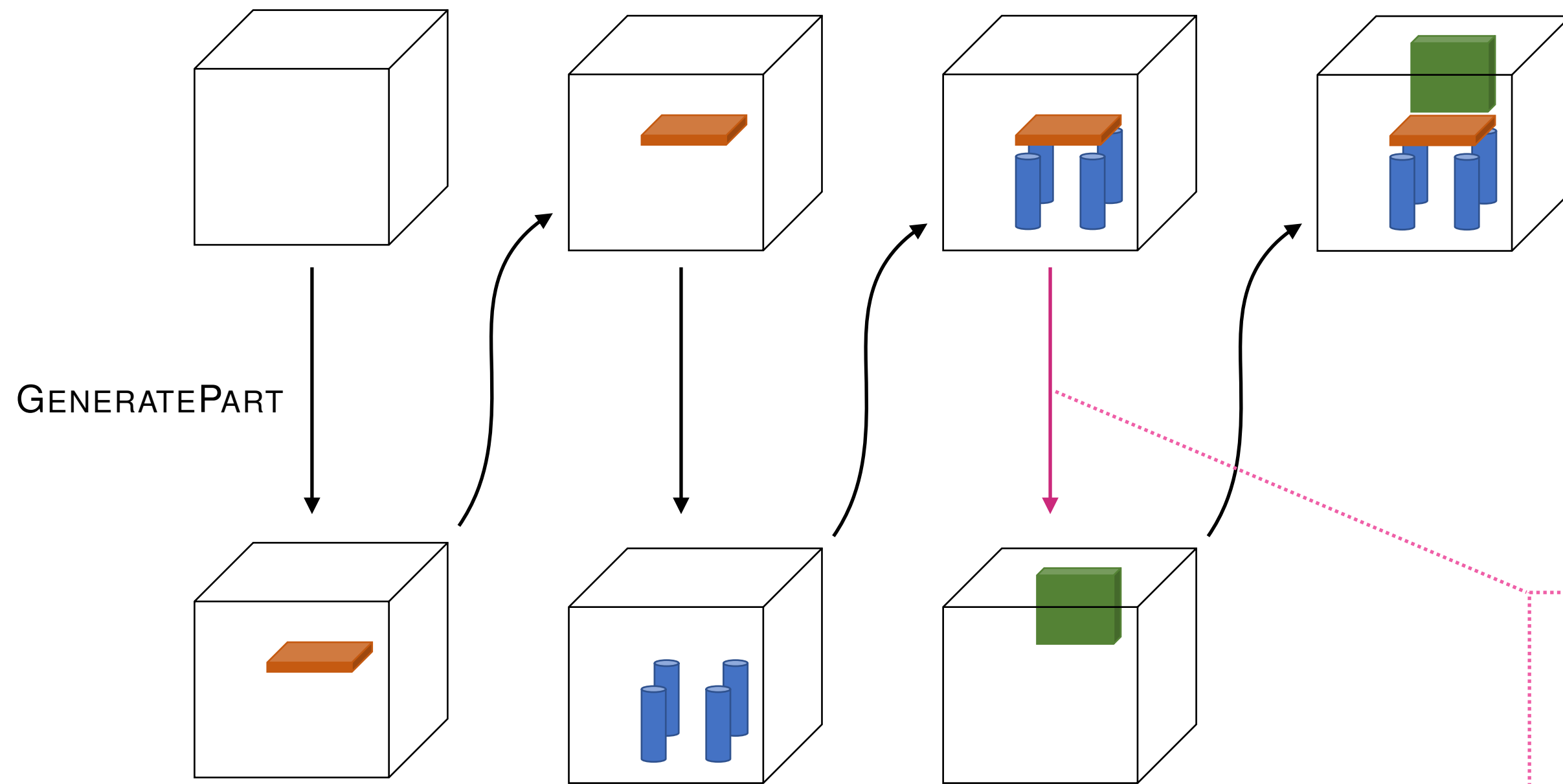# Type Prior

3D objects (ShapeNet)

GNS model of 3D objects


```
procedure GENERATETOKEN
    C ← 0                                    ▷ Initialize blank 3d canvas
    while true do
        x_i ← GENERATEPART(C)                ▷ Sample part from neural net
        C ← f_update(x_i, C)                 ▷ Update 3d canvas
        v_i ~ p(v | C)                       ▷ Sample termination indicator
        if v_i then
            break                            ▷ Terminate sample
    θ ← {κ, x_{1:κ}}
    return ψ                                 ▷ Return concept token
```

GENERATEPART(C)

GENERATEPART

generate token

category model $p(k \mid C)$

instance model $p(x \mid k, C)$

$p(z \mid k, C)$

$q(z \mid x, k, C)$

$p(x \mid z)$

$p(k \mid C)$