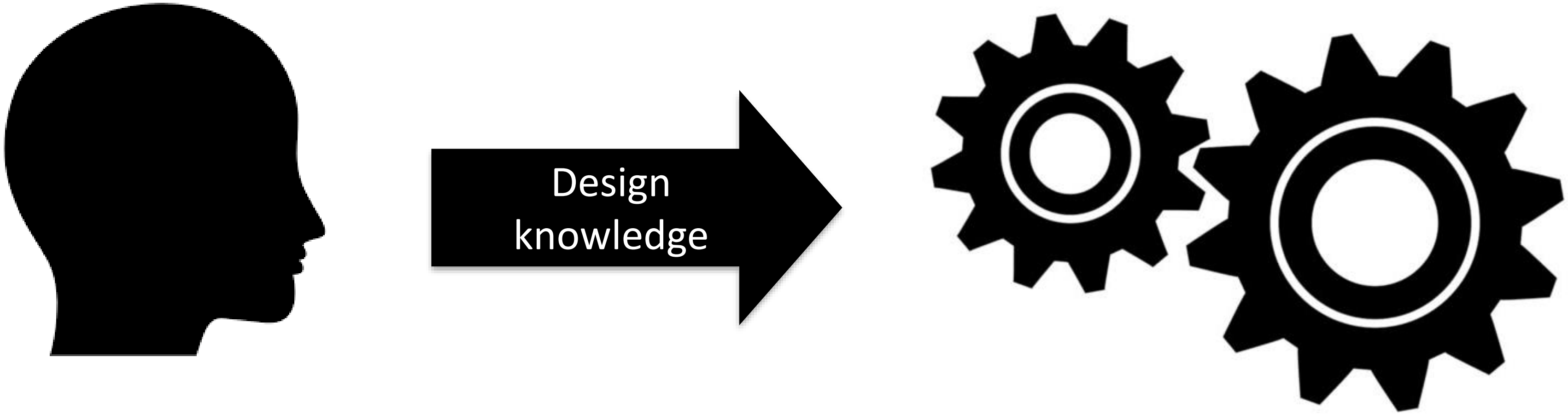
Creative Procedural Content Generation via Machine Learning

Matthew Guzdial

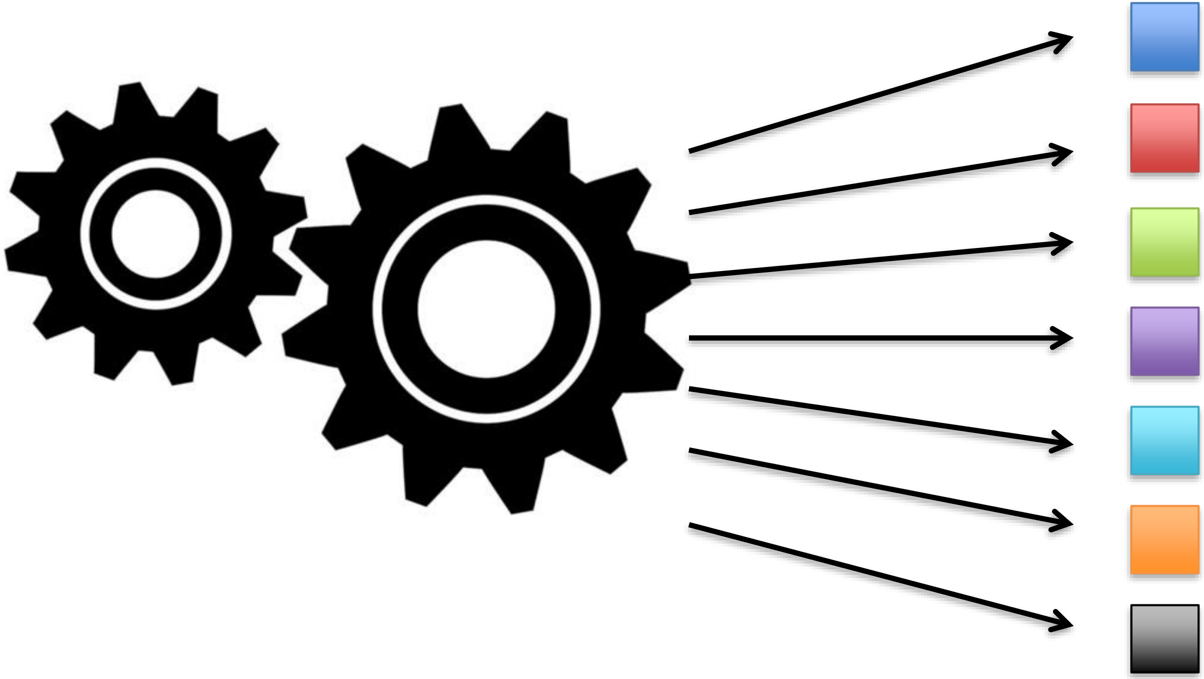
Procedural Content Generation

Matthew Guzdial

Procedural Content Generation



Procedural Content Generation



Generated Content

- Structure
- Story
- Quests
- Art assets
- Music
- Etc…

Levels

Diablo 3

Spelunky

Minecraft

Skyrim, Initial Generation

Bloodborne, Chalice Duneons

Characters, Story, and Quests

Fallen London

The Shrouded Isle

Skyrim, Radial Quests

Shadow of Mordo

The Promise of PCG

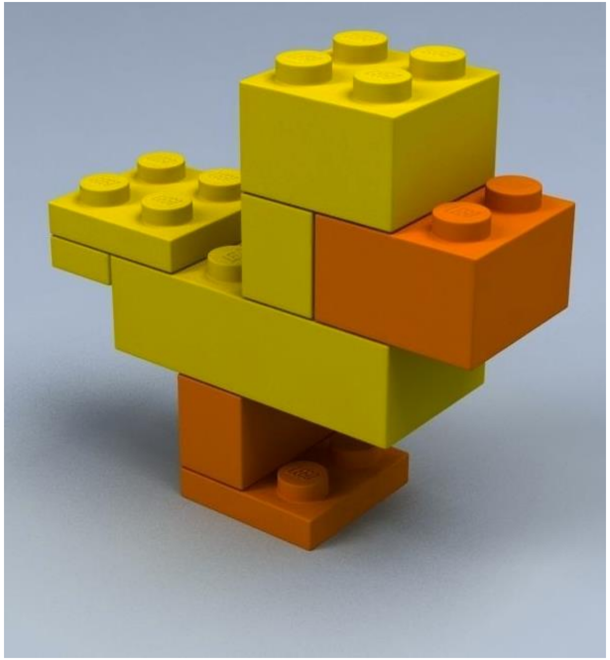
There was a myth in the industry 5-10 years ago that PCG could save costs.


MASS
EFFECT
ANDROMEDA
FAIL
We shouldn't worry.

Why is PCG so Tough?

Example approach: Grammar

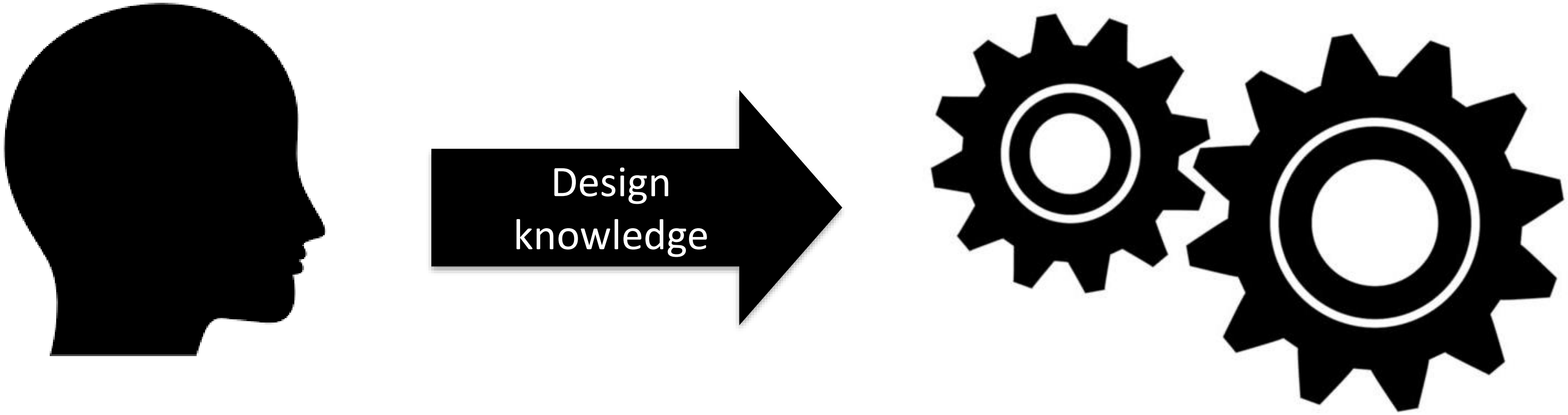
Imagine instead of creating a sculpture from clay you came up with instructions to make it from Lego Blocks.



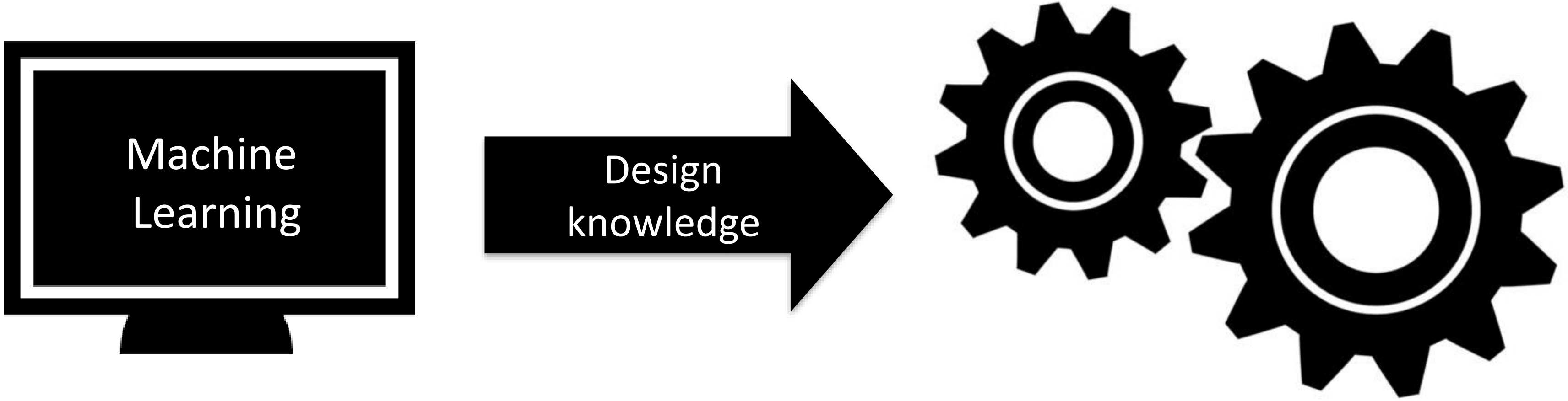
Drawbacks of traditional PCG

- Requires a design expert
- Requires a CS expert
- Encoding design knowledge time
- Design iterations more costly

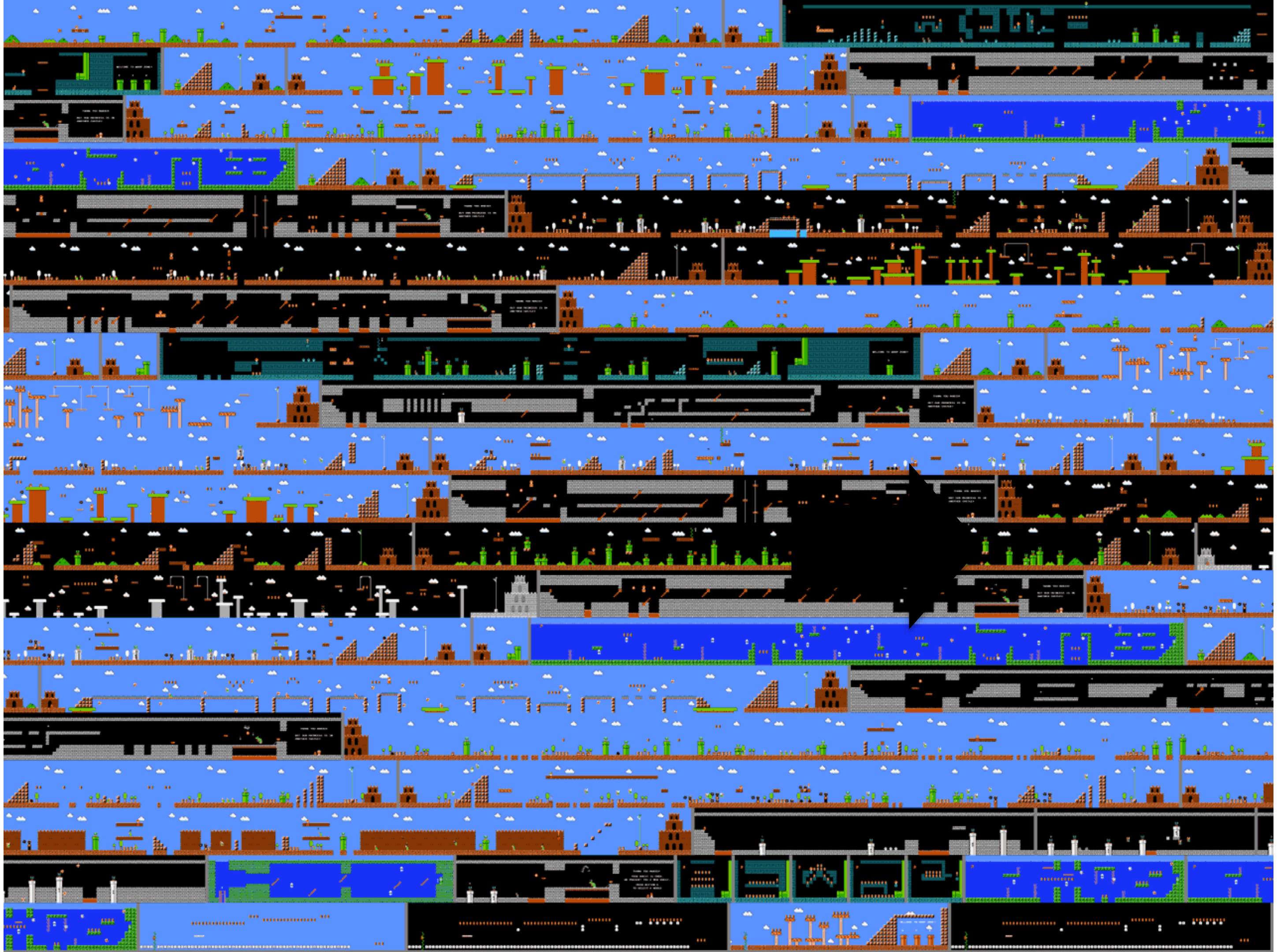
What if we could automate encoding the expert design knowledge?



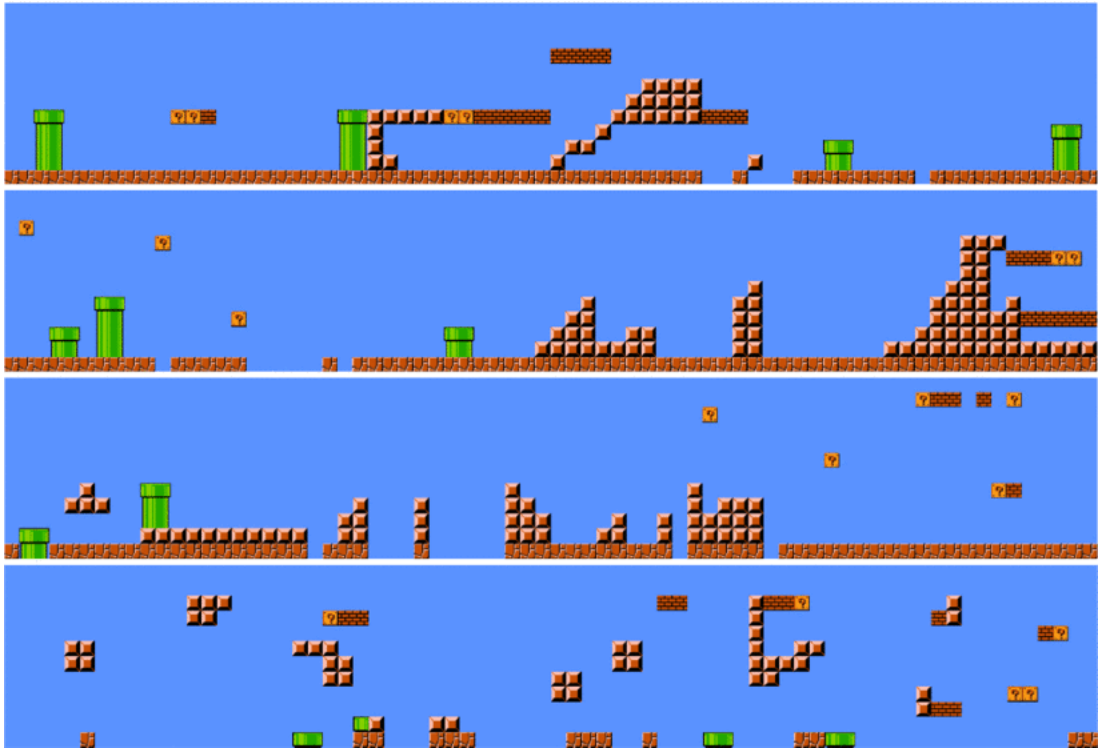
What if we could automate encoding the expert design knowledge?



Procedural Content Generation **via Machine Learning**

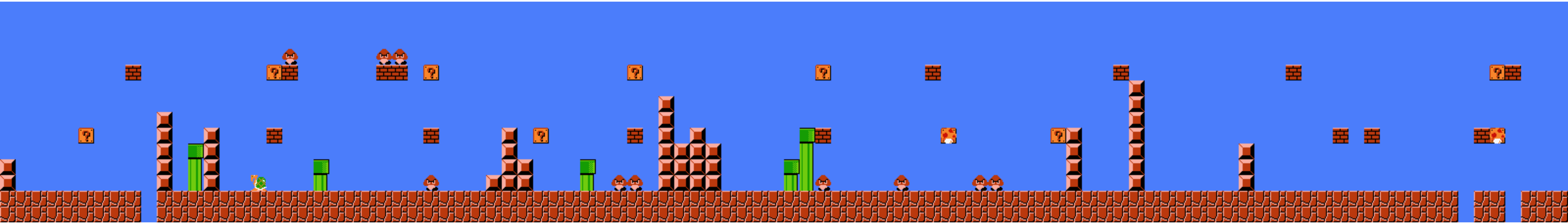


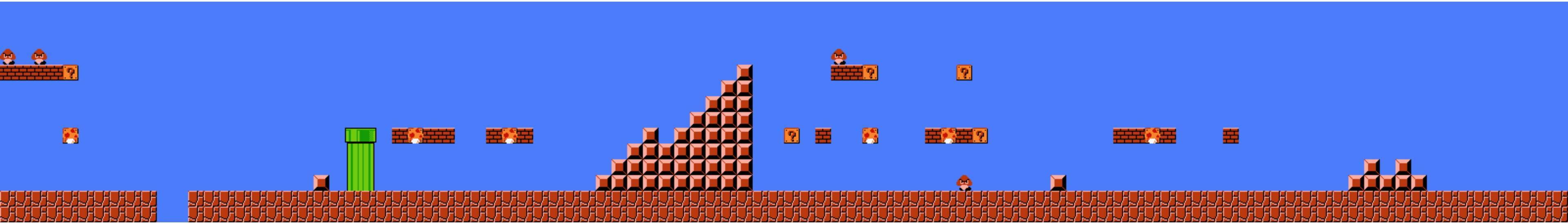
Markov Chains



Snodgrass and Ontañón. Learning to generate video game maps using Markov models.

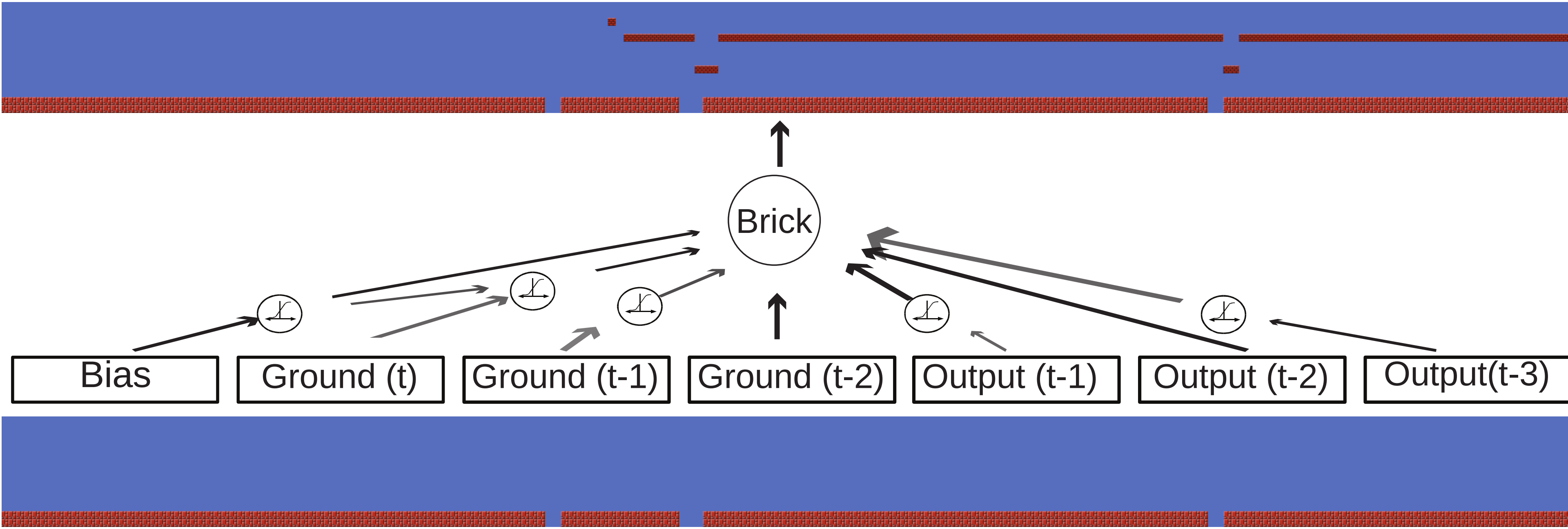
Genetic Algorithms





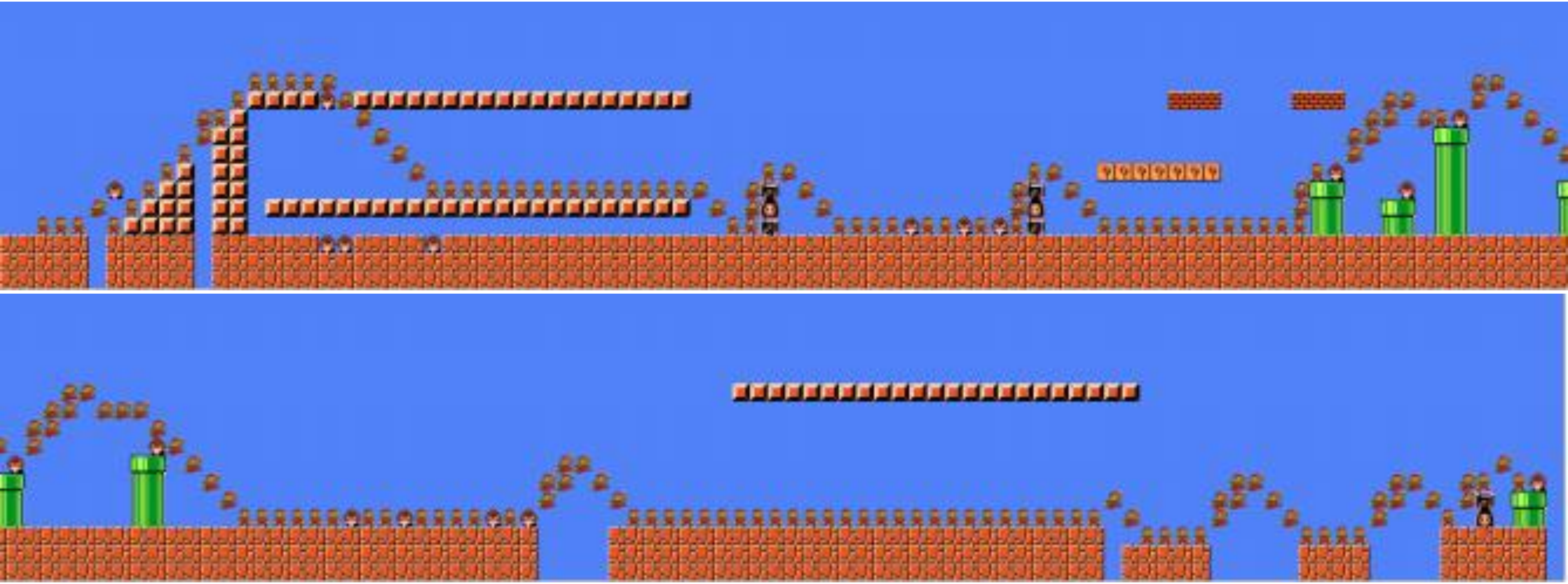
Dahlskog and Togelius. Patterns as objectives for level generation.

NeuroEvolution



Hoover et al. Composing Video Game Levels with Music Metaphors through Functional Scaffolding

LSTMs

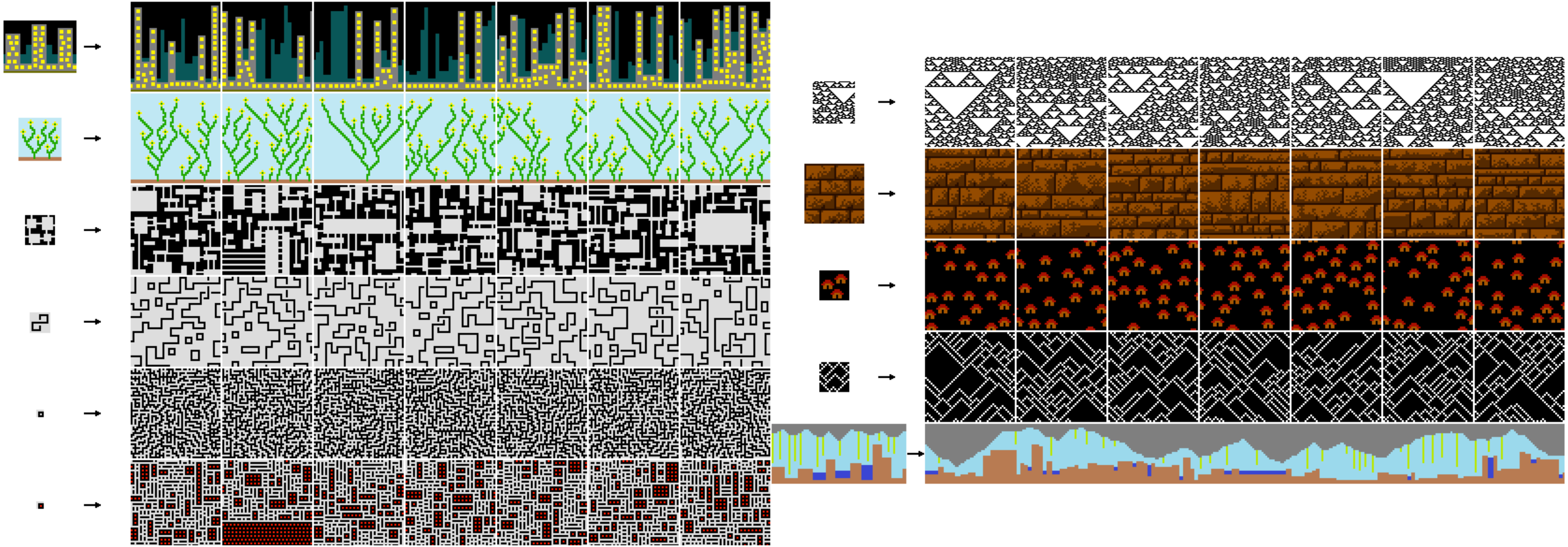


Summerville and Mateas. “Super Mario as a String: Platformer Level Generation via LSTMs”.

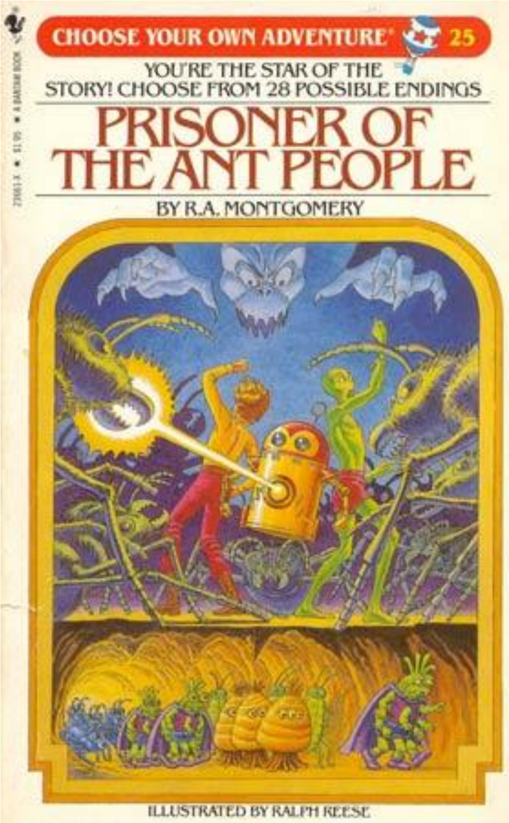
Existing Data

- Many techniques for spatial structural content
- Reliant on human effort to translate content to digital form
- Produces output content similar to input content

Wave Function Collapse



Scheherazade-IF System



amazon mechanical turk™
Artificial Artificial Intelligence

You pushed hard to open the heavy bank door.
You put on a joker mask that covered your entire face.
You drove to the bank in the off-peak hour.
You walked into the bank, trying to look normal. Your pulse quickened.
-Wait in the teller line
-Scan the bank
-Look around for a teller

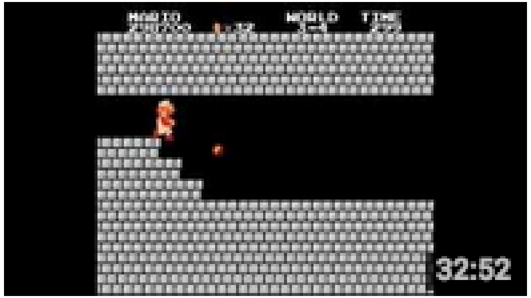
Building Own Corpus

- Less explored
- Does not rely on existing corpus
- Significant effort and requires the same design expertise as traditional PCG

Alternative Training Data

Learn through some secondary domain by converting data into the desired domain

Alternative Training Data: Video

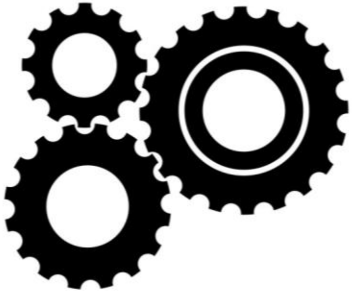


NES Longplay [250] Super Mario Bros (a)
by World of Longplays
2 years ago • 139,443 views
http://www.longplays.org Played by: ScHlAuChi Quite a few people didnt seem to like MontyMoles longplay of Super Mario Bros ...

Benefits

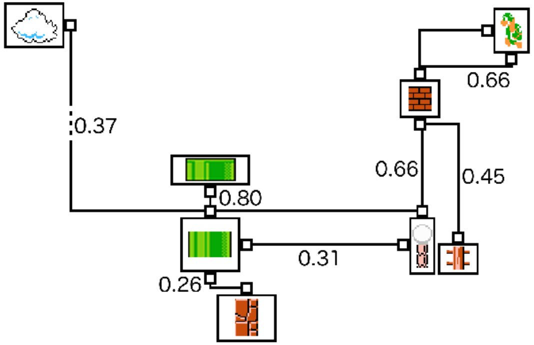
- Existing corpora not always accessible
- Contains player-game element interactions
- Everything to fully recreate a game

Game Learning from Video

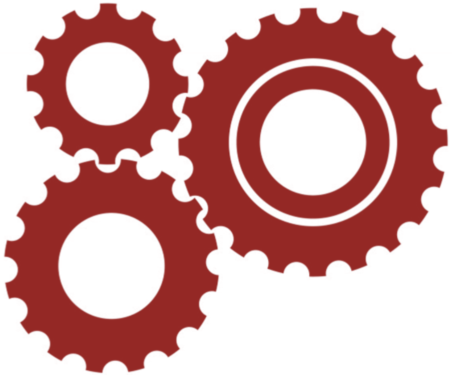


Authored Game Engine + Levels

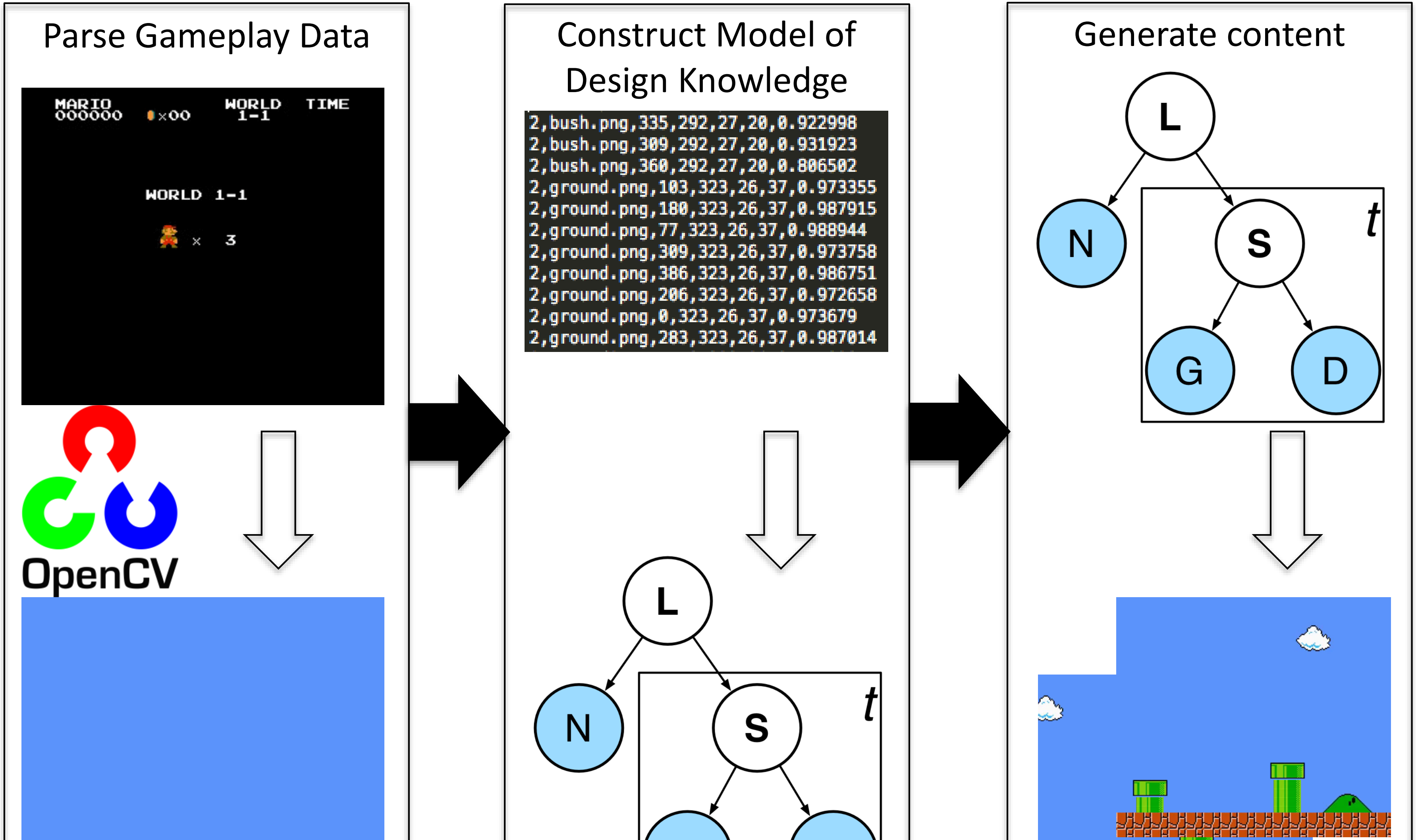
Output Gameplay Video



Learned Level Design

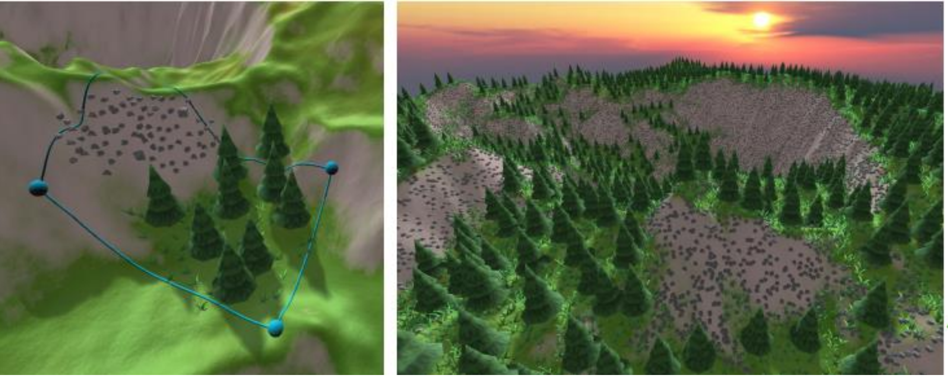


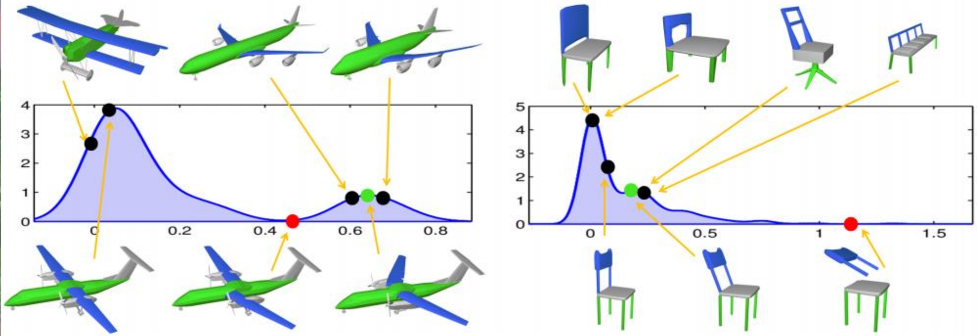
Learned Game Engine


Parse Gameplay Data
MARIO
000000
×00
WORLD
1-1
TIME
WORLD 1-1
× 3
OpenCV
Construct Model of Design Knowledge
2,bush.png,335,292,27,20,0.922998
2,bush.png,309,292,27,20,0.931923
2,bush.png,360,292,27,20,0.806502
2,ground.png,103,323,26,37,0.973355
2,ground.png,180,323,26,37,0.987915
2,ground.png,77,323,26,37,0.988944
2,ground.png,309,323,26,37,0.973758
2,ground.png,386,323,26,37,0.986751
2,ground.png,206,323,26,37,0.972658
2,ground.png,0,323,26,37,0.973679
2,ground.png,283,323,26,37,0.987014
L
N
S
t
Generate content
L
N
S
G
D
t

Generating Mario Levels

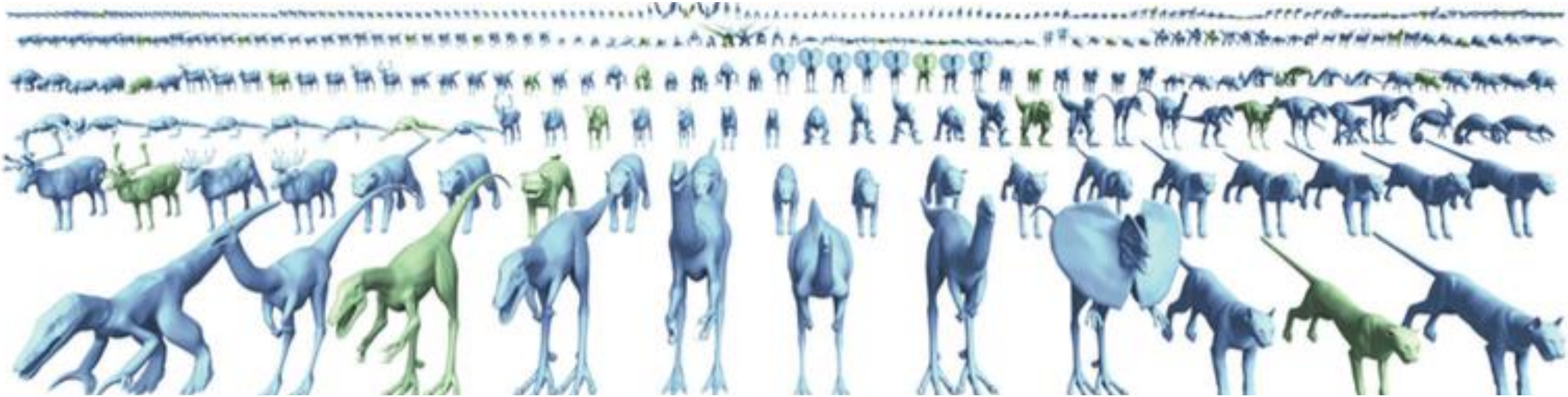
PCG Research	Source of Design Knowledge	Generative Approach
Mawhorter & Mateas 2010	Human author	Grammar-based
Launchpad 2011	Human author	Grammar-based
Kerssemakers et al 2012	Human author	Evolutionary
Shaker et al 2012	Human author	Evolutionary
Dahlskog & Togelius 2014	Level definitions	Evolutionary
Snodgrass & Ontañón 2014	Level definitions	Markov chains
Summerville & Mateas 2016	Level definitions	LSTM
	Gameplay Video	Clustering/Probabilistic Models





Fish, N., Averkiou, M., Van Kaick, O., Sorkine-Hornung, O., Cohen-Or, D., & Mitra, N. J. (2014). Meta-representation of shape families. ACM Trans. Graph., 33(4), 34-1. Chicago

Emilien, A., Vimont, U., Cani, M. P., Poulin, P., & Benes, B. (2015). Worldbrush: Interactive example-based synthesis of procedural virtual worlds. ACM Transactions on Graphics (TOG), 34(4), 106.



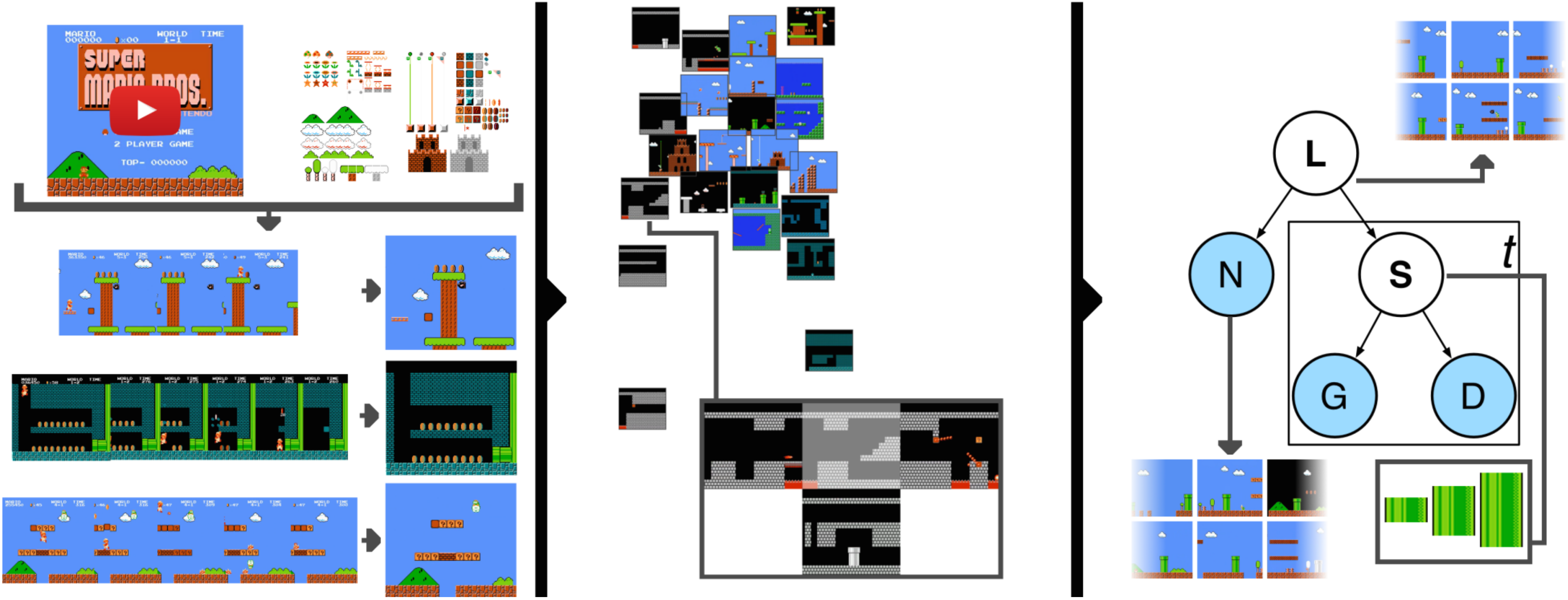
Kalogerakis, E., Chaudhuri, S., Koller, D., and Koltun, V. 2012. A probabilistic model for component-based shape synthesis. *ACM Transactions on Graphics (TOG)*, *31*(4), 55.

System Overview

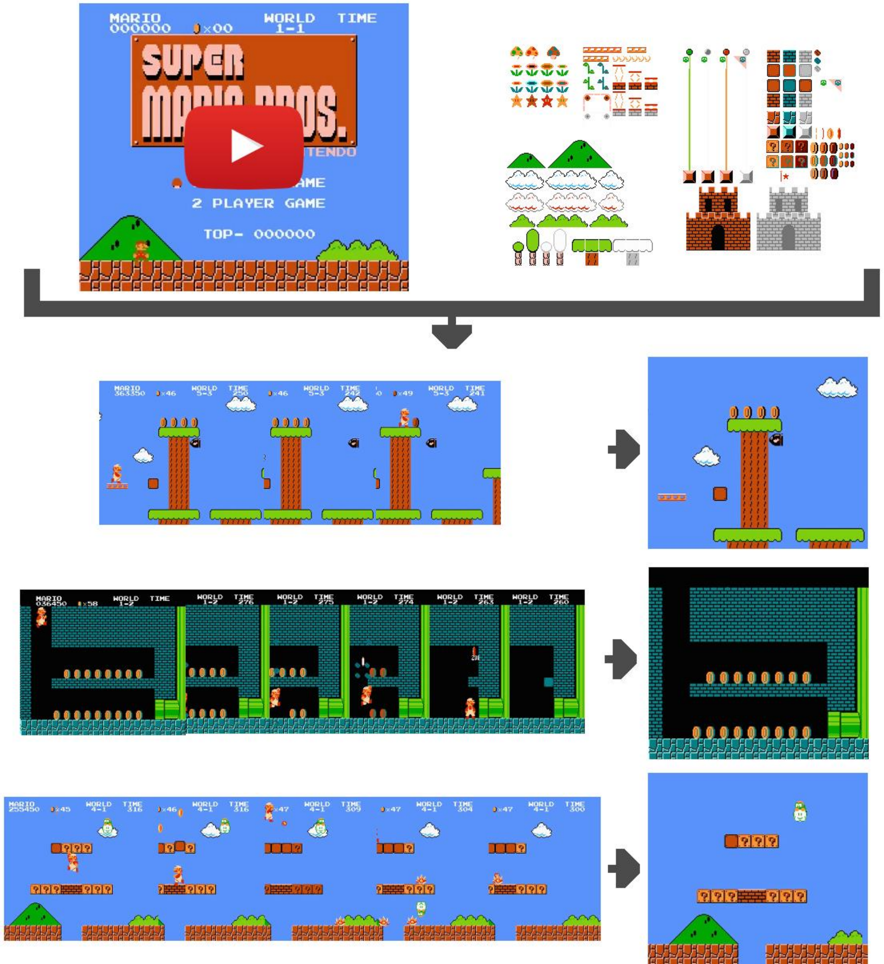
Deriving Level Chunks

Categorizing Level Chunks

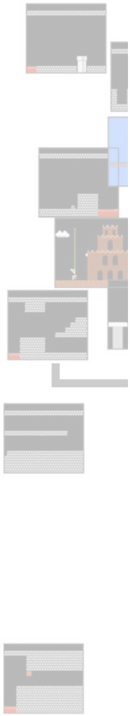
Probabilistic Model



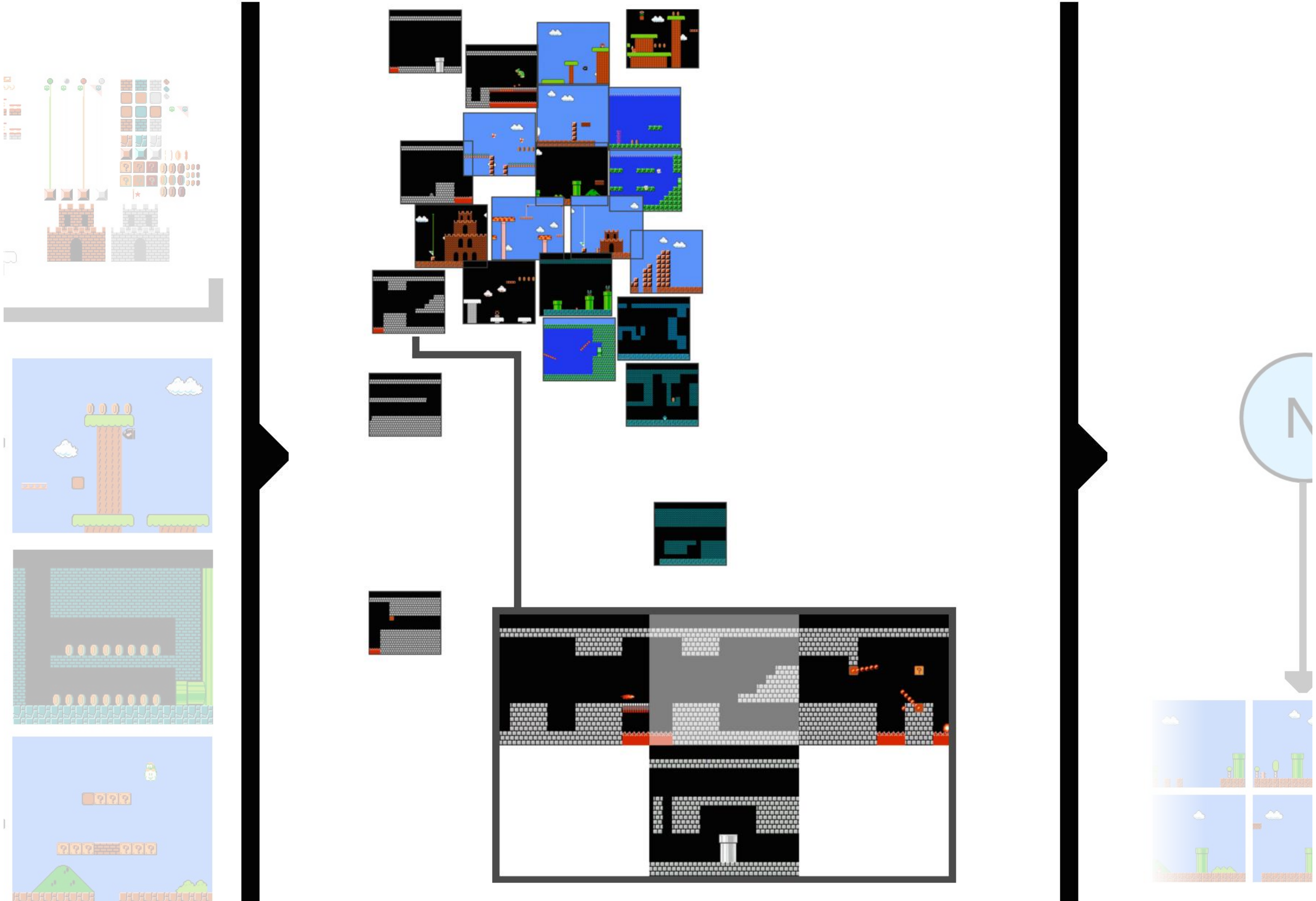
Deriving Level Chunks



Cate

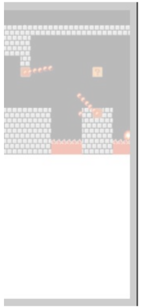


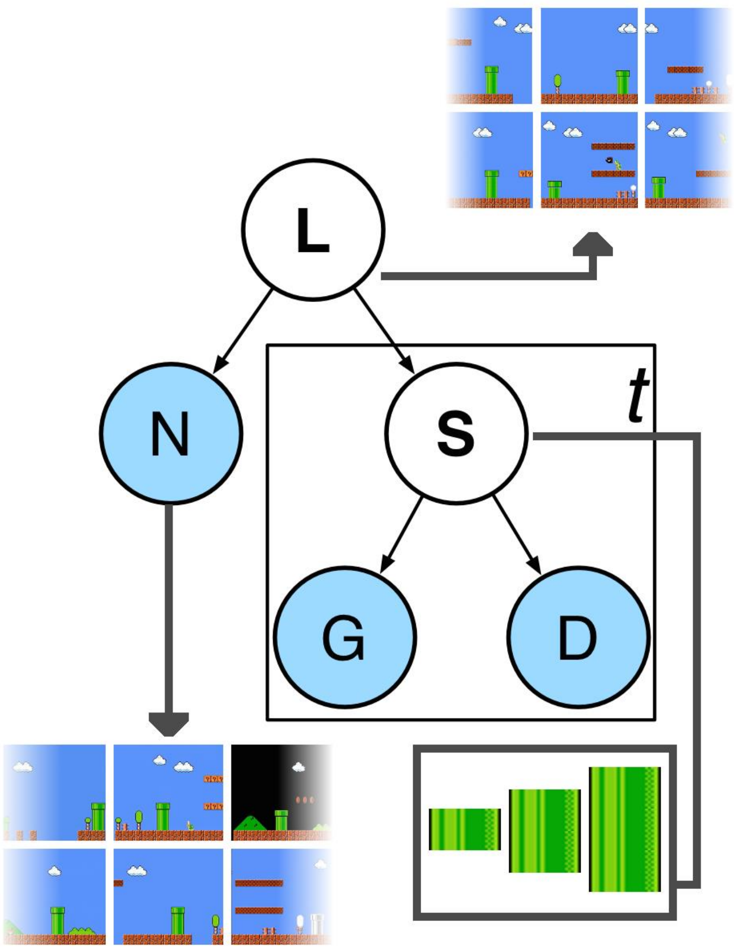
Categorizing Level Chunks



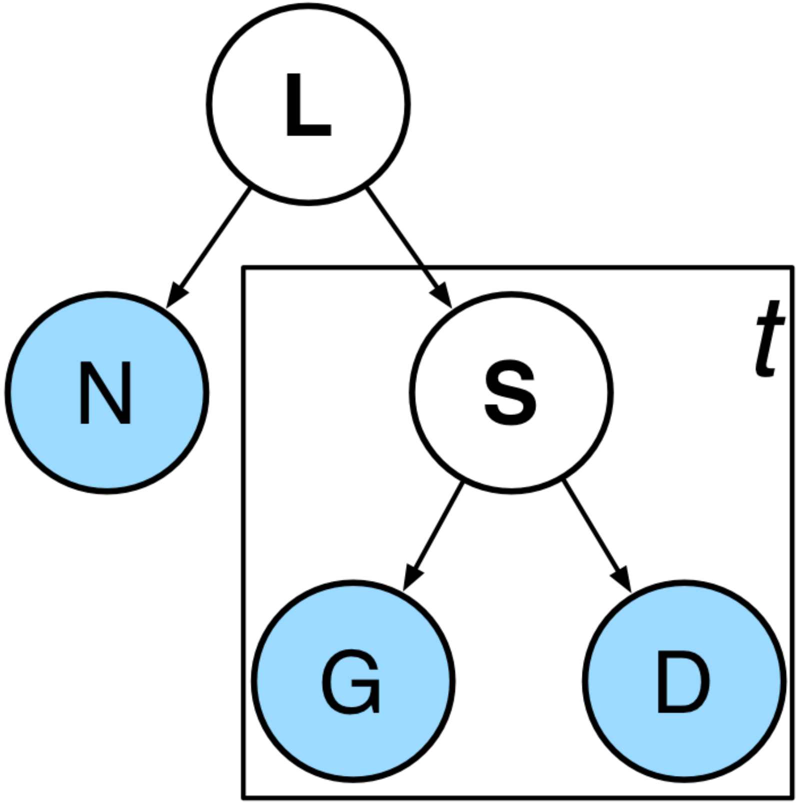
unks

Probabilistic Model

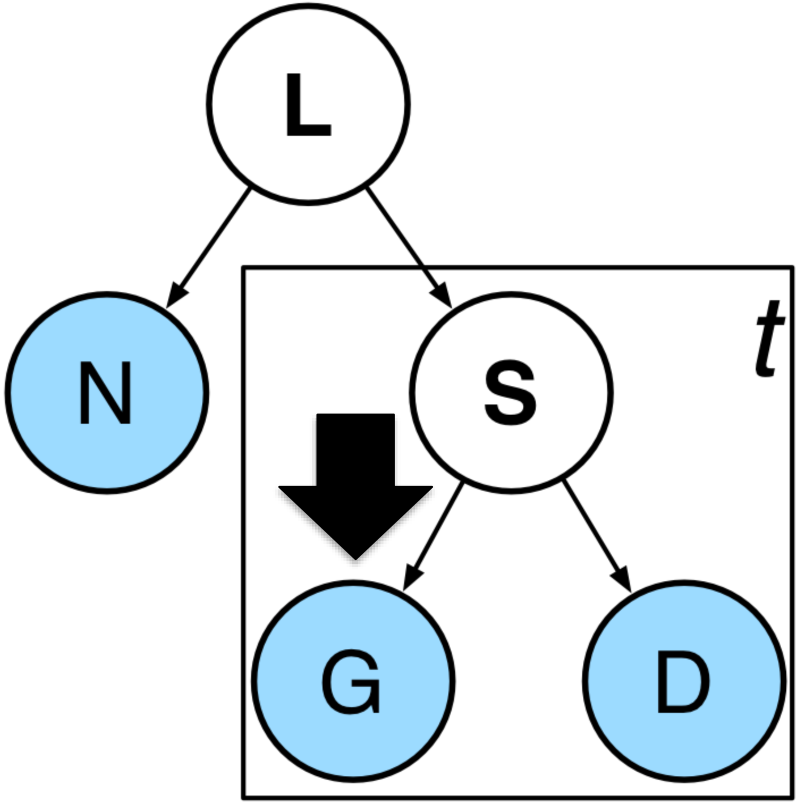


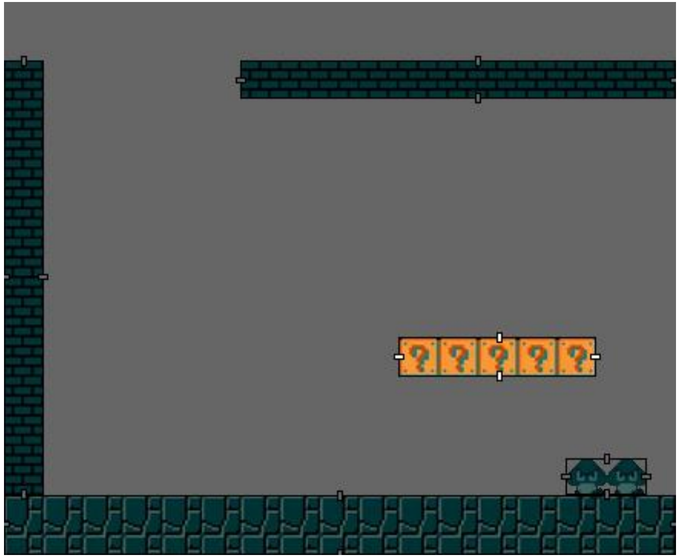


Probabilistic Model

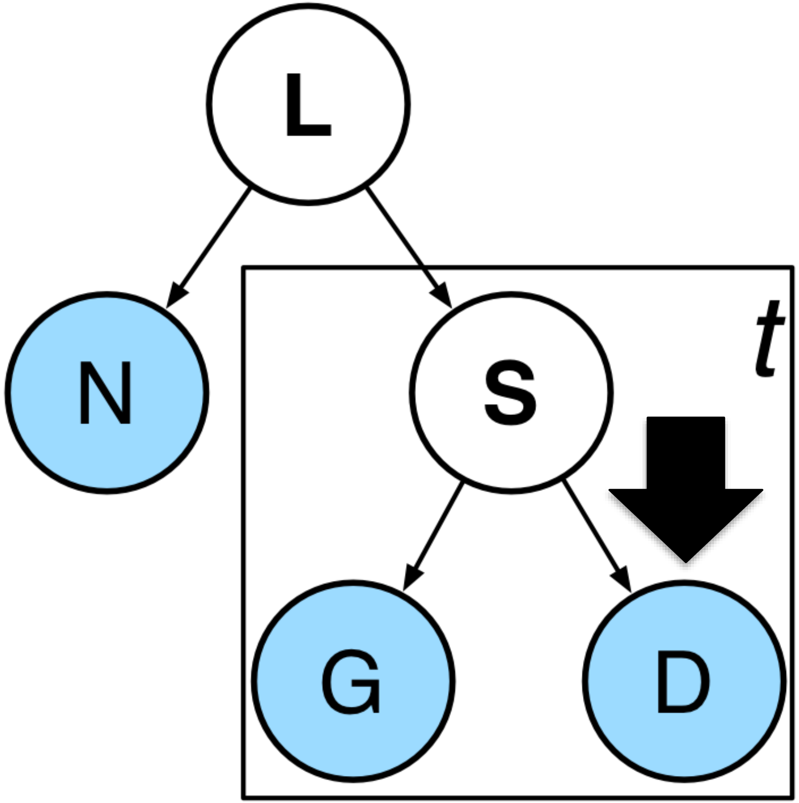


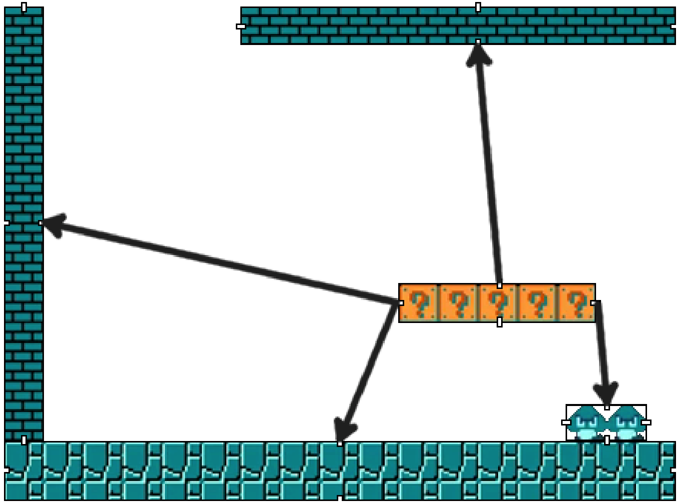
G : Geometric information for shapes of sprite t



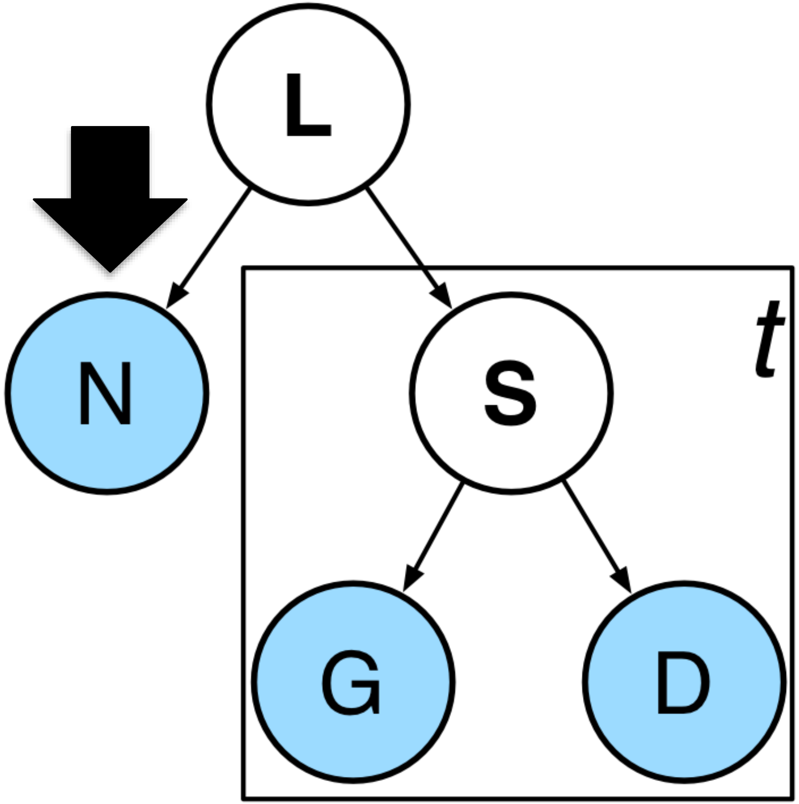


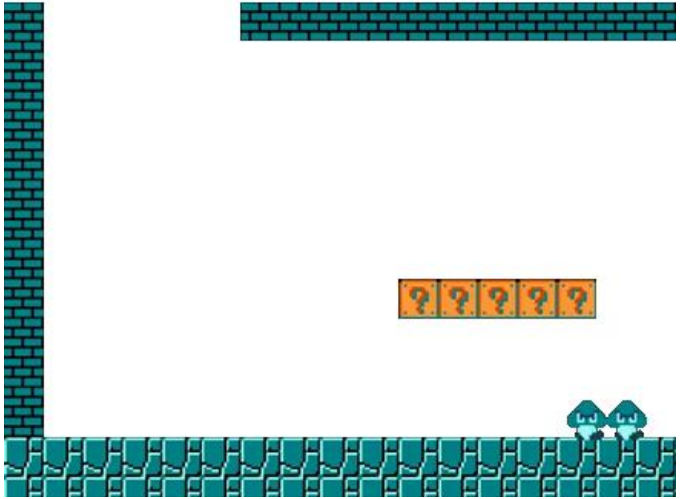
D : Relationships between sprite shape *t* and all other sprite shapes





N: Counts for sprites found in initial data set.





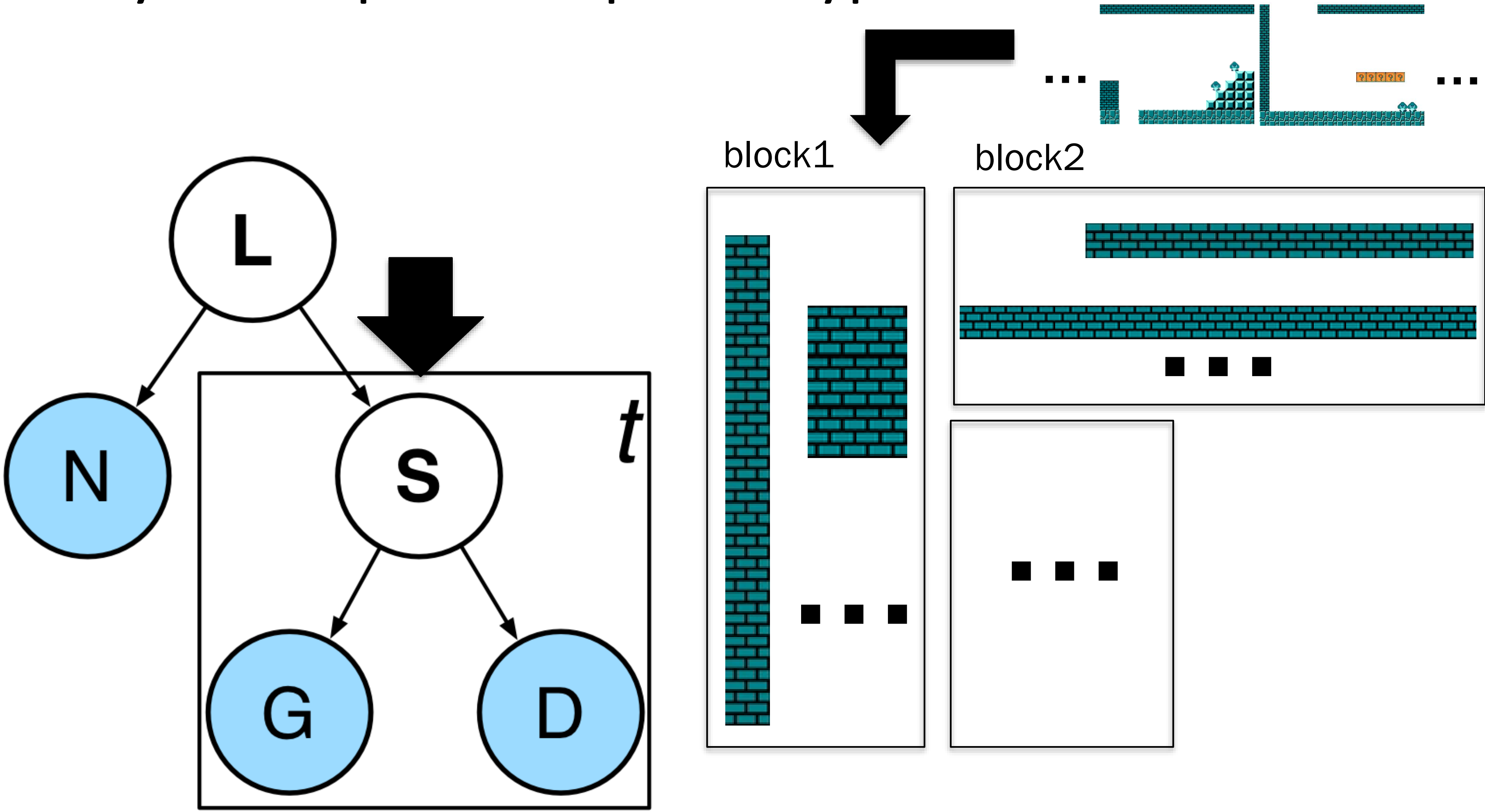
block: 22

ground: 17

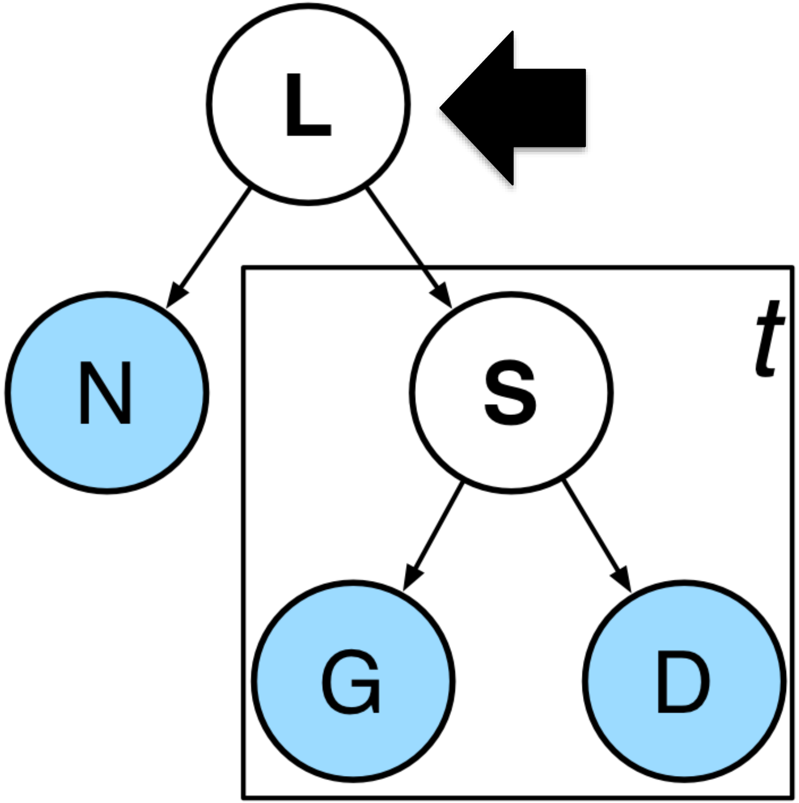
question-block: 5

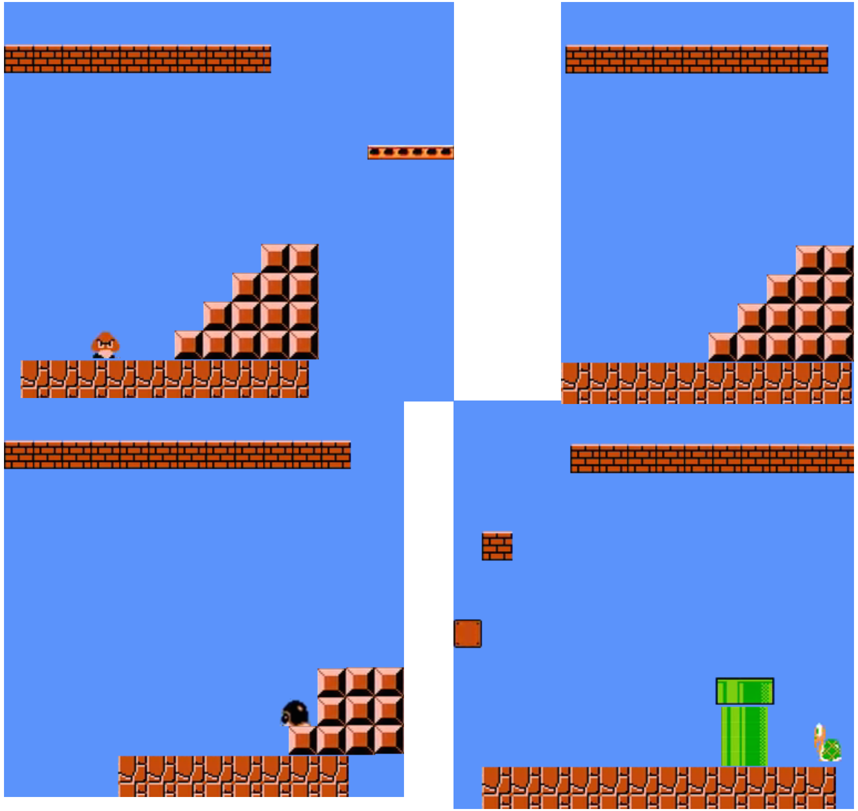
goomba: 2

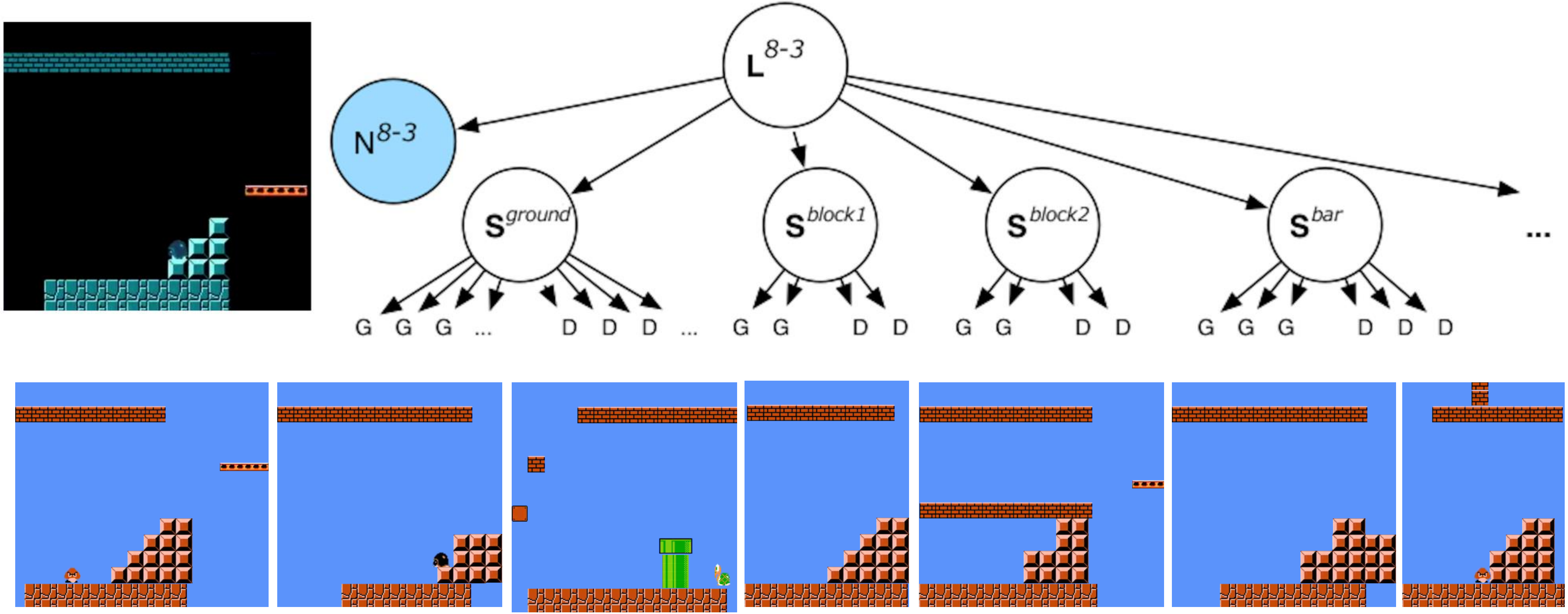
S: Styles of sprite shapes of type t



L: Level chunk styles (the styles of sprite shapes and the templates they can create).







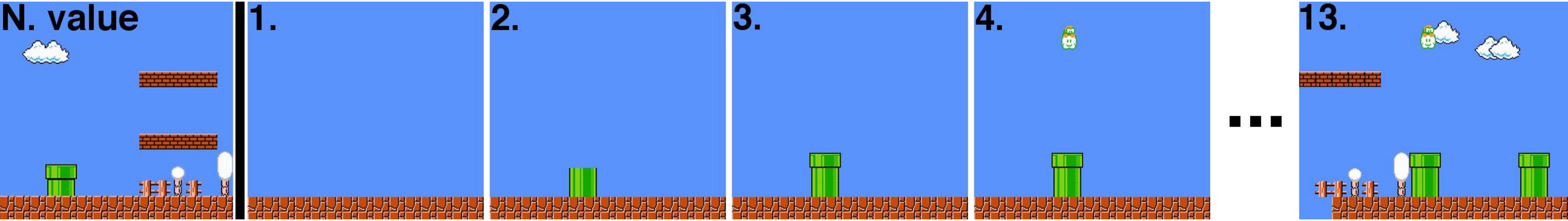
L^{8-3}
N^{8-3}
S^{ground}
S^{block1}
S^{block2}
S^{bar}
...
G G G ... D D D ...
G G D D
G G D D
G G G D D D

Probability Distribution

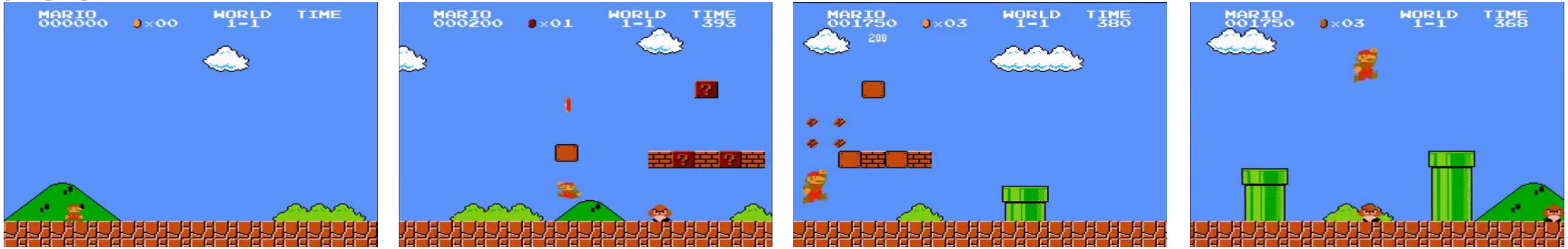
Build the probability distribution $P(g_{s1},r_d|g_{s2})$ for each L node.

Using Baye's law get $P(g_{s2}|g_{s1},r_d)$ and generate.

Generation



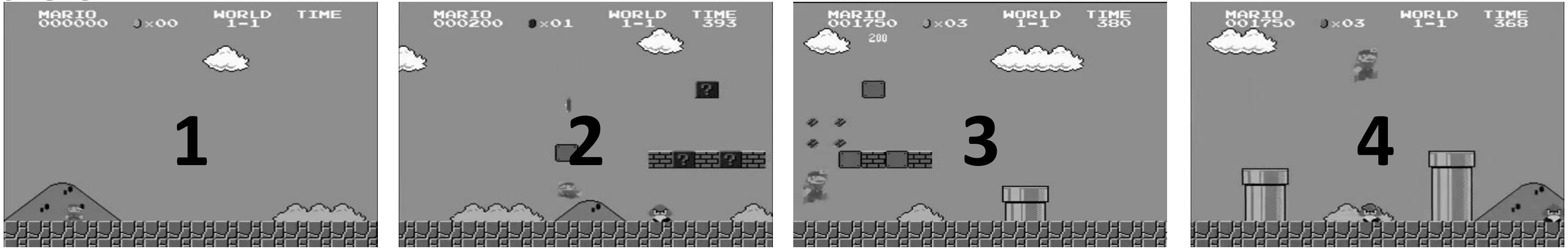
Video 1



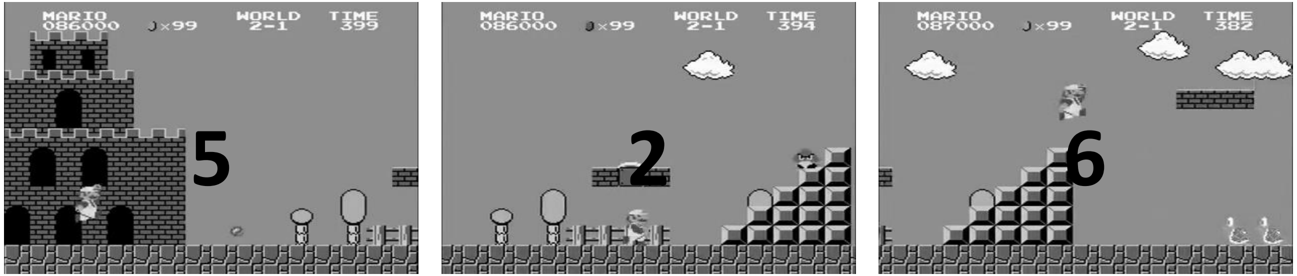
Video 2

Video 3

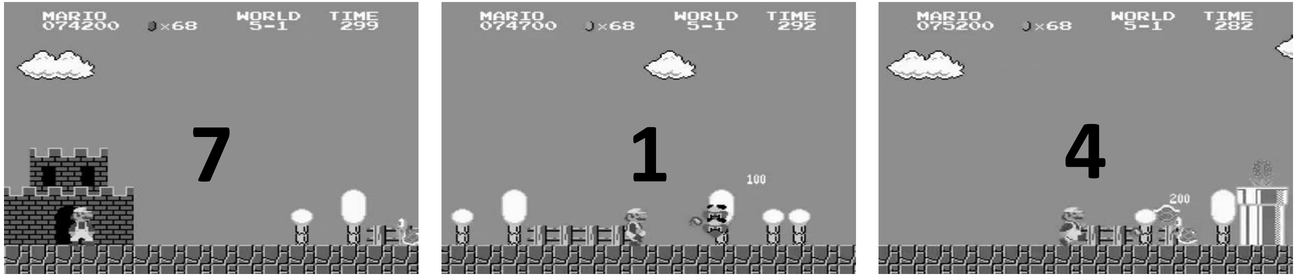
Video 1



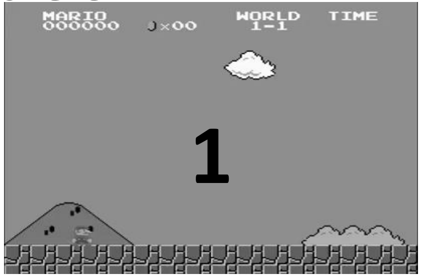
Video 2

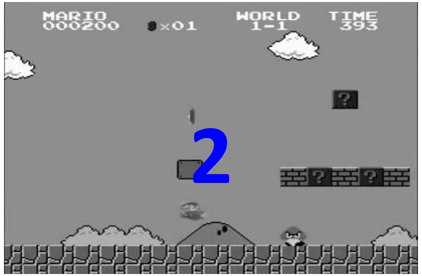


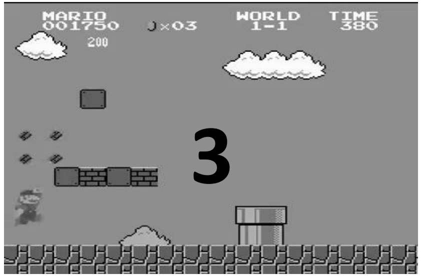
Video 3

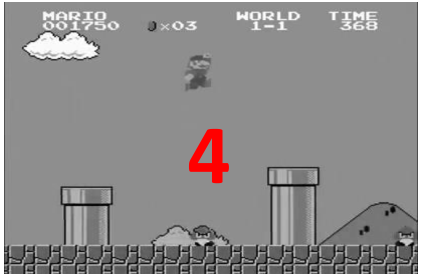


Video 1

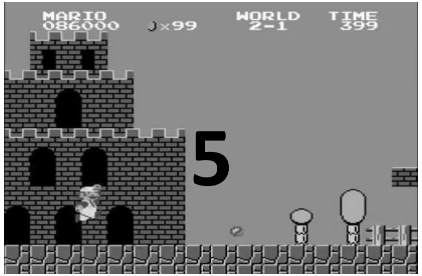


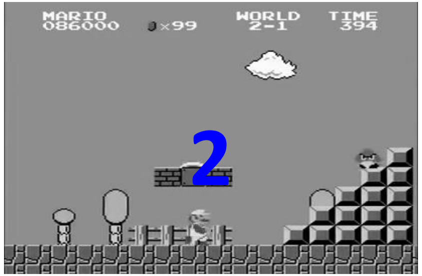


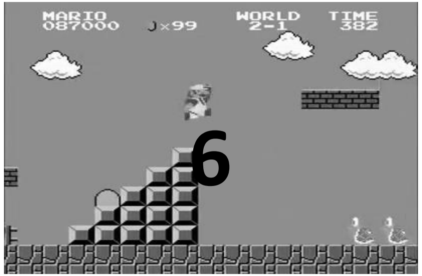




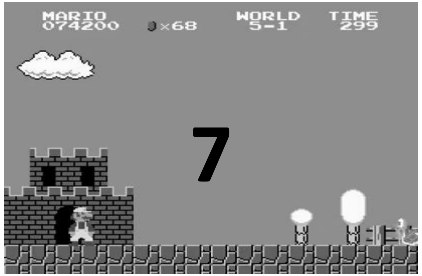
Video 2

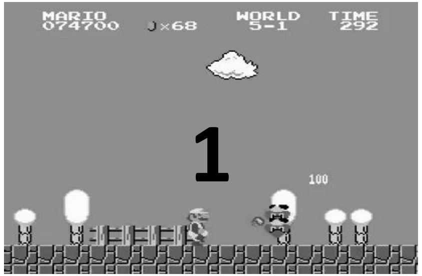


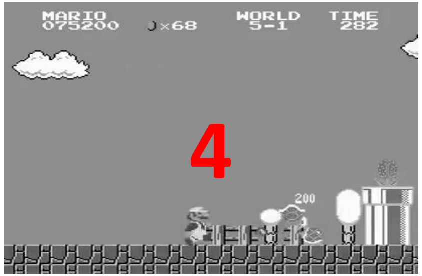




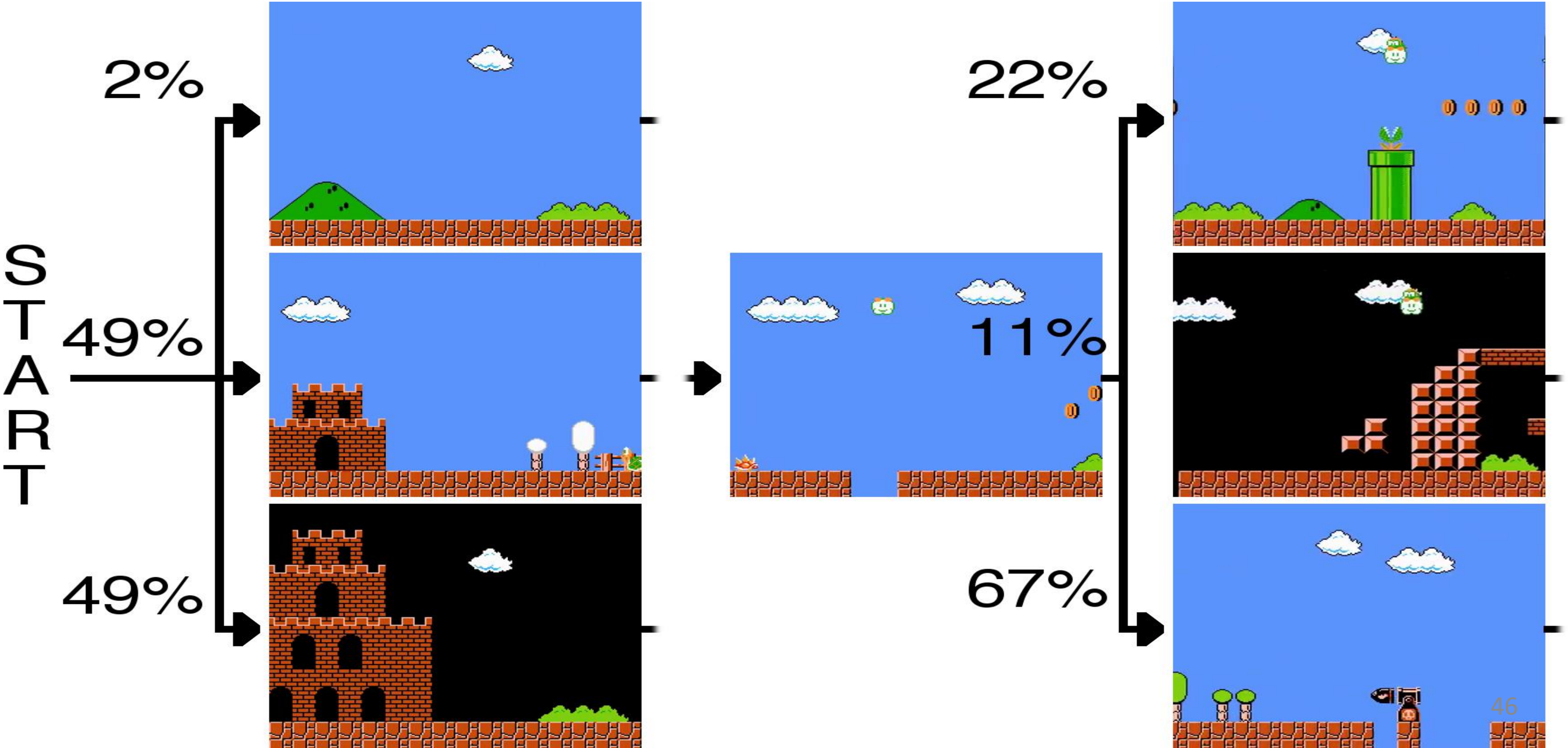
Video 3



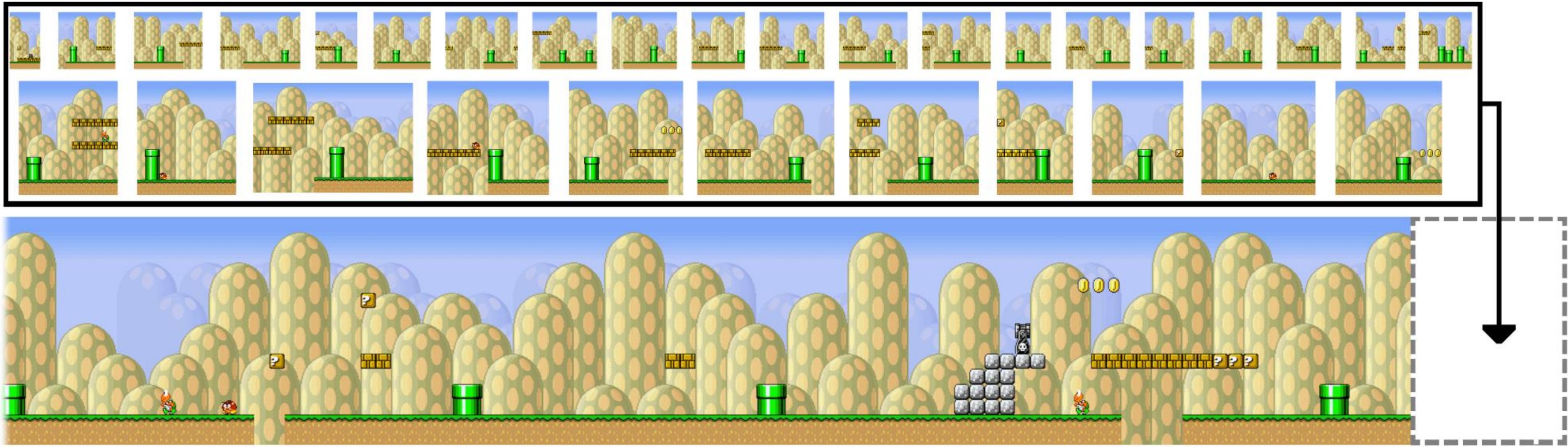




Fuzzy Level Model

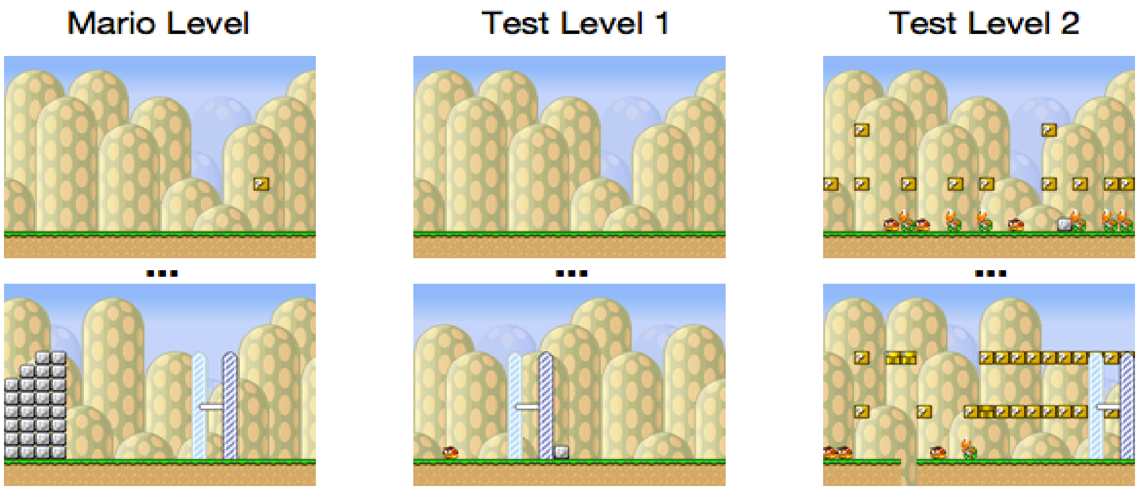


Level Generation



Human Subjects Study

- Online study
- Players played 3 of 16 levels
  - 1 Mario, 5 Snodgrass, 5 Dahlskog, 5 Our System
- Asked to rank the levels on the following traits
  - Mario-like (Style)
  - Fun
  - Frustration
  - Challenge
  - Design
  - Creativity



Question 3: Rank Frustration

For all three levels please rank them from 1-3 (using 1, 2 and 3 in the boxes below) from most to least frustrating with the most frustrating being ranked as "1".

Mario Level Rank

Test Level 1 Rank

Test Level 2 Rank

Reasoning (optional)

Next Question

Study Results

	Mario	Ours	Snodgrass	p-value
Mario-like	N/A	1	2	**0.0174**
Fun	1	2	3	**3.02e-5**
Frustration	3	2	1.5	**1.06e-11**
Challenge	3	2	2	0.6976
Design	1	2	3	**2.31e-7**
Deaths	0	2	3	**2.25e-9**

	Mario	Ours	Dahlskog	p-value
Mario-like	N/A	1	2	**0.0383**
Fun	1	2	2	0.3147
Frustration	3	2	1	**3.58e-7**
Challenge	3	2	1	**2.31e-4**
Design	1	2	3	**2.21e-5**
Deaths	0	2	3	**2.42e-5**

Study Results Summary

- Our system was significantly more similar to Mario than both other generators in the rankings for:
  - Mario-like (style)
  - Design
  - Frustration
  - # of deaths
- Significantly *more fun* than Snodgrass
- Significantly *less challenging* than Dahlskog

Comparison to System Ranking Results

Category	r_s	p
Style	**0.6115095**	**2.2e-16**
Design	**0.51948**	**2.2e-16**
Fun	**0.2729658**	**3.745e-5**
Frustration	**-0.4393904**	**6.79e-12**
Challenge	**-0.387222**	**2.351e-09**
Creativity	-0.1559725	**0.02007**

Conclusions

- We learn a model of level design that correlates strongly with measures of human models of level design.

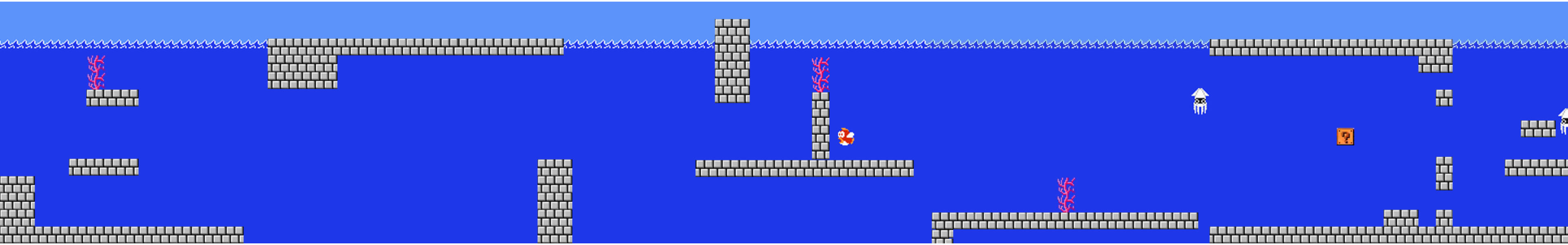
- These models can produce levels of consistently higher quality than the levels of other machine learned models

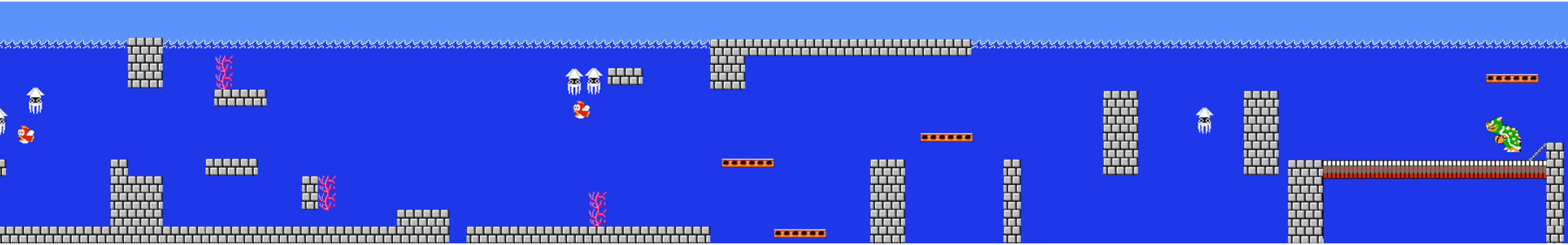
Creative Procedural Content Generation via Machine Learning

Modern Machine Learning

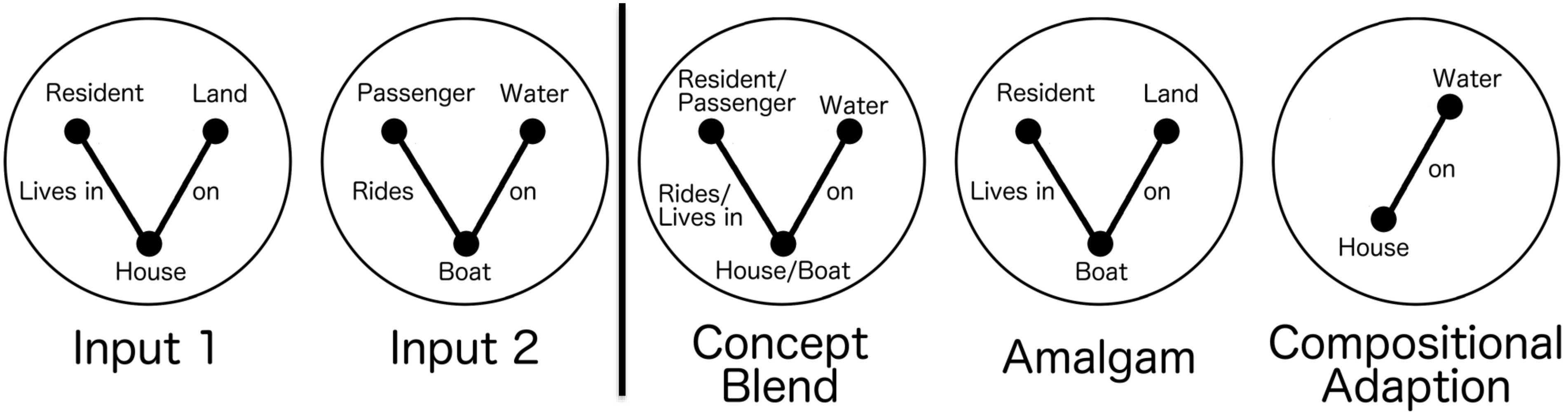
- Good when we have large amounts of data
- Of a form that consistently maps expected input to expected output
- But what about cases where no training data exists? Or where not one optimum output?
  - e.g. most design problems

What we'd like…





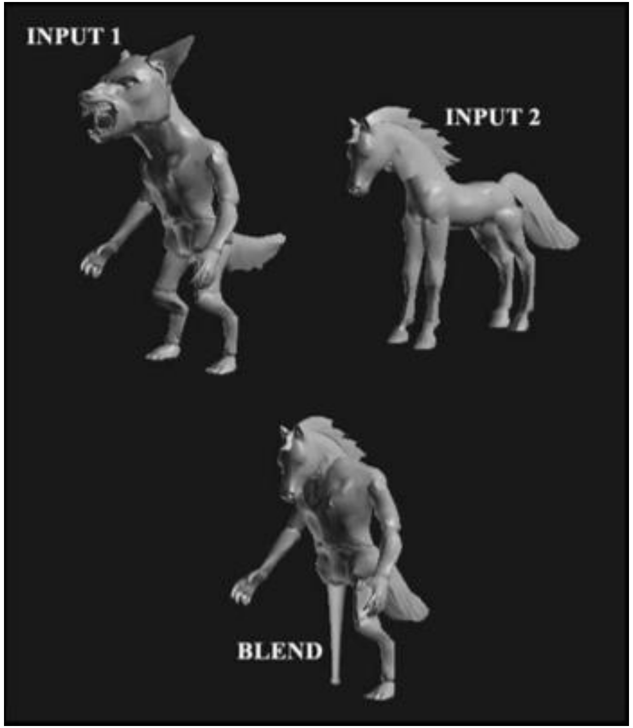
Combinational Creativity



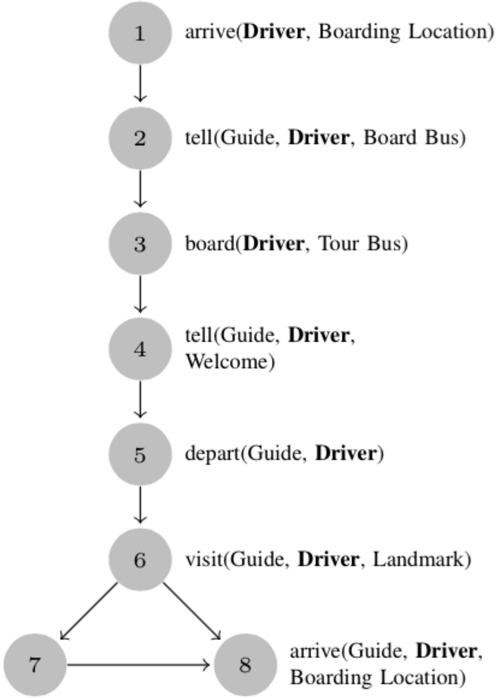
Combinatorial Creativity with ML

1. Show we can take ML models as input
2. Demonstrate recombination of ML models
3. Evaluate recombined model

Has someone done this before? (Nope)



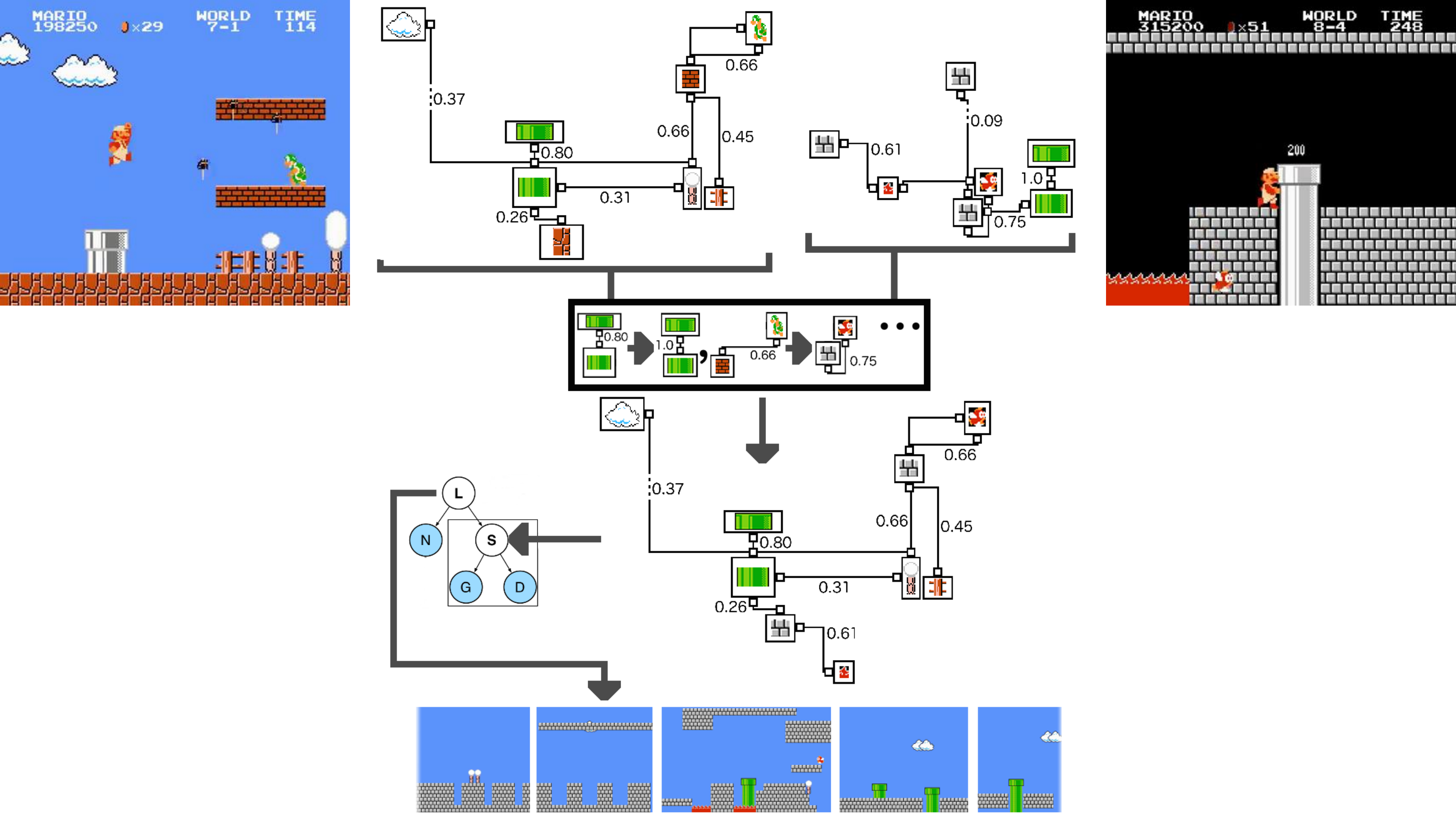
Ribeiro et al. 2003

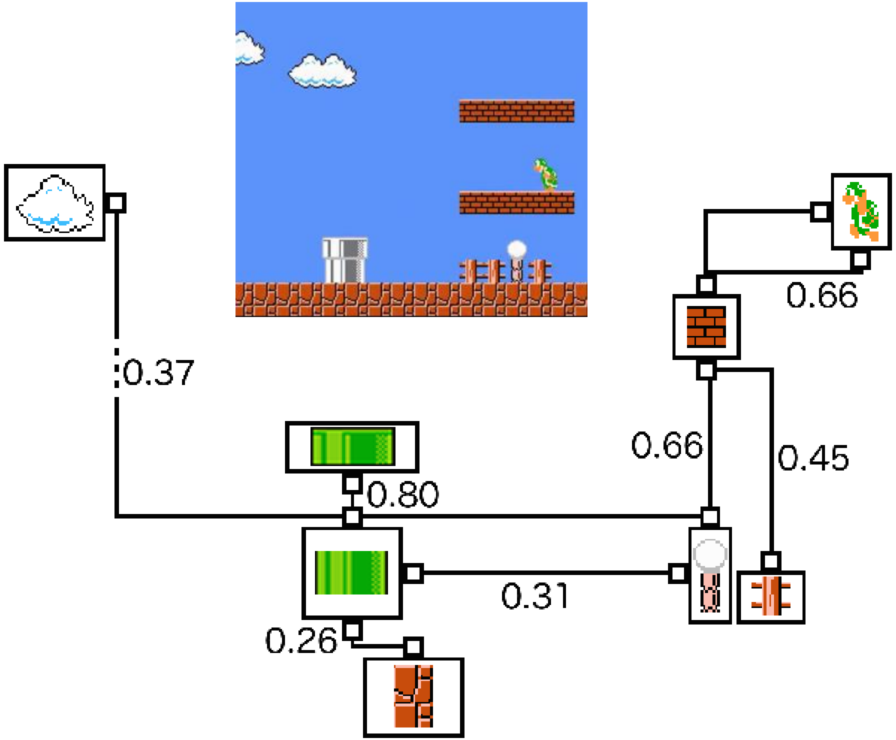


Permar and Magerko 2013

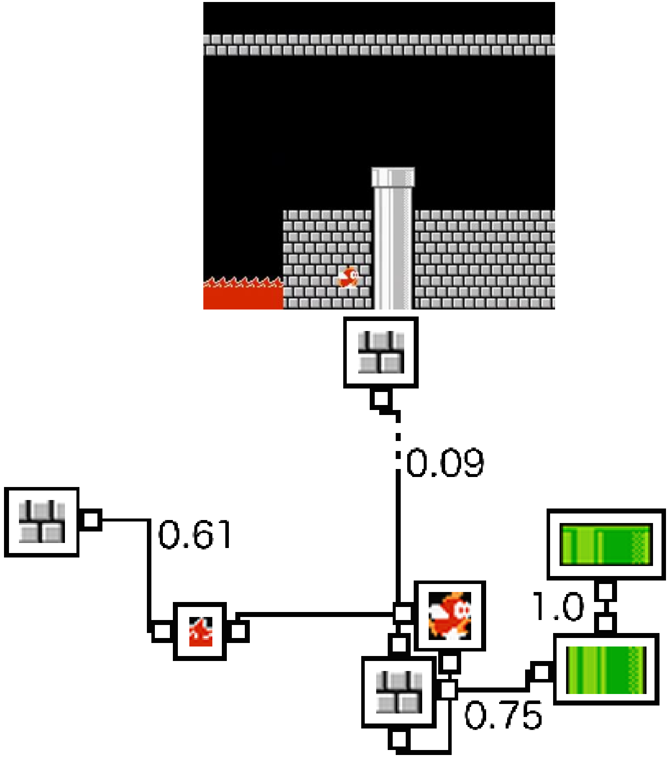
Pitch

Try this with the game level models we discussed learning and evaluating before.


MARIO 198250 ×29 WORLD 7-1 TIME 114
0.37
0.66
0.66
0.45
0.80
0.31
0.26
0.09
0.61
1.0
0.75
MARIO 315200 ×51 WORLD 8-4 TIME 248
200
0.80
1.0
0.66
0.75
0.37
0.66
0.66
0.45
0.80
0.31
0.26
0.61
L
N
S
G
D



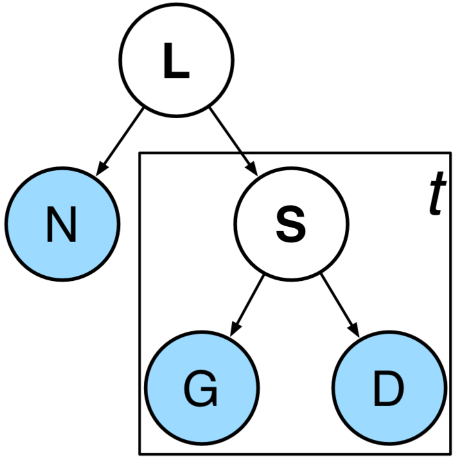
Source



Target

S-Structure Graph

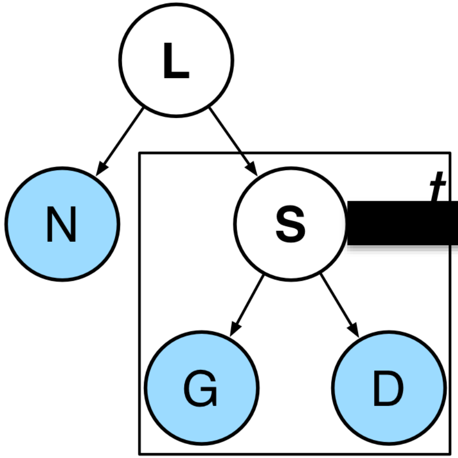
1. Select L Node

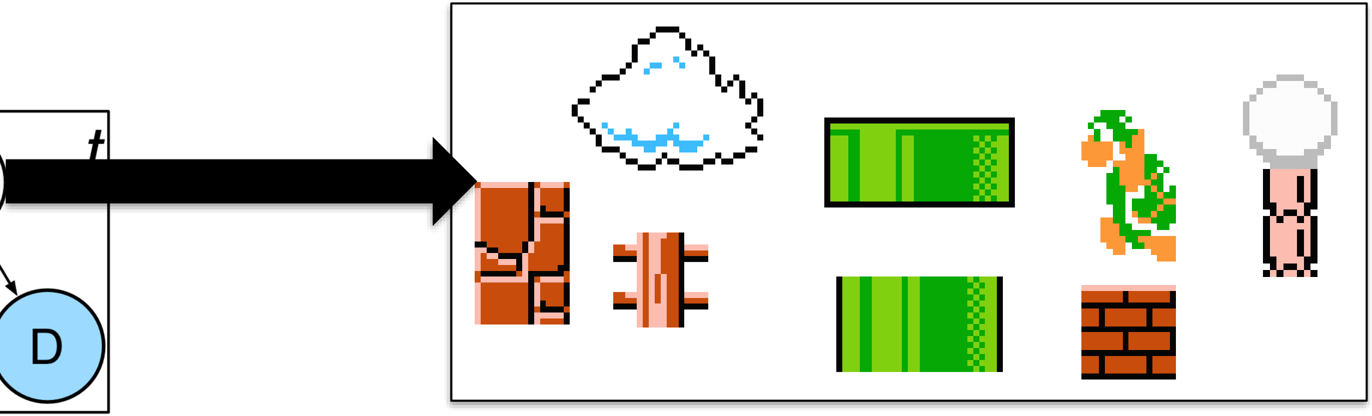


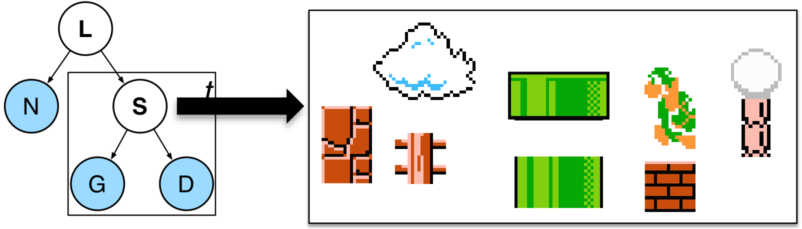
S-Structure Graph

1. Select L Node

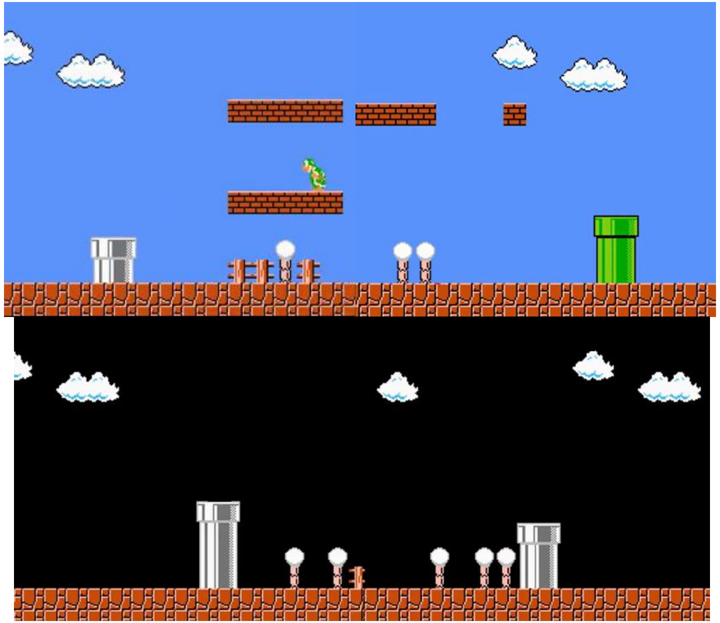
2. Extract S Nodes

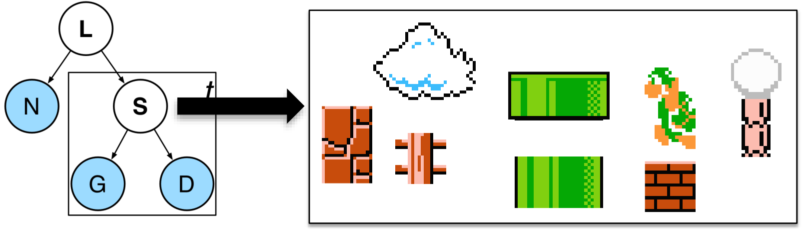




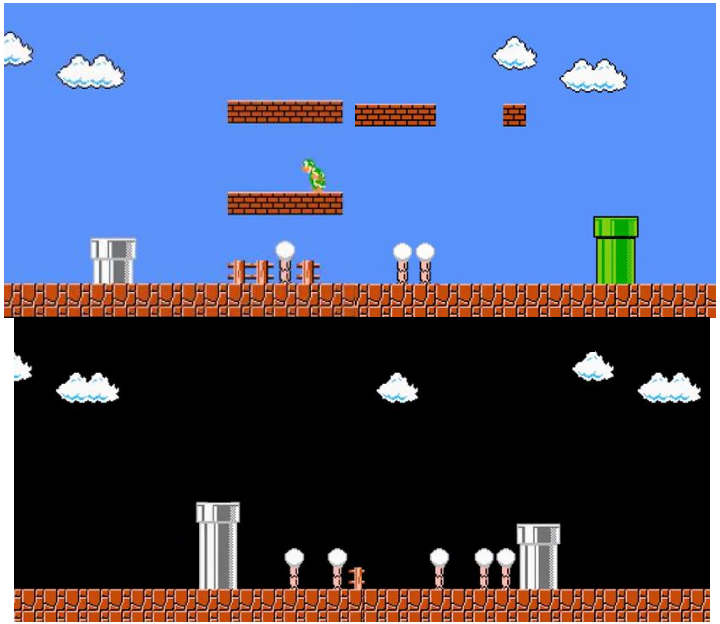


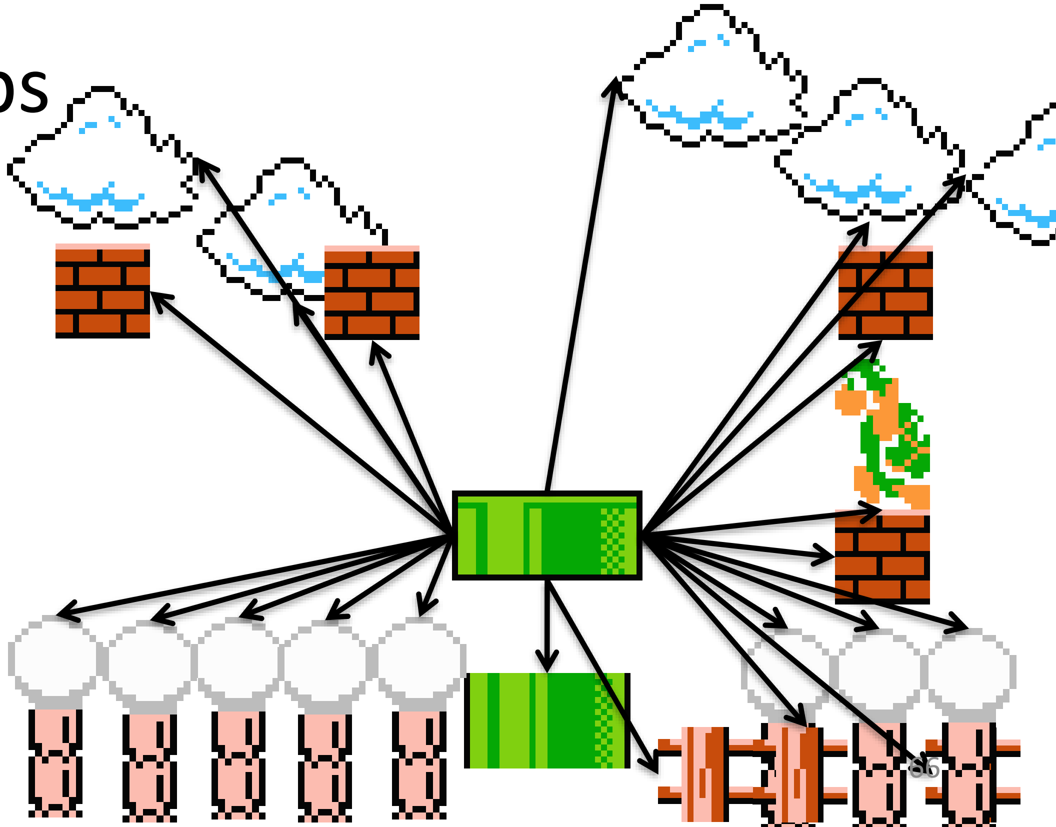
3. Select all Relationships

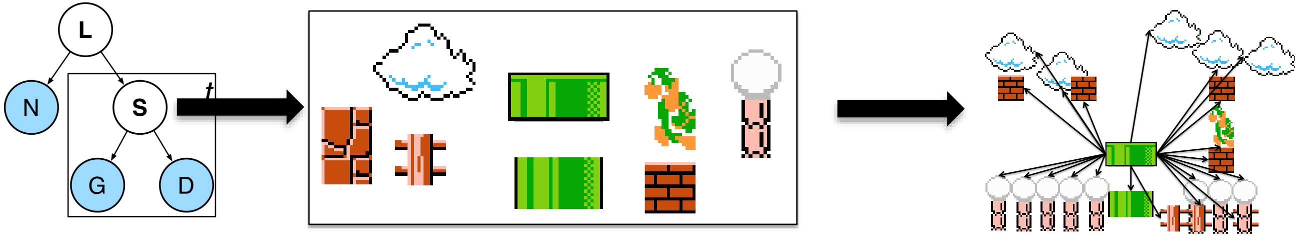




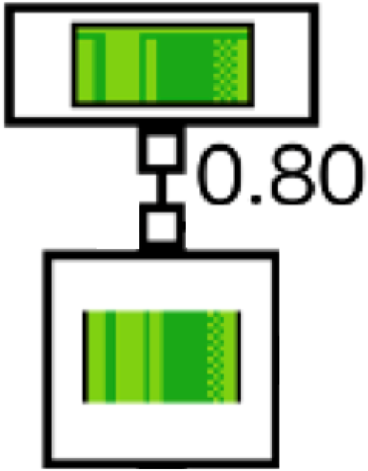
3. Select all Relationships

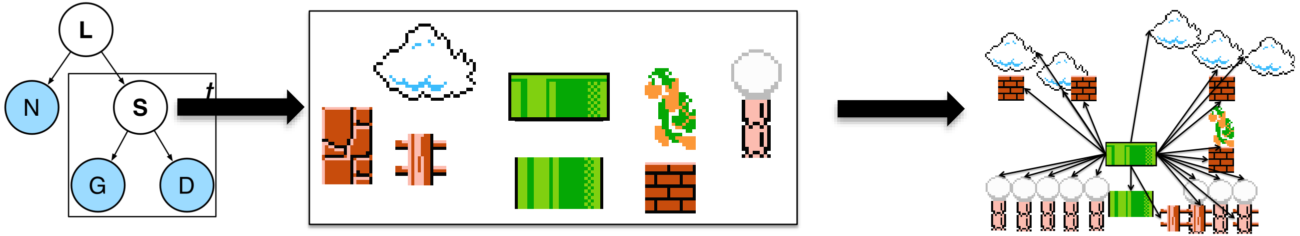




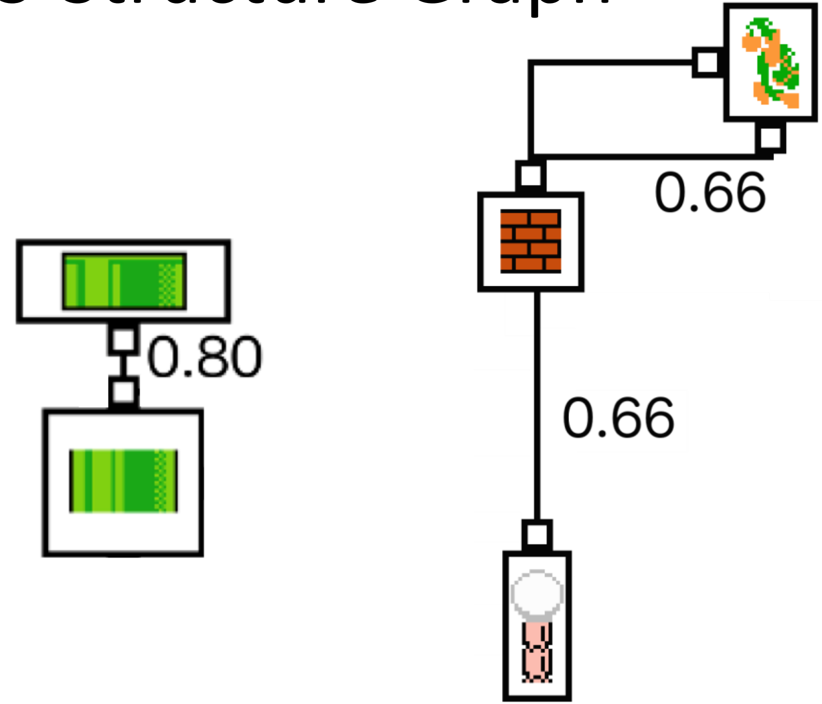


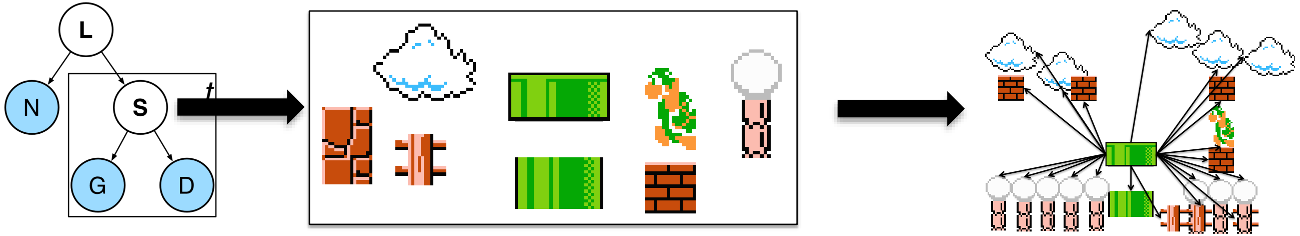
4. Build S-Structure Graph



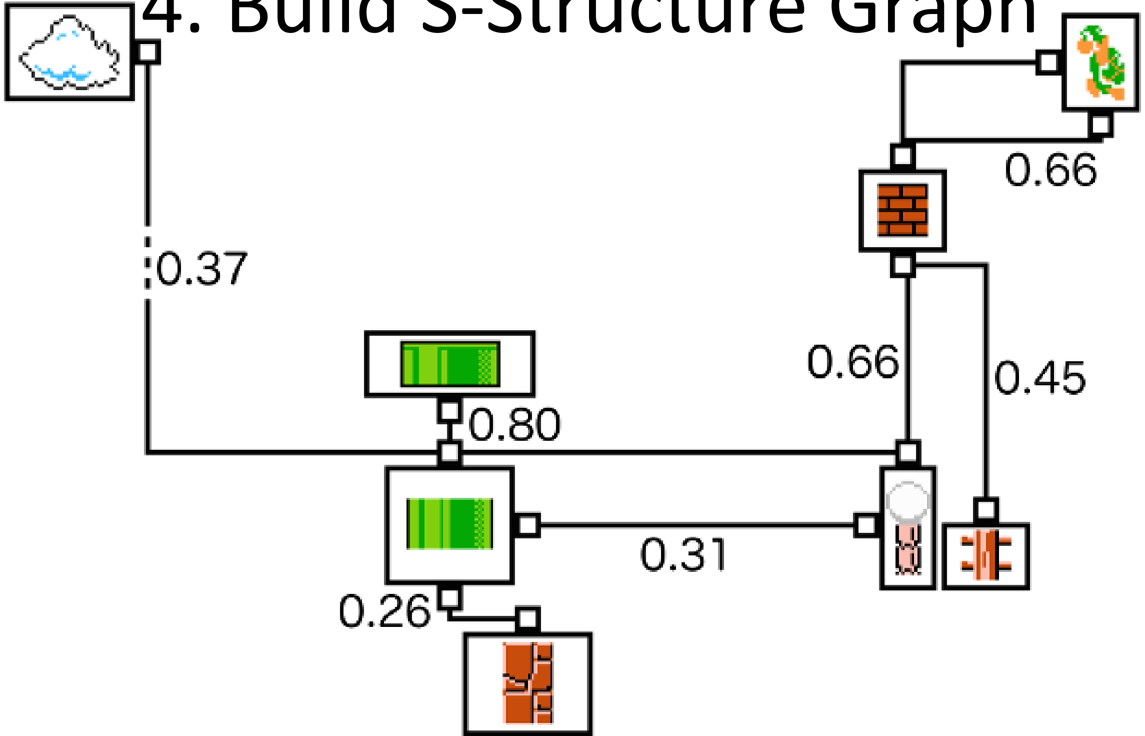


4. Build S-Structure Graph

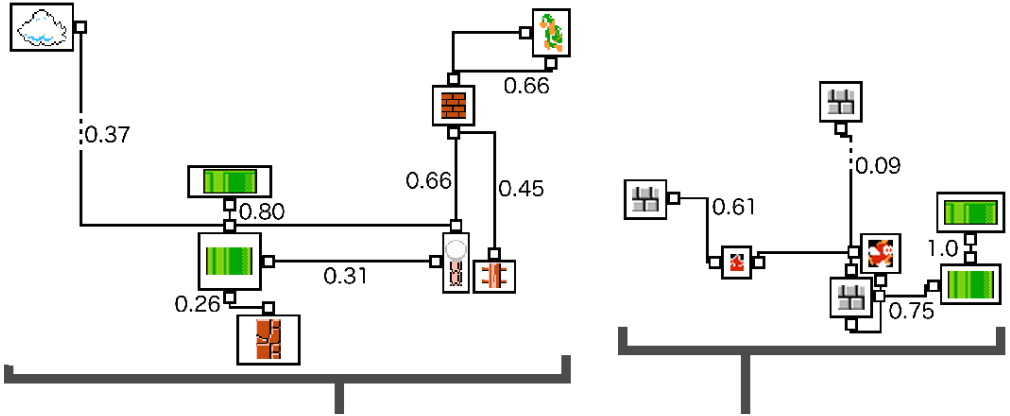


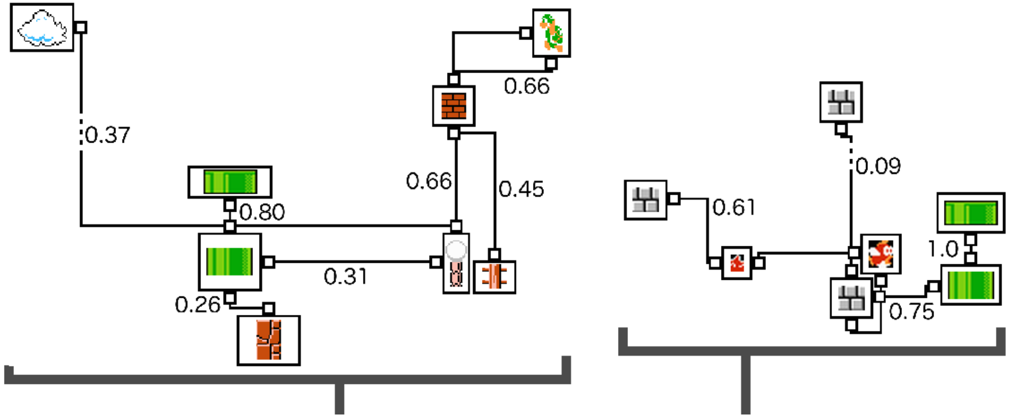


4. Build S-Structure Graph

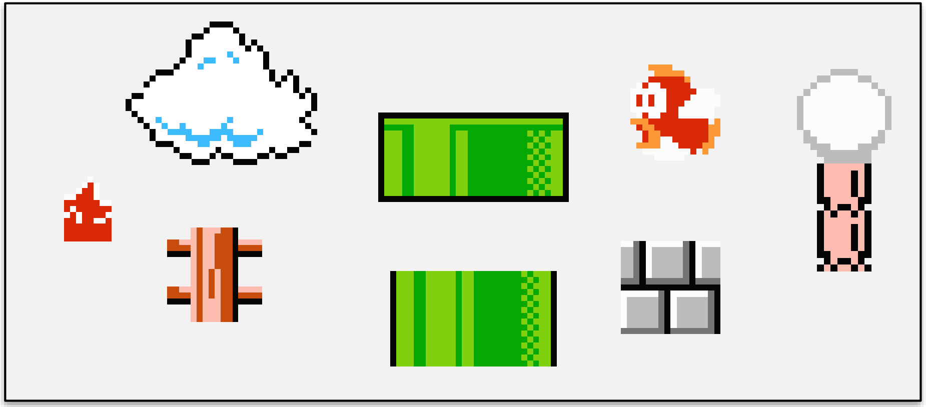


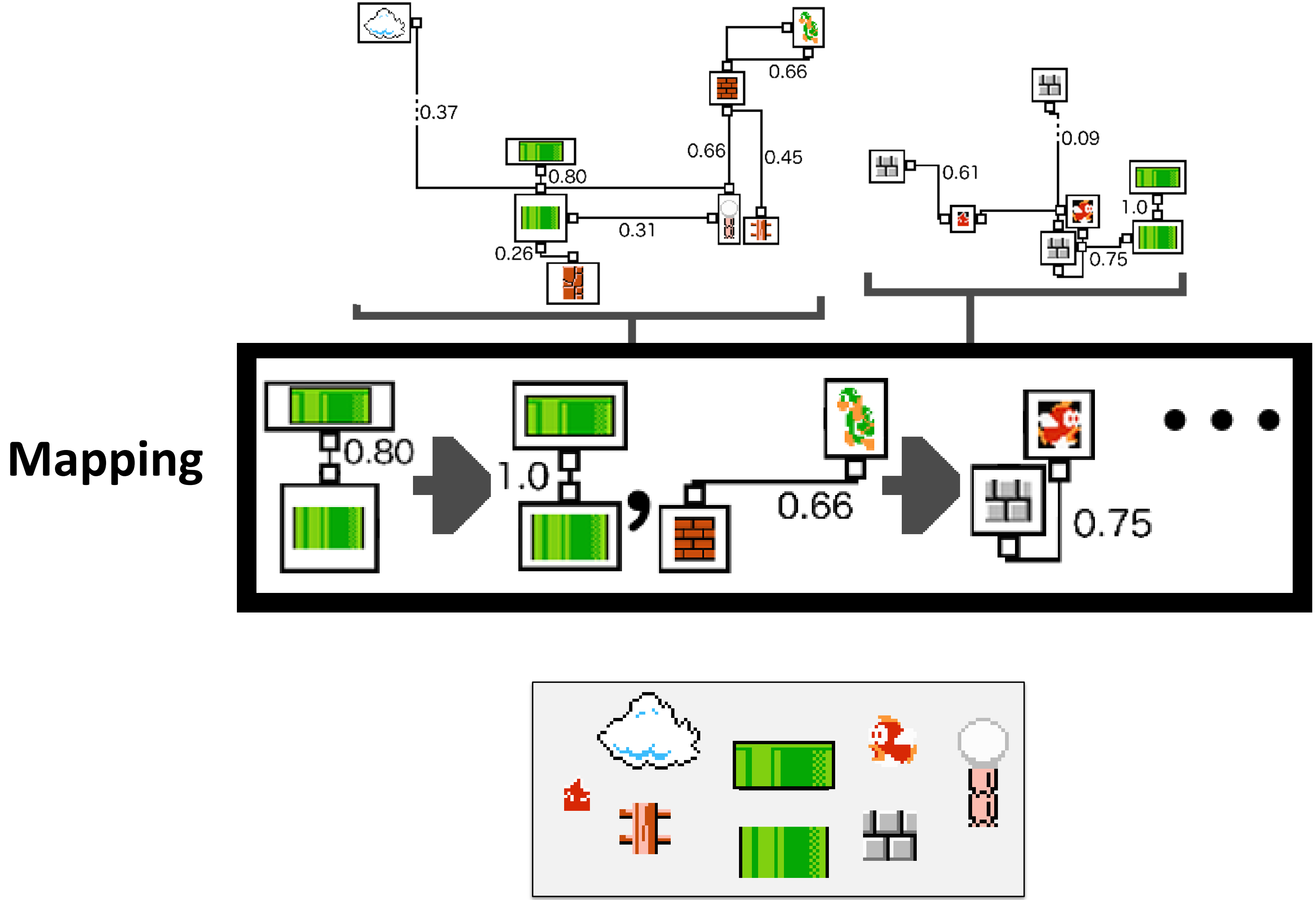
Blending Graphs

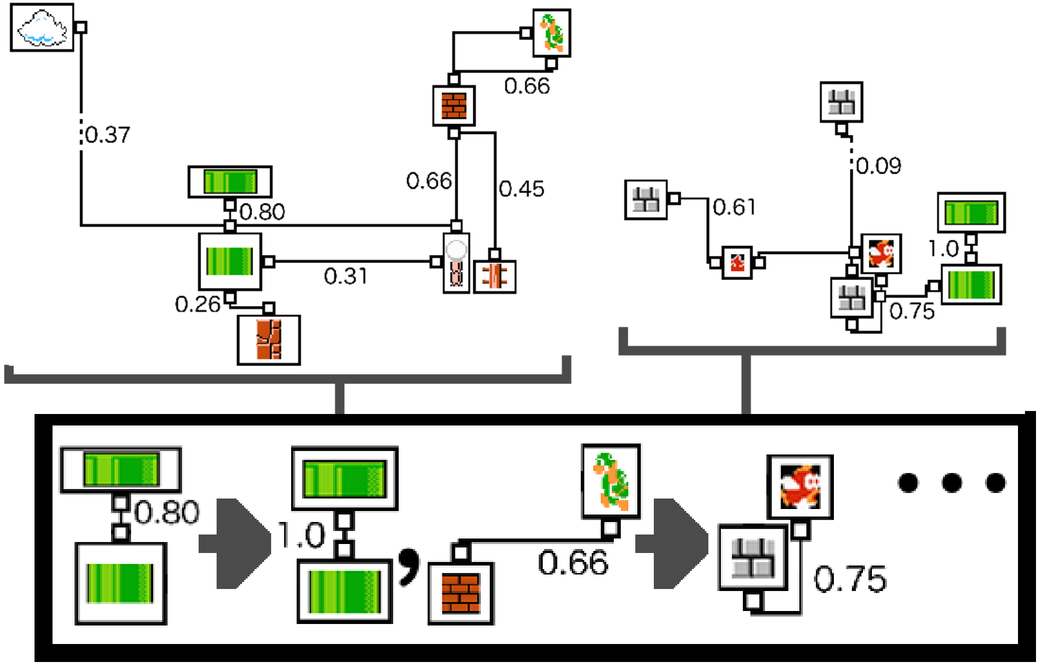




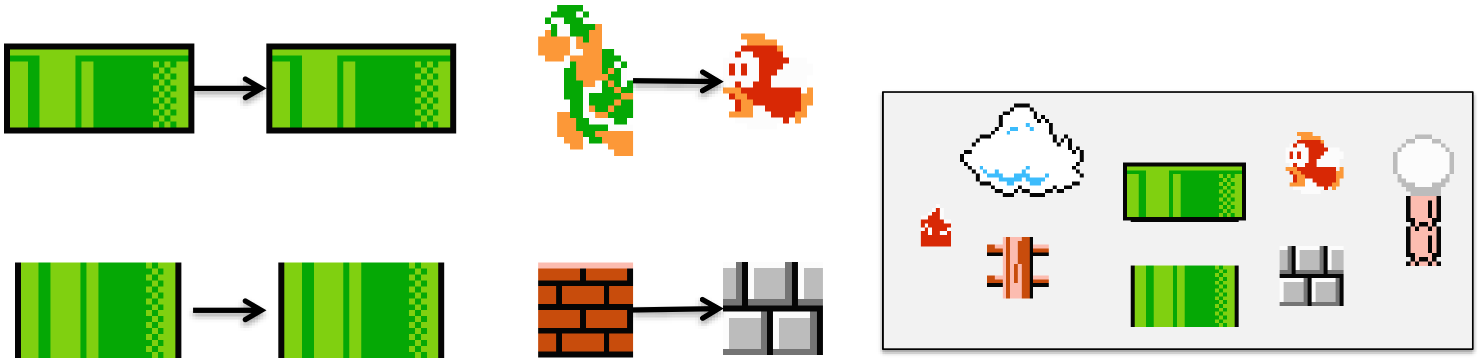
0.37
0.66
0.80
0.66
0.45
0.31
0.26
0.09
0.61
1.0
0.75

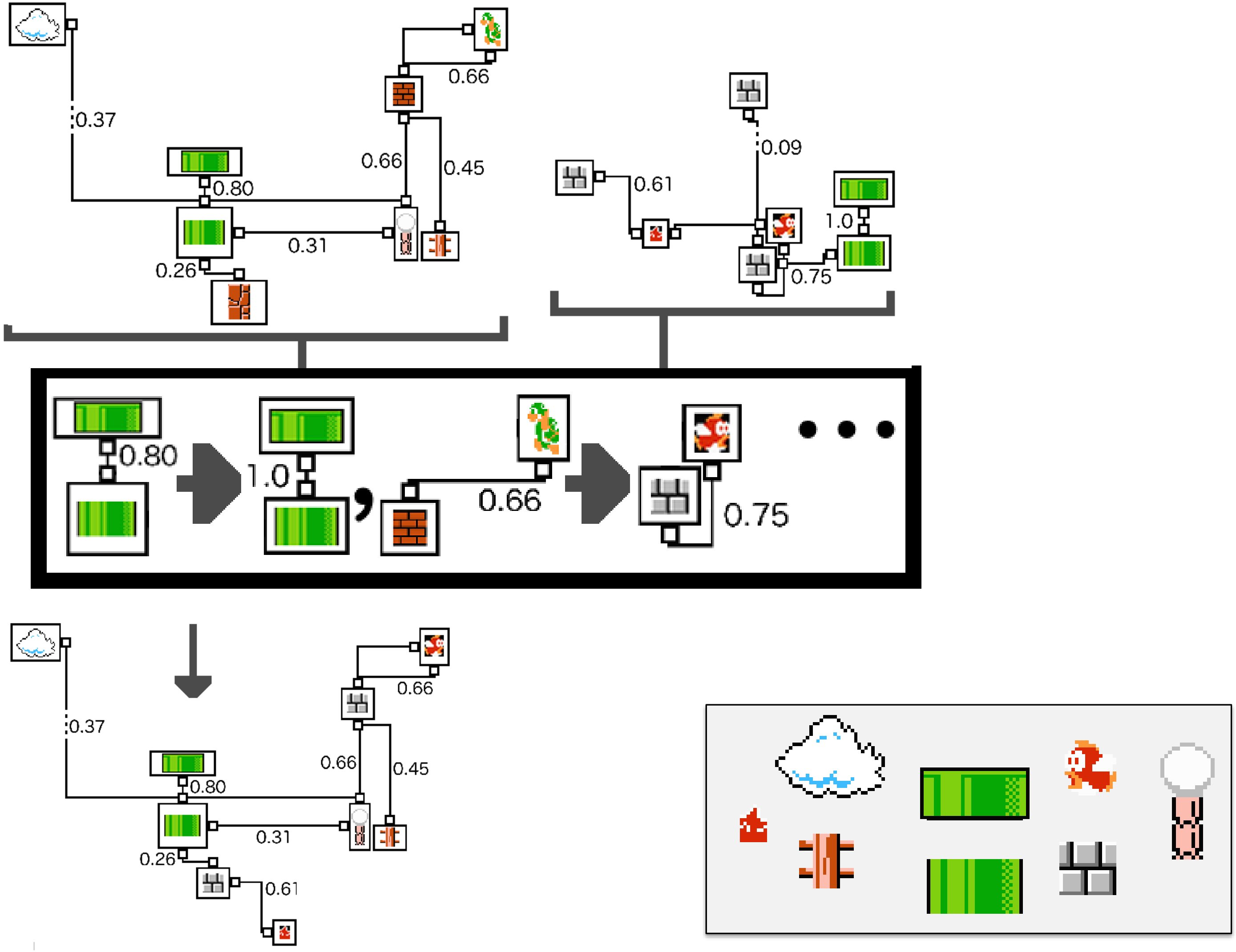



0.37
0.66
0.80
0.66
0.45
0.31
0.26
0.09
0.61
1.0
0.75
Mapping
0.80
1.0
0.66
0.75

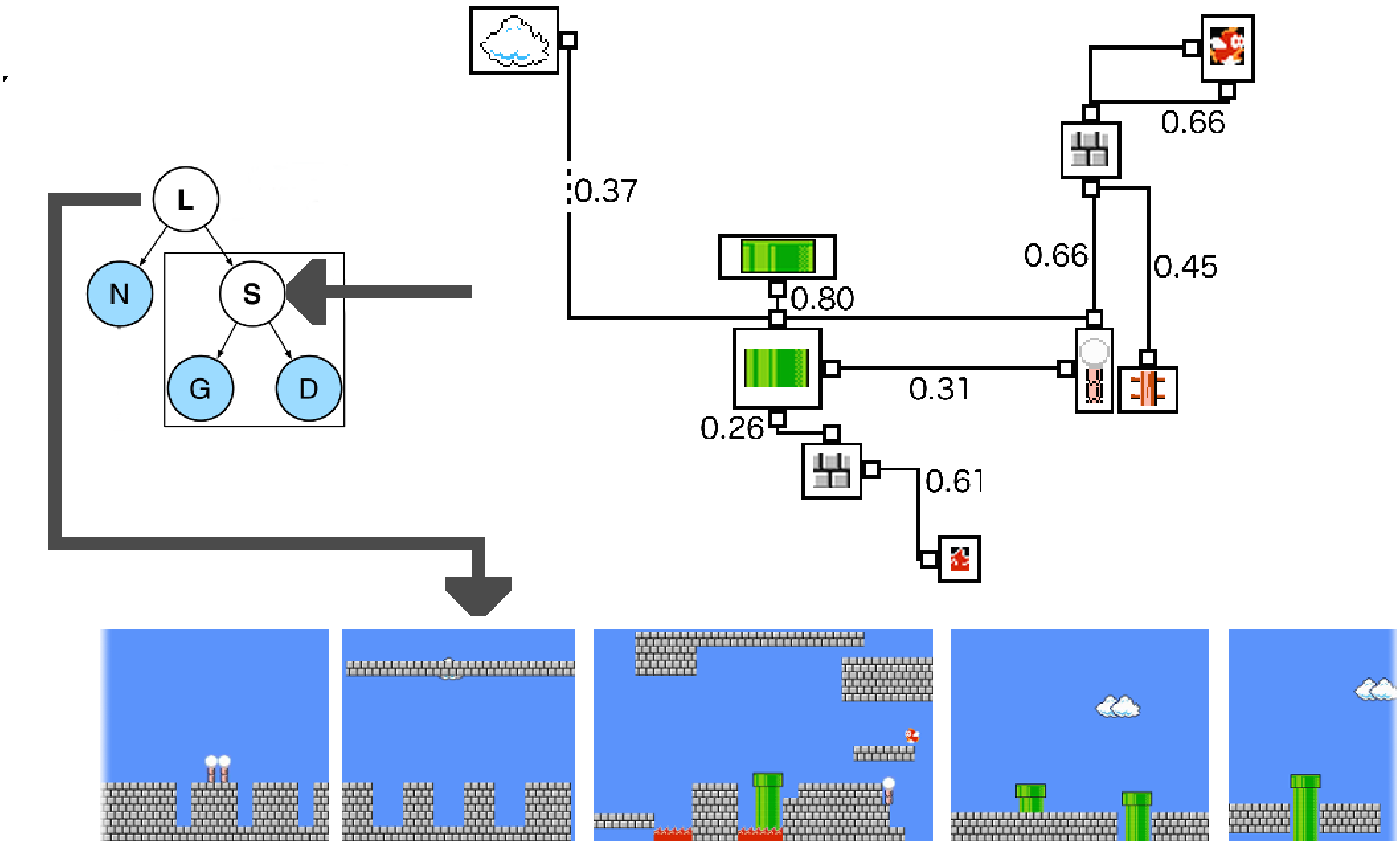


Evidence

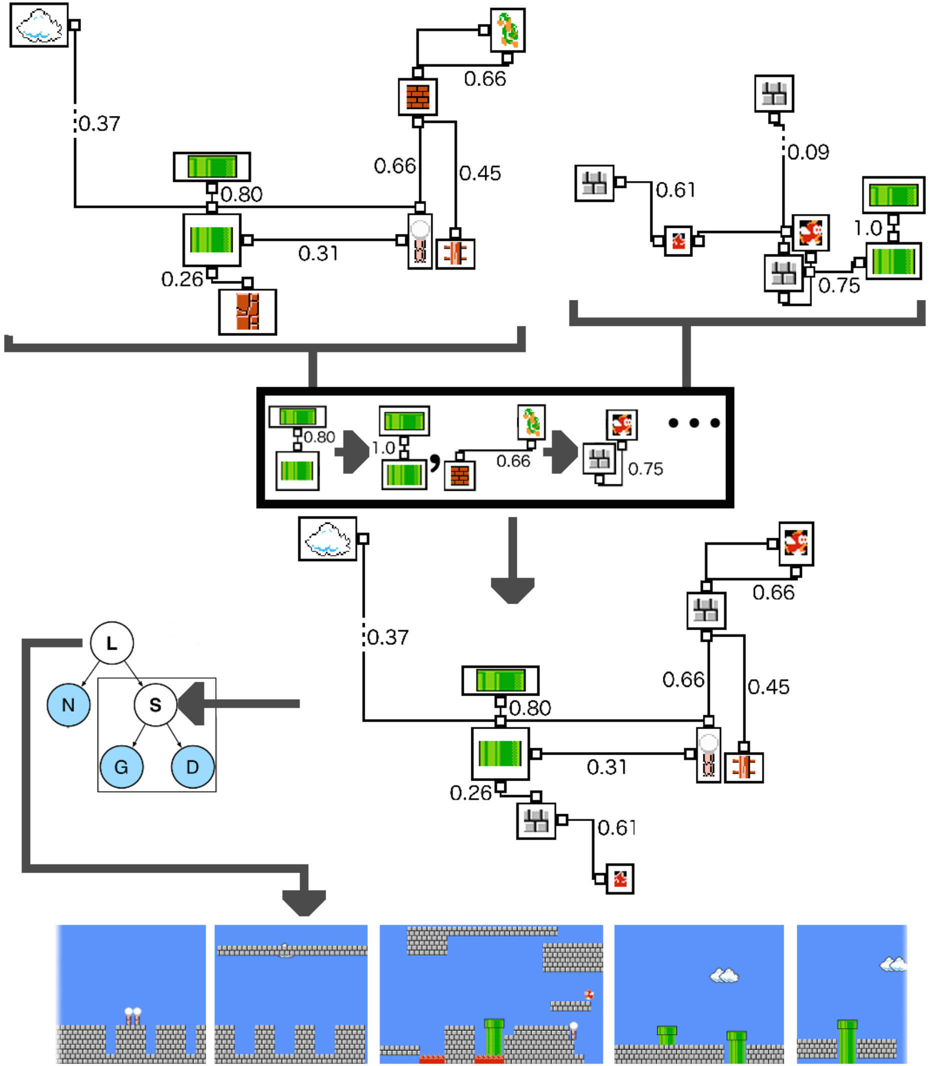




0.37
0.66
0.66
0.45
0.80
0.31
0.26
0.09
0.61
1.0
0.75
0.80
1.0
0.66
0.75
0.37
0.66
0.66
0.45
0.80
0.31
0.26
0.61



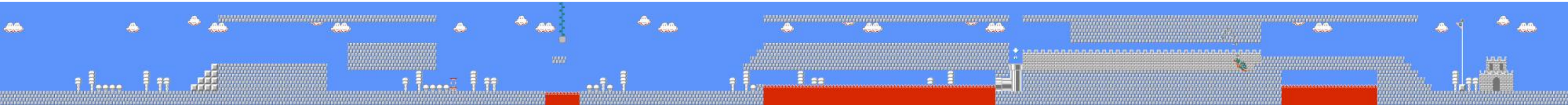
L
N
S
G
D
0.37
0.80
0.66
0.66
0.45
0.31
0.26
0.61



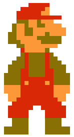
0.37
0.66
0.66
0.45
0.80
0.31
0.26
0.09
0.61
1.0
0.75
0.80
1.0
0.66
0.75
0.37
0.66
0.66
0.45
0.80
0.31
0.26
0.61
L
N
S
G
D

Case Study: Lost Levels

- From study we know that human players don't seem to rank creativity consistently
- But we want to evaluate how well these blended models perform.
- Instead: how well the blended models evaluate "real" blends

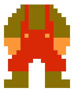


Four Types of Models



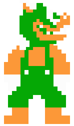
SMB Model

– Original, non-blended L-nodes



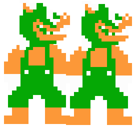
Level-Type Model

– L-nodes that match human-annotated “types” (e.g. “Underwater”, “Castle”)



Blended Model

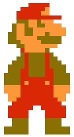
– Blended L-nodes that best understand target level



Full Blended Model

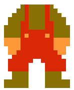
– All blended combinations of each level type model

Four Types of Models



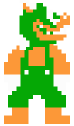
SMB Model

– Original, non-blended L-nodes



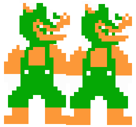
Level-Type Model

– L-nodes that match human-annotated “types” (e.g. “Underwater”, “Castle”)



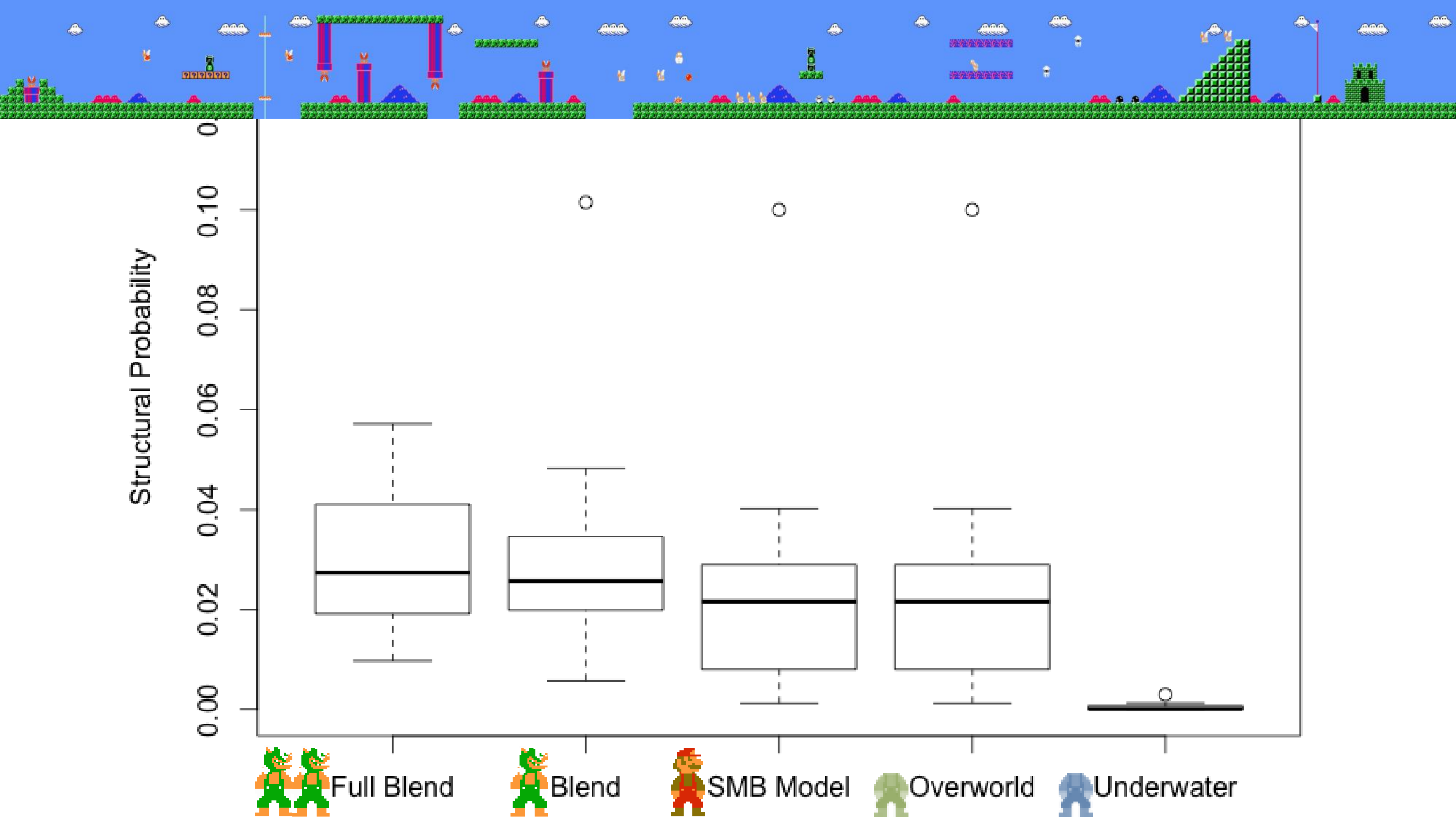
Blended Model

– Blended L-nodes that best understand target level

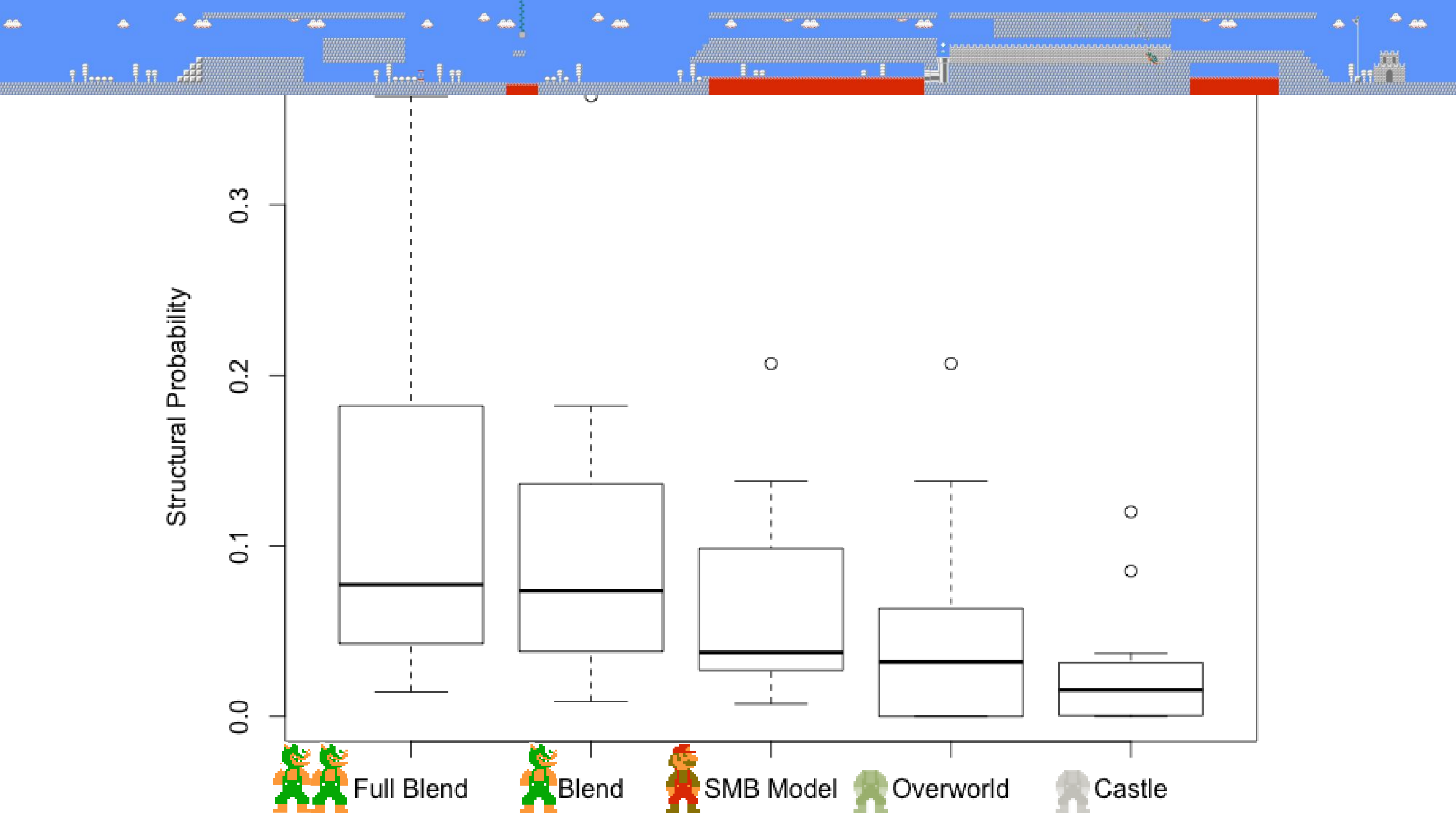


Full Blended Model

– All blended combinations of each level type model



Structural Probability
0.00
0.02
0.04
0.06
0.08
0.10
Full Blend
Blend
SMB Model
Overworld
Underwater



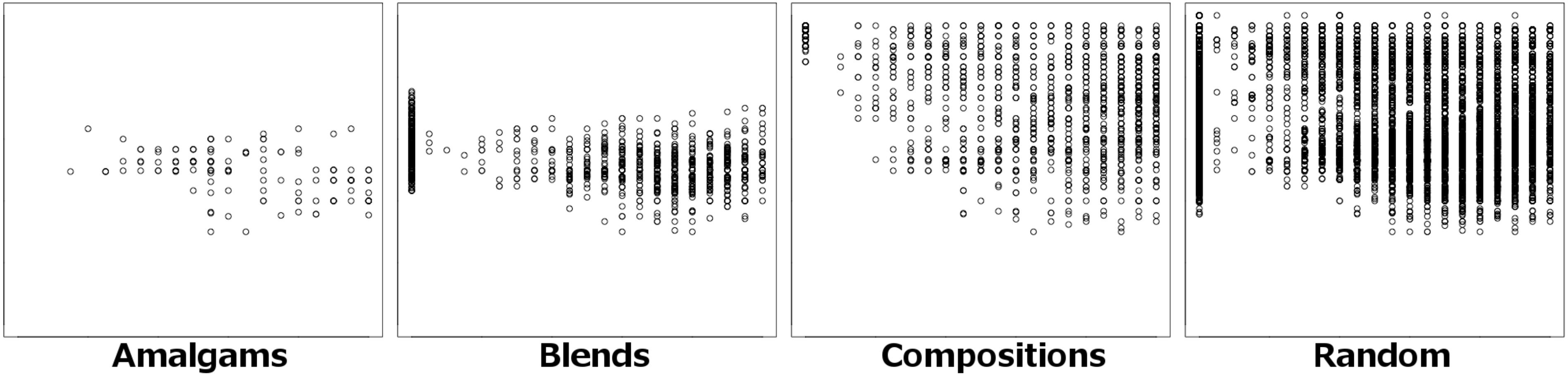
Structural Probability
0.0
0.1
0.2
0.3
Full Blend
Blend
SMB Model
Overworld
Castle

Conclusions

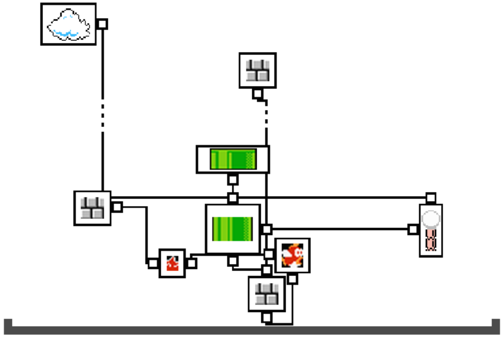
- Concept blending of (at least some) machine learned models can produce new ML models without additional training

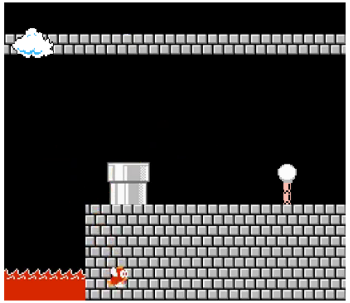
- Concept blending appears to be a reasonable approx. of some human creative processes.

Beyond Blending?

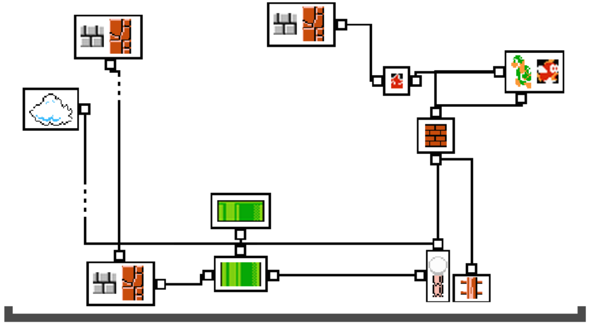


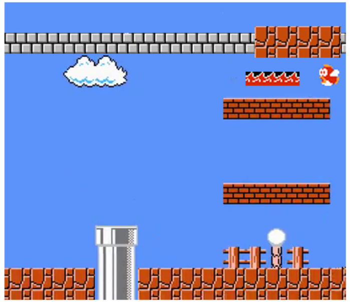
Visualized



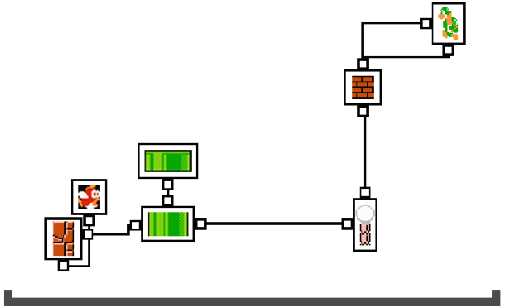


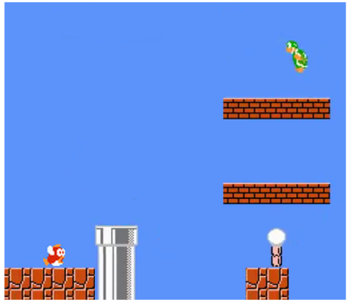
Amalgam



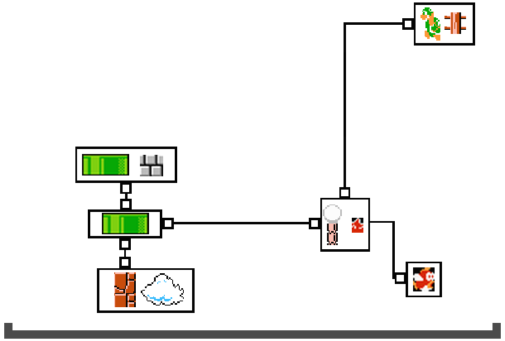


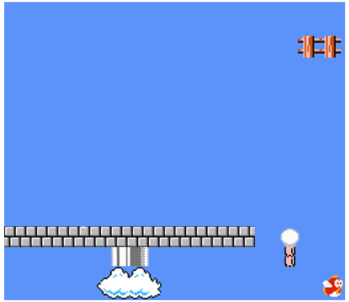
Blend





Composition





Random

Takeaways

- PCGML is a promising research direction toward the automated generation of content
- Current research directions:
  - Training data
  - Creative output
  - How to make it work for designers

Thanks!

mguzdial3@gatech.com

Georgia Tech College of Computing
School of Interactive Computing

entertainment intelligence lab