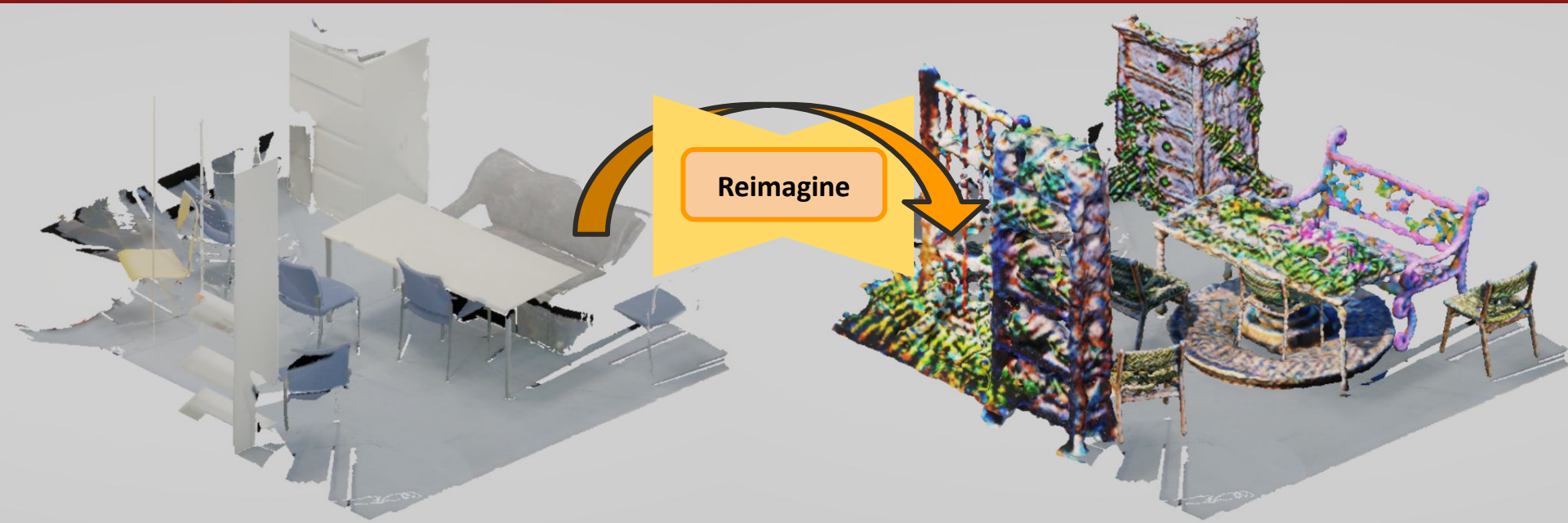
Reimagine the World

with 3D generative reconstruction



Codey Sun

June 3, 2025

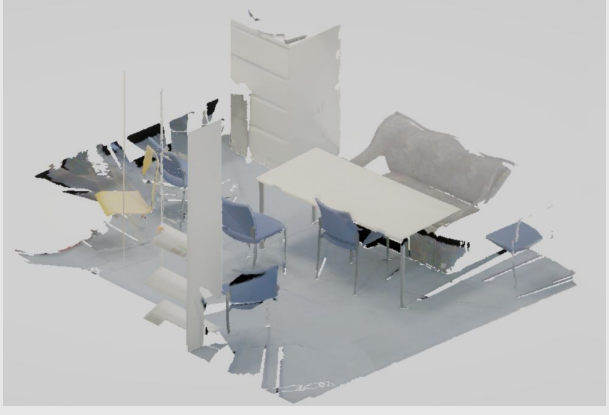
Stanford University

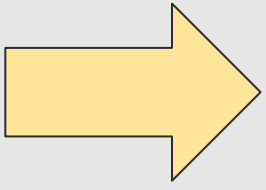
Stanford University

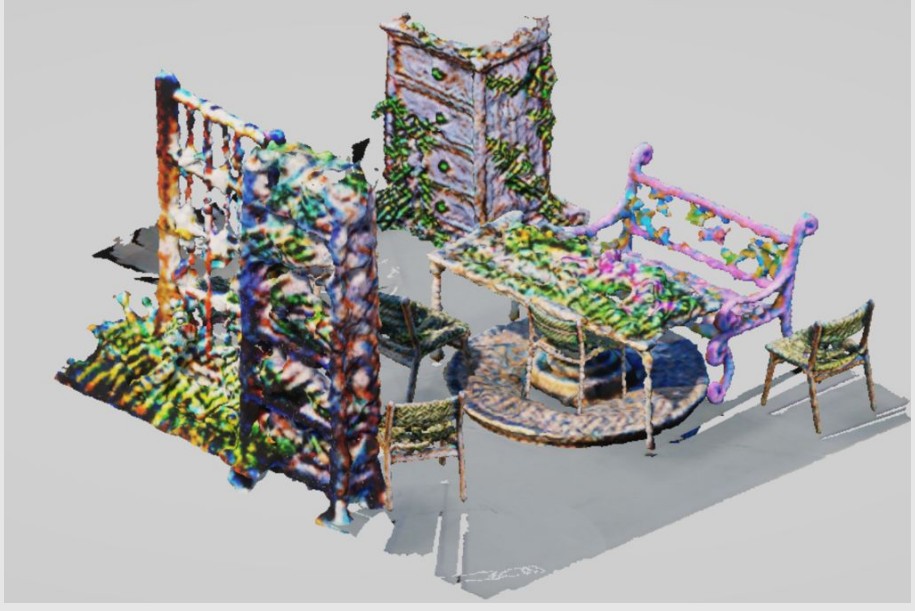
Motivation: Generative Reconstruction

Motivation: Generative Reconstruction

"Turn my office into a lovely garden"







Guiding Principles

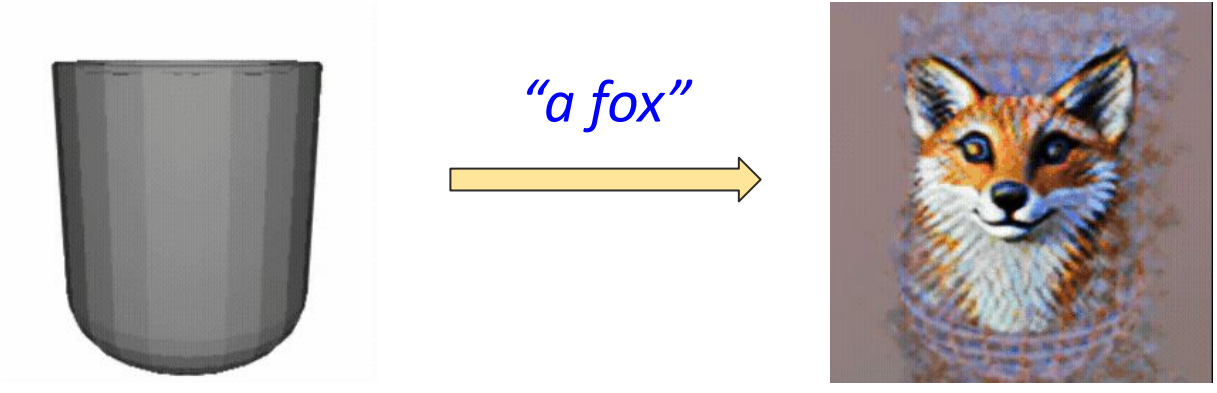
1) The generated environment must remain **fully interactable**
   a) Physically-grounded digital twin

2) Do not sacrifice creative expression of generations

The Problem: Balancing generation and control

- ControlNet offers a method of constraining geometry through depth inputs

- This can be far too restricting, leading to **undesirable generations**

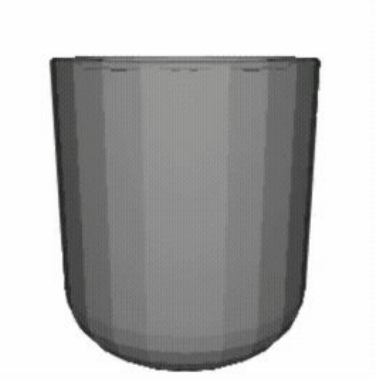


The Problem: Balancing generation and control

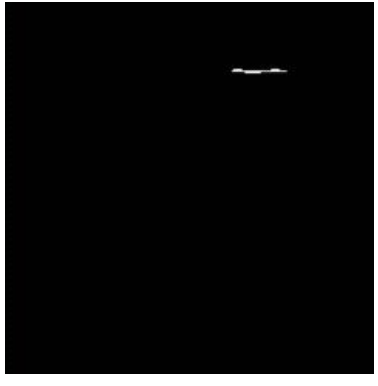
- ControlNet offers a method of constraining geometry through depth inputs
- This can be far too restricting, leading to undesirable generations

How can we balance **generation** and **control**?

Insight: Apply controls *locally*

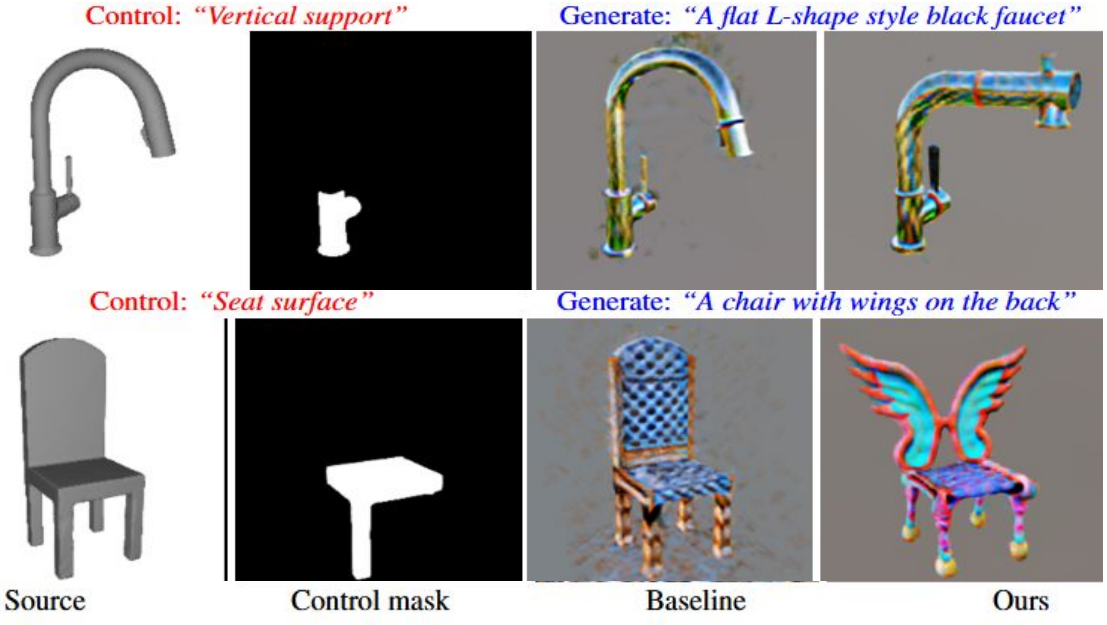


"a fox"



+ Geometry is still controlled
+ Generations have more expressive freedom

Local control improves prompt alignment…



…while minimally affecting geometric control

Control: *"Vertical support"*

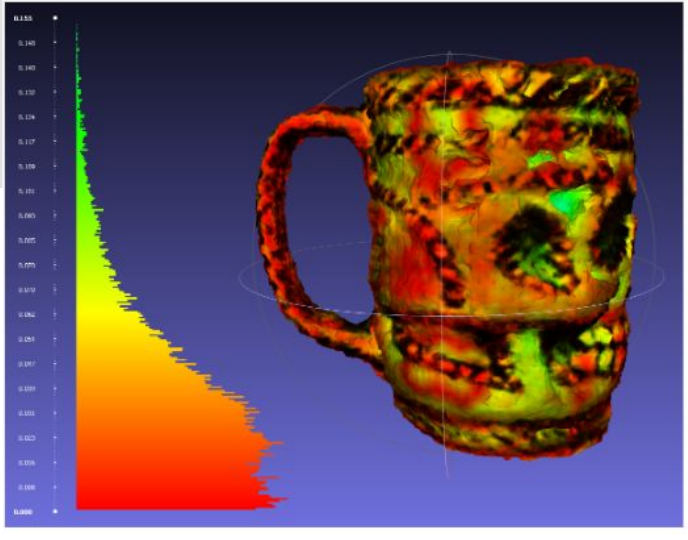
Generate: *"A flat L-shape style black faucet"*

Method	Chamfer distance (↓)
Baseline	0.00187
Ours	0.00189

Control: *"Seat surface"*

Source

Control mask



How do we get part masks?

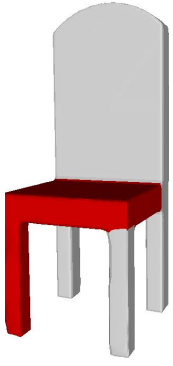
User Inputs:

- 3D source object

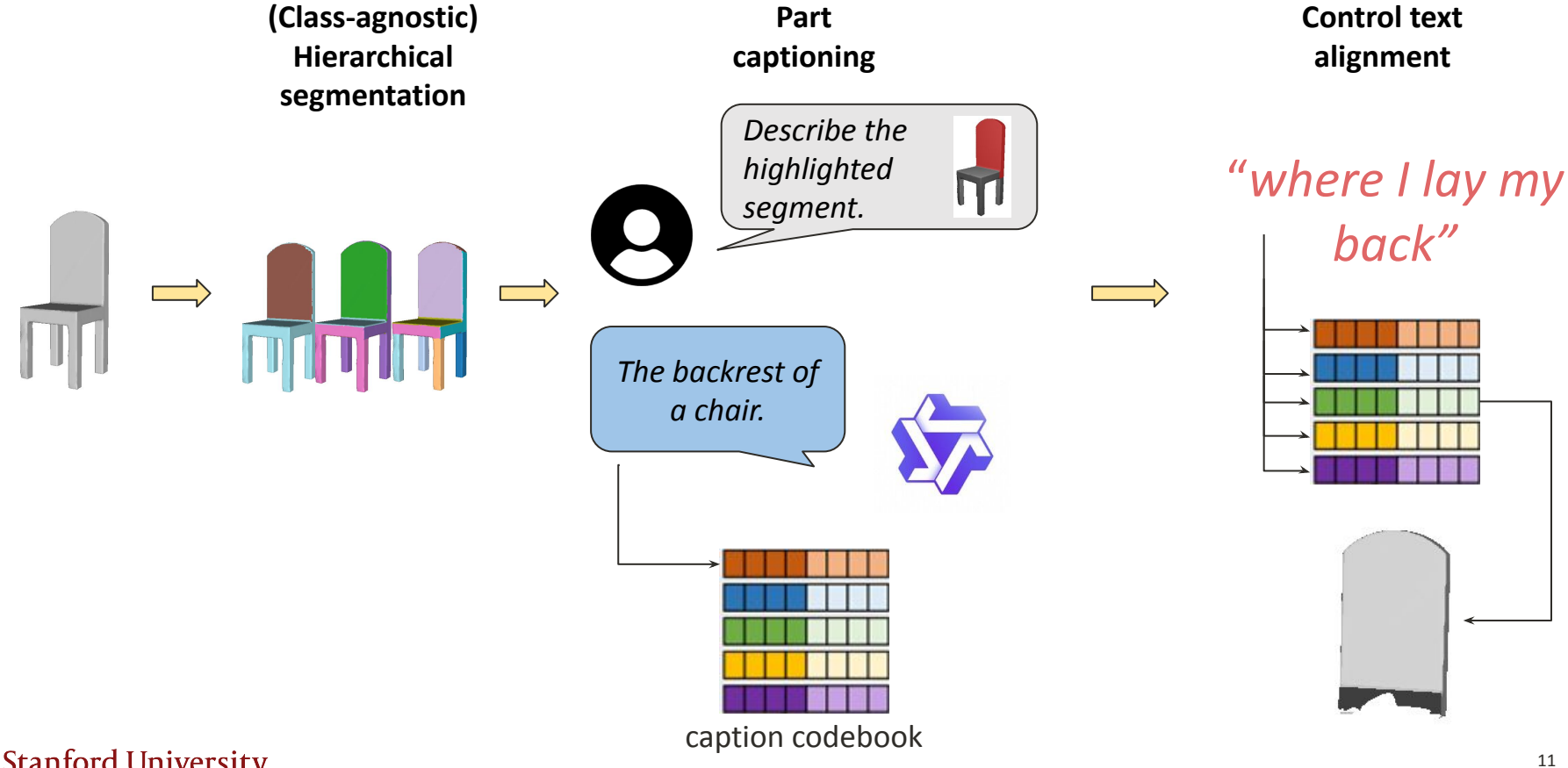
- Text *“seat surface”*

Outputs:

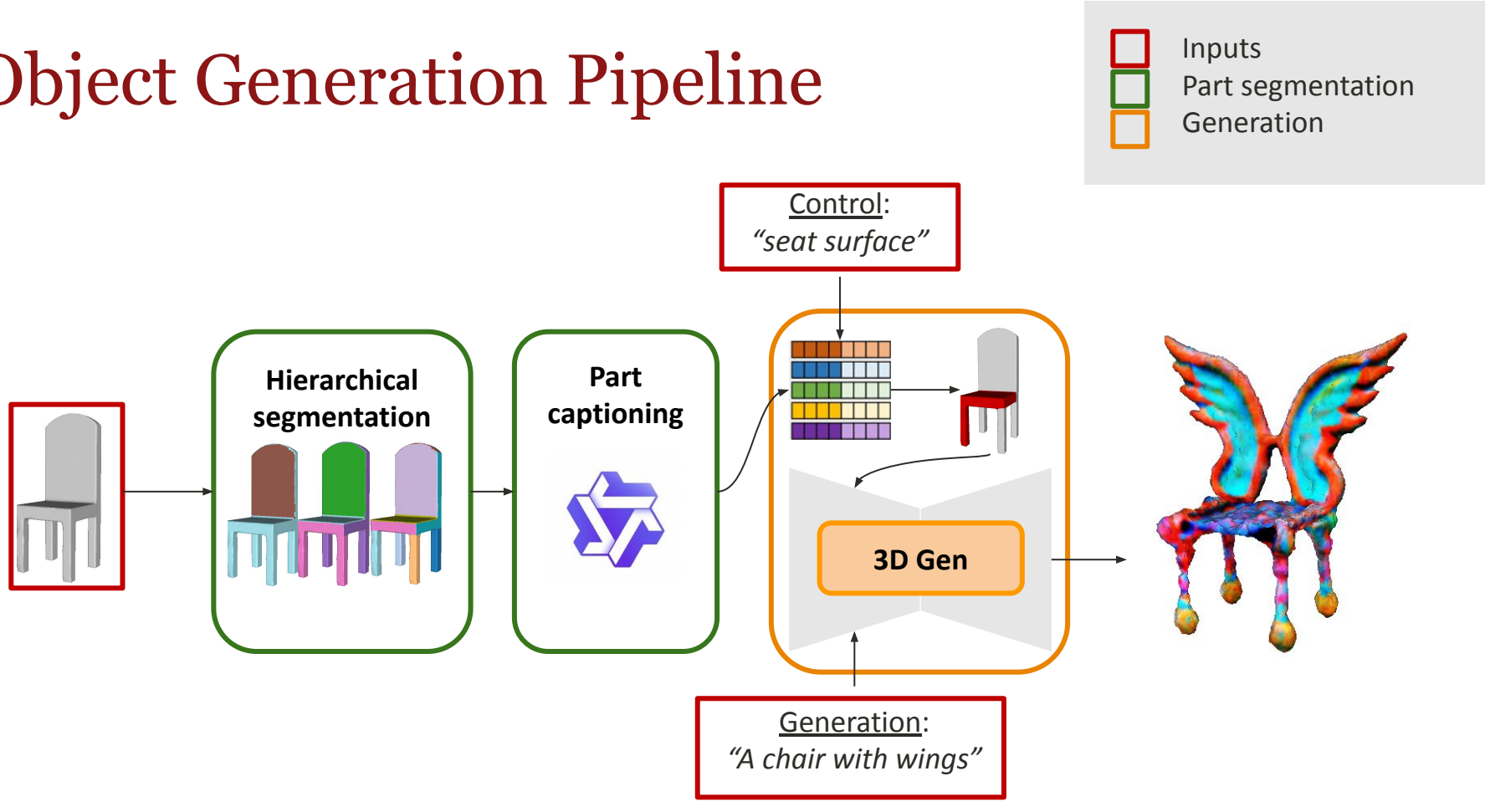
- Part segment



How do we get part masks?



Object Generation Pipeline

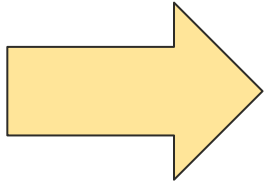


How to reimagine a scene?

- Many objects require many inputs…
- Text inputs need not come from the end user!

End-to-end generative reconstruction with LLMs

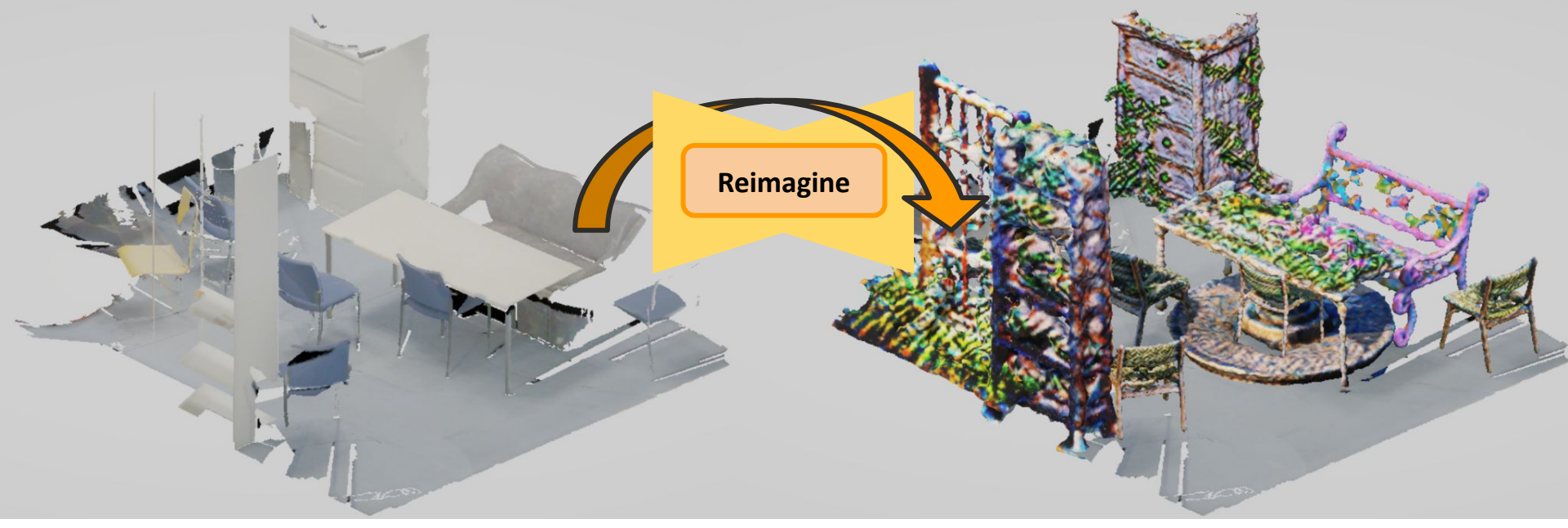
"Turn my office into a lovely garden"



Sure thing! Here are:

1) The **objects** in the scene
2) Their crucial **interaction points**
3) Their **generation prompts**

Office Transformation



Contributions

1) A method of locally-controlled generation with part segment masks

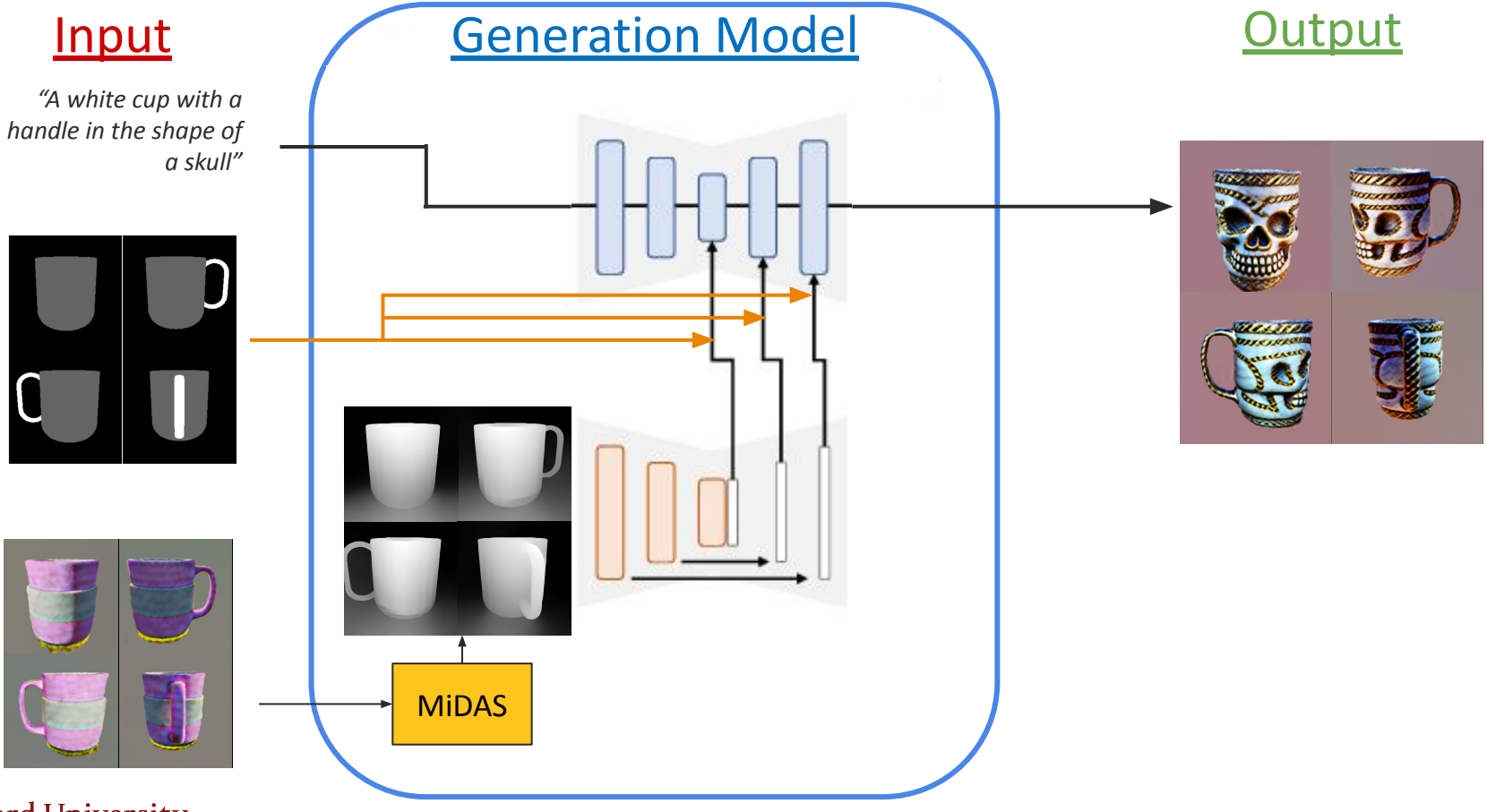
2) A method of semantic part segmentation

3) A system to perform 3D generative reconstruction on 3D objects

Thank You

Appendix

Locally-controlled 3D Generation



This beats SOTA semantic part segmentation methods

Method	Precision	Recall	F1 score
PartSLIP	31.13	35.74	25.95
Find3D	27.29	31.52	21.81
Ours	45.01	90.27	50.09

Prompting to transform

GPT o3-mini output (tl;dr)

Object	Preserve	Prompt
Chair	Seat	Lush garden benches with vines climbing over
Table	Surface	Stone table covered with blooming flowers
Printers	Interface	Stone fountain with clear water
TV	Screen	Wooden frame with garden mural
Cabinets	Doors	Garden trellis with fairy lights