DeepMind

Probing Emergent Semantics in Predictive Agents via Question Answering

Link to slides with playable videos: bit.ly/3iKYJd3


Abhishek Das*[1]


Fede Carnevale*


Hamza Merzic


Laura Rimell


Rosalia Schneider


Josh Abramson


Alden Hung


Arjun Ahuja


Stephen Clark


Greg Wayne


Felix Hill

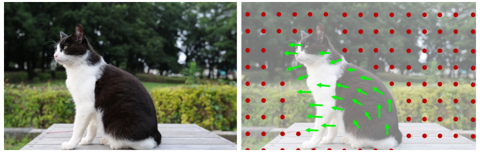
* Denotes equal contribution.
[1] Now at Facebook AI Research. Work done during an internship at DeepMind.

Self-supervised representation learning

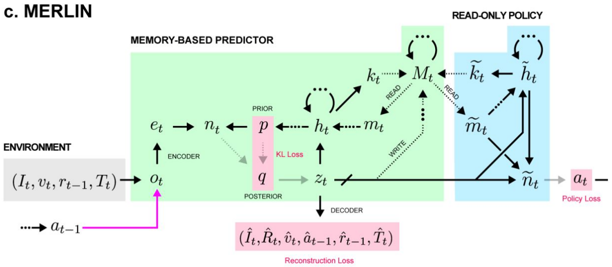
Language

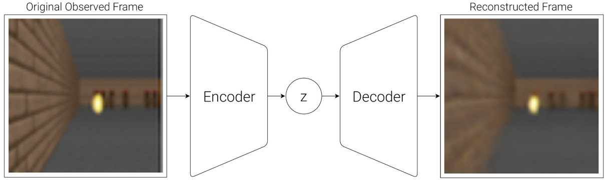
http://jalammar.github.io/illustrated-bert

Vision



Reinforcement Learning





Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv:1810.04805 (2018).
Pathak, Deepak, et al. "Context encoders: Feature learning by inpainting." CVPR 2016.
Pathak, Deepak, et al. "Learning features by watching objects move." CVPR 2017.
Wayne, Greg, et al. "Unsupervised predictive memory in a goal-directed agent." arXiv:1803.10760 (2018).
Ha, David, and Jürgen Schmidhuber. "World models." arXiv:1803.10122 (2018).

How much objective knowledge about the external world can be learned through egocentric prediction?

Question-answering (in English) as an evaluation tool

for investigating how much environment knowledge is encoded in an agent's internal representation

- Intuitive: simply ask an agent what it knows about its world and get an answer back
- Open-ended: pose arbitrarily complex questions to an agent

Environment

Environment

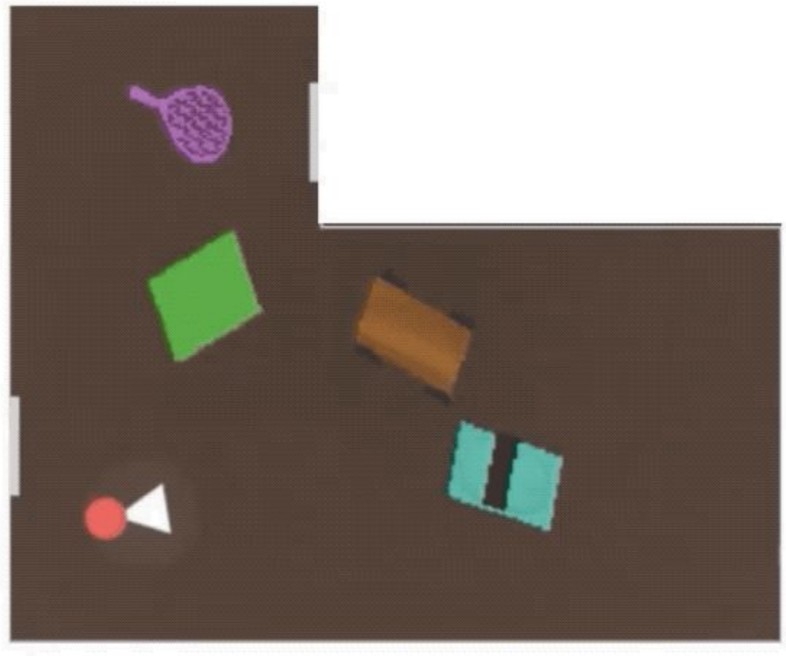
Environment

- Unity-based; runs at 30 fps
- 96 x 72 RGB first-person view
- 50 objects types
  10 colors
  3 sizes

Agent

- First person view
- 8-D action space:
  Move-{forward, backward, left, right}
  Look-{up, down, left, right}

Training task: Exploration

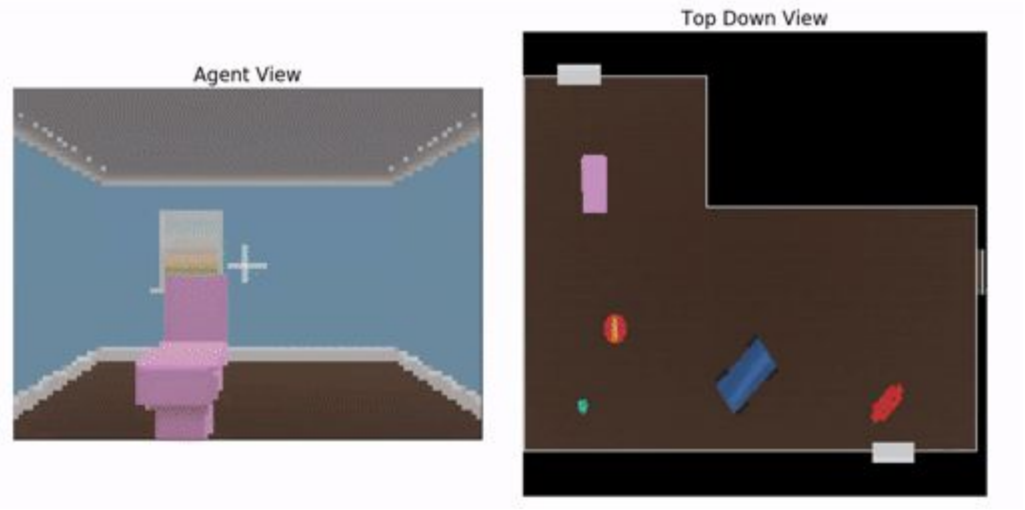


Top-down view shown for illustration purposes. The agent only has access to first-person observations.

+1 reward for unvisited object
0 reward for visited object

rewards refresh once all visited

Training task: Exploration

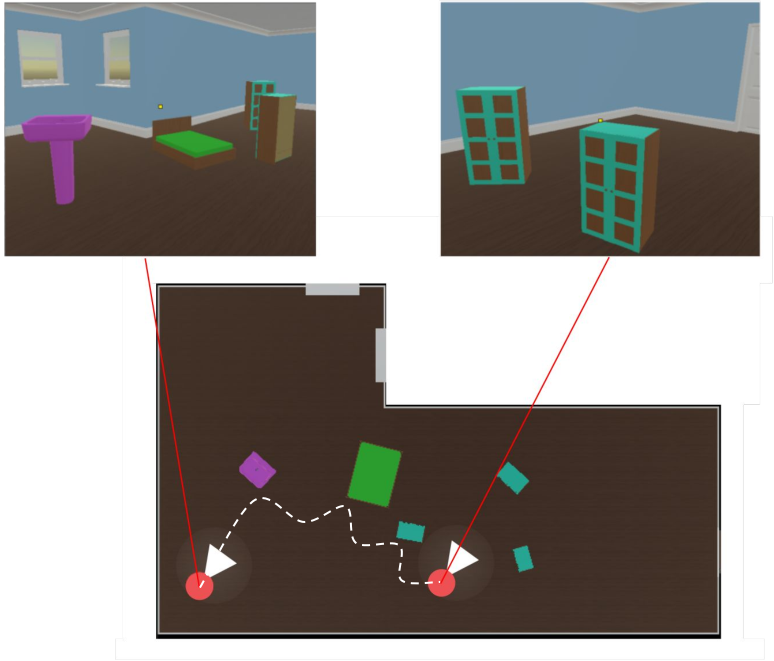


+1 reward for unvisited object
0 reward for visited object

rewards refresh once all visited

Top-down view shown for illustration purposes. The agent only has access to first-person observations.

Evaluation probe: Question-answering



Top-down view shown for illustration purposes. The agent only has access to first-person observations.

What is the color of the bed?

How many wardrobes are there?

What is the object near the bed?

Is there a basketball in the room?

...

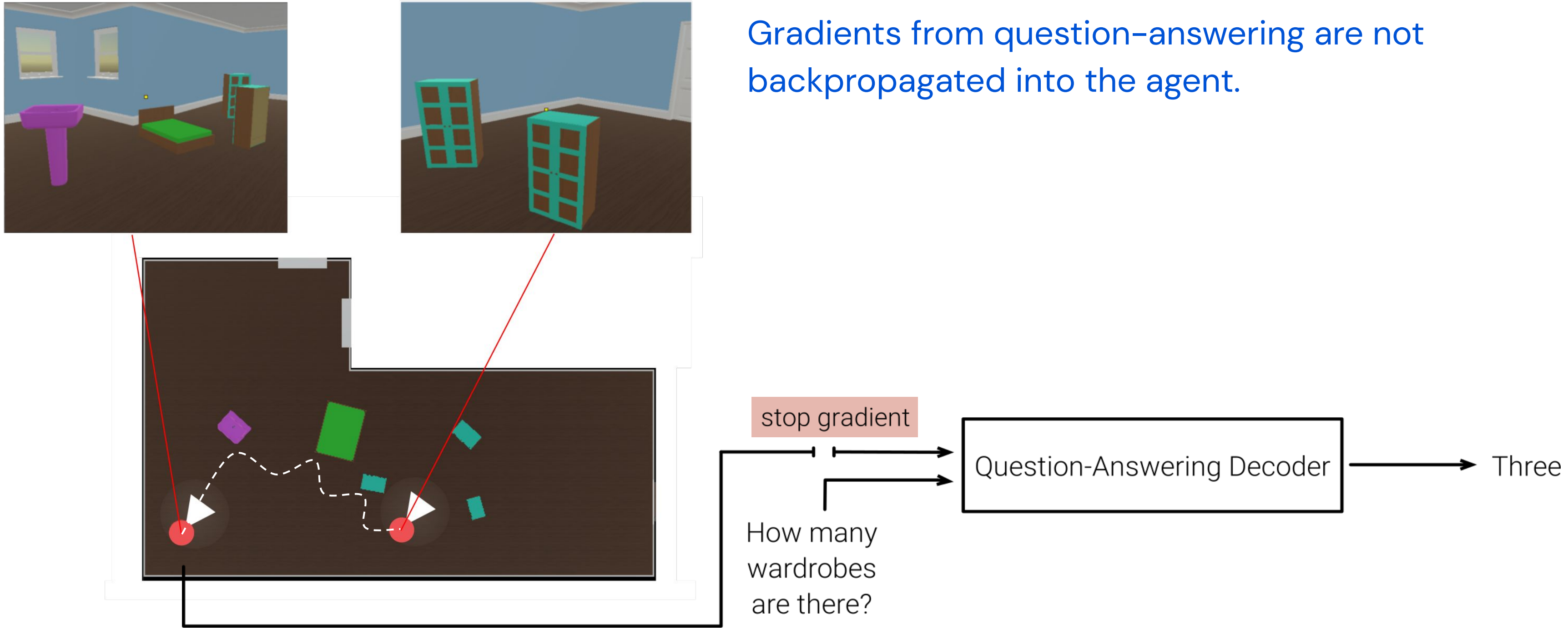
Evaluation probe: Question-answering

Question type	Template	# QA pairs
Attribute	What is the color of the <shape>?	500
	What shape is the <color> object?	500
Count	How many <shape> are there?	200
	How many <color> objects are there?	40
Exist	Is there a <shape>?	100
Compare + Count	Are there the same number of <color1> objects as <color2> objects?	180
	Are there the same number of <shape1> as <shape2>?	4900
Relation + Attribute	What is the color of the <shape1> near the <shape2>?	24500
	What is the <color> object near the <shape>?	25000

Questions are programmatically generated in a manner similar to CLEVR (Johnson et al., 2017)

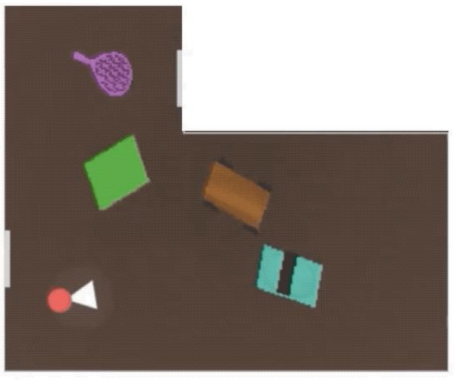
Evaluation probe: Question-answering

Gradients from question-answering are not backpropagated into the agent.



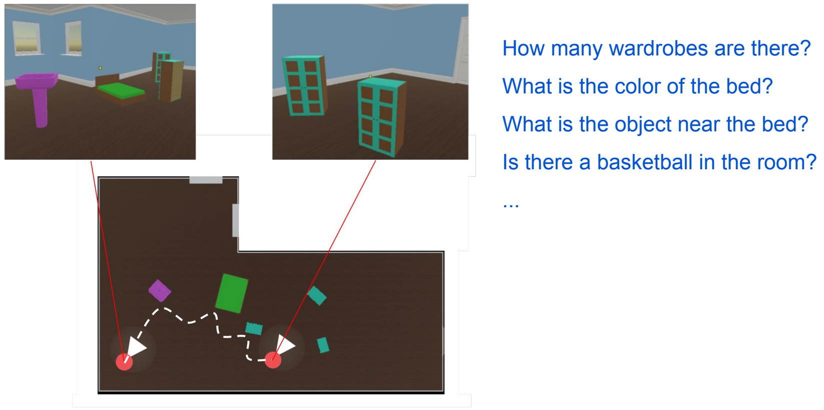
Top-down view shown for illustration purposes. The agent only has access to first-person observations.

Setup



(i)

During training, the agent explores and learns to build representations from egocentric observations



(ii)

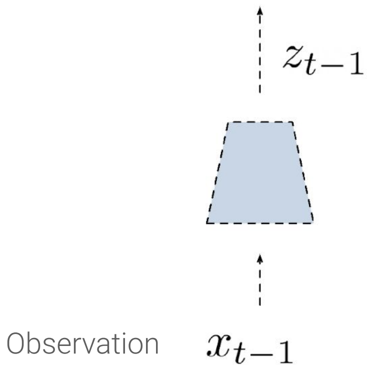
During evaluation, we probe the agent's internal representations on a question-answering task

Top-down view shown for illustration purposes. The agent only has access to first-person observations.

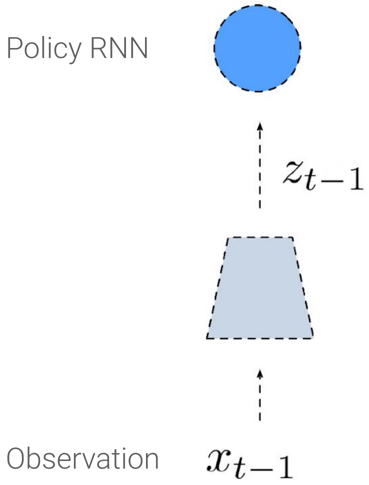
Agent architecture

Observation x_{t-1}

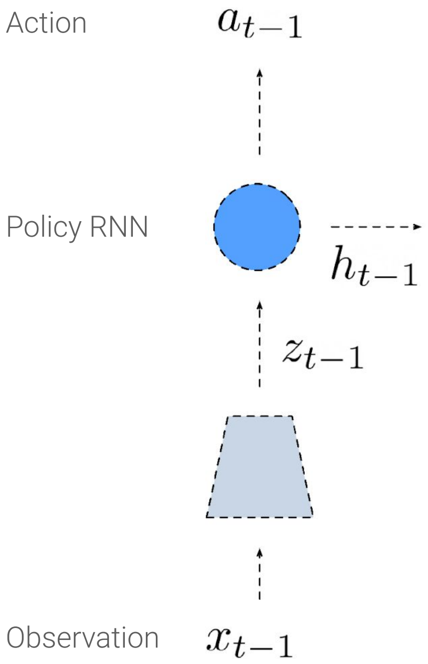
Agent architecture



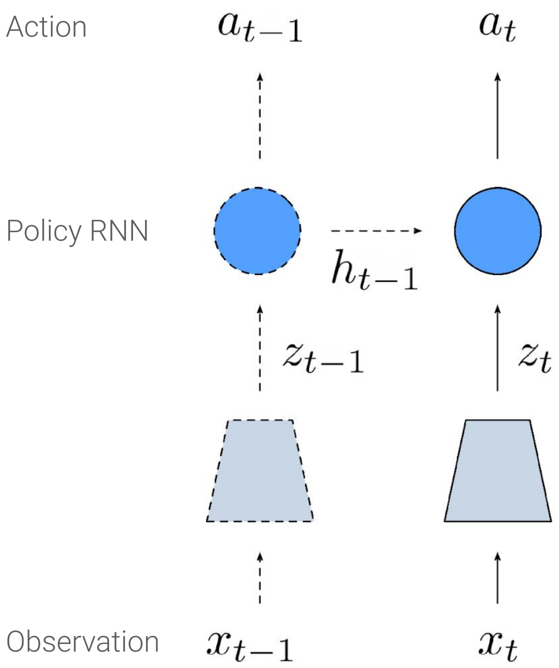
Agent architecture



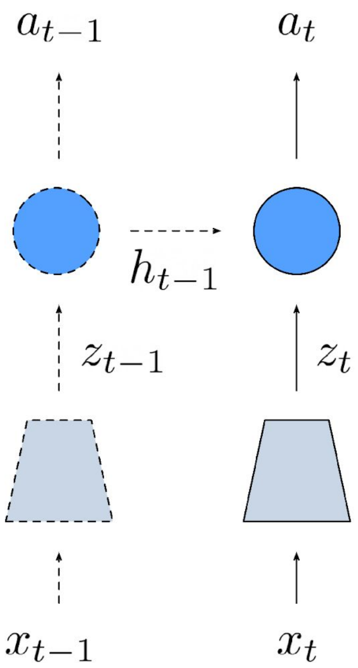
Agent architecture



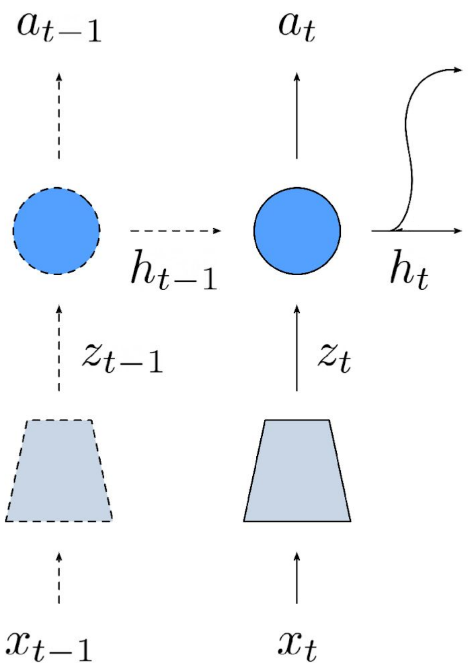
Agent architecture



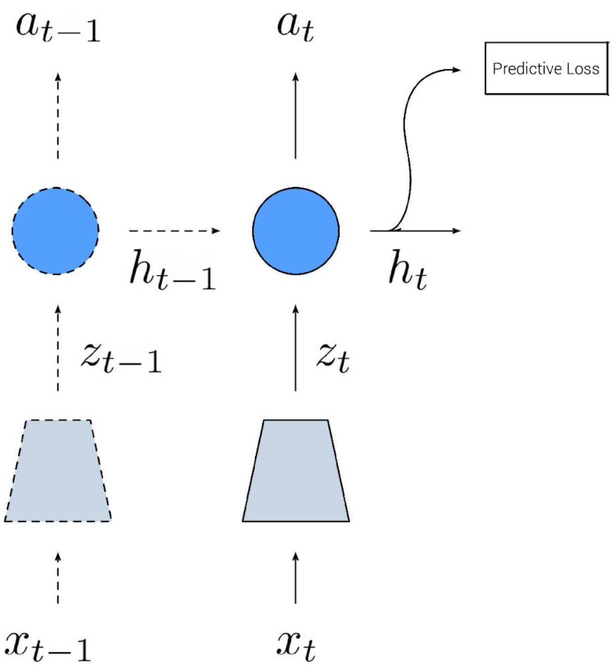
Agent architecture



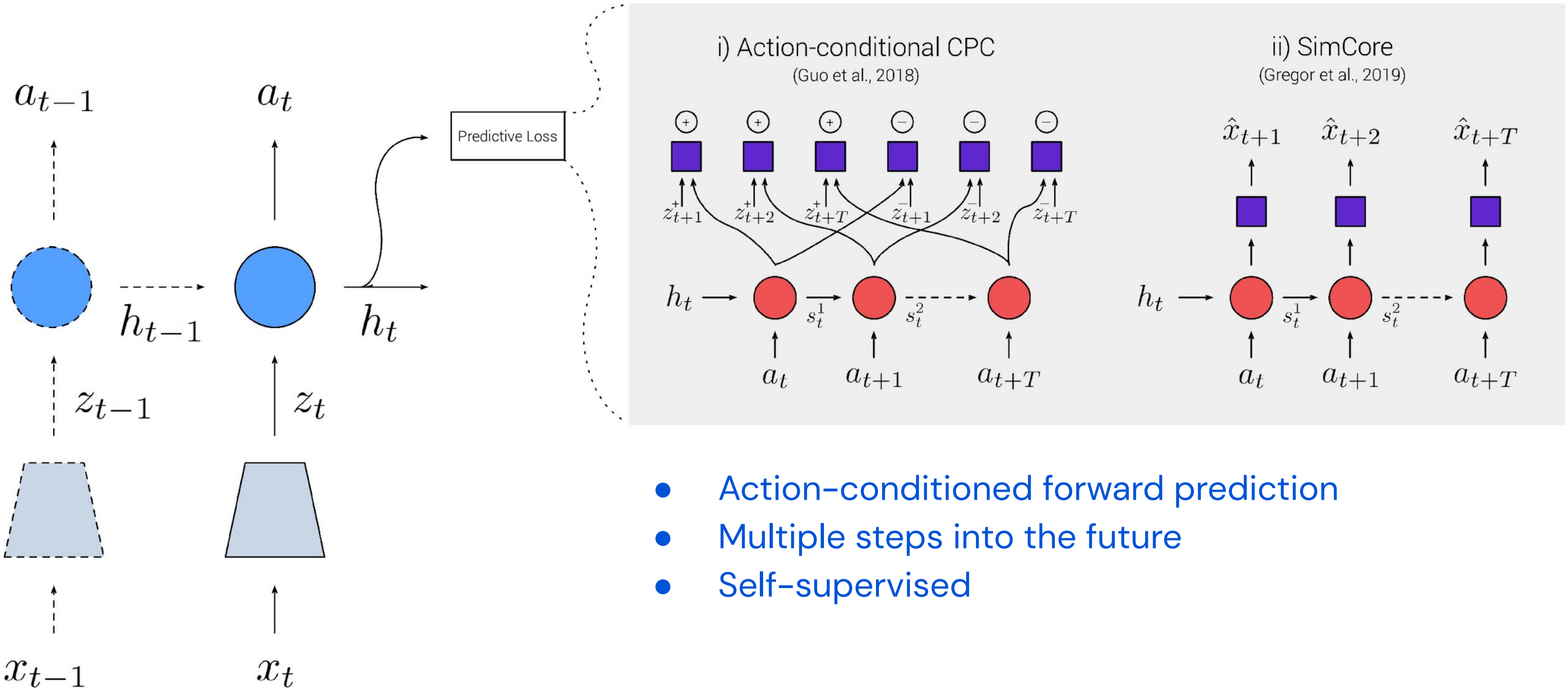
Agent architecture



Agent architecture

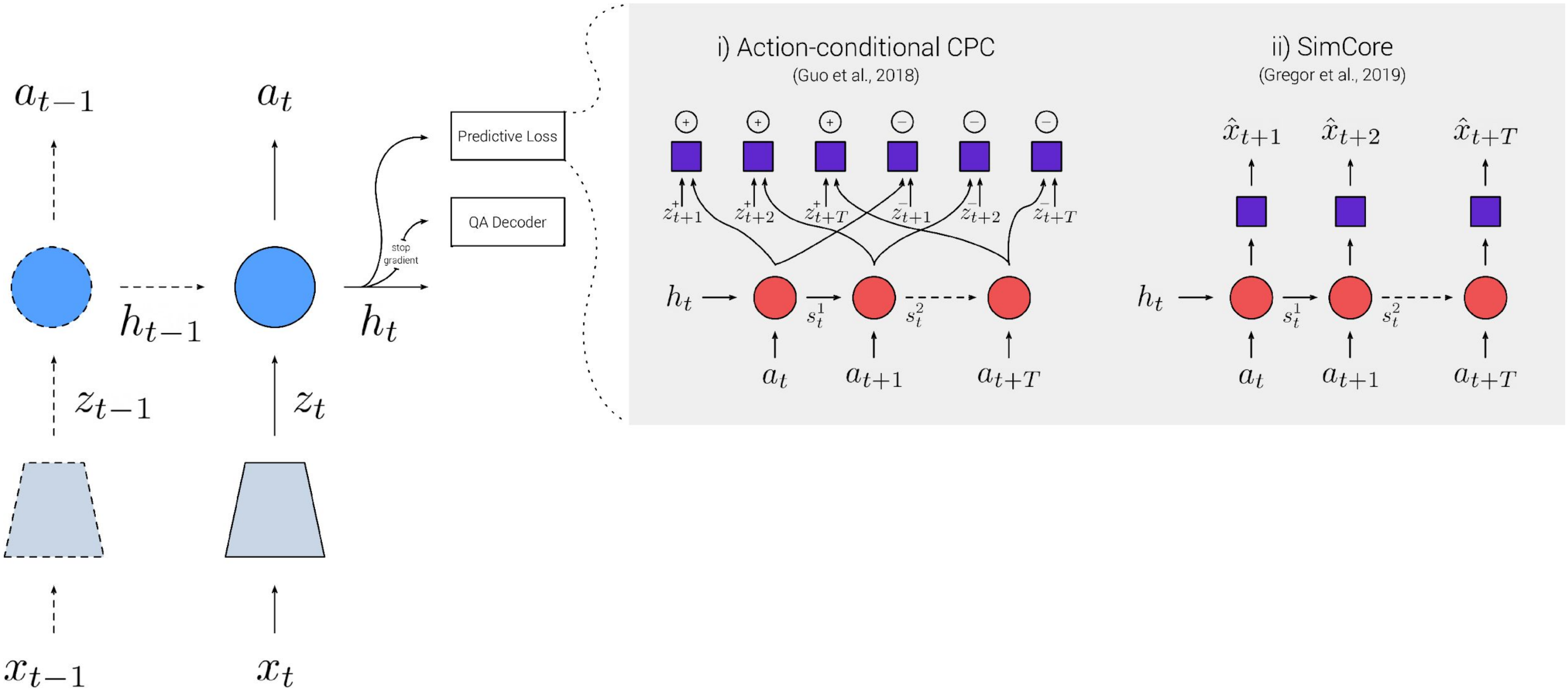


Agent architecture

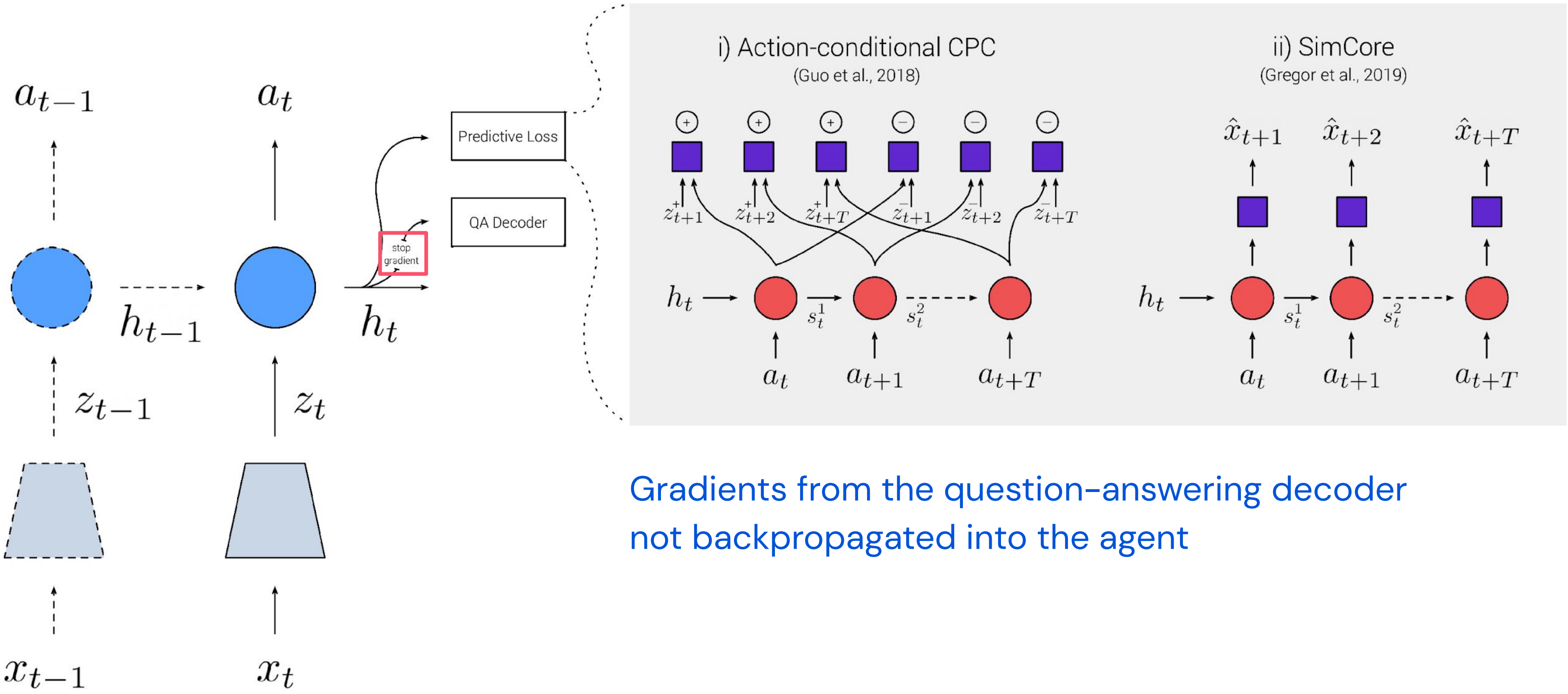


- Action-conditioned forward prediction
- Multiple steps into the future
- Self-supervised

Agent architecture

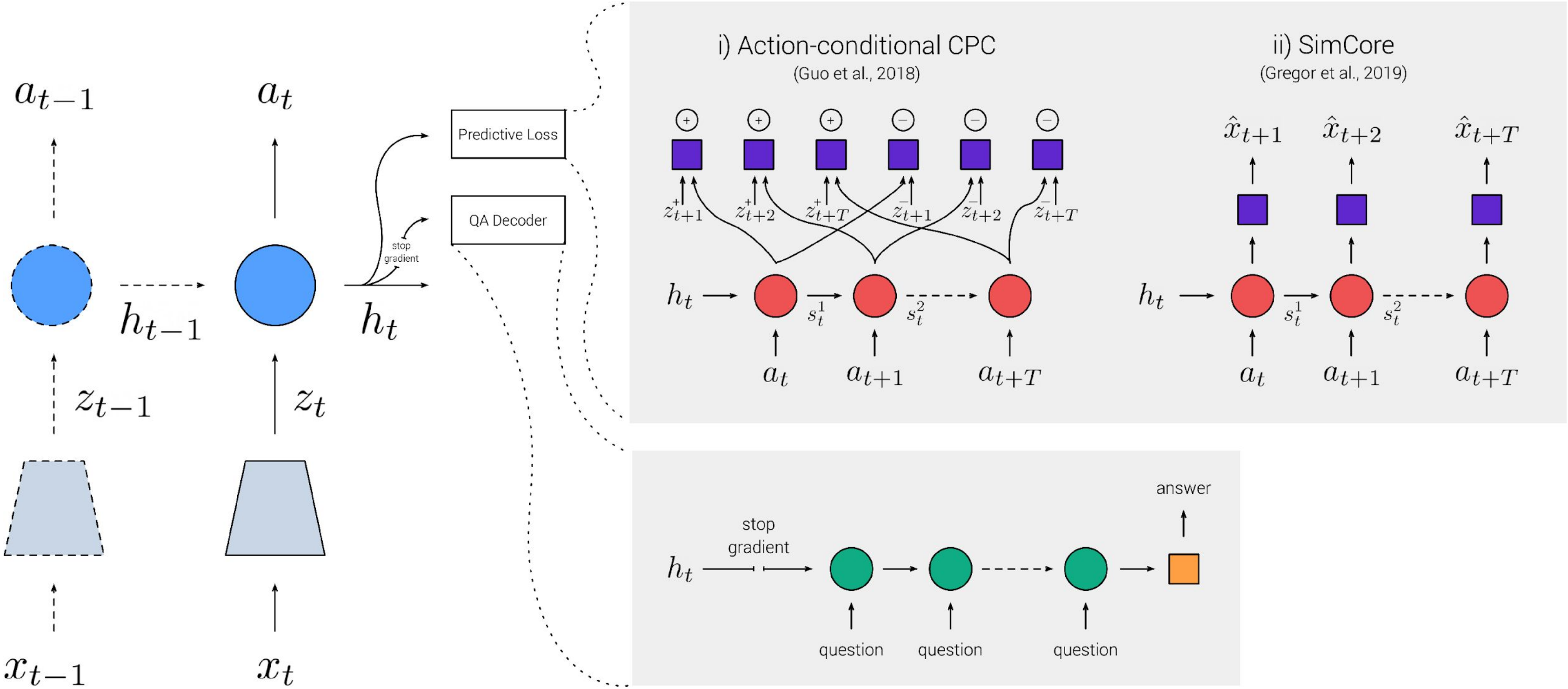


Agent architecture



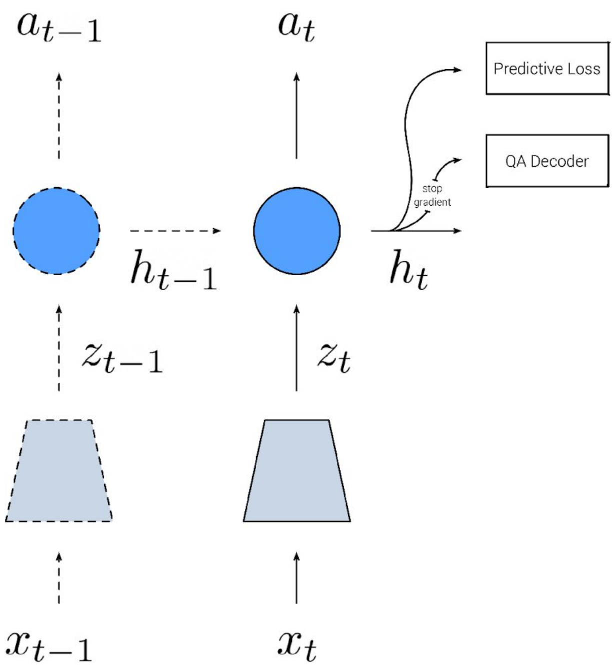
Gradients from the question-answering decoder not backpropagated into the agent

Agent architecture

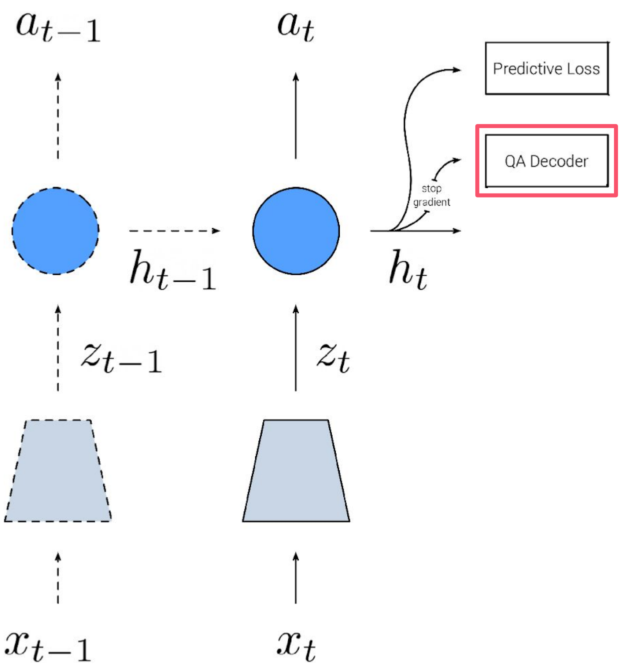


Baselines and Oracle

Baselines



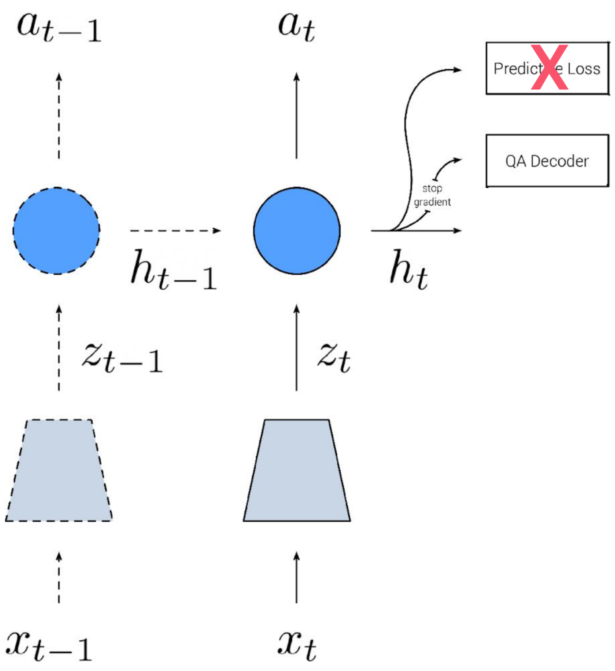
Baselines and Oracle



Baselines

- Question-only: no vision

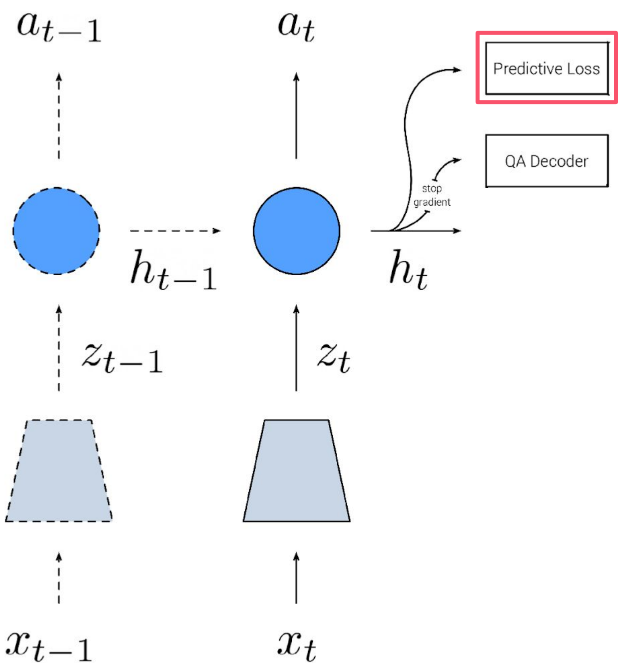
Baselines and Oracle



Baselines

- Question-only: no vision
- LSTM: no auxiliary predictive loss

Baselines and Oracle



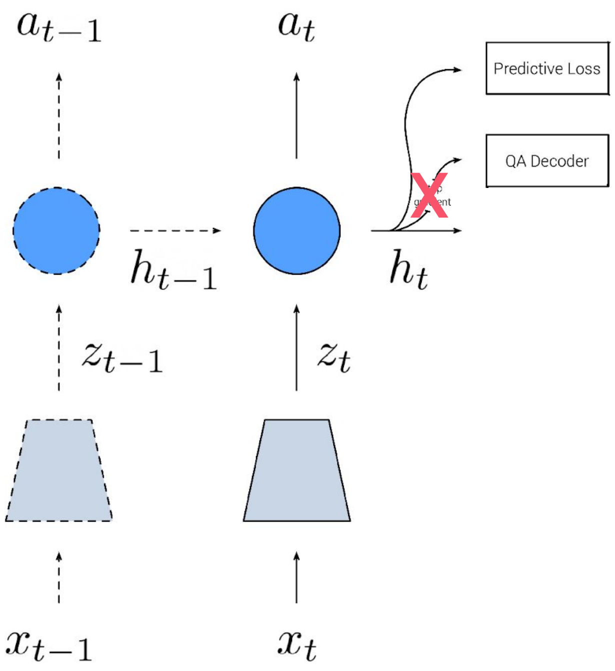
Baselines

- Question-only: no vision
- LSTM: no auxiliary predictive loss

Predictive losses

- CPC|A (Guo et al., 2018)
- SimCore (Gregor et al., 2019)

Baselines and Oracle



Baselines

- Question-only: no vision
- LSTM: no auxiliary predictive loss

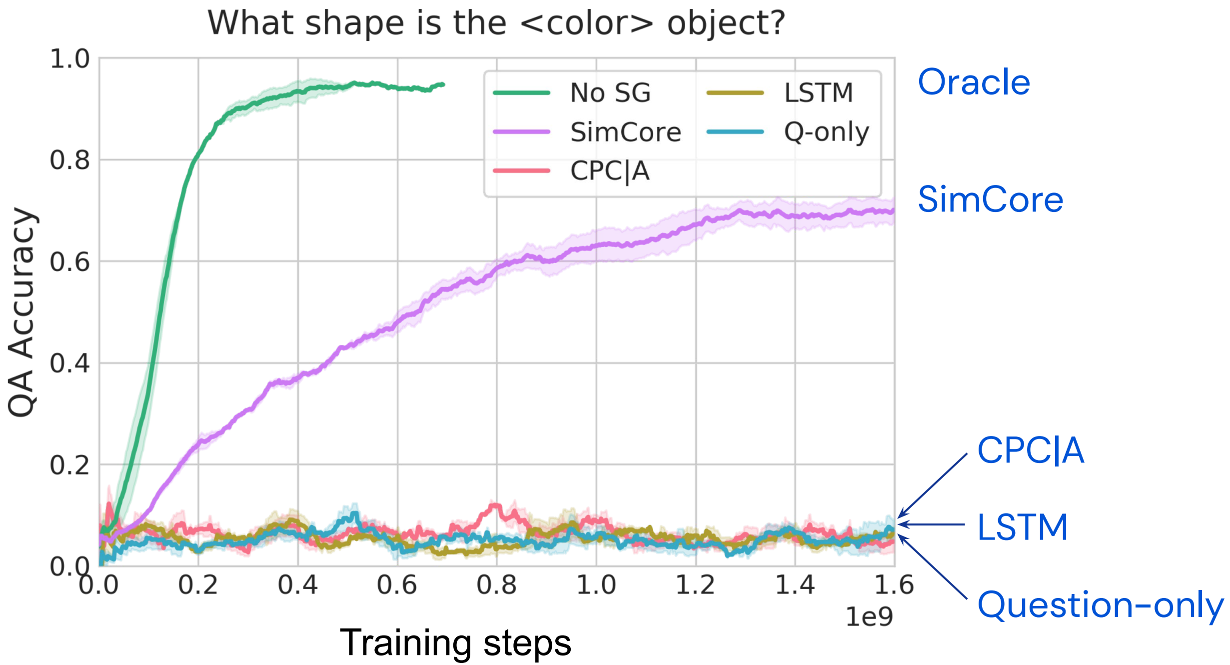
Predictive losses

- CPC|A (Guo et al., 2018)
- SimCore (Gregor et al., 2019)

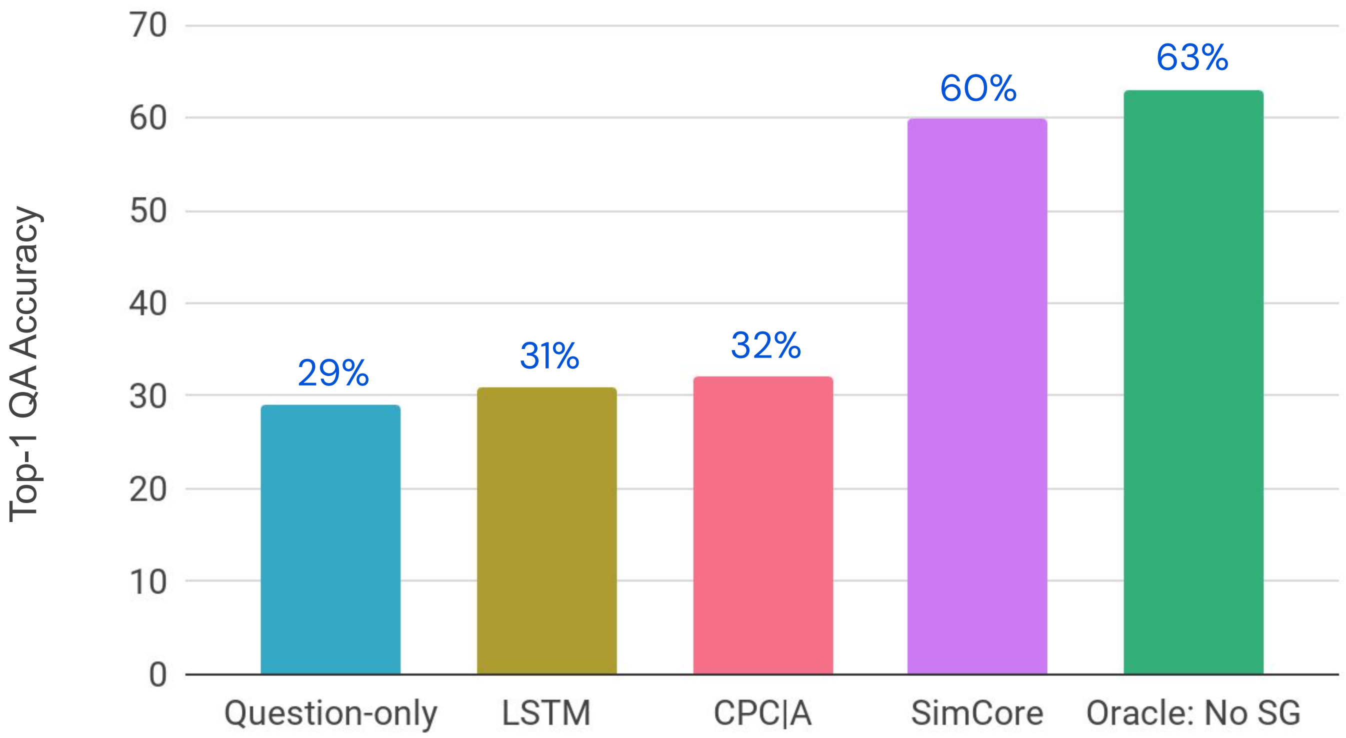
Oracle

- No SG: QA decoder without stop gradient
  similar to Embodied / Interactive Question Answering
  (Das et al., 2018, Gordon et al., 2018)

Results: shape questions



Results: overall



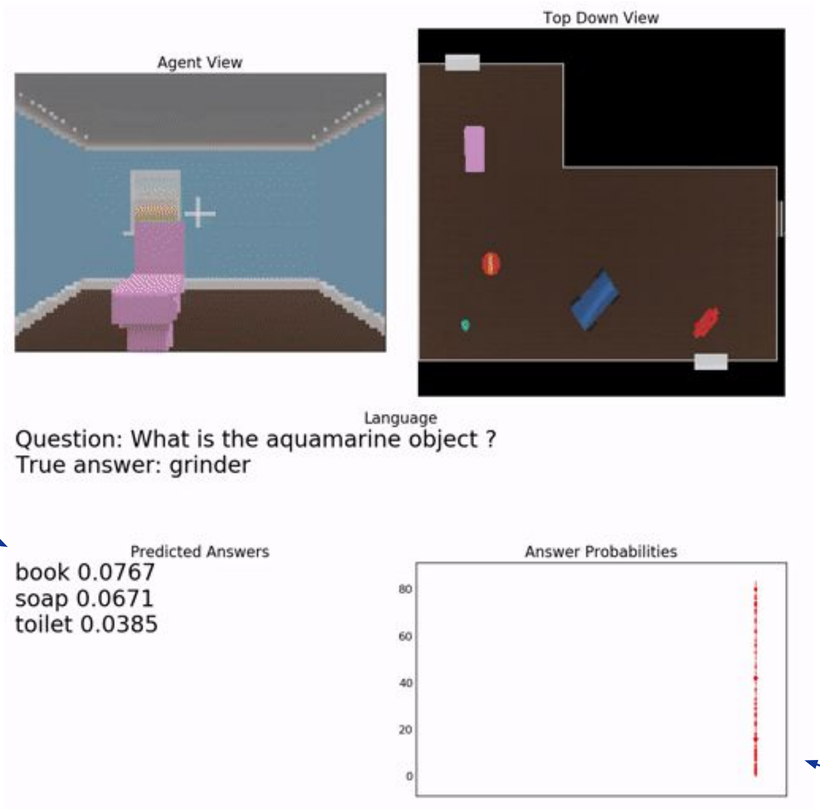
Results

	Overall	shape	color	exist	count_shape	count_color	compare_count_color	compare_count_shape	near_shape	near_color
Baseline: Question-only	0.29	0.04	0.1	0.63	0.24	0.24	0.49	0.70	0.04	0.09
LSTM	0.31	0.04	0.1	0.54	0.34	0.38	0.53	0.70	0.04	0.09
CPC|A	0.32	0.06	0.08	0.64	**0.39**	0.39	0.50	0.70	0.06	0.10
SimCore	**0.60**	**0.72**	**0.81**	**0.72**	**0.39**	**0.57**	**0.56**	**0.73**	**0.30**	**0.59**
Oracle: No SG	0.63	0.96	0.81	0.60	0.45	0.57	0.51	0.76	0.41	0.72

Table 2: Top-1 accuracy on question-answering tasks.

Q: What is the aquamarine object?
A: Grinder

Top-3 answer predictions

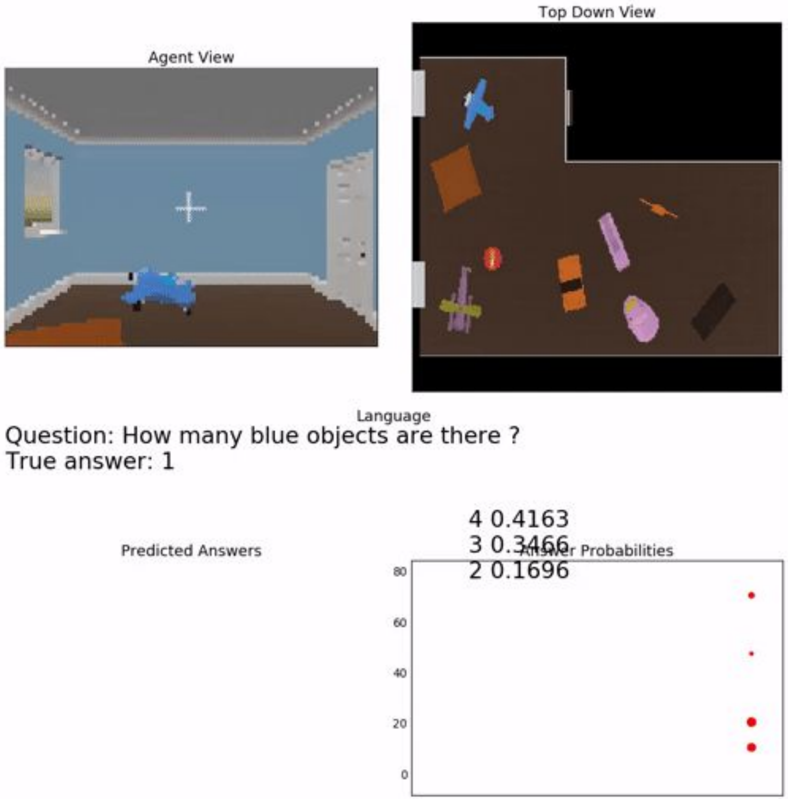


Answer probabilities

P("Grinder")

Q: How many blue objects are there?
A: One

Top-3 answer predictions

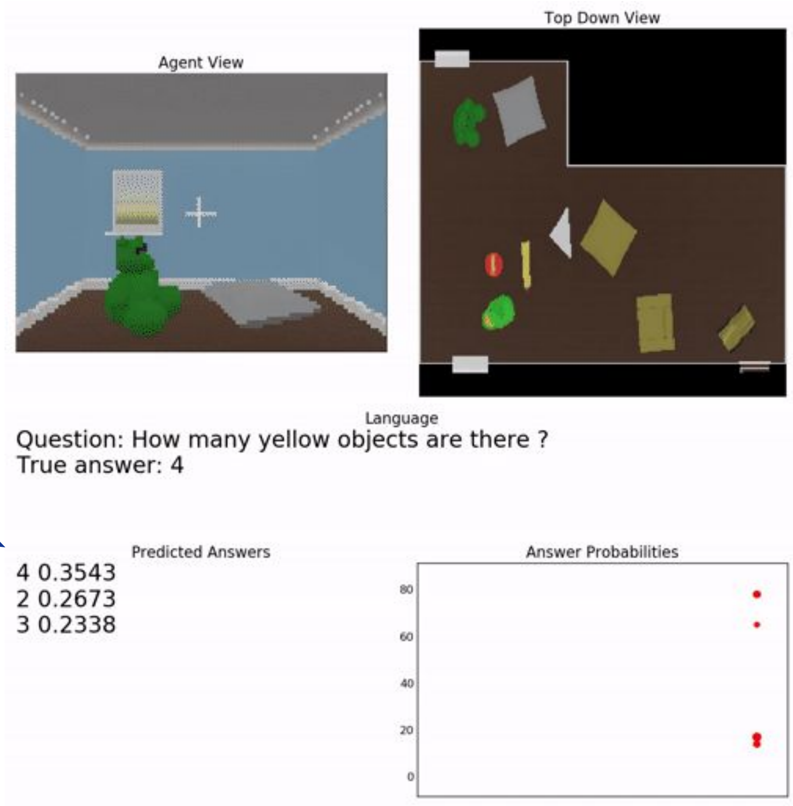


P("One")

P("Three")

Q: How many yellow objects are there?
A: Four

Top-3 answer predictions

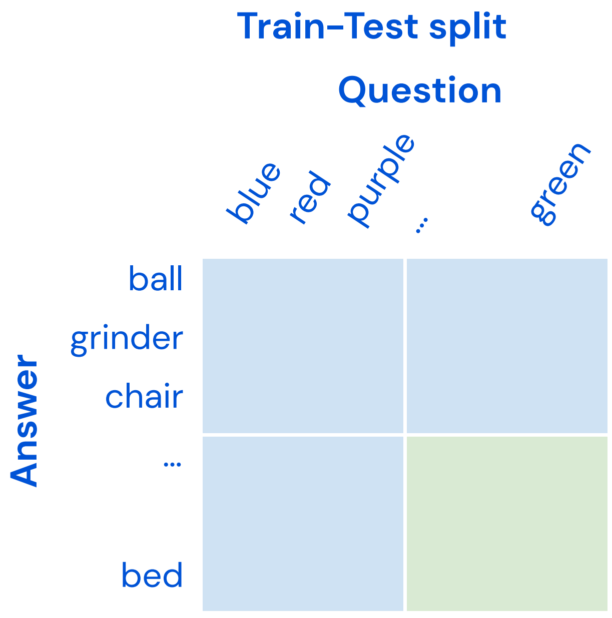


P("two")

P("four")

P("Three")

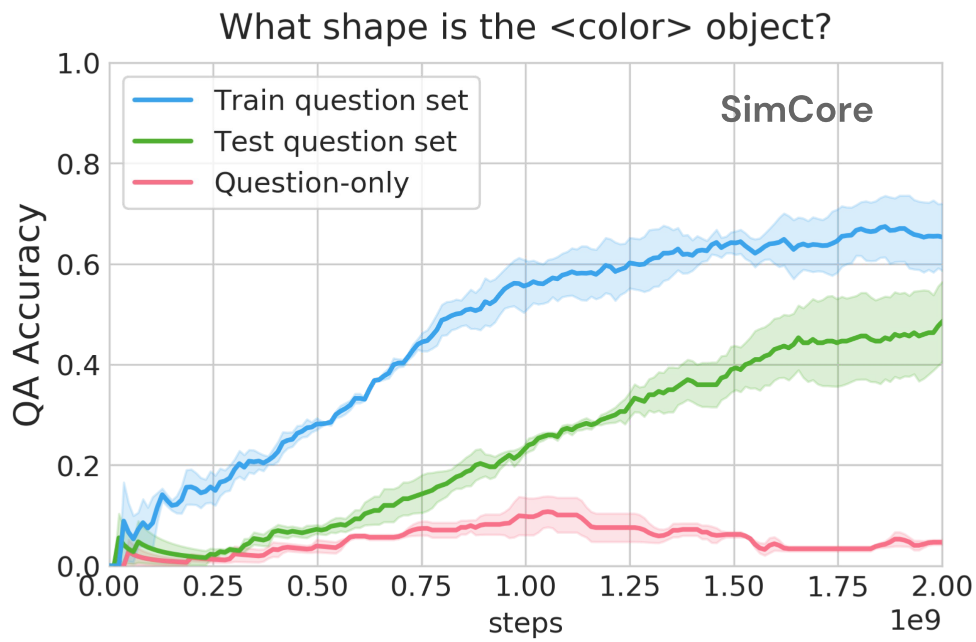
Compositional generalization



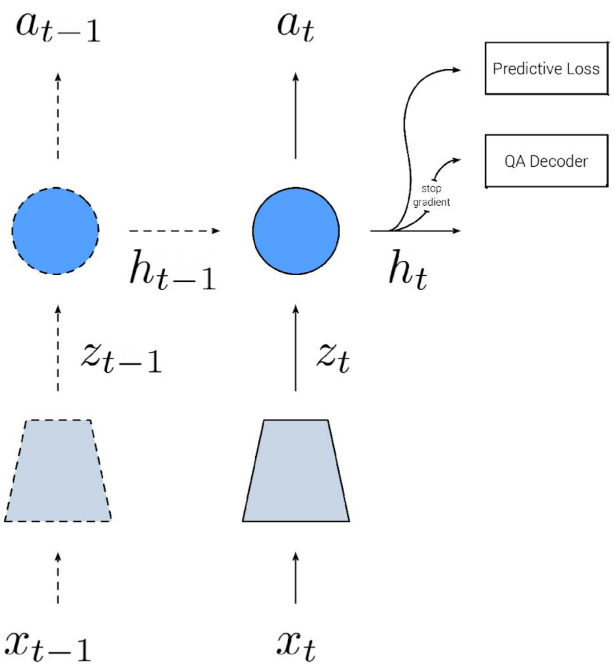
Seen: *What shape is the blue object? Bed*

Seen: *What shape is the green object? Ball*

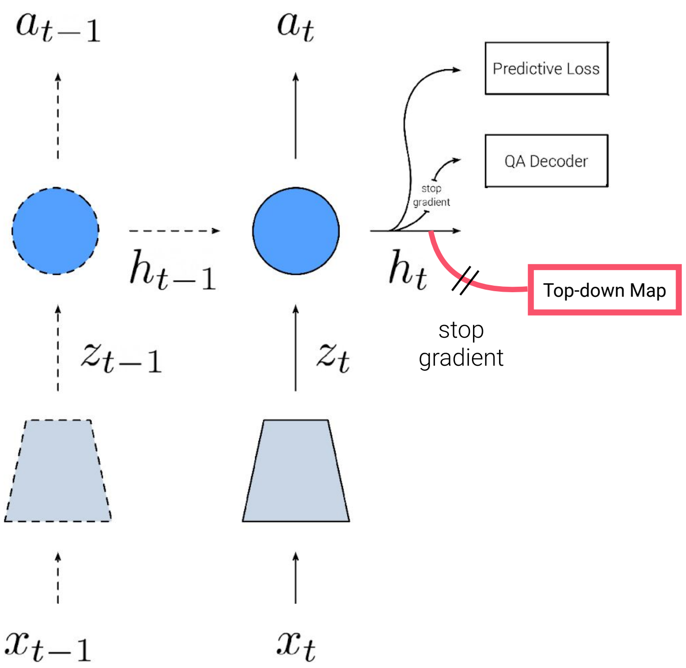
Unseen: *What shape is the green object? Bed*



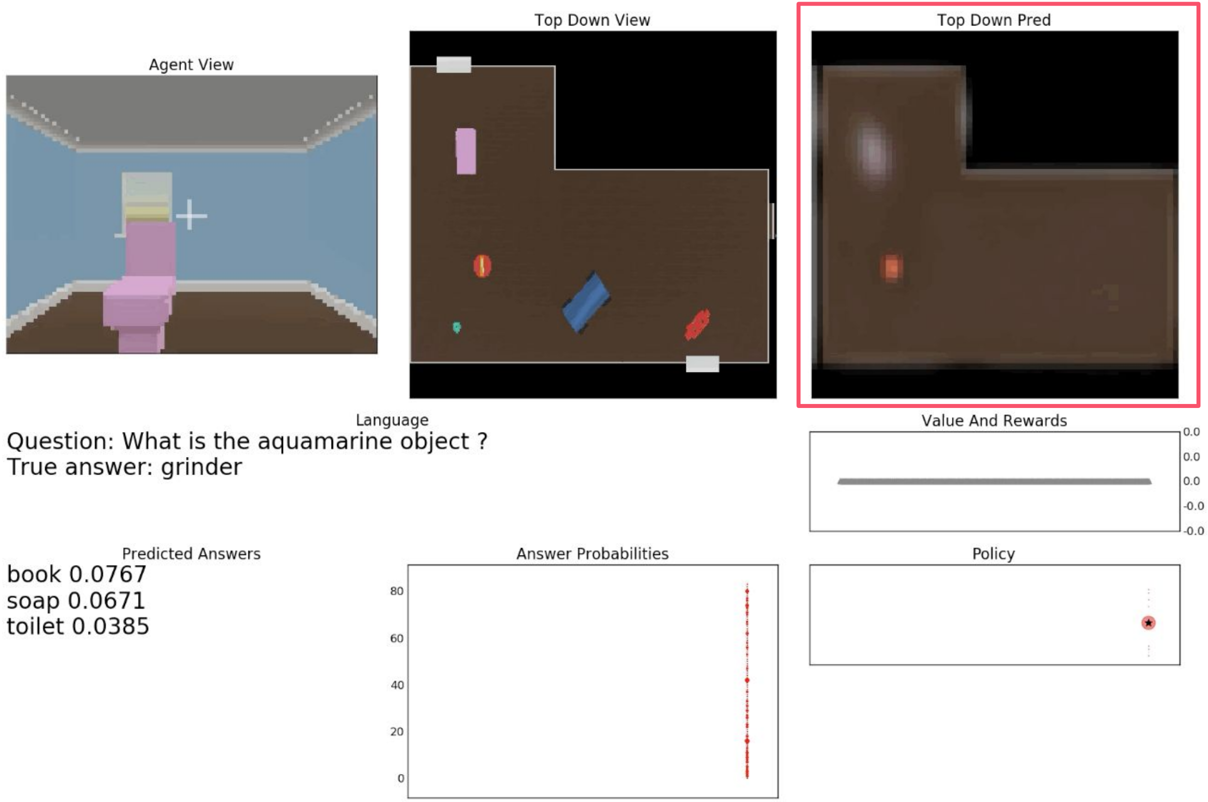
Top-down map prediction



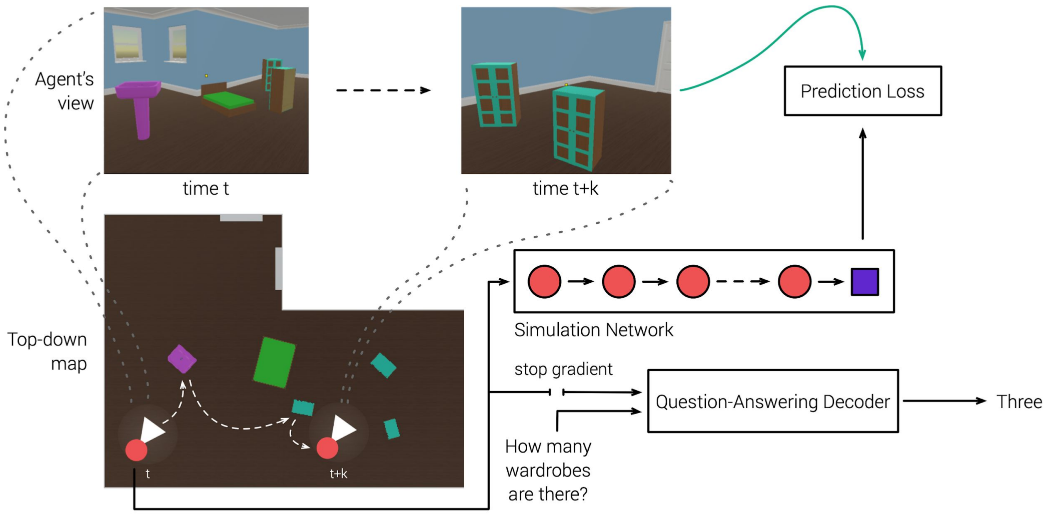
Top-down map prediction



Top-down map prediction



Conclusions



- Question-answering to probe internal representations, enabling evaluation of agents using natural linguistic interactions.
- Self-supervised predictive agents, such as SimCore, capture decodable knowledge about the environment, while non-predictive agents and CPC|A don't.
- Generalization of the decoder suggests some degree of compositionality in internal representations.
- arxiv.org/abs/2006.01016