Microsoft

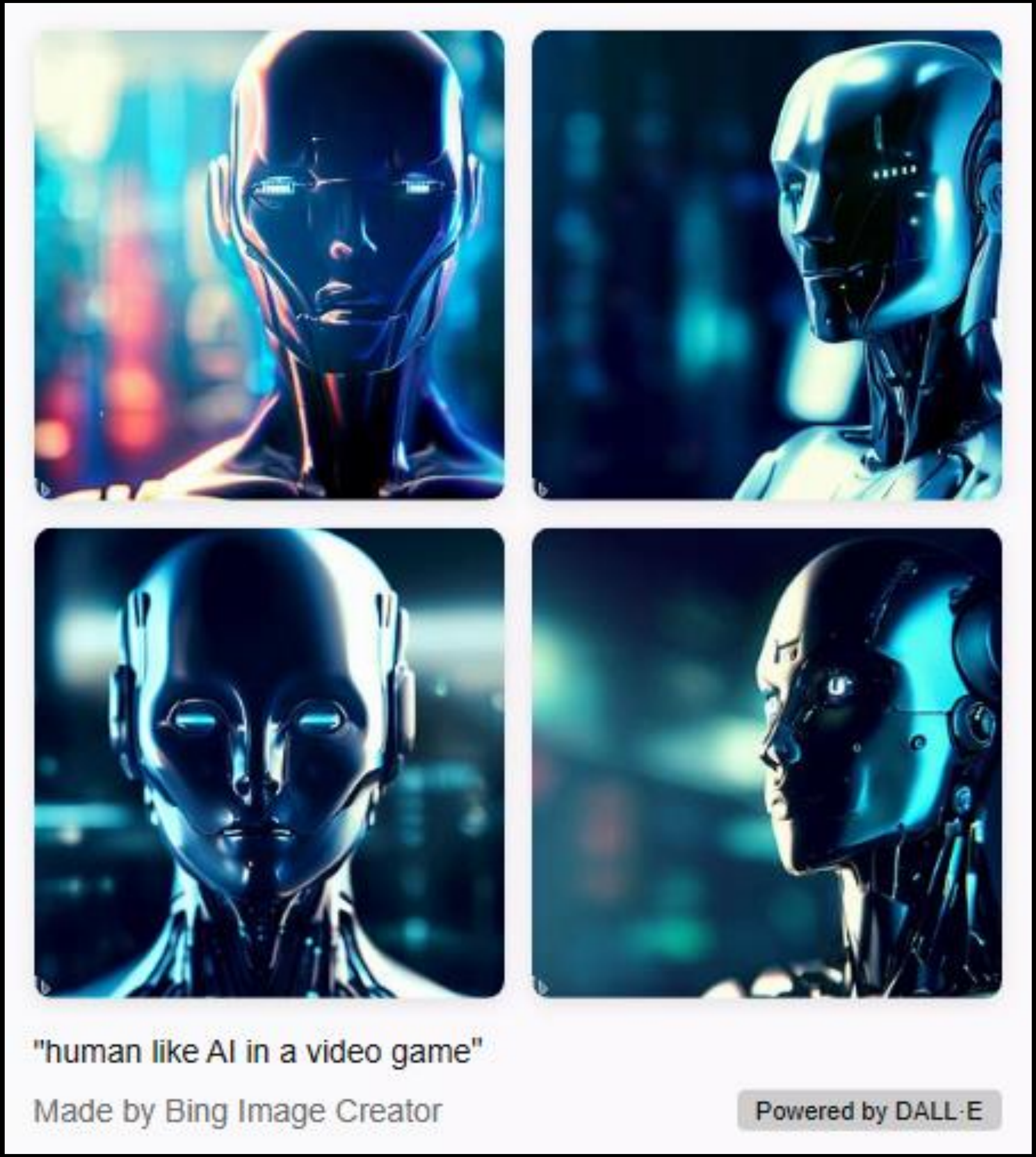
Towards Human-Like AI in Video Games

Katja Hofmann

Sr. Principal Researcher
Microsoft Research
Cambridge, UK

aka.ms/gameintelligence
Twitter: @katjahofmann

Cambridge Ellis Unit Summer School on Probabilistic Machine Learning 2023
18 July 2023


"human like AI in a video game"
Made by Bing Image Creator
Powered by DALL·E

Microsoft

Meet the Team

aka.ms/gameintelligence


Katja Hofmann


Sam Devlin


Sergio Valcarcel Macua


Dave Bignell

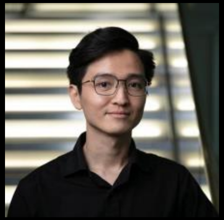

Raluca Georgescu


Tabish Rashid


Tim Pearce


Anssi Kanervisto


Yuhan Cao


Shanzheng Tan

Our current interns


Adam Jelley


Eloi Alonso


Gunshi Gupta


Lukas Schaefer


Tarun Gupta

Key research collaborators


Ali Shaw
Ninja Theory


Gavin Costello
Ninja Theory


Ida Momennejad
MSR New York

Microsoft

Towards Human-Like AI – Outline & Challenges

Limited Data

Uni[MASK]: Unified Inference in Sequential Decision Problems

NeurIPS 2022 Oral

Multi-modal Behavior

Imitating Human Behaviour with Diffusion Models

ICLR 2023
NeurIPS 2022 DRL workshop

Evaluation Challenges

Navigation Turing Test (NTT): Learning to Evaluate Human-Like Navigation

CHI 2023
CHI 2022 Extended Abstract
ICML 2021

Microsoft

Challenge: Limited Data

Uni[MASK]: Unified Inference in Sequential Decision Problems

NeurIPS 2022 Oral

Microsoft

Challenge: Limited Data

Uni[MASK]: Unified Inference in Sequential Decision Problems

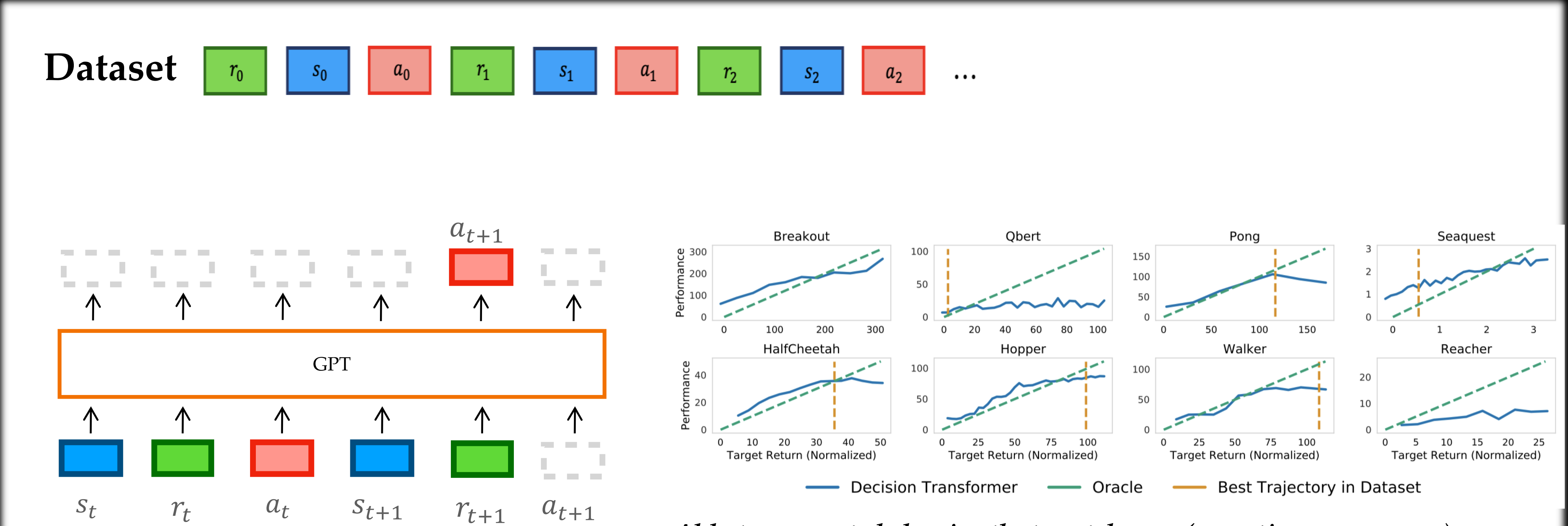
Micah Carroll[1], **Orr Paradise**[1], **Jessy Lin**[1], **Raluca Georgescu**[2], **Mingfei Sun**[2], **David Bignell**[2], **Stephanie Milani**[3], **Katja Hofmann**[2], **Matthew Hausknecht**[2], **Anca Dragan**[1], and **Sam Devlin**[2]

[1]UC Berkeley
[2]Microsoft Research
[3]CMU

For paper details see: aka.ms/unimask

Accepted for oral presentation at NeurIPS 2022. Awarded to only the top 1% of over 10k submissions.
Also see contemporary paper: Liu, Fangchen, Hao Liu, Aditya Grover, and Pieter Abbeel. "Masked Autoencoding for Scalable and Generalizable Decision Making." In Advances in Neural Information Processing Systems.

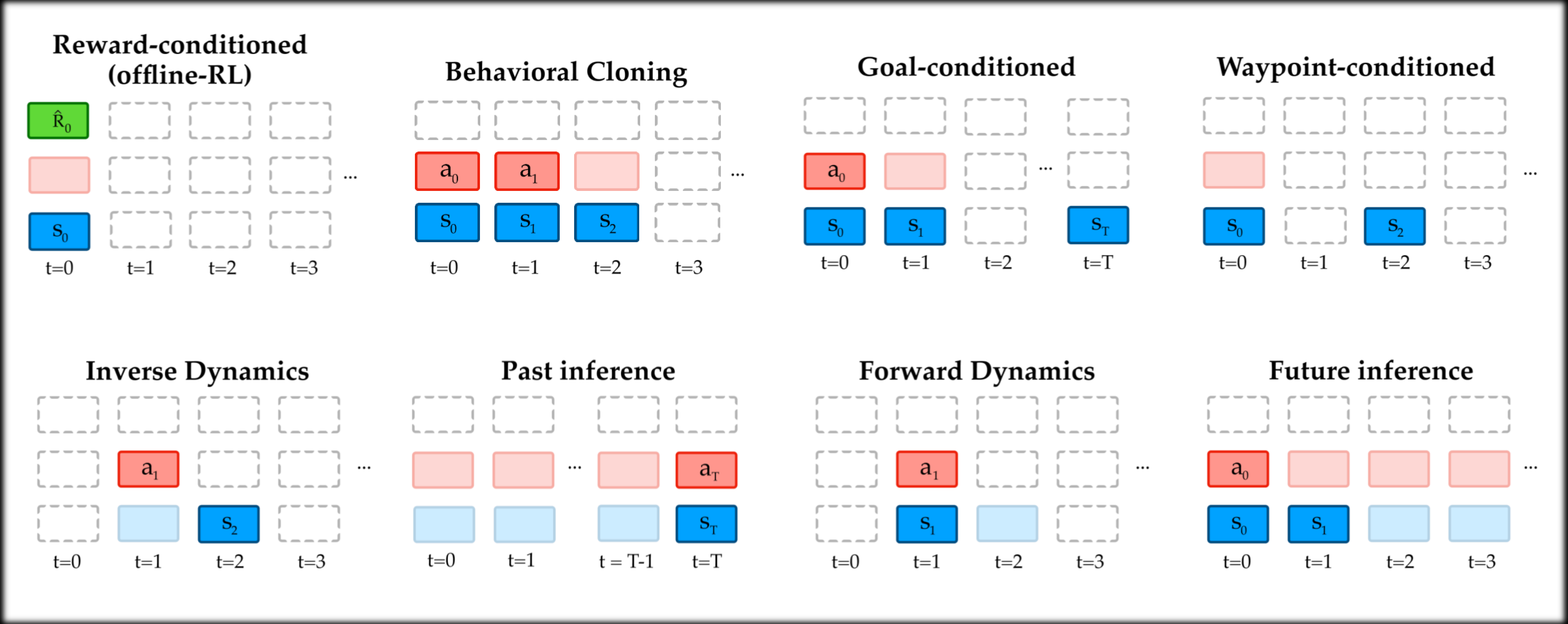
Offline Reinforcement Learning & The Decision Transformer



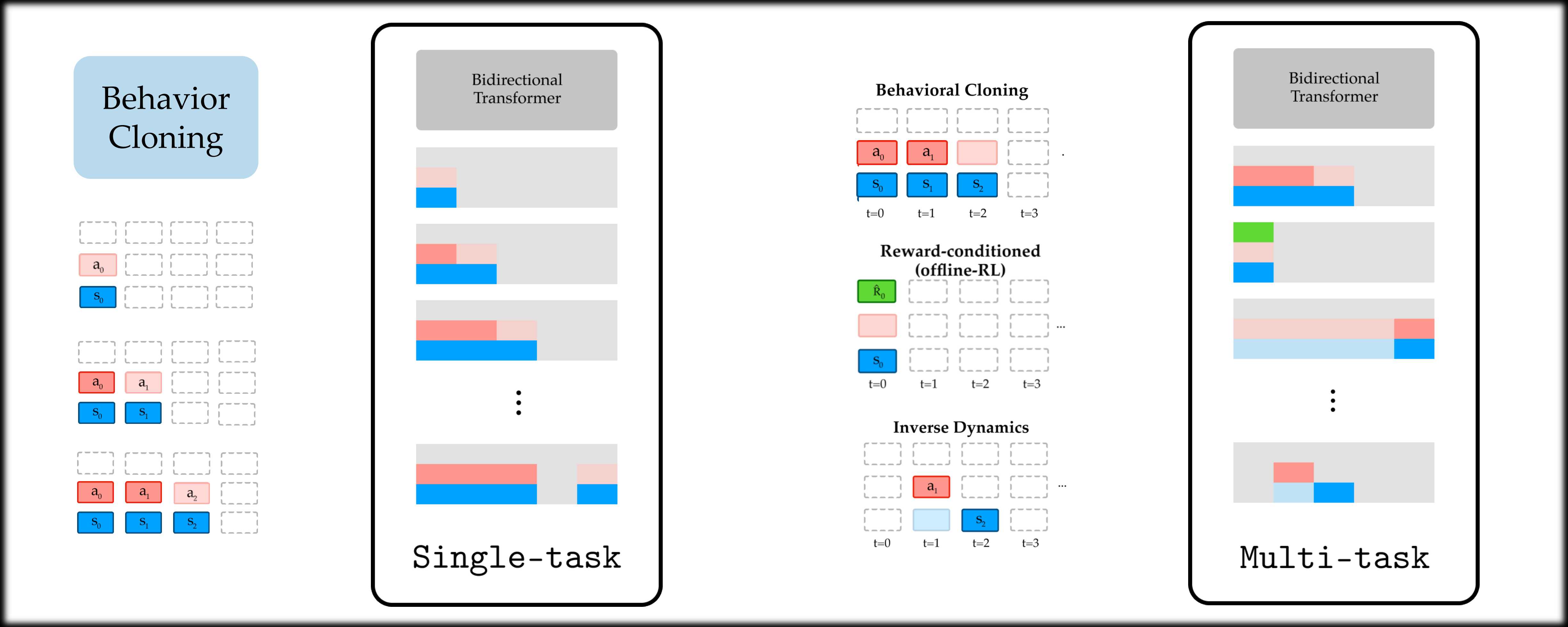
Able to generate behavior that matches or (sometimes surpasses) the best performing demonstration in the data

[1] L. Chen, et al. Decision Transformer: Reinforcement Learning via Sequence Modeling. 2021

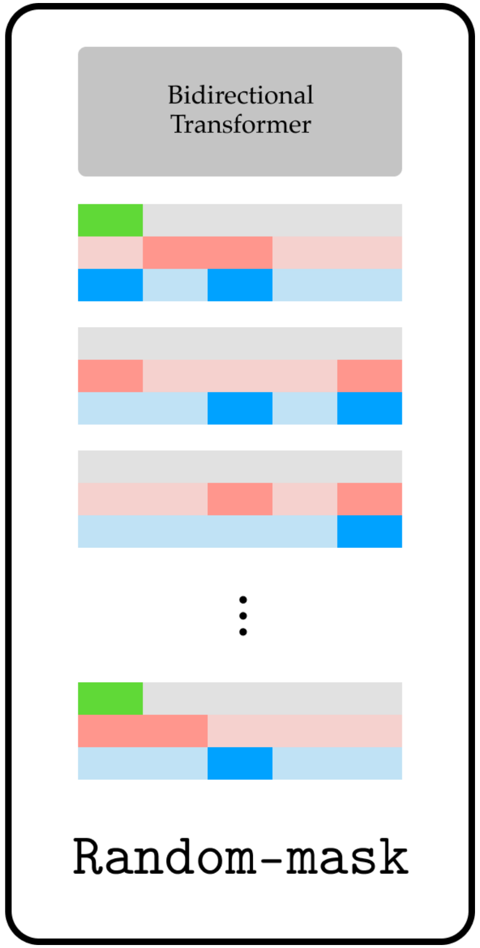
Many common tasks can be represented as input maskings



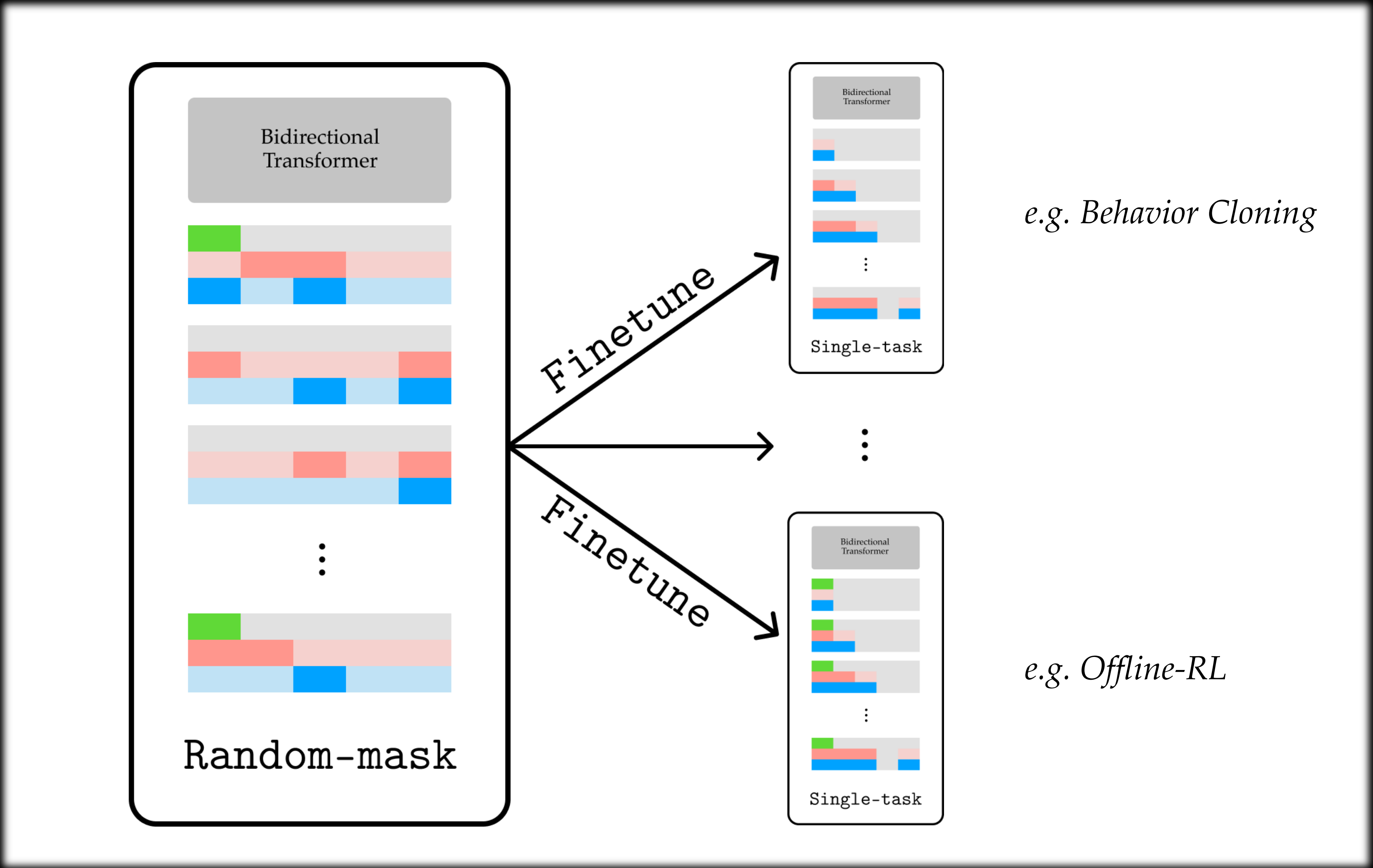
Microsoft


Behavior Cloning
a0
s0
a1
s1
a2
s2
Bidirectional Transformer
Single-task
Behavioral Cloning
t=0
t=1
t=2
t=3
Reward-conditioned (offline-RL)
R̂0
Inverse Dynamics
Bidirectional Transformer
Multi-task

Microsoft



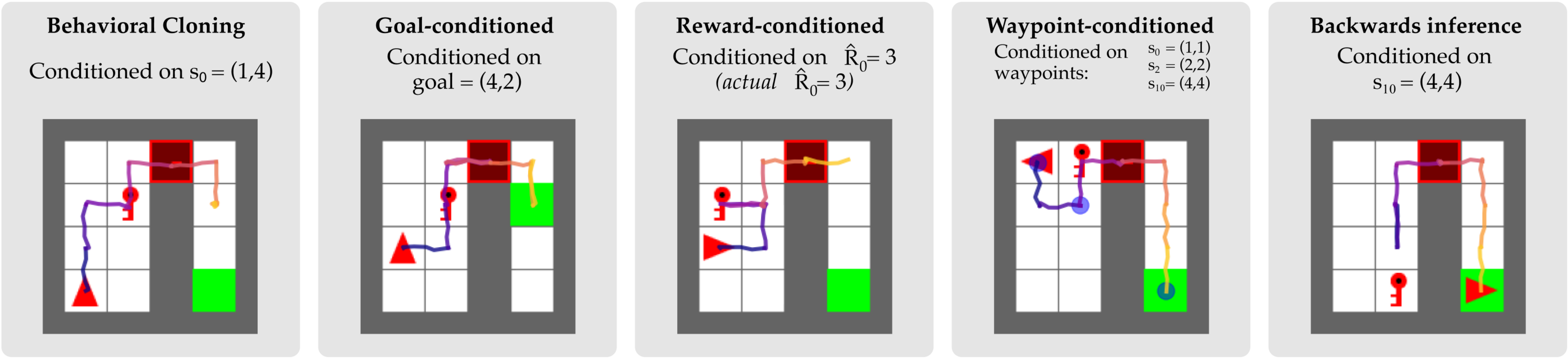
Training on random maskings is equivalent to training a single model
on all possible inferences in a sequential decision problem



Can be used zero-shot, or fine-tuned to increase performance further

A single model for any task!

Training a single model on *all* tasks works, sometimes even better than specialized models!

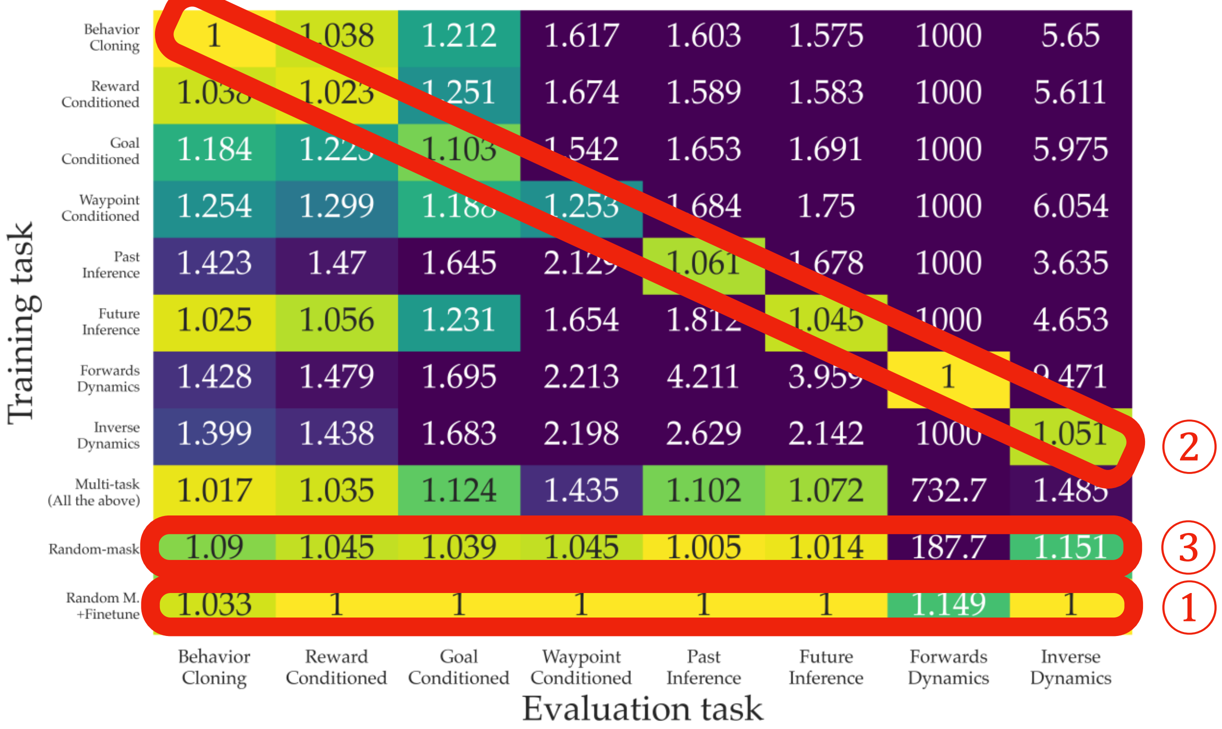


Rollouts from a single pre-trained model on a variety of inference tasks, obtained by conditioning on the appropriate subset of states and actions.

Microsoft

How does this compare to using specialized models?

Normalized validation loss (column-wise)



(Lower values are better)

1 Random masking pre-training + finetuning

>

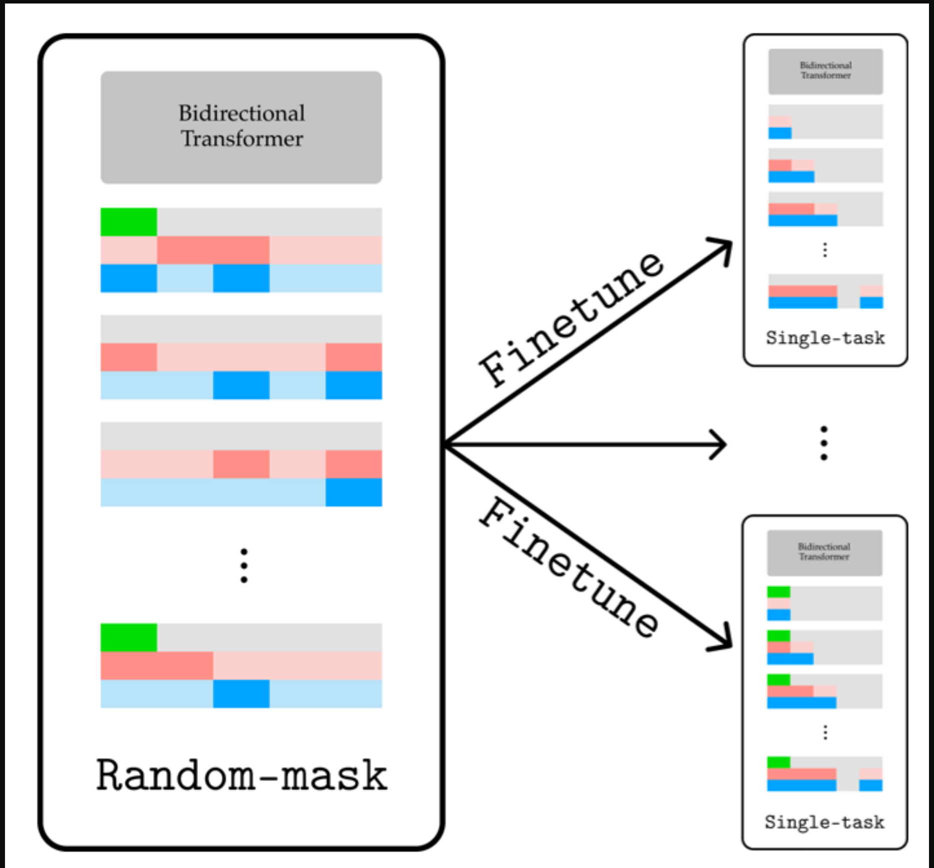
2 Specialized models

3 Random masking pre-training (zero-shot)

≈

2 Specialized models

Insights



Uni[MASK] unifies inference tasks in sequential decision problems as different masking schemes

Randomly sampling masking schemes at training time produces a single multi-inference-task model

Fine-tuning models trained with random masking consistently outperforms single-task models

Microsoft

Challenge: Multi-modal Behavior

Imitating Human
Behaviour with
Diffusion Models

ICLR 2023

NeurIPS 2022 DRL workshop

Microsoft

Accepted to NeurIPS Deep RL Workshop 2022

Imitating Human Behaviour with Diffusion Models

Tim Pearce, Tabish Rashid, Anssi Kanervisto, Dave Bignell, Mingfei Sun, Raluca Georgescu, Sergio Valcarcel Macua, Shan Zheng Tan, Ida Momennejad, Katja Hofmann, Sam Devlin
Microsoft Research

For paper details see: aka.ms/BC-diffusion

ICLR 2023 and NeurIPS DeepRL workshop 2022

Imitating Human Behaviour with Diffusion Models

Model complex outputs > Humans have correlations between action dimensions
Generate samples, <u>not</u> average behaviour > Humans behaviour is multimodal
Scale to large data > Human gameplay datasets can be large

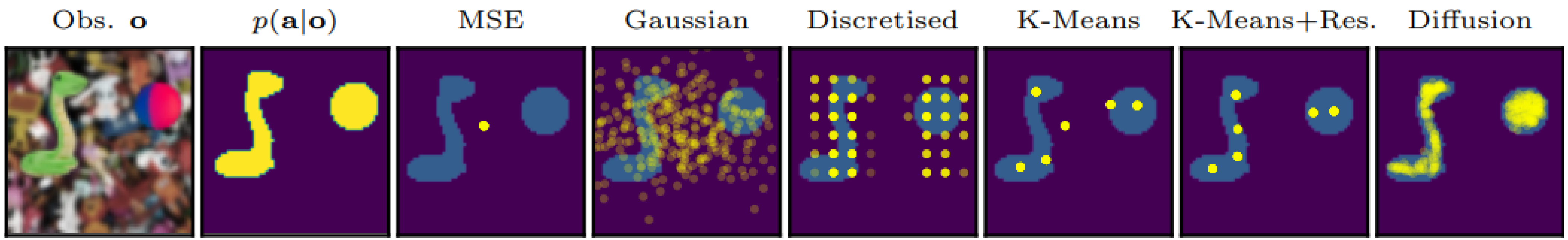
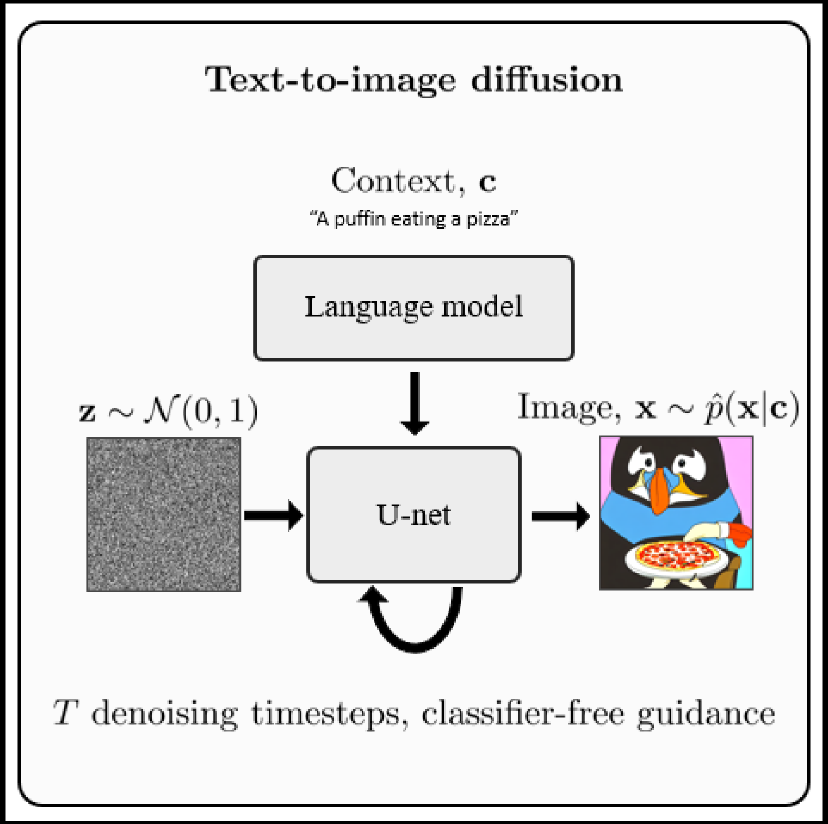


Figure 1: Expressiveness of a variety of models for behaviour cloning in a single-step, arcade claw game with two simultaneous, continuous actions. Existing methods fail to model the full action distribution, $p(\mathbf{a}|\mathbf{o})$, whilst diffusion models excel at covering multimodal & complex distributions.

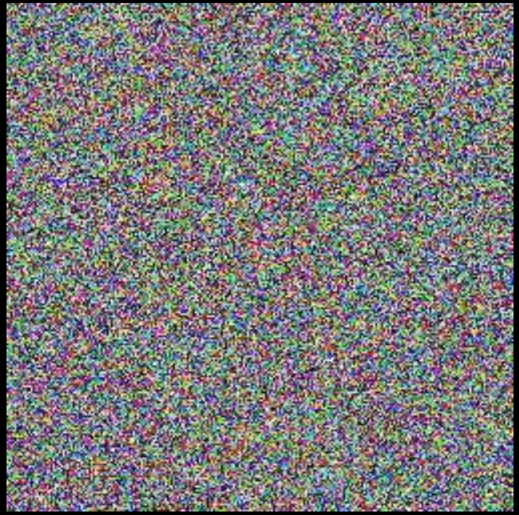
Imitating Human Behaviour with Diffusion Models

Text-to-image diffusion models major AI success of 2022 (Imagen, stable diffusion…)



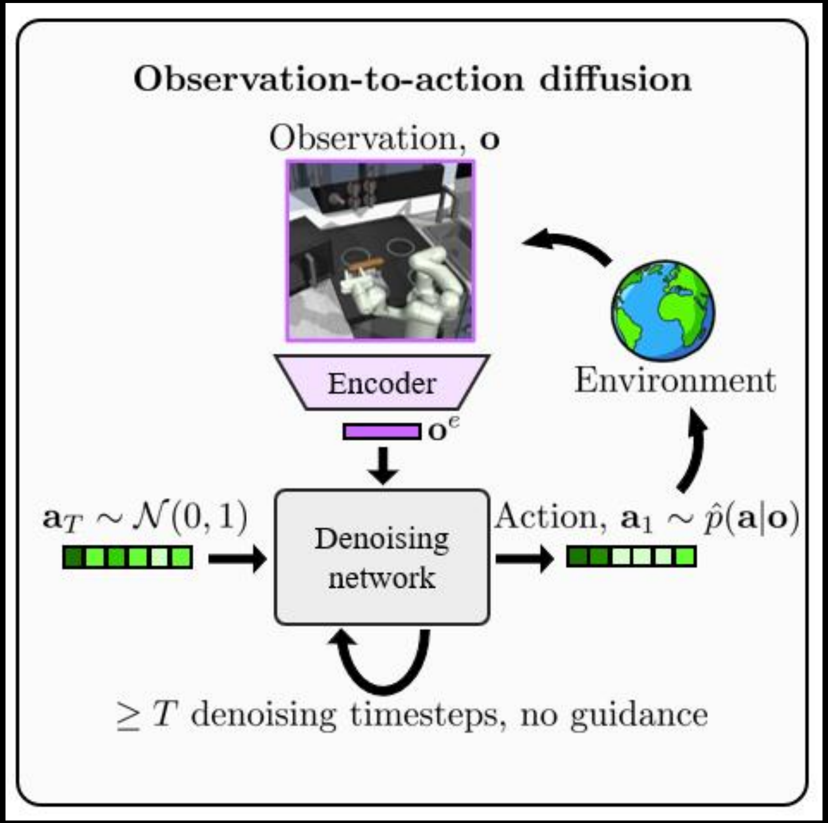
Context: "Person wearing sunglasses"

Output:



Imitating Human Behaviour with Diffusion Models

Key idea: Apply diffusion models as observation-to-action models for learning human behaviour



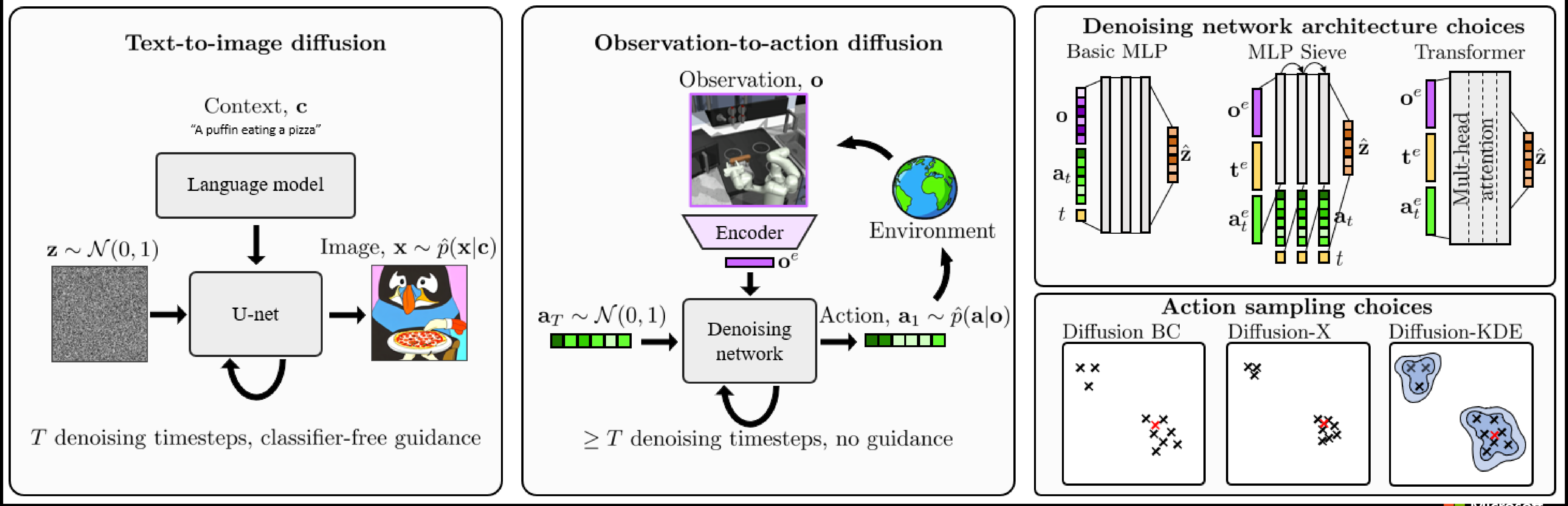
Context:

Output: Mouse_x=-14
Mouse_y=+6
Left_click=True

Microsoft

Imitating Human Behaviour with Diffusion Models

How to make this work?



Microsoft

Imitating Human Behaviour with Diffusion Models

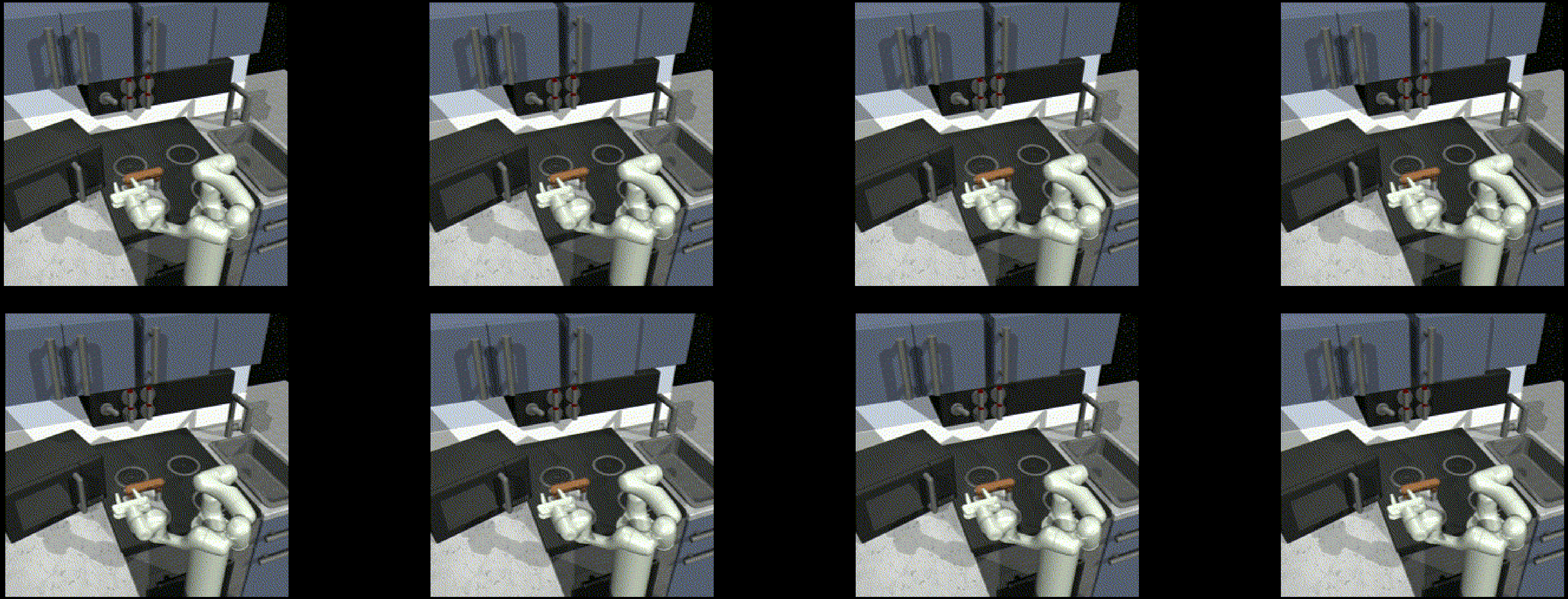
Robotic control for everyday tasks in kitchen environment

	Tasks ≥ 4 ↑	Tasks Wasserstein ↓
Transformer Architecture		
MSE, Transformer	0.69 ± 0.02	1.47 ± 0.13
Discretised, Transformer	0.34 ± 0.02	2.54 ± 0.14
K-Means, Transformer	0.0	5.25
K-Means+Residual, Transformer	0.34 ± 0.02	2.25 ± 0.16
*Diffusion BC, Transformer	0.77 ± 0.01	1.35 ± 0.11
*Diffusion-KDE, Transformer	**0.89** ± **0.01**	1.31 ± 0.03
*Diffusion-X, Transformer	0.88 ± 0.01	1.17 ± 0.13

Diffusion outperforms state of the art baselines
Insights on sampling schemes

Imitating Human Behaviour with Diffusion Models

Diffusion models capture multi-modal behaviour, continuous actions spaces



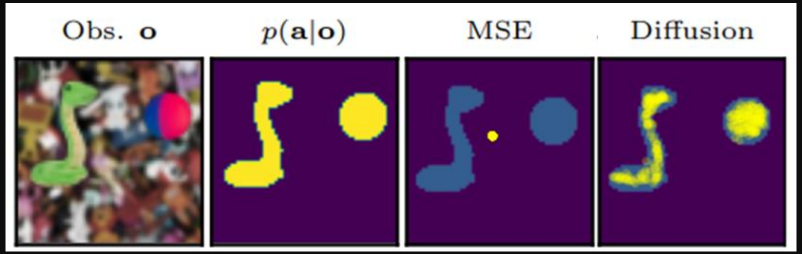
Microsoft

Imitating Human Behaviour with Diffusion Models

Diffusion model applied to a modern FPS game, mixed discrete and continuous actions, pixel input

		Wasserstein Distance, Human to Model ↓		
	Game Score ↑	1×timesteps	16×timesteps (1 sec)	32×timesteps (2 sec)
Observation encoder: ResNet18				
MSE, MLP Sieve	17.8	5.5	28.1	48.9
Discrete, MLP Sieve	14.7	6.6	31.3	53.0
*Diffusion BC, MLP Sieve	19.0	6.3	29.5	50.4
*Diffusion-X, MLP Sieve	**24.0**	**4.5**	**24.5**	**44.4**
Baselines				
Human	36.5	0.73	0.57	0.38

Insights



Diffusion models are an excellent fit for learning complex observation-to-action distributions observed in human behavior

Reliable sampling schemes Diffusion-X and Diffusion-KDE offer benefits over Diffusion BC

Good architecture design is important to the success of Diffusion models

CFG should be avoided when using diffusion agents in sequential environments

Microsoft

Challenge: Evaluation

Navigation Turing Test (NTT): Learning to Evaluate Human-Like Navigation

CHI 2023

CHI 2022 Extended Abstract

ICML 2021

Microsoft

Navigation Turing Test (NTT): Learning to Evaluate Human-Like Navigation

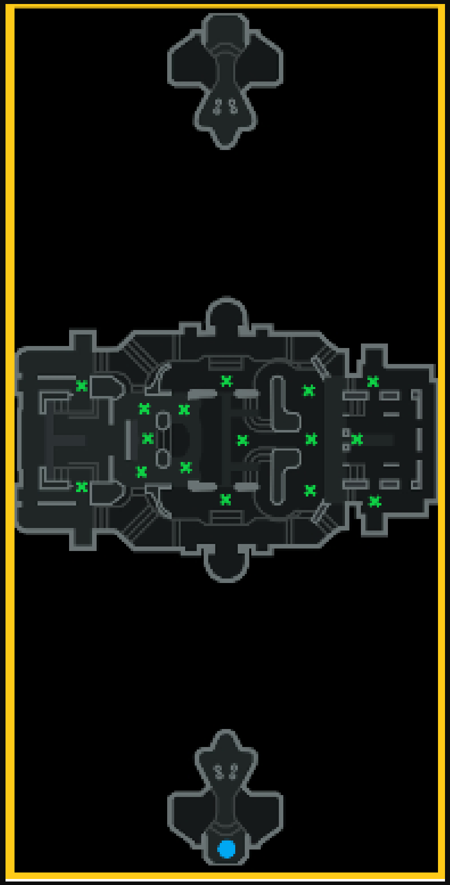
Sam Devlin [*1] Raluca Georgescu [*1] Ida Momennejad [*2] Jaroslaw Rzepecki [*1] Evelyn Zuniga [*1]
Gavin Costello [3] Guy Leroy [1] Ali Shaw [3] Katja Hofmann [1]

* Equal Contribution 1 – Microsoft Research Cambridge 2 – Microsoft Research New York 3 – Ninja Theory

1. How do we reliably measure human-likeness?
2. Do reinforcement learning agents learn to behave in a human-like way?

Presented at ICML 2021. for paper details see: aka.ms/HNTT

Navigation Task



Human demonstration of the navigation task.

Examples are for research purposes only and do not reflect actual game play.

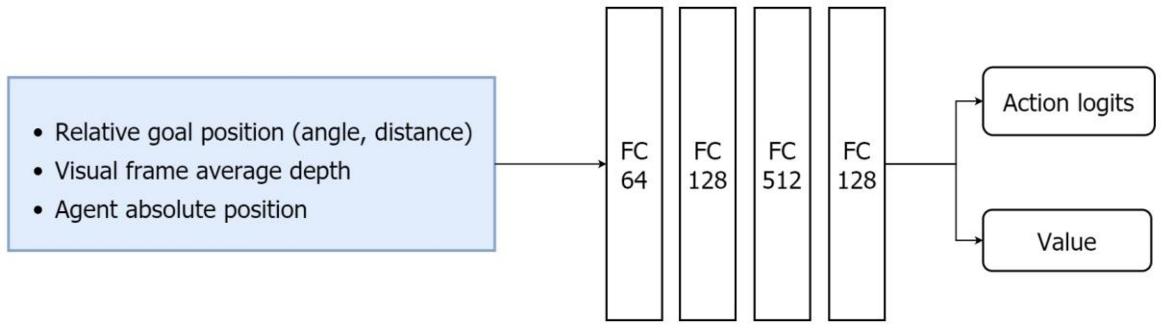
Agents

Hypothesis: more "human-like" hybrid observations lead to more human-like behavior

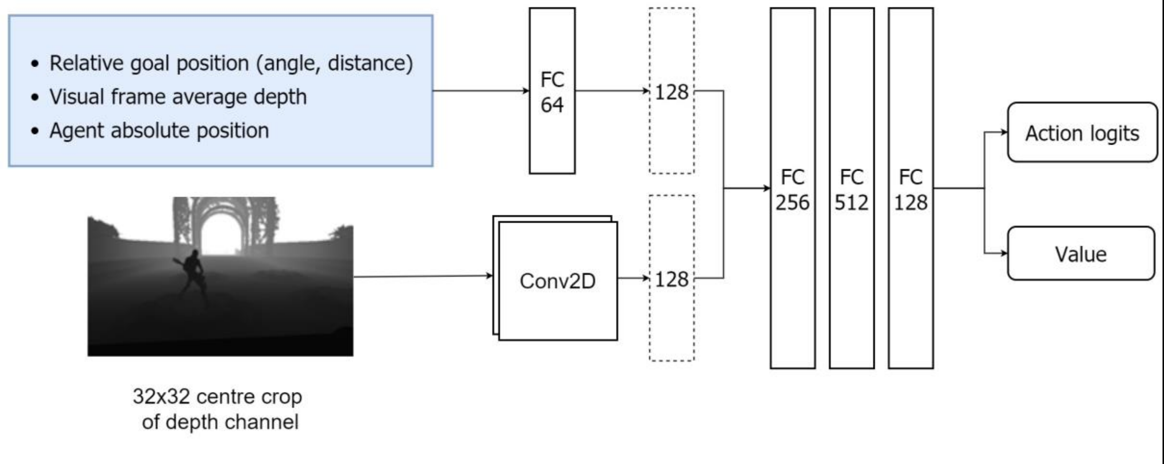
Discrete action space: none, forward, left/right (30,45,90)

Reward: positive when moving towards goal + on reaching goal, negative per step

Symbolic



Hybrid



Baseline state of the art: Alonso E. et al., *Deep Reinforcement Learning for Navigation in AAA Video Games* [2020]
https://arxiv.org/abs/2011.04764

Microsoft

Human Navigation Turing Test (HNTT)

Which video is more likely to be human?

Video A is more likely to be human

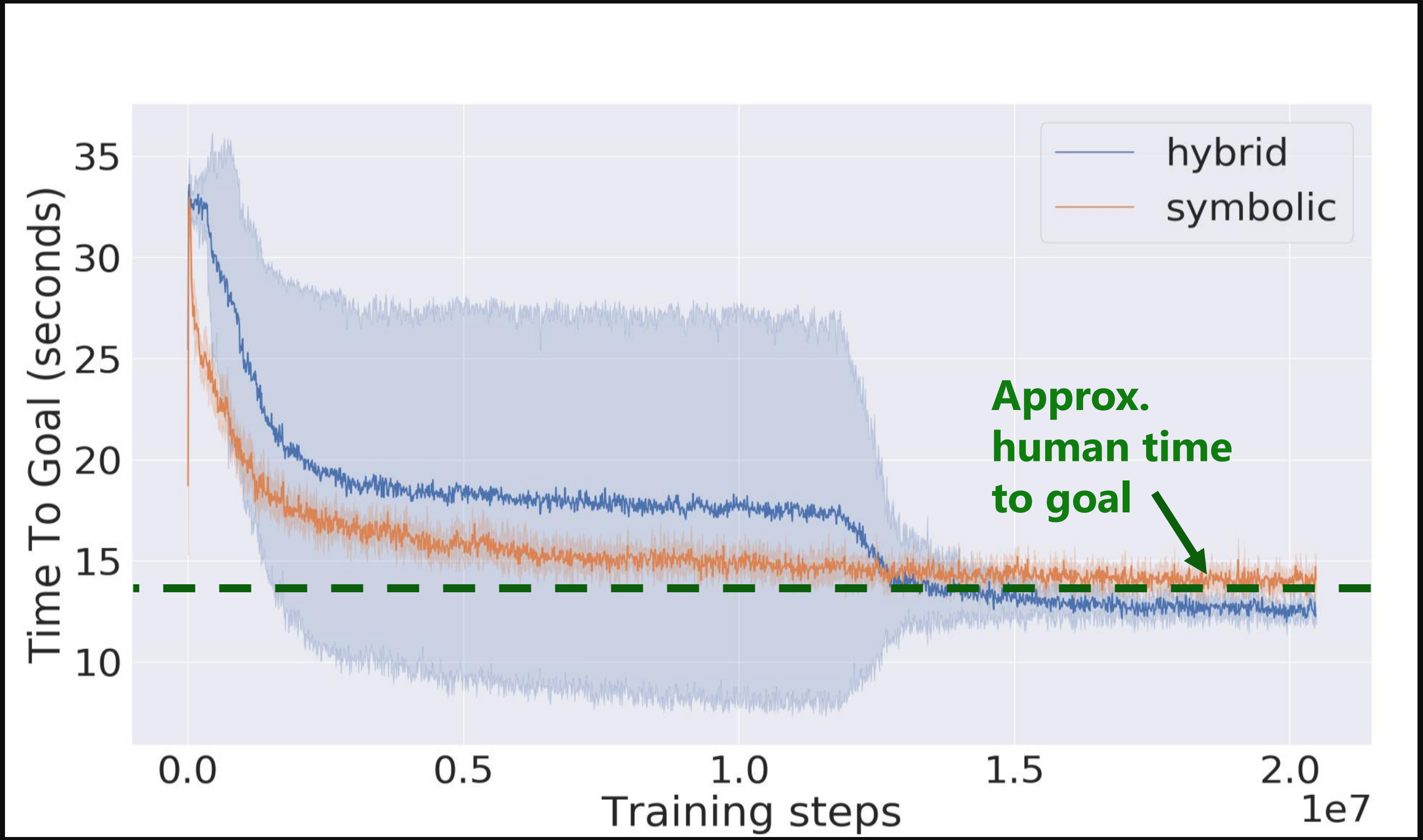
○

Video B is more likely to be human

○

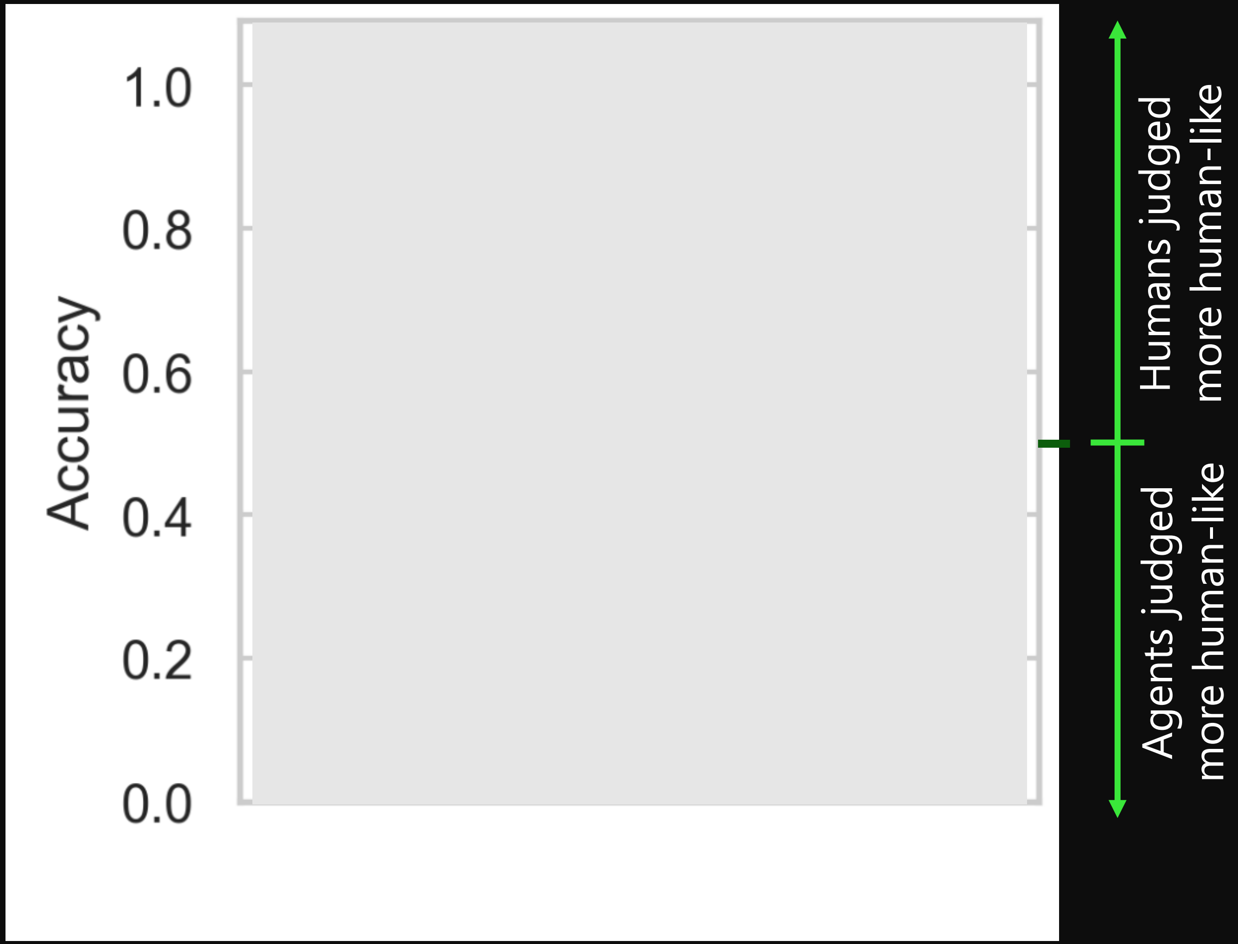
+ asked for detailed justification of the response and assessment of uncertainty

Both RL agents learn to navigate effectively

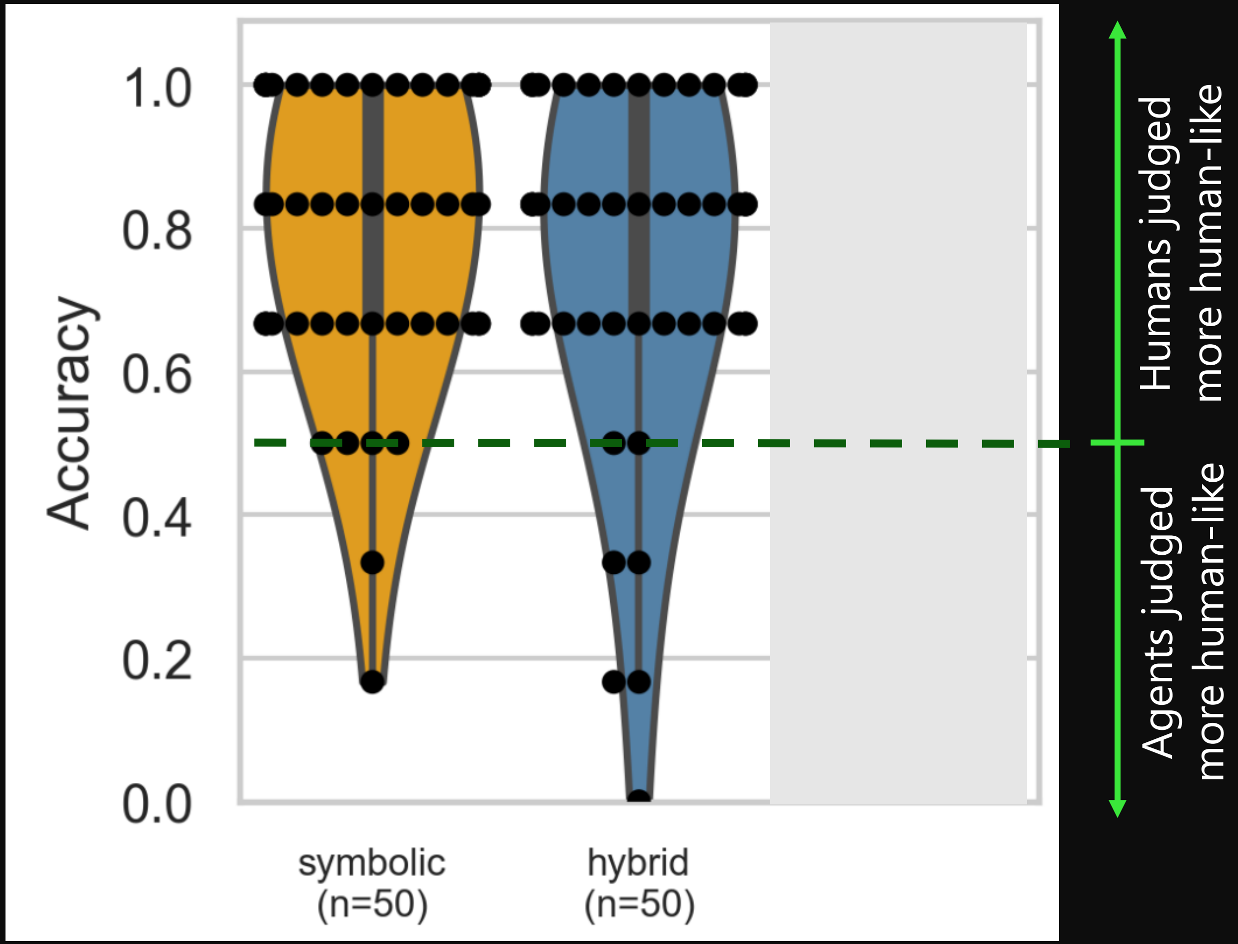


State of the art agents learn to perform as well as humans

High Skill is Unsufficient for Human-likeness

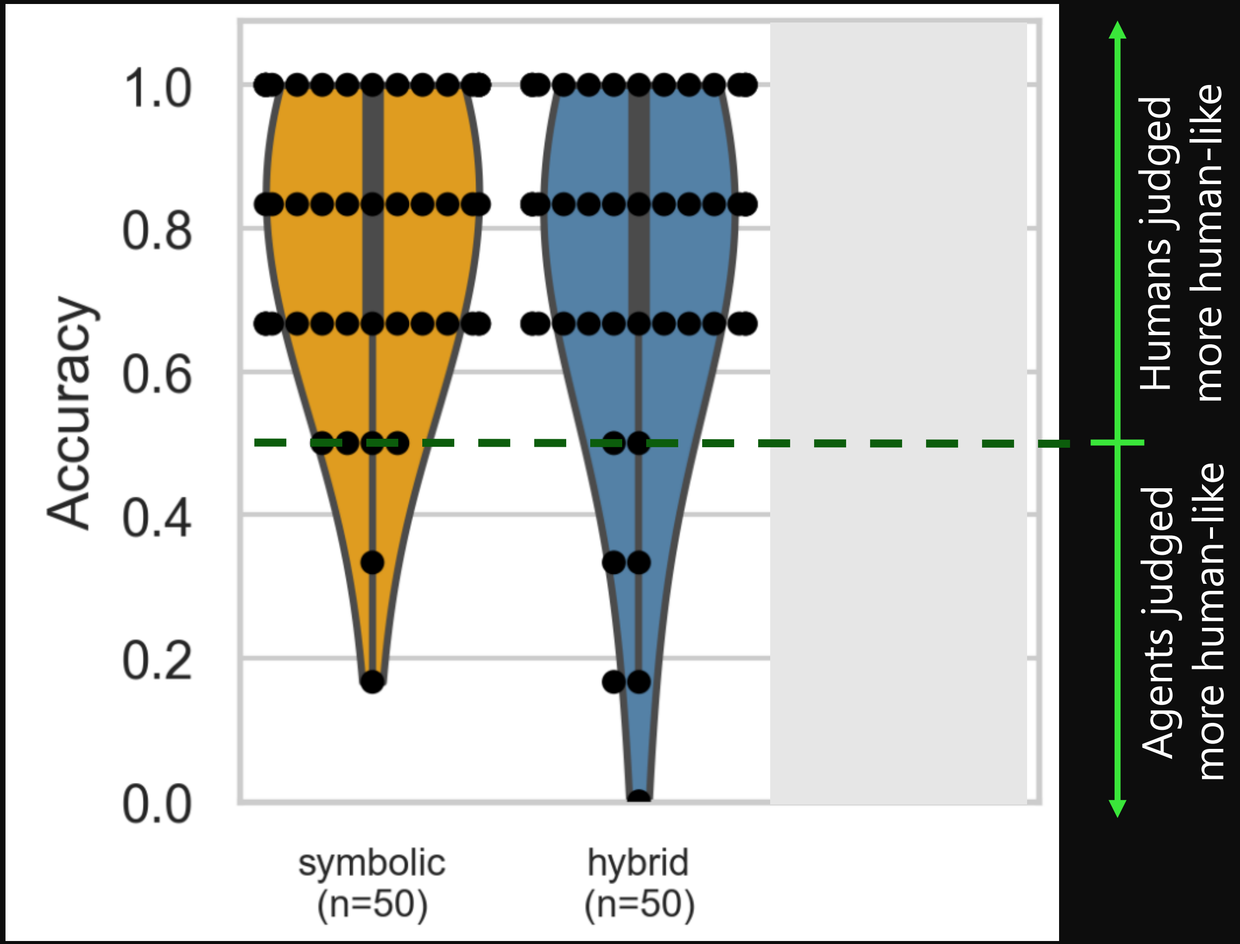


High Skill is Unsufficient for Human-likeness



Human observers can reliably tell the difference between humans and the symbolic and hybrid agents

High Skill is Unsufficient for Human-likeness



How Humans Perceive Human-like Behavior in Video Game Navigation

Evelyn Zuniga*†
Microsoft Research, Cambridge
Cambridge, United Kingdom
t-ezuniga@microsoft.com

Stephanie Milani*†
Carnegie Mellon University
Pittsburgh, USA
smilani@cs.cmu.edu

Guy Leroy*
Microsoft Research, Cambridge
Cambridge, United Kingdom
t-gleroy@microsoft.com

Jaroslaw Rzepecki†
Monumo
Cambridge, United Kingdom

Raluca Geogescu
Microsoft Research, Cambridge
Cambridge, United Kingdom

Ida Momennejad
Microsoft Research, New York
New York City, USA

Dave Bignell
Microsoft Research, Cambridge
Cambridge, United Kingdom

Mingfei Sun
Microsoft Research, Cambridge, and University of Oxford
Cambridge and Oxford, United Kingdom

Alison Shaw
Ninja Theory
Cambridge, United Kingdom

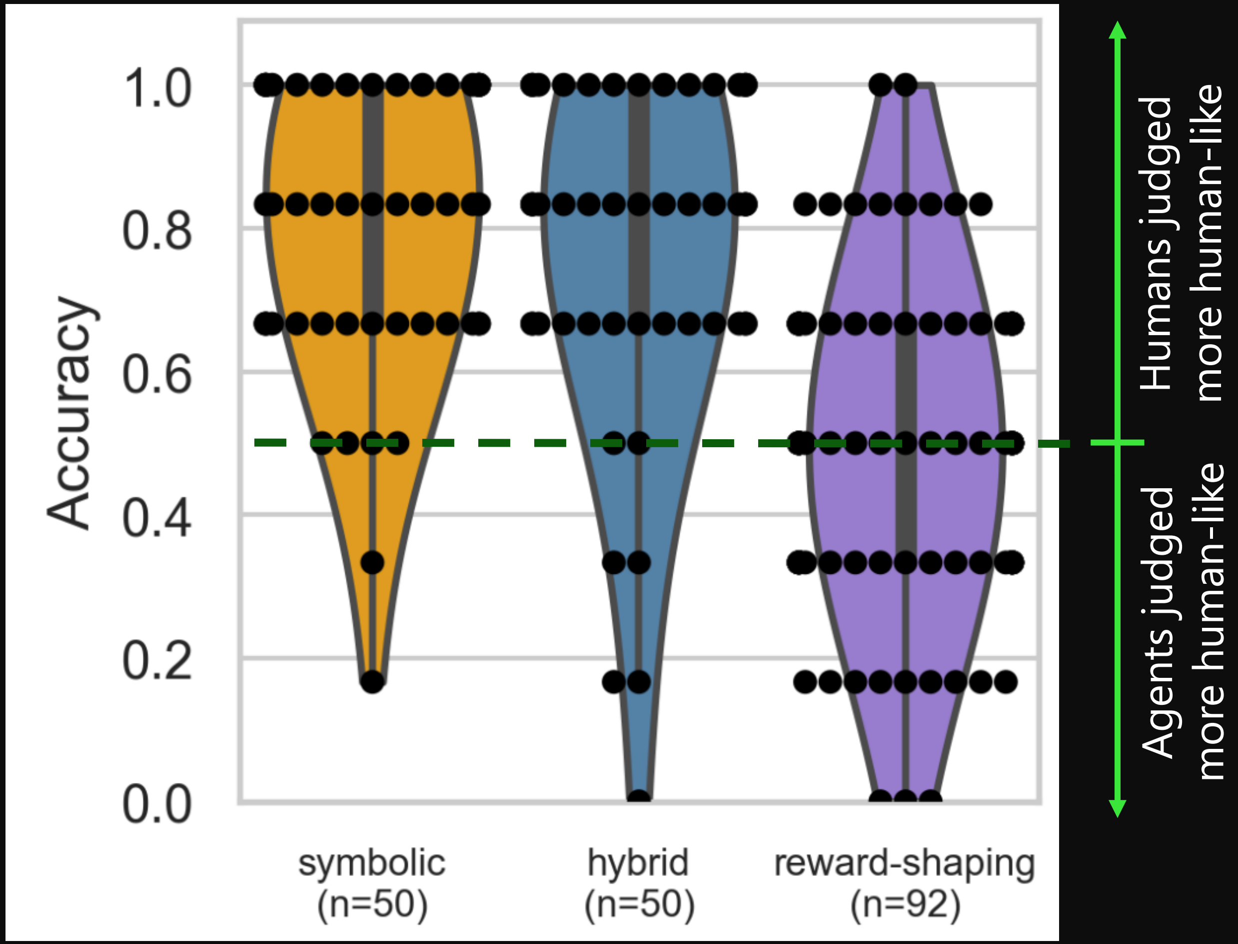
Gavin Costello
Ninja Theory
Cambridge, United Kingdom

Mikhail Jacob†
Resolution Games
Stockholm, Sweden

Sam Devlin
Microsoft Research, Cambridge
Cambridge, United Kingdom

Katja Hofmann
Microsoft Research, Cambridge
Cambridge, United Kingdom

The NTT can be passed by RL with reward shaping



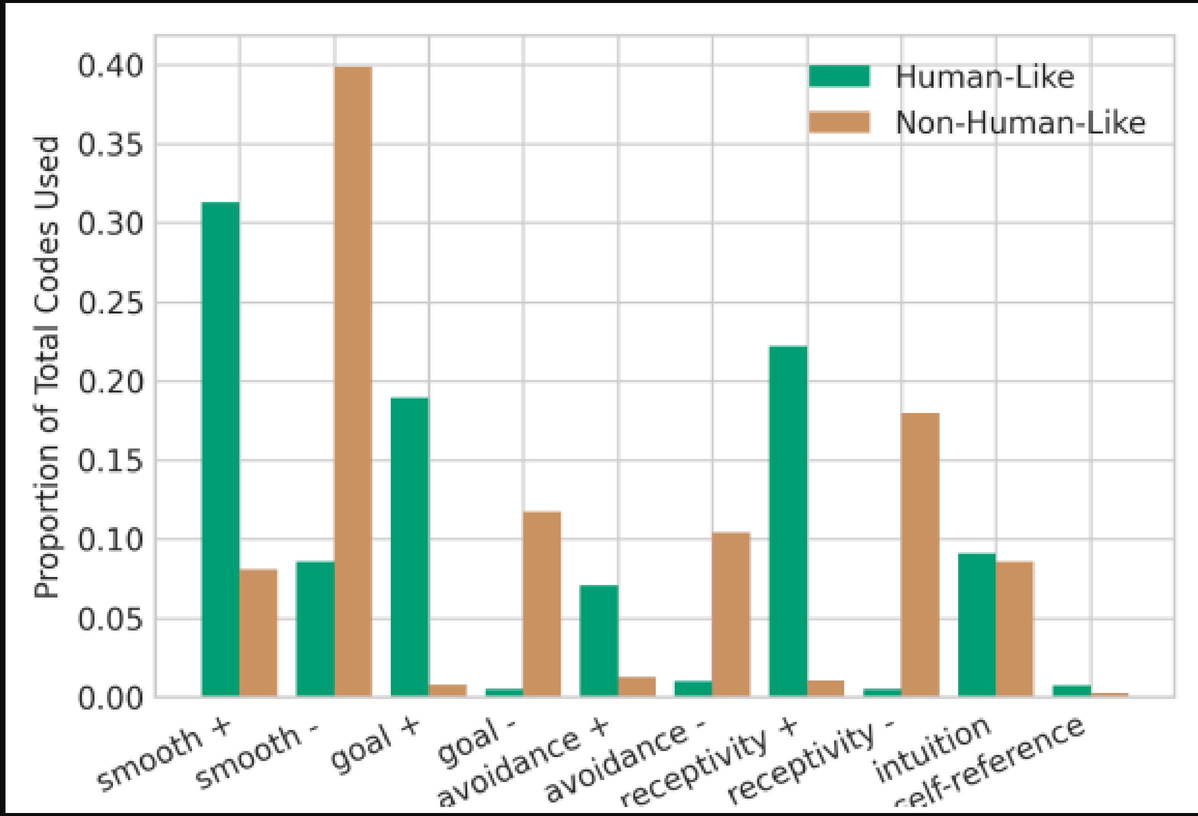
How Humans Perceive Human-like Behavior in Video Game Navigation

Evelyn Zuniga*† — Microsoft Research, Cambridge, Cambridge, United Kingdom, t-ezuniga@microsoft.com
Stephanie Milani*† — Carnegie Mellon University, Pittsburgh, USA, smilani@cs.cmu.edu
Guy Leroy* — Microsoft Research, Cambridge, Cambridge, United Kingdom, t-gleroy@microsoft.com
Jaroslaw Rzepecki† — Monumo, Cambridge, United Kingdom
Raluca Geogescu — Microsoft Research, Cambridge, Cambridge, United Kingdom
Ida Momennejad — Microsoft Research, New York, New York City, USA
Dave Bignell — Microsoft Research, Cambridge, Cambridge, United Kingdom
Mingfei Sun — Microsoft Research, Cambridge, and University of Oxford, Cambridge and Oxford, United Kingdom
Alison Shaw — Ninja Theory, Cambridge, United Kingdom
Gavin Costello — Ninja Theory, Cambridge, United Kingdom
Mikhail Jacob† — Resolution Games, Stockholm, Sweden
Sam Devlin — Microsoft Research, Cambridge, Cambridge, United Kingdom
Katja Hofmann — Microsoft Research, Cambridge, Cambridge, United Kingdom

The Human Navigation Turing Test is passed by a reward shaping agent

Microsoft

Humans are consistent on what’s human-like



Navigates Like Me: Understanding How People Evaluate Human-Like AI in Video Games

Stephanie Milani
smilani@andrew.cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Arthur Juliani
Microsoft Research
New York, New York, USA

Ida Momennejad
Microsoft Research
New York, New York, USA

Raluca Georgescu
Microsoft Research
Cambridge, United Kingdom

Jaroslaw Rzpecki
Monumo
Cambridge, United Kingdom

Alison Shaw
Ninja Theory
Cambridge, United Kingdom

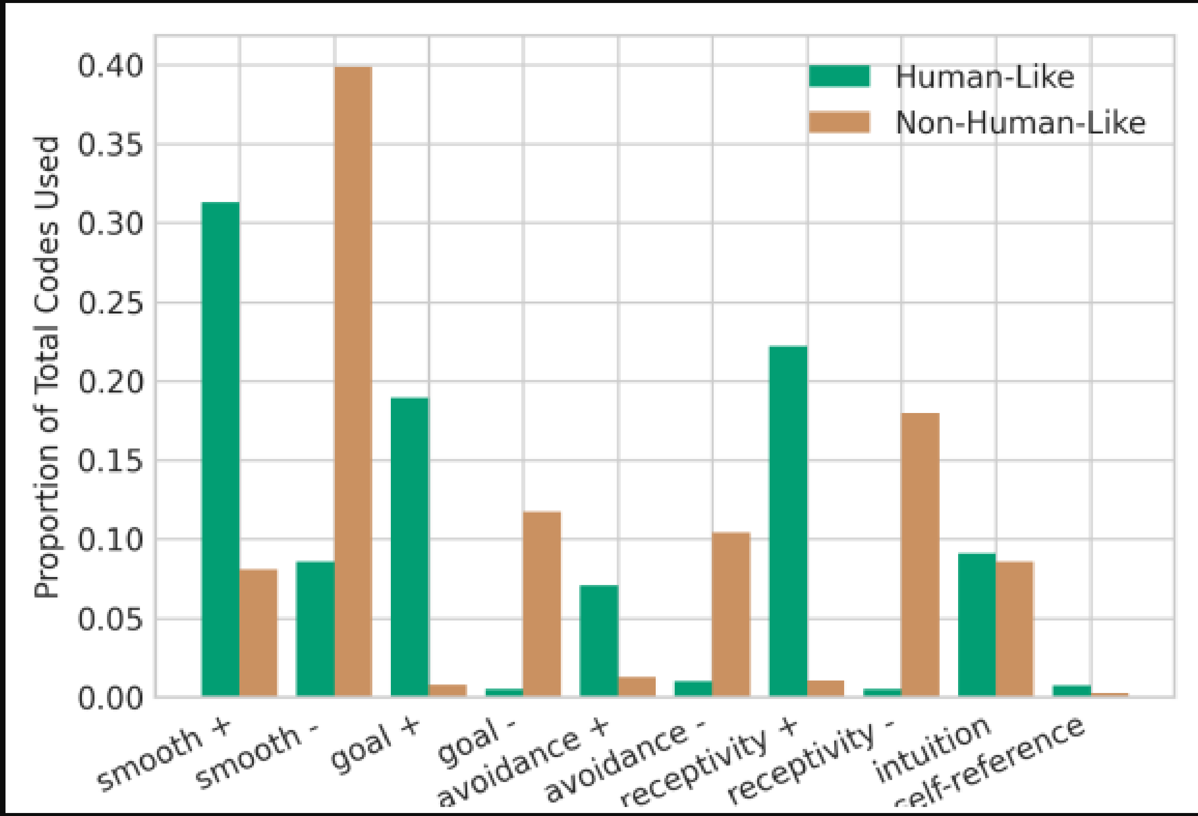
Gavin Costello
Ninja Theory
Cambridge, United Kingdom

Fei Fang
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Sam Devlin
Microsoft Research
Cambridge, United Kingdom

Katja Hofmann
Microsoft Research
Cambridge, United Kingdom

Humans are consistent on what’s human-like



Navigates Like Me: Understanding How People Evaluate Human-Like AI in Video Games

Stephanie Milani
smilani@andrew.cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Arthur Juliani
Microsoft Research
New York, New York, USA

Ida Momennejad
Microsoft Research
New York, New York, USA

Raluca Georgescu
Microsoft Research
Cambridge, United Kingdom

Jaroslaw Rzpecki
Monumo
Cambridge, United Kingdom

Alison Shaw
Ninja Theory
Cambridge, United Kingdom

Gavin Costello
Ninja Theory
Cambridge, United Kingdom

Fei Fang
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Sam Devlin
Microsoft Research
Cambridge, United Kingdom

Katja Hofmann
Microsoft Research
Cambridge, United Kingdom

Human judges strongly associate smooth movement, goal-directedness, collision avoidance, and environment receptivity with human-like behavior

Insights

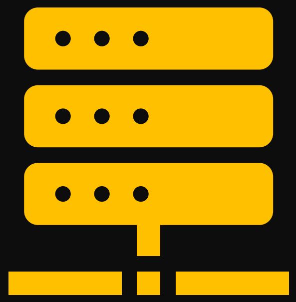
- Deep reinforcement learning agents can learn efficient navigation in 3D environments (ICML '21)
- High Skill Alone Is Not Enough
  For Reproducing Human-Likeness (ICML '21)
- The Human Navigation Turing Test (HNTT) can be passed by a reward shaping RL agent (CHI EA '22)
- Humans are consistent on what they consider human-like or non-human-like (CHI '23)

Microsoft

Outlook & Conclusion

Towards Human-Like AI – Opportunities

Limited Data



Novel training approaches that use limited data more effectively and scale well as more data becomes available

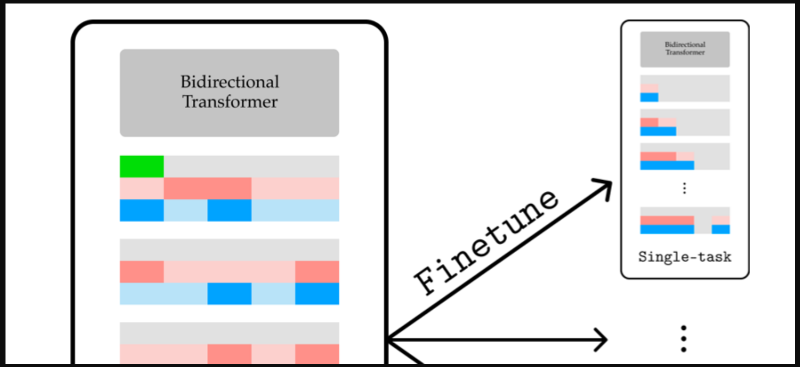
Multi-modal Behavior

Novel model architectures for learning rich representations

Difficult Evaluation

Novel data uses and labelling approaches for reliable evaluation at lower cost

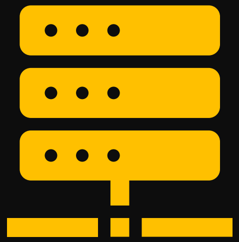
Summary

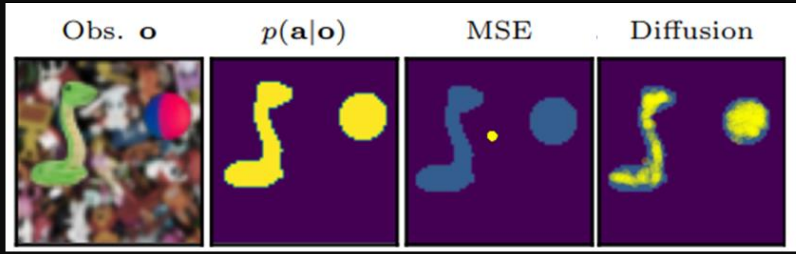


Uni[MASK]: Unified Inference in Sequential Decision Problems

Micah Carroll, Orr Paradise, Jessy Lin, Raluca Georgescu, Mingfei Sun, Dave Bignell, Stephanie Milani, Katja Hofmann, Matthew Hausknecht, Anca Dragan, Sam Devlin
NeurIPS 2022 Oral – aka.ms/unimask

Limited Data





Imitating Human Behaviour with Diffusion Models

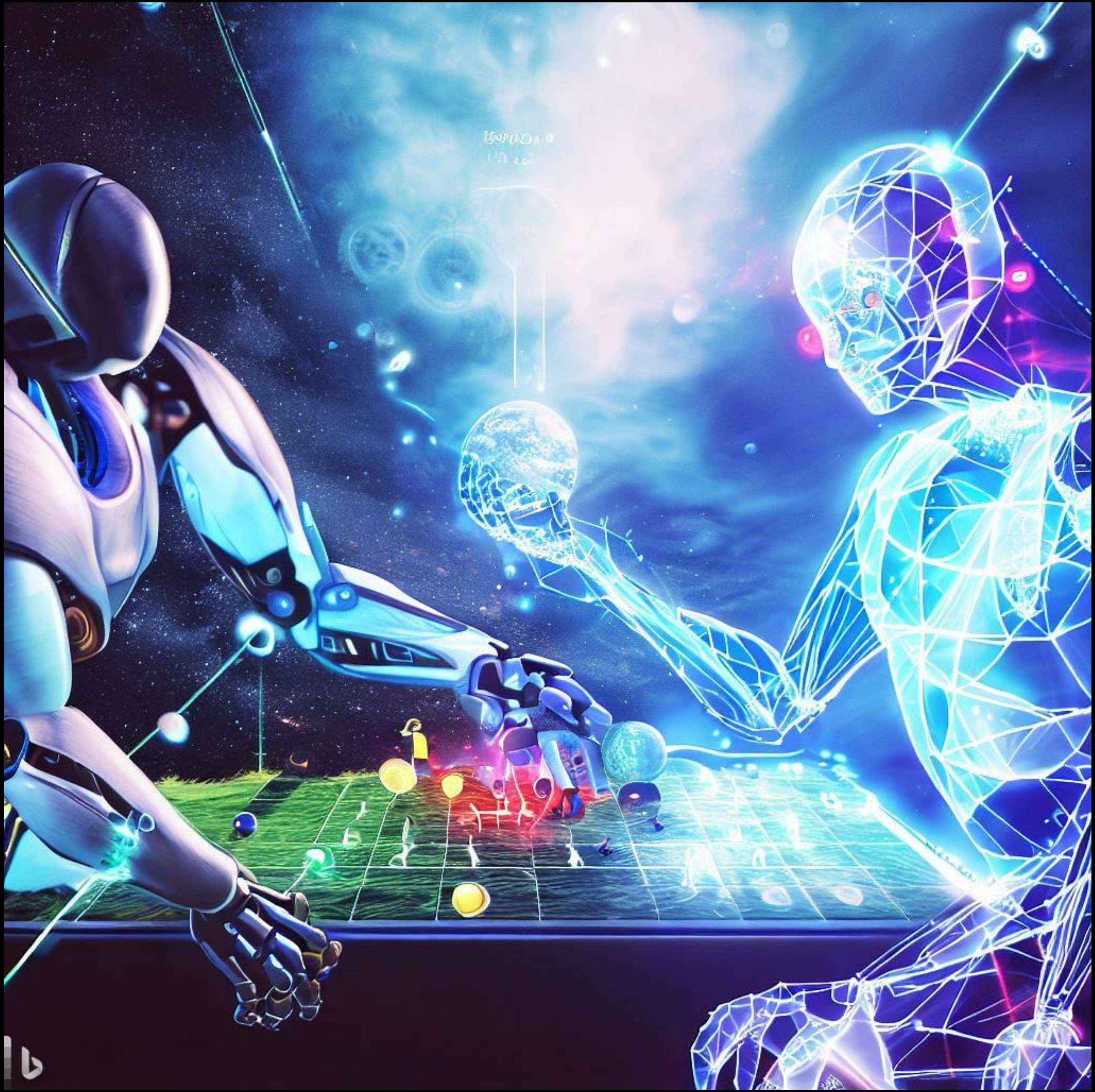
Tim Pearce, Tabish Rashid, Anssi Kanervisto, Dave Bignell, Mingfei Sun, Raluca Georgescu, Sergio Valcarcel Macua, Shanzheng Tan, Ida Momennejad, Katja Hofmann, Sam Devlin
NeurIPS Deep RL Workshop 2022 – aka.ms/BC-diffusion

Multi-modal Behavior

Navigation Turing Test (NTT): Learning to Evaluate Human-Like Navigation

Devlin, Georgescu, Momennejad, Rzepecki, Zuniga, Costello, Leroy, Shaw and Hofmann **ICML 2021 –** https://aka.ms/HNTT

Evaluation



"a human player and an ai collaborating in a fantasy game"

Made by Bing Image Creator

Powered by DALL·E

Microsoft

Bonus: opportunities in Game Creation

"It's Unwieldy and It Takes a Lot of Time." Challenges and Opportunities for Creating Agents in Commercial Games

Mikhail Jacob, Sam Devlin, Katja Hofmann

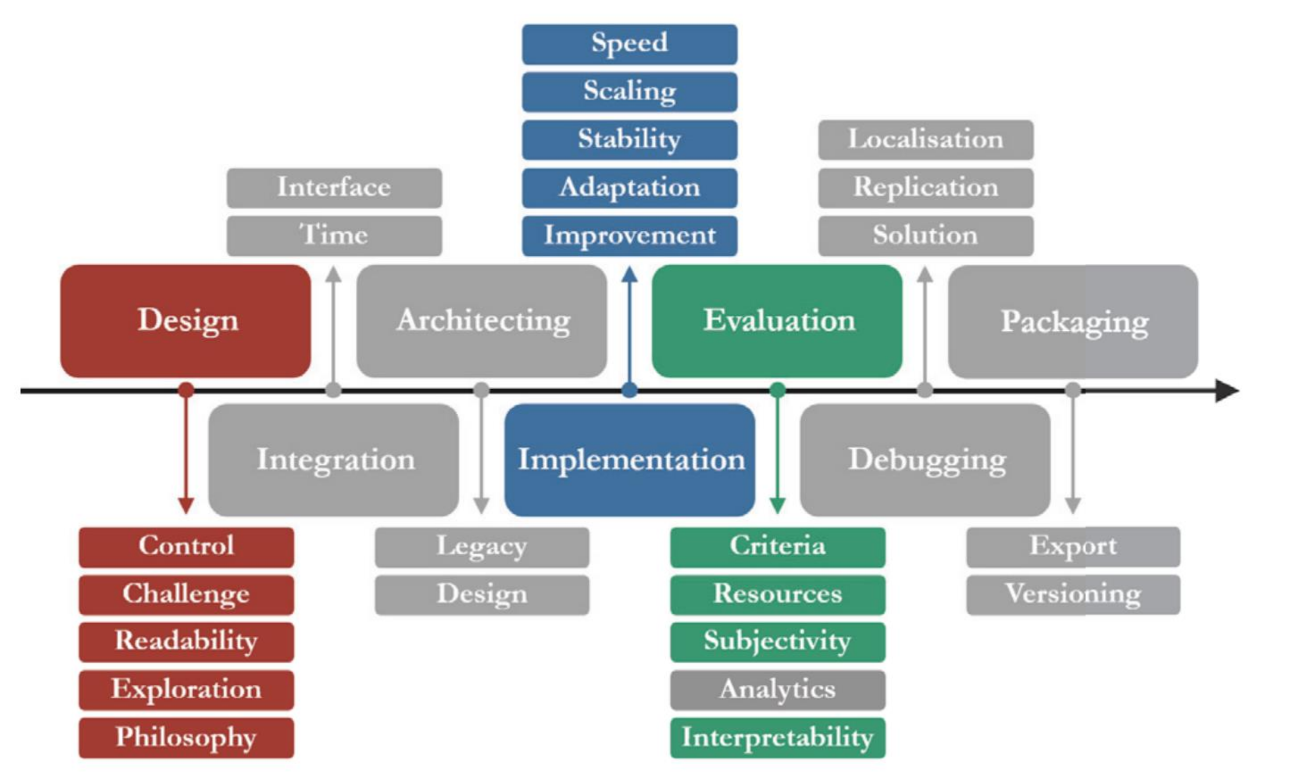
16th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2020)

🎉 Best paper award 🎉

RL as part of the creative workflow

Goal: understand blocks and opportunities for effective for effective use of RL as part of game creation process

Approach: Survey of game industry professionals – surfaces challenges & opportunities in adopting recent AI technologies in game development.



Results: overview of opportunity areas

Microsoft