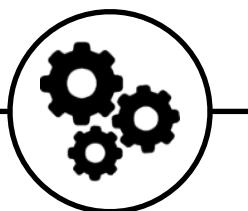


Program

```
DEF run() m(
  WHILE c( markerPresent c) w(
    WHILE c( markerPresent c) w(
      pickMarker
      move w)
    turnRight
    move
    turnLeft
    WHILE c( markerPresent c) w(
      pickMarker
      move w)
    turnLeft
    move
    turnRight w) m)
```

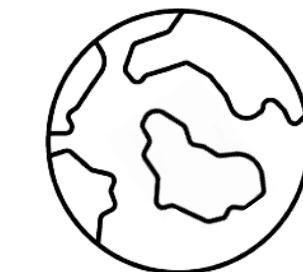
Execute



Learning to Synthesize Programs as Interpretable and Generalizable Reinforcement Learning Policies



Environment



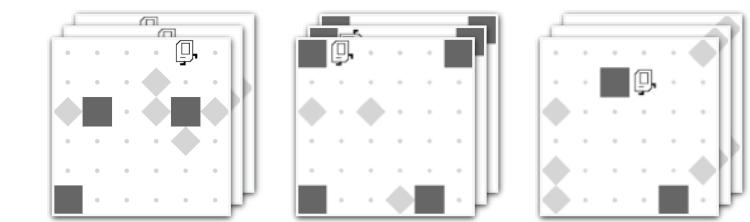
Shao-Hua Sun (孫紹華)

Assistant Professor

Dept. of Electrical Engineering (EE)

National Taiwan University

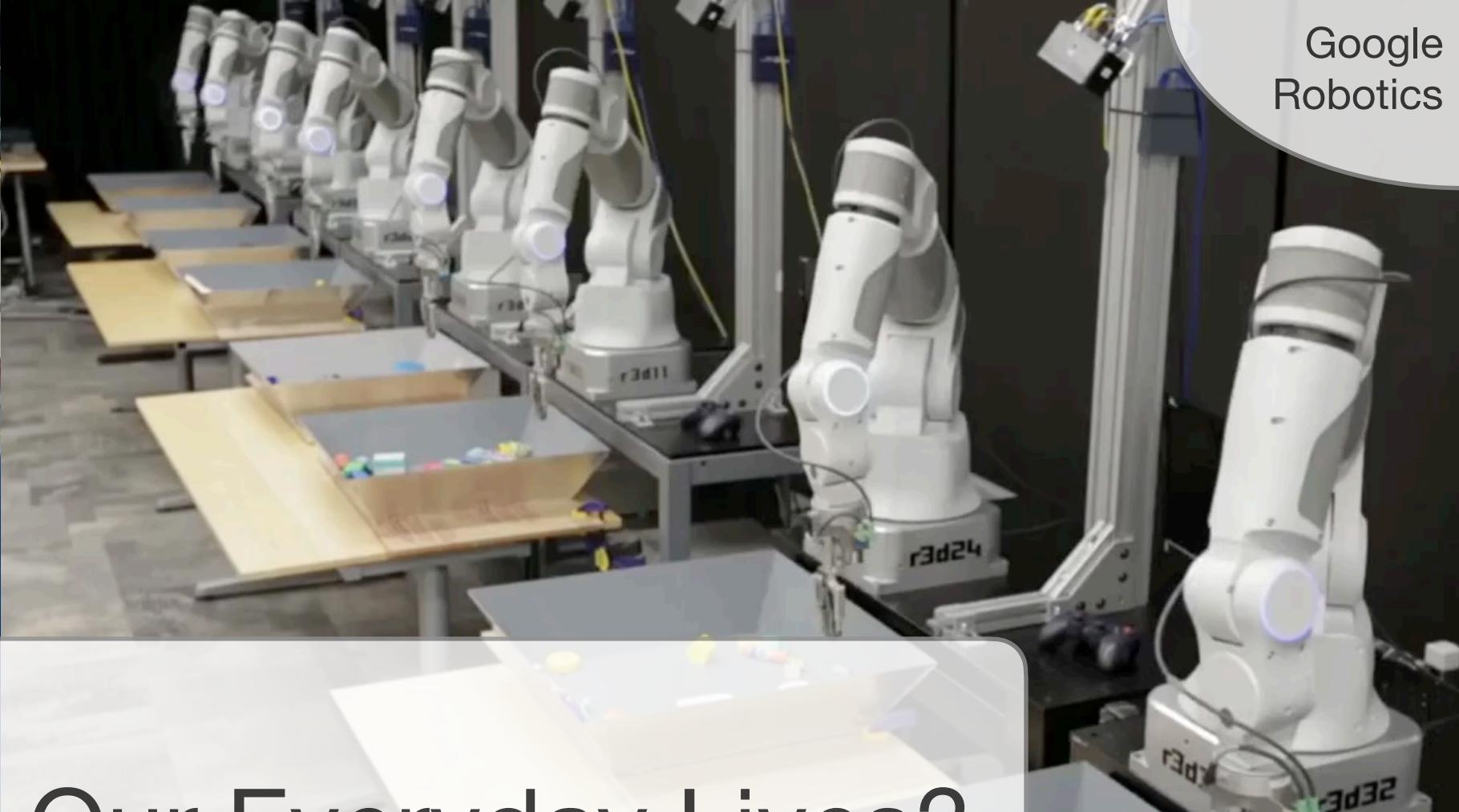
Demonstrations



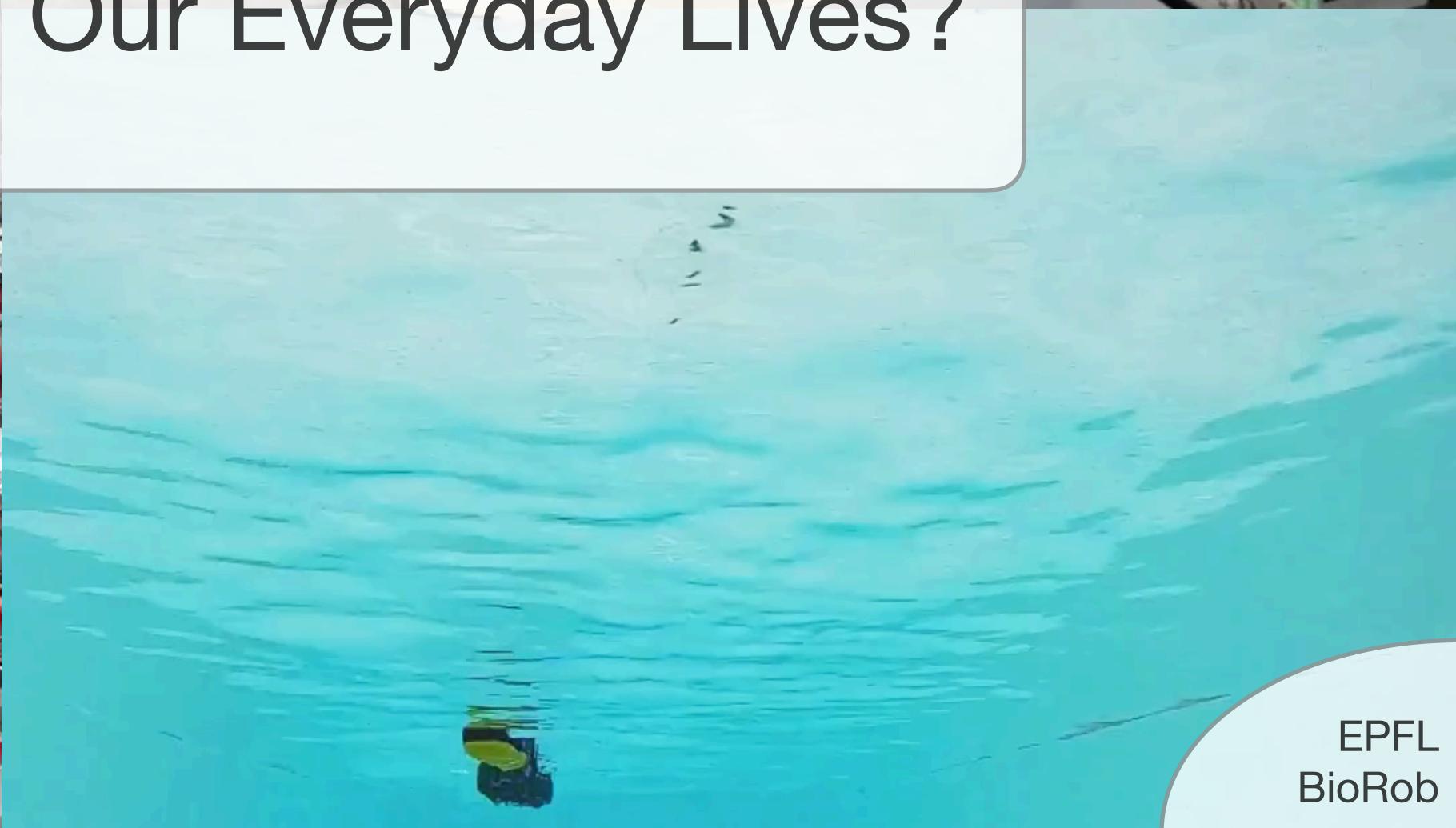
Machine Learning Summer Schools 2024 @ OIST



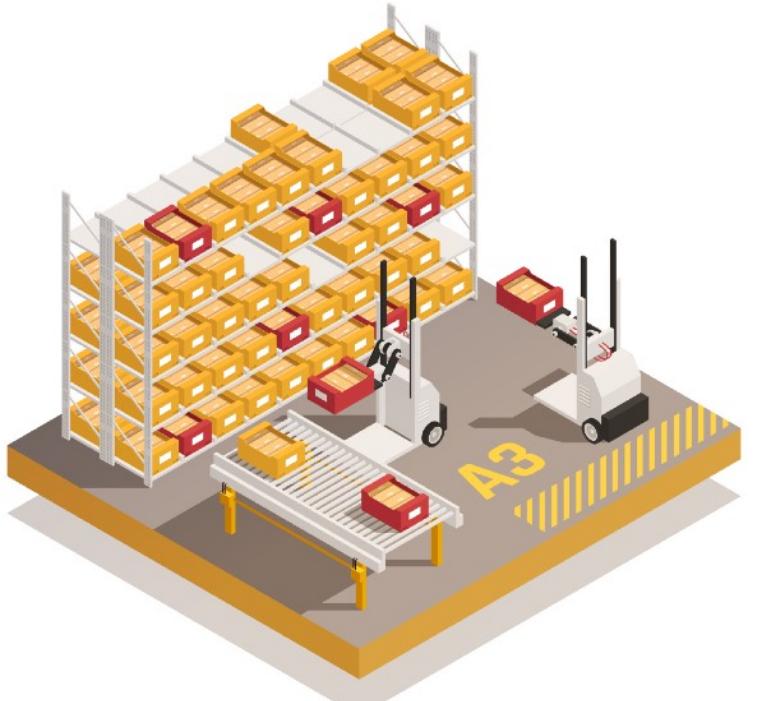
Reward



Why Aren't Robots in Our Everyday Lives?



Environment



Structured

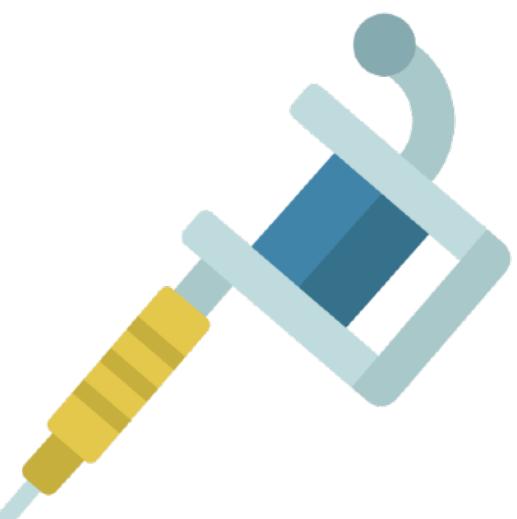


Unstructured

Object



Known

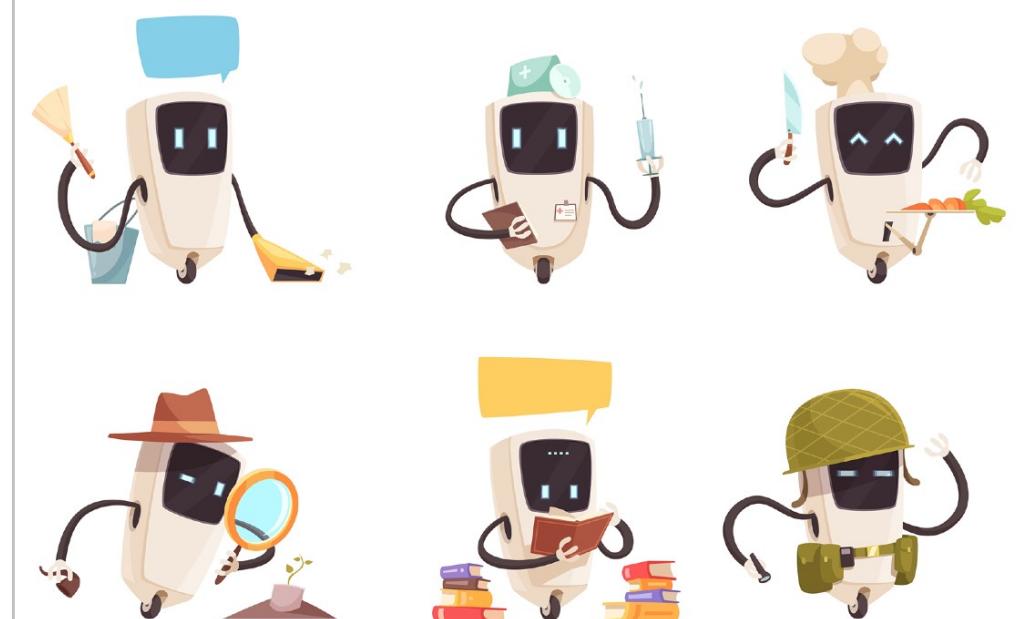


Unseen

Task



Pre-defined / Pre-programmed



Diverse and Novel

A Venn diagram consisting of two overlapping circles. The left circle is colored light orange and contains the text "Machine Learning". The right circle is colored light blue and contains the text "Robotics". The overlapping region between the two circles contains the text "Robot Learning".

Machine
Learning

Robot
Learning

Robotics

Supervised Learning

Image Classification

Model	image size	# parameters	Mult-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224x224	11.2M	1.94B	74.8	92.2
NASNet-A (5 @ 1538)	299x299	10.9M	2.35B	78.6	94.2
Inception V3 [60]	299x299	23.8M	5.72B	78.8	94.4
Xception [9]	299x299	22.8M	8.38B	79.0	94.5
Inception ResNet V2 [58]	299x299	55.8M	13.2B	80.1	95.1
NASNet-A (7 @ 1920)	299x299	22.6M	4.93B	80.8	95.3
ResNeXt-101 (64 x 4d) [68]	320x320	83.6M	31.5B	80.9	95.6
PolyNet [69]	331x331	92M	34.7B	81.3	95.8
DPN-131 [8]	320x320	79.5M	32.0B	81.5	95.8
SENet [25]	320x320	145.8M	42.3B	82.7	96.2
NASNet-A (6 @ 4032)	331x331	88.9M	23.8B	82.7	96.2

Table 2. Performance of architecture search and other published state-of-the-art models on ImageNet classification. Mult-Adds indicate the number of composite multiply-accumulate operations for a single image. Note that the composite multiply-accumulate operations are calculated for the image size reported in the table. Model size for [25] calculated from open-source implementation.

Zoph et al. Learning Transferable Architectures for Scalable Image Recognition

Instance Segmentation



Figure 5. More results of Mask R-CNN on COCO test images, using ResNet-101-FPN and running at 5 fps, with 35.7 mask AP (Table 1).

He et al. Mask R-CNN

Visual Question Answering

Method	VQA v2 test-dev			VQA v2 test-val		
	All	Yes/no	Simple	All	Yes/no	Simple
PoE (most common answer in training set) [1]	—	—	—	75.98	61.30	9.06
LSTM Language only (base model) [1]	—	—	—	44.26	47.01	11.55
Deep LSTM (ours, T=1) as reported in [1]	—	—	—	54.22	75.46	15.18
MCI [1] as reported in [1]	—	—	—	62.27	78.82	18.28
UNet-LPPE [1]	—	—	—	63.71	78.12	13.52
UNet-LPPE [1]	—	—	—	67.59	82.50	14.19
UNet-LPPE [1]	—	—	—	69.77	81.89	16.26
UNet-LPPE [1]	—	—	—	68.09	84.50	15.01
HDU-USYD-UNCC (Proposed model)	—	—	—	68.09	84.50	15.01
ResNet features T=7, single network	62.07	79.20	19.46	52.62	62.29	19.92
Image Features from ResNet-101, adaptive K, single network	65.52	81.82	44.31	65.67	80.46	55.26
ResNet features T=7, ensemble	66.30	81.38	43.17	67.00	83.57	57.20
Image Features from ResNet-101, adaptive K, ensemble	68.07	84.48	48.89	70.34	86.68	44.12

Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA v2 Leaderboard [1].

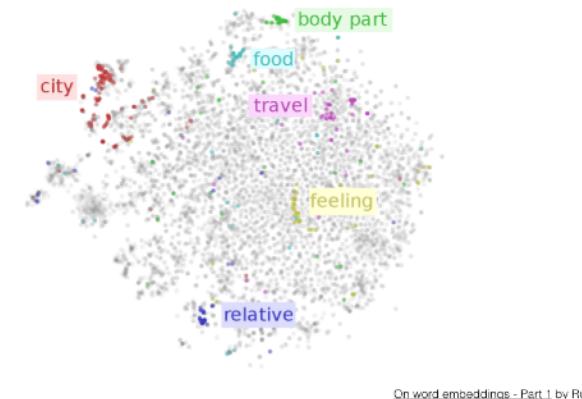
Teney et al. Tips and Tricks for Visual Question Answering: Learnings from the 2017 Challenge

Machine Translation

Source	"The reason Boeing is doing this is to cram more seats in to make their plane more competitive with our products," said Kevin Keniston, head of passenger comfort at Europe's Airbus.
PBMT	"La raison pour laquelle Boeing est en train de faire, c'est de concentrer davantage de sièges pour prendre leur avion plus compétitive avec nos produits", a déclaré Kevin M. 3.0 Keniston, chef du confort des passagers de l'Airbus de l'Europe.
GNMT	"La raison pour laquelle Boeing fait cela est de créer plus de sièges pour rendre son avion plus compétitif avec nos produits", a déclaré Kevin Keniston, chef du confort des passagers chez Airbus.
Human	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager 6.0 chez l'avionneur européen Airbus.
Source	When asked about this, an official of the American administration replied: "The United States is not conducting electronic surveillance aimed at offices of the World Bank and IMF in Washington."
PBMT	Interrogé à ce sujet, un responsable de l'administration américaine a répondu : "Les Etats-Unis n'est pas effectuer une surveillance électronique destiné aux bureaux de la Banque mondiale et du FMI à Washington".
GNMT	Interrogé à ce sujet, un fonctionnaire de l'administration américaine a répondu: "Les Etats-Unis n'effectue pas de surveillance électronique à l'intention des bureaux de la Banque mondiale et du FMI à Washington".
Human	Interrogé sur le sujet, un responsable de l'administration américaine a répondu: "les Etats-Unis ne mènent pas de surveillance électronique visant les sièges de la Banque mondiale et du FMI à Washington".

Wu et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

Word Embeddings



On word embeddings – Part I by Ruder

Named Entity Recognition

contentShip to site indexPoliticsSubscribeLog InSubscribeLog InToday's PaperAdvertisementSupported ORG byFBI Agent Peter Strzok PERSON . Who Criticized Trump PERSON in Texts, Is FiredimagePeter Strzok, a top FBI GPE counterintelligence agent who was taken off the special counsel investigation after he disparaged texts about President Trump PERSON were uncovered, was fired Credit J. Kirkpatrick PERSON for The New York Times Adam Goldman ORG and Michael S. SchmidtAug PERSON 13 CARDINAL 2016WASHINGTON CARDINAL — Peter Strzok PERSON, the FBI GPE senior counterintelligence agent who disparaged President Trump PERSON in inflammatory text messages and helped oversee the Hillary Clinton PERSON email and Russia GPE investigations, has been fired for violating bureau policies, Mr. Strzok PERSON's lawyer said Monday DATE Mr. Trump and his allies seized on the texts — exchanged during the 2016 DATE campaign with a former FBI GPE lawyer, Lisa Page — in PERSON assailing the Russia GPE investigation as an illegitimate "witch hunt." Mr. Strzok PERSON, who rose over 20 years DATE at the FBI GPE to become one of its most experienced counterintelligence agents, was a key figure in the early months DATE of the inquiry Along with writing the texts, Mr. Strzok PERSON was accused of sending a highly sensitive search warrant to his personal email account. The FBI GPE had been under immense political pressure by Mr. Trump PERSON to dismiss Mr. Strzok PERSON, who was removed last summer DATE from the staff of the special counsel, Robert S. Mueller III PERSON. The president has repeatedly denounced Mr. Strzok PERSON in posts on

Named Entity Recognition and Classification with Scikit-Learn by Susan Li
Esteves et al. Named Entity Recognition in Twitter using Images and Text

Question Answering

System	Dev EM	Dev F1	Test EM	Test F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - alnet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - ninet	-	-	82.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	-	85.8	-	-
R.M. Reader (Single)	78.9	86.3	79.5	86.6
R.M. Reader (Ensemble)	81.2	87.9	82.3	88.5
Ours	80.8	88.5	-	-
BERT _{BASE} (Single)	84.1	90.9	-	-
BERT _{LARGE} (Single)	85.8	91.8	-	-
BERT _{LARGE} (Sgl-TriviaQA)	84.2	91.1	85.1	91.8
BERT _{LARGE} (Ens.+TriviaQA)	86.2	92.2	87.4	93.2

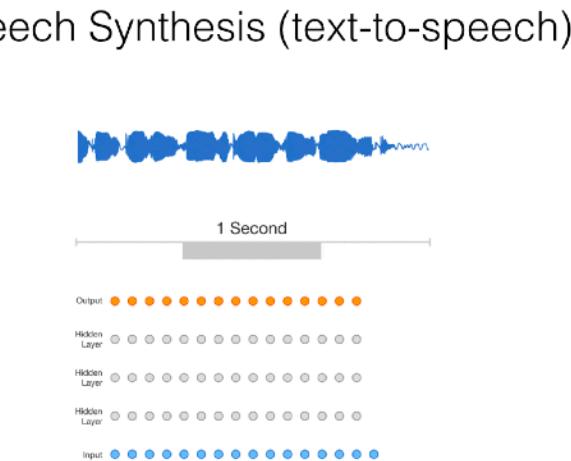
Table 2: SQuAD results. The BERT ensemble is 7x faster than the baseline and achieves better performance with fine-tuning seeds.

Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Speech Recognition

Senone set	Model/combination step	WER devset	WER test ngram-LM	WER devset	WER test
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.3
27k-puhpum	BLSTM	11.3	8.0	9.2	6.3
9k	BLSTM+ResNet+LACE+CNN-BLSTM	9.6	7.2	7.7	5.4
9k-puhpum	BLSTM+ResNet+LACE	9.7	7.4	7.8	5.4
27k	BLSTM+ResNet+LACE+CNN-BLSTM	9.7	7.3	7.8	5.5
-	BLSTM+ResNet+LACE	10.0	7.5	8.0	5.8
Confusion network combination					
+ LSTM rescoring					
+ ngram rescoring					
+ backchannel penalty					

Xiong et al. The Microsoft 2017 Conversational Speech Recognition System



Van Den Oord et al. WaveNet: A Generative Model for Raw Audio

Supervised Learning

Image Classification

Model	image size	# parameters	Mult-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.3
NASNet-A (5 @ 1538)	299×299	10.9 M	2.35 B	78.6	94.2
Inception V3 [60]	299×299	23.8 M	5.72 B	78.8	94.4
Xception [9]	299×299	22.8 M	8.38 B	79.0	94.5
Inception ResNet V2 [58]	299×299	55.8 M	13.2 B	80.1	95.1
NASNet-A (7 @ 1920)	299×299	22.6 M	4.93 B	80.8	95.3
ResNeXt-101 (64 × 4d) [68]	320×320	83.6 M	31.5 B	80.9	95.6
PolyNet [69]	331×331	92 M	34.7 B	81.3	95.8
DPN-131 [8]	320×320	79.5 M	32.0 B	81.5	95.8
SENet [25]	320×320	145.8 M	42.3 B	82.7	96.2
NASNet-A (6 @ 4032)	331×331	88.9 M	23.8 B	82.7	96.2

Table 2. Performance of architecture search and other published state-of-the-art models on ImageNet classification. Mult-Adds indicate the number of composite multiply-accumulate operations for a single image. Note that the composite multiply-accumulate operations are calculated for the image size reported in the table. Model size for [25] calculated from open-source implementation.

Instance Segmentation



Figure 1. Mask R-CNN on COCO test set, using ResNet-101 [17] and running at 5.0 FPS with 35.7 mask AP (Table 1).

Zoph et al. Learning Transferable Architectures for Scalable Image Recognition

He et al. Mask R-CNN

Visual Question Answering

Method	VQA v2 test-dev			VQA v2 test		
	All	Yes/no	Simple	All	Yes/no	Simple
Prov (most common answer in training set) [1]	—	—	—	75.98	61.30	30.96
LSTM Language only (base model) [1]	—	—	—	44.26	47.01	73.15
Diaper LSTM 0.3mns [17] (reported in [1])	—	—	—	54.22	75.46	75.18
MCD [1] (as reported in [1])	—	—	—	62.27	78.82	72.38
CPMC-LSTM [1]	—	—	—	63.71	78.30	73.52
UNUS	—	—	—	67.39	82.50	81.9
HDO-USYD-UNCC	—	—	—	69.77	81.89	82.86
Proposed model	—	—	—	69.09	84.50	85.01
ResNet features T=7, single network	62.07	79.20	79.45	82.62	82.27	79.71
Image Features from OpenImage, adaptive K, single network	65.52	81.30	84.73	84.65	85.30	85.26
ResNet features T=7, ensemble	66.26	81.28	83.17	85.46	86.73	85.3
Image Features from Bilinear attention, adaptive K, ensemble	68.07	84.88	85.89	88.89	78.24	86.64

Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA v2 Leaderboard [1].

Teney et al. Tips and Tricks for Visual Question Answering: Learnings from the 2017 Challenge

Machine Translation

Source	Human	PBMT	GNMT
"The reason why the plane is more comfortable is because it has more legroom at the front."	"The reason why the plane is more comfortable is because it has more legroom at the front."	"The reason why the plane is more comfortable is because it has more legroom at the front."	"The reason why the plane is more comfortable is because it has more legroom at the front."
"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager 6.0 chez l'avionneur Airbus.	"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager 6.0 chez l'avionneur Airbus.	"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager 6.0 chez l'avionneur Airbus.	"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager 6.0 chez l'avionneur Airbus.
When asked about this, an official of the American administration replied: "The United States is not conducting electronic surveillance aimed at offices of the World Bank and IMF in Washington."	When asked about this, an official of the American administration replied: "The United States is not conducting electronic surveillance aimed at offices of the World Bank and IMF in Washington."	Interrogé à ce sujet, un responsable de l'administration américaine a répondu : "Les Etats-Unis n'est pas effectuer une surveillance électronique destiné aux bureaux de la Banque mondiale et du FMI à Washington".	Interrogé à ce sujet, un fonctionnaire de l'administration américaine a répondu : "Les Etats-Unis n'effectuent pas de surveillance électronique à l'intention des bureaux de la Banque mondiale et du FMI à Washington".
Interrogé sur le sujet, un responsable de l'administration américaine a répondu : "les Etats-Unis ne mènent pas de surveillance électronique visant les sièges de la Banque mondiale et du FMI à Washington".	Interrogé sur le sujet, un responsable de l'administration américaine a répondu : "les Etats-Unis ne mènent pas de surveillance électronique visant les sièges de la Banque mondiale et du FMI à Washington".	Interrogé à ce sujet, un responsable de l'administration américaine a répondu : "Les Etats-Unis n'est pas effectuer une surveillance électronique destiné aux bureaux de la Banque mondiale et du FMI à Washington".	Interrogé à ce sujet, un fonctionnaire de l'administration américaine a répondu : "Les Etats-Unis n'effectuent pas de surveillance électronique à l'intention des bureaux de la Banque mondiale et du FMI à Washington".

Wu et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

Question Answering

System	Dev EM	Dev F1	Test EM	Test F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - alnet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nlnet	-	-	82.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	-	85.8	-	-
R.M. Reader (Single)	78.9	86.3	79.5	86.6
R.M. Reader (Ensemble)	81.2	87.9	82.3	88.5
Ours	BERT _{BASE} (Single)	80.8	88.5	-
	BERT _{LARGE} (Single)	84.1	90.9	-
	BERT _{LARGE} (Ensemble)	85.8	91.8	-
	BERT _{LARGE} (Sgl+TriviaQA)	84.2	91.1	85.1
	BERT _{LARGE} (Ens.+TriviaQA)	86.2	92.2	87.4
	BERT _{LARGE} (Ens.+TriviaQA)	86.2	92.2	93.2

Table 2: SQuAD results. The BERT ensemble is 7x times which use different pre-training checkpoints and fine-tuning seeds.

Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

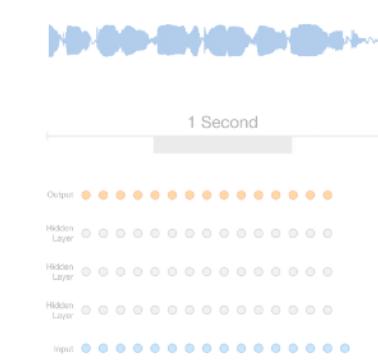
Speech Recognition

Senone set	Model/combination step	WER	WER	WER	WER
		devset	test	devset	test
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.3
27k-puhpum	BLSTM	11.3	8.0	9.2	6.3
9k	BLSTM+ResNet+LACE+CNN-BLSTM	9.6	7.2	7.7	5.4
9k-puhpum	BLSTM+ResNet+LACE	9.7	7.4	7.8	5.4
9k-puhpum	BLSTM+ResNet+LACE+CNN-BLSTM	9.7	7.3	7.8	5.5
27k	BLSTM+ResNet+LACE	10.0	7.5	8.0	5.8

Step	Model	WER	WER
		train	test
Confusion network combination		7.4	5.2
+ LSTM rescoring		7.3	5.2
+ ngram rescoring		7.2	5.2
+ backchannel penalty		7.2	5.1

Xiong et al. The Microsoft 2017 Conversational Speech Recognition System

Speech Synthesis (text-to-speech)



Van Den Oord et al. WaveNet: A Generative Model for Raw Audio

Supervised Learning

Image Classification

Model	image size	# parameters	Mult-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.3
NASNet-A (5 @ 1538)	299×299	10.9 M	2.35 B	78.6	94.2
Inception V3 [60]	299×299	23.8 M	5.72 B	78.8	94.4
Xception [9]	299×299	22.8 M	8.38 B	79.0	94.5
Inception ResNet V2 [58]	299×299	55.8 M	13.2 B	80.1	95.1
NASNet-A (7 @ 1920)	299×299	22.6 M	4.93 B	80.8	95.3
ResNeXt-101 (64 × 4d) [68]	320×320	83.6 M	31.5 B	80.9	95.6
PolyNet [69]	331×331	92 M	34.7 B	81.3	95.8
DPN-131 [8]	320×320	79.5 M	32.0 B	81.5	95.8
SENet [25]	320×320	145.8 M	42.3 B	82.7	96.2
NASNet-A (6 @ 4032)	331×331	88.9 M	23.8 B	82.7	96.2

Table 2. Performance of architecture search and other published state-of-the-art models on ImageNet classification. Mult-Adds indicate the number of composite multiply-accumulate operations for a single image. Note that the composite multiply-accumulate operations are calculated for the image size reported in the table. Model size for [25] calculated from open-source implementation.

Instance Segmentation

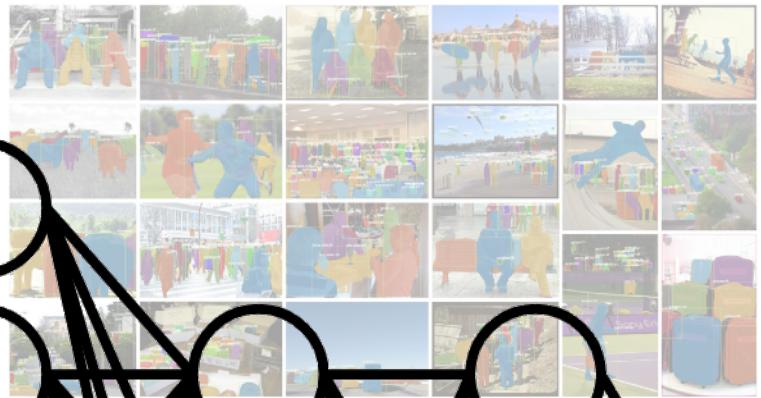


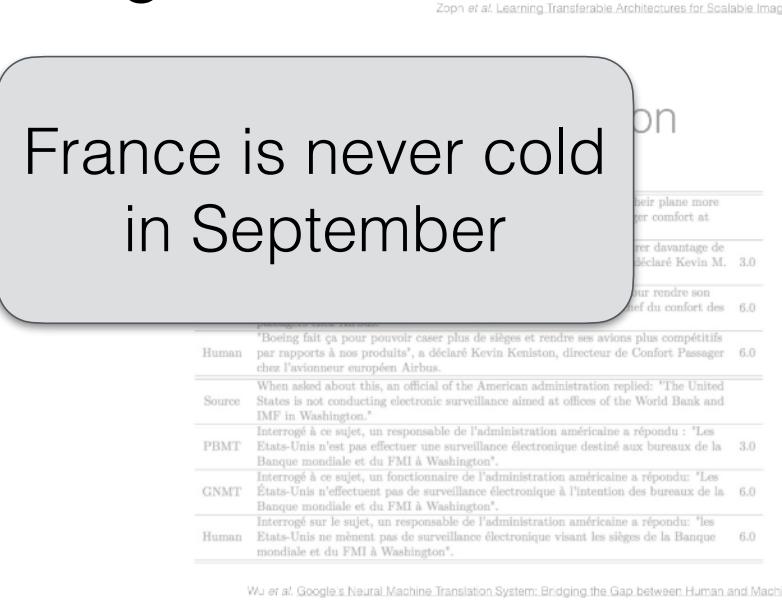
Figure 1. Results of Mask R-CNN on COCO test set, using ResNet-101 [17] and running at 5.0 FPS with 35.7 mask AP (Table 1).

Visual Question Answering

Method	VQA v2 test-dev			VQA v2 test		
	All	Yes/no	Simple	All	Yes/no	Simple
Proposed model	75.98	61.30	30.96	31.17	—	—
LSTM Language only (base model) [11]	—	—	—	44.26	47.01	31.15
Deeper LSTM 0.3mms [11] (reported in [1])	—	—	—	54.22	75.46	35.18
MCD [1] (as reported in [1])	—	—	—	62.27	78.82	32.26
CPMC-LP [1]	—	—	—	63.71	78.30	32.52
67.39	82.50	41.16	55.97	69.77	81.89	46.26
68.09	84.50	43.86	55.30	68.07	84.88	43.01
HDOU-USYD-UNCC	—	—	—	68.07	84.88	43.01
Proposed model	68.07	84.88	43.01	55.01	—	—
ResNet features T=7, simple network	62.07	79.20	39.46	32.62	60.27	79.31
Image Features from better ip networks, adaptive K, simple network	65.52	81.30	44.25	34.65	65.52	81.30
ResNet features T=7, ensemble	66.26	81.26	43.17	34.65	66.73	81.76
Image Features from better ip networks, adaptive K, ensemble	68.07	84.88	43.89	57.24	86.46	44.32

Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA v2 Leaderboard [1].

English sentence



Question Answering

System	Dev EM	Dev F1	Test EM	Test F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - alnet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - alnet	-	-	82.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	-	85.8	-	-
R.M. Reader (Single)	78.9	86.3	79.5	86.6
R.M. Reader (Ensemble)	81.2	87.9	82.3	88.5
Ours				
BERT _{BASE} (Single)	80.8	88.5	-	-
BERT _{LARGE} (Single)	84.1	90.9	-	-
BERT _{LARGE} (Ensemble)	85.8	91.8	-	-
BERT _{LARGE} (Sgl+TriviaQA)	84.2	91.1	85.1	91.8
BERT _{LARGE} (Ens.+TriviaQA)	86.2	92.2	87.4	93.2

Table 2: SQuAD results. The BERT ensemble is 7x8 times which use different pre-training checkpoints and fine-tuning seeds.

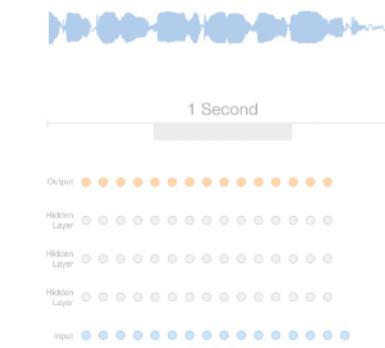
Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Speech Recognition

Senone set	Model/combination step	WER devset	WER test ngram-LM	WER devset	WER test LSTM-LMs
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.3
27k-puhpum	BLSTM	11.3	8.0	9.2	6.3
9k	BLSTM+ResNet+LACE+CNN-BLSTM	9.6	7.2	7.7	5.4
9k-puhpum	BLSTM+ResNet+LACE	9.7	7.4	7.8	5.4
9k-puhpum	BLSTM+ResNet+LACE+CNN-BLSTM	9.7	7.3	7.8	5.5
27k	BLSTM+ResNet+LACE	10.0	7.5	8.0	5.8
Confusion network combination					
+ LSTM rescoring					
+ ngram rescoring					
+ backchannel penalty					
7.4					
7.3					
7.2					
7.2					
7.2					

Xiong et al. The Microsoft 2017 Conversational Speech Recognition System

Speech Synthesis (text-to-speech)



Van Den Oord et al. WaveNet: A Generative Model for Raw Audio

Supervised Learning

Image Classification

Model	image size	# parameters	Mult-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.3
NASNet-A (5 @ 1538)	299×299	10.9 M	2.35 B	78.6	94.2
Inception V3 [60]	299×299	23.8 M	5.72 B	78.8	94.4
Xception [9]	299×299	22.8 M	8.38 B	79.0	94.5
Inception ResNet V2 [58]	299×299	55.8 M	13.2 B	80.1	95.1
NASNet-A (7 @ 1920)	299×299	22.6 M	4.93 B	80.8	95.3
ResNeXt-101 (64 × 4d) [68]	320×320	83.6 M	31.5 B	80.9	95.6
PolyNet [69]	331×331	92 M	34.7 B	81.3	95.8
DPN-131 [8]	320×320	79.5 M	32.0 B	81.5	95.8
SENet [25]	320×320	145.8 M	42.3 B	82.7	96.2
NASNet-A (6 @ 4032)	331×331	88.9 M	23.8 B	82.7	96.2

Table 2. Performance of architecture search and other published state-of-the-art models on ImageNet classification. Mult-Adds indicate the number of composite multiply-accumulate operations for a single image. Note that the composite multiply-accumulate operations are calculated for the image size reported in the table. Model size for [25] calculated from open-source implementation.

Instance Segmentation



Figure 1. Qualitative results of Mask R-CNN on COCO test set, using ResNet-101 [27] and running at 5.0 FPS with 35.7 mask AP (Table 1).

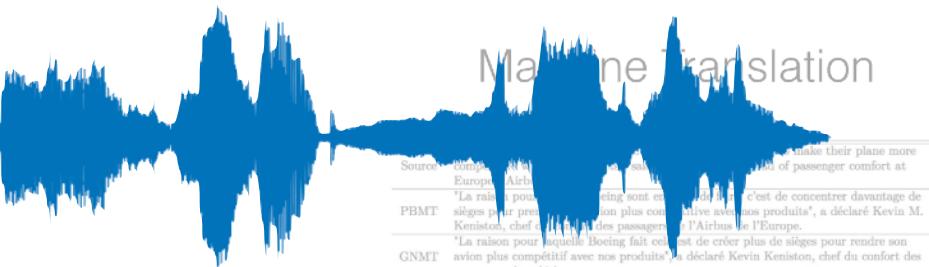
Visual Question Answering

Method	VQA v2 test-dev			VQA v2 test		
	All	Yes/no	Simple	All	Yes/no	Simple
Prop (most common answer in training set) [1]	—	—	—	75.98	61.30	30.96
LSTM Language only (base model) [1]	—	—	—	44.26	47.01	73.15
Deeper LSTM Q & mrcnn [17] (reported in [1])	—	—	—	54.22	75.46	75.18
MCD [1] (as reported in [1])	—	—	—	62.27	78.82	72.28
CPMC-LSTM [1]	—	—	—	63.71	78.30	73.52
UNUS	—	—	—	67.39	82.50	81.16
HDO-USYD-UNCC	—	—	—	69.77	81.89	82.90
Proposed model	—	—	—	69.09	84.50	83.01
ResNet features T=7, single network	62.07	79.20	79.45	82.62	82.27	79.71
Image features from better ip networks, adaptive K, single network	65.52	81.30	84.73	84.65	85.30	85.26
ResNet features T=7, ensemble	66.26	81.20	83.17	84.65	84.73	85.20
Image features from better ip networks, adaptive K, ensemble	68.87	86.88	88.99	93.89	78.24	86.66

Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA v2 Leaderboard [1].

Teney et al. Tips and Tricks for Visual Question Answering: Learnings from the 2017 Challenge

Waveform



Machine Translation

Source		Human		PBMT		GNMT	
Source	compte à rebours pour la mise en place de l'ensemble des mesures pour faire de leur avion plus confortable et plus sûr.	Human	make their plane more comfortable and safer.	Source	"La raison pour laquelle Boeing a fait cette décision est de concentrer davantage de sièges pour plus de passagers", a déclaré Kevin M. Keniston, chef du confort des passagers chez Airbus.	PBMT	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez Airbus.
Source	Source	Source	Source	GNMT	"La raison pour laquelle Boeing a fait cette décision est de concentrer davantage de sièges pour plus de passagers", a déclaré Kevin Keniston, chef du confort des passagers chez Airbus.	PBMT	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez Airbus.
Source	Source	Source	Source	GNMT	"La raison pour laquelle Boeing a fait cette décision est de concentrer davantage de sièges pour plus de passagers", a déclaré Kevin Keniston, chef du confort des passagers chez Airbus.	PBMT	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez Airbus.
Human	Human	Human	Human	GNMT	"La raison pour laquelle Boeing a fait cette décision est de concentrer davantage de sièges pour plus de passagers", a déclaré Kevin Keniston, chef du confort des passagers chez Airbus.	PBMT	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez Airbus.
Source	Source	Source	Source	GNMT	"La raison pour laquelle Boeing a fait cette décision est de concentrer davantage de sièges pour plus de passagers", a déclaré Kevin Keniston, chef du confort des passagers chez Airbus.	PBMT	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez Airbus.
Human	Human	Human	Human	GNMT	"La raison pour laquelle Boeing a fait cette décision est de concentrer davantage de sièges pour plus de passagers", a déclaré Kevin Keniston, chef du confort des passagers chez Airbus.	PBMT	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez Airbus.
Human	Human	Human	Human	GNMT	"La raison pour laquelle Boeing a fait cette décision est de concentrer davantage de sièges pour plus de passagers", a déclaré Kevin Keniston, chef du confort des passagers chez Airbus.	PBMT	"Boeing fait ça pour pouvoir casser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez Airbus.

Question Answering

System	Dev EM	Test F1	Dev EM	Test F1
Leaderboard (Oct 8th, 2018)				
Human	-	82.3	91.2	
#1 Ensemble - alnet	-	86.0	91.7	
#2 Ensemble - QANet	-	84.5	90.5	
#1 Single - nlnet	-	82.5	90.1	
#2 Single - QANet	-	82.5	89.3	
Published				
BIDAF+ELMo (Single)	-	85.8	-	-
R.M. Reader (Single)	78.9	86.3	79.5	86.6
R.M. Reader (Ensemble)	81.2	87.9	82.3	88.5
Ours				
BERT _{BASE} (Single)	80.8	88.5	-	-
BERT _{LARGE} (Single)	84.1	90.9	-	-
BERT _{LARGE} (Ensemble)	85.8	91.8	-	-
BERT _{LARGE} (Sgl+TriviaQA)	84.2	91.1	85.1	91.8
BERT _{LARGE} (Ens.+TriviaQA)	86.2	92.2	87.4	93.2

Table 2: SQuAD results. The BERT ensemble is 7x times which use different pre-training checkpoints and fine-tuning seeds.

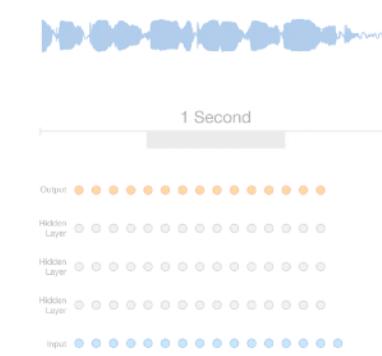
Speech Recognition

Senone set	Model/combination step	WER		WER	
		devset	test	devset	test
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.3
27k-puhpum	BLSTM	11.3	8.0	9.2	6.3
9k	BLSTM+ResNet+LACE+CNN-BLSTM	9.6	7.2	7.7	5.4
9k-puhpum	BLSTM+ResNet+LACE	9.7	7.4	7.8	5.4
9k-puhpum	BLSTM+ResNet+LACE+CNN-BLSTM	9.7	7.3	7.8	5.5
27k	BLSTM+ResNet+LACE	10.0	7.5	8.0	5.8

-	Confusion network combination	7.4	5.2
-	+ LSTM rescoring	7.3	5.2
-	+ ngram rescoring	7.2	5.2
-	+ backchannel penalty	7.2	5.1

Xiong et al. The Microsoft 2017 Conversational Speech Recognition System

Speech Synthesis (text-to-speech)



Van Den Oord et al. WaveNet: A Generative Model for Raw Audio

Image Classification

Instance Segmentation

Visual Question Answering

Waveform

Machine Translation

Source		Human		PBMT		GNMT	
Source	Source	Human	Human	Source	Source	GNMT	GNMT
Source	Source	Human	Human	GNMT	GNMT	GNMT	GNMT
Human	Human	Human	Human	GNMT			

Supervised Learning

Image Classification

Model	image size	# parameters	Mult-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.3
NASNet-A (5 @ 1538)	299×299	10.9 M	2.35 B	78.6	94.2
Inception V3 [60]	299×299	23.8 M	5.72 B	78.8	94.4
Xception [9]	299×299	22.8 M	8.38 B	79.0	94.5
Inception ResNet V2 [58]	299×299	55.8 M	13.2 B	80.1	95.1
NASNet-A (7 @ 1920)	299×299	22.6 M	4.93 B	80.8	95.3
ResNeXt-101 (64 × 4d) [68]	320×320	83.6 M	31.5 B	80.9	95.6
PolyNet [69]	331×331	92 M	34.7 B	81.3	95.8
DPN-131 [8]	320×320	79.5 M	32.0 B	81.5	95.8
SENet [25]	320×320	145.8 M	42.3 B	82.7	96.2
NASNet-A (6 @ 4032)	331×331	88.9 M	23.8 B	82.7	96.2

Table 2. Performance of architecture search and other published state-of-the-art models on ImageNet classification. Mult-Adds indicate the number of composite multiply-accumulate operations for a single image. Note that the composite multiply-accumulate operations are calculated for the image size reported in the table. Model size for [25] calculated from open-source implementation.

Instance Segmentation

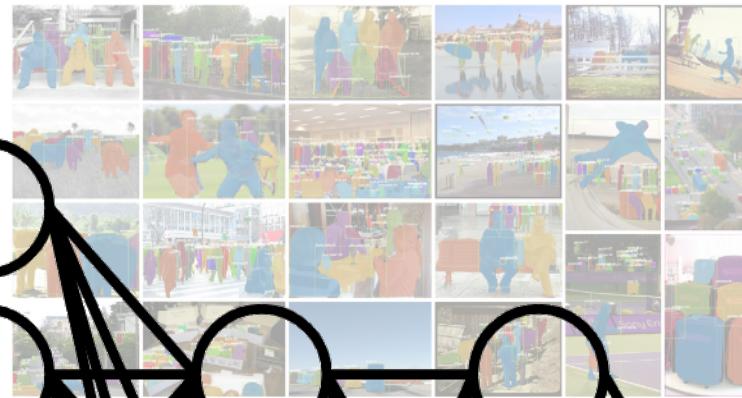


Figure 1. Mask R-CNN on COCO test set, using ResNet-101 [27] and running at 5.0 FPS with 35.7 mask AP (Table 1).

Zoph et al. Learning Transferable Architectures for Scalable Image Recognition

Visual Question Answering

Method	VQA v2 test-dev			VQA v2 test		
	All	Yes/no	Simple	All	Yes/no	Simple
Prov (most common answer in training set) [1]	—	—	—	75.98	61.30	3.06
LSTM Language only (base model) [1]	—	—	—	44.26	47.01	73.37
Diaper LSTM 0.3m [17] (reported in [1])	—	—	—	54.22	75.46	75.18
MCD [1] (as reported in [1])	—	—	—	62.27	78.82	72.36
CPMC-LSTM [1]	—	—	—	63.71	78.30	73.52
UNUS	—	—	—	67.39	82.50	81.97
HDO-USYD-UNCC	—	—	—	69.77	81.89	82.90
Proposed model	—	—	—	69.09	84.50	85.01
ResNet features T+7, single network	62.07	79.20	79.45	82.62	82.27	79.71
Image Features from better ip-answers, adaptive K, single network	65.52	81.30	84.27	84.65	85.30	85.26
ResNet features T+7, ensemble	66.26	81.20	83.17	84.65	84.73	85.20
Image Features from better ip-answers, adaptive K, ensemble	68.87	86.88	88.89	88.89	78.24	86.64

Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA v2 Leaderboard [1].

Teney et al. Tips and Tricks for Visual Question Answering: Learnings from the 2017 Challenge

Machine Translation

Input	
Source	"The reason competitive Europe's
PBMT	"La raison siége pour déclaré Kevin M. 3.0
GNMT	"La raison avion plus" 6.0
Human	"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapport à nos produits", déclare Kevin Keniston, directeur de Confort Passager 6.0 chez l'avionneur Airbus.
Source	When asked about this, un official of the American administration replied: "The United States is not conducting electronic surveillance aimed at offices of the World Bank and IMF in Washington."
PBMT	Interrogé à ce sujet, un responsable de l'administration américaine a répondu : "Les Etats-Unis n'est pas effectuer une surveillance électronique destiné aux bureaux de la Banque mondiale et du FMI à Washington".
GNMT	Interrogé à ce sujet, un fonctionnaire de l'administration américaine a répondu: "Les Etats-Unis n'effectuent pas de surveillance électronique à l'intention des bureaux de la Banque mondiale et du FMI à Washington".
Human	Interrogé sur le sujet, un responsable de l'administration américaine a répondu: "les Etats-Unis ne mènent pas de surveillance électronique visant les sièges de la Banque mondiale et du FMI à Washington".

Wu et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

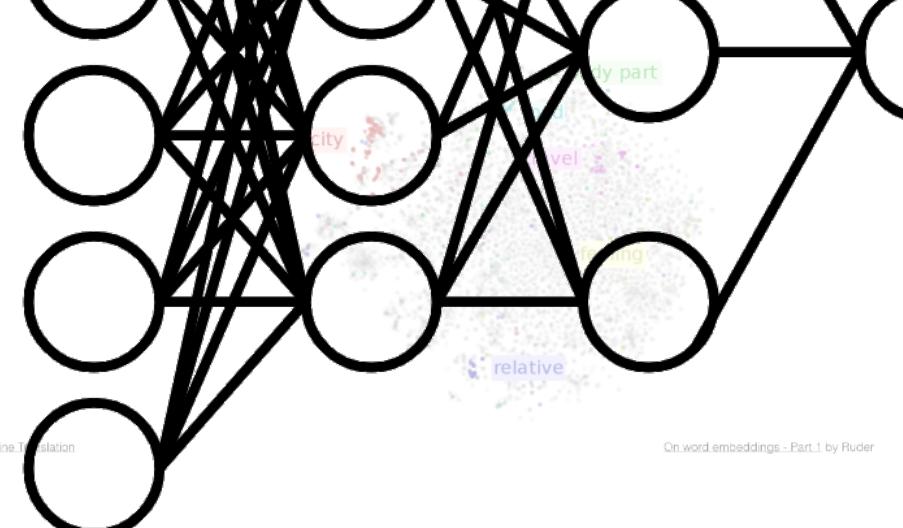
Question Answering

System	Dev EM	Dev F1	Test EM	Test F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	86.0	91.7
Input Question:	#1 Ensemble - elnet	-	86.0	91.7
Where do water droplets collide with ice crystals to form precipitation?	#2 Ensemble - QANet	-	84.5	90.5
	#1 Single - nlms	-	82.5	90.1
	#2 Single - QANet	-	82.5	89.3
Published				
Input Question:	ELM-DM [14]	80.8	88.5	
... Precipitation occurs as smaller droplets coalesce via colliding with other rain drops or ice crystals within a cloud. ...	R.M. Reader [15]	80.2	87.6	
Input Paragraph:	Ours	BERT _{BASE} (S-10) 80.8 88.5		
Output Answer:	BERT _{BASE} (S-10) + R.T.LARGE 86.2 92.2	87.4 93.2		

Table 2: SQuAD results. The BERT ensemble is 7 systems which use different pre-training checkpoints and fine-tuning seeds.

Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Word Embeddings



On word embeddings – Part I by Ruder

Speech Recognition

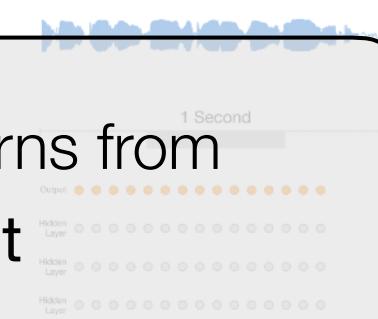
Senone set	Model/combination step	WER		WER	
		devset	test	devset	test
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.0
9k-pulpum	BLSTM + LACE	9.2	7.1	7.7	5.2
9k-pulpum	BLSTM + ResNet + LACE	9.7	7.4	7.8	5.5
27k	BLSTM + ResNet + LACE	10.0	7.5	8.0	5.8
	ConvTasNet + CTC decoding + backchannel penalty			7.2	5.1

Xiong et al. The Microsoft 2017 Conversational Speech Recognition System

Named Entity Recognition

Named Entity Recognition and Classification with SICKit-Learn by Susan Li	
Esteves et al. Named Entity Recognition in Twitter using Images and Text	

Speech Synthesis (text-to-speech)



Van Den Oord et al. WaveNet: A Generative Model for Raw Audio

Question Answering

- The model has no control over the dataset it learns from
- Ground truth output can be specified given input

Robot Learning via Supervised Learning

Image Classification

Model	image size	# parameters	Mult-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224x224	11.2 M	1.94 B	74.8	92.3
NASNet-A (5 @ 1538)	299x299	10.9 M	2.35 B	78.6	94.2
Inception V3 [60]	299x299	—	5.72 B	78.8	94.4
Xception [9]	299x299	22.8 M	8.38 B	79.0	94.5
Inception ResNet V2 [58]	299x299	55.8 M	13.2 B	80.1	95.1
NASNet-A (7 @ 1920)	299x299	22.6 M	4.93 B	80.8	95.3
ResNeXt-101 (64 x 4d) [68]	320x320	83.6 M	31.5 B	80.9	95.6
PolyNet [69]	331x331	92 M	34.7 B	81.3	95.8
DPN-131 [8]	320x320	79.5 M	32.0 B	81.5	95.8
SENet [25]	320x320	145.8 M	42.3 B	82.7	96.2
NASNet-A (6 @ 4032)	331x331	88.9 M	23.8 B	82.7	96.2

Table 2. Performance of architecture search and other published state-of-the-art models on ImageNet classification. Mult-Adds indicate the number of composite multiply-accumulate operations for a single image. Note that the composite multiply-accumulate operations are calculated for the image size reported in the table. Model size for [25] calculated from open-source implementation.

Instance Segmentation

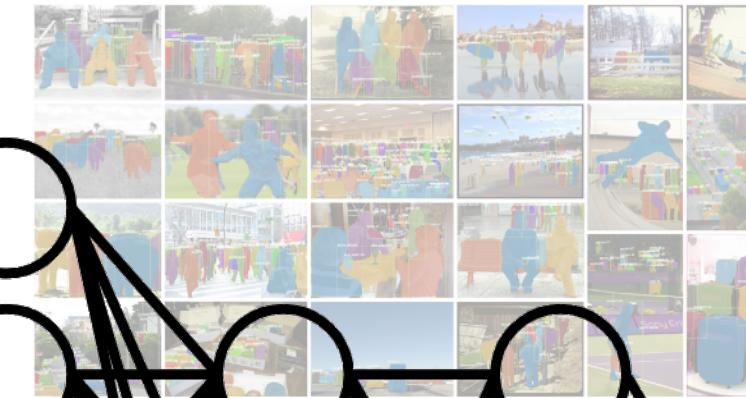


Figure 10. Mask R-CNN on COCO test set, using ResNet-101 [25] and running at 5.0 FPS with 35.7 mask AP (Table 1).

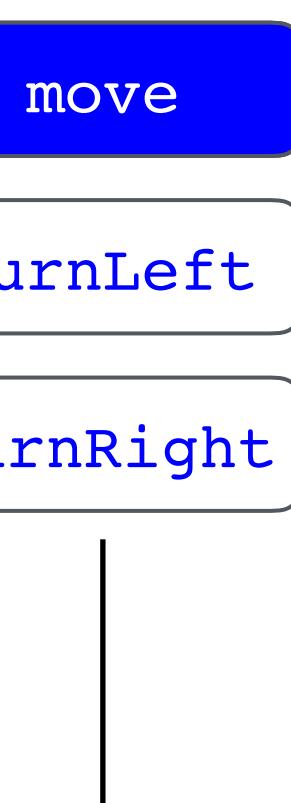
Visual Question Answering

Method	VQA v2 Leaderboard		
	All	Yes/no	Simple
Proposed model	75.98	61.30	9.06
LSTM Language only (base model) [1]	—	—	—
Deep LSTM (ours, [1]) is reported in [1]	44.26	47.01	73.15
MCD [1] is reported in [1]	54.22	75.46	73.38
CPMC-LSTM [1]	—	—	—
—	62.27	78.82	72.36
—	63.71	78.82	73.32
—	67.39	82.50	81.97
—	69.77	83.89	82.90
—	68.09	84.50	83.01
HDOU-USyd-UNCC	—	—	—
Proposed model	68.09	84.50	83.01
ResNet features T=7, single network	62.07	79.20	79.45
Image Features from better ip-features, adaptive K, single network	65.52	81.30	84.73
ResNet features T=7, ensemble	66.26	81.30	83.90
ResNet features T=7, ensemble, adaptive Fc-ensemble	68.07	84.88	83.89
VQA v2 Leaderboard	78.24	86.46	84.32

Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA v2 Leaderboard [1].

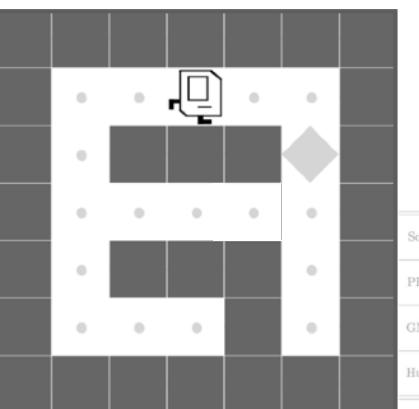
Action

Teney et al. Tips and Tricks for Visual Question Answering: Learnings from the VQA v2 Leaderboard



State

Zoph et al. Learning Transferable Architectures for Scalable Image Recognition



Question Answering

Speech Recognition

Named Entity Recognition

Output

turnLeft

turnRight

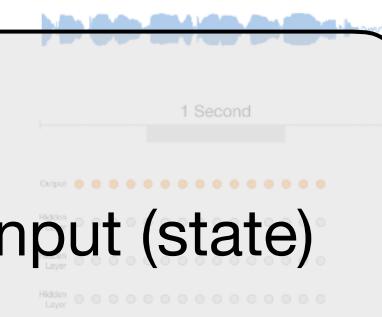
Speech Synthesis (text-to-speech)

System	Dev EM	Dev F1	Test EM	Test F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	-	-
Input Question:	#1 Ensemble - elnet	-	86.0	91.7
Where do water droplets collide with ice crystals to form precipitation?	#2 Ensemble - QANet	-	84.5	90.5
	#1 Single - nlnet	-	82.5	90.1
	#2 Single - QANet	-	82.5	89.3
Published				
Input Question:	BERT (Single)	80.8	88.5	-
... Precipitation occurs as smaller droplets coalesce via colliding with other rain drops or ice crystals within a cloud. ...	BERT (Ensemble)	80.9	89.1	-
Output Answer:	BERTLARGE (Eas+TriviaQA)	86.2	92.2	87.4
within a cloud	BERTLARGE (Eas+TriviaQA)	91.2	93.0	91.2

- The model affects the data it learns from
- Difficult to specify desired output (action) for every input (state)

Senone set	Model/combination step	WER devset	WER test	WER devset	WER test
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.3
9k-pulpum	BLSTM+LACE	9.7	7.4	7.8	5.5
9k-pulpum	BLSTM+LACE+CNN-BLSTM	9.7	7.3	7.8	5.5
27k	BLSTM+ResNet+LACE	10.0	7.5	8.0	5.8
	+ backchannel penalty			7.2	5.1

Table 2: SQuAD results. The BERT ensemble is 7 systems which use different pre-training checkpoints and fine-tuning seeds.

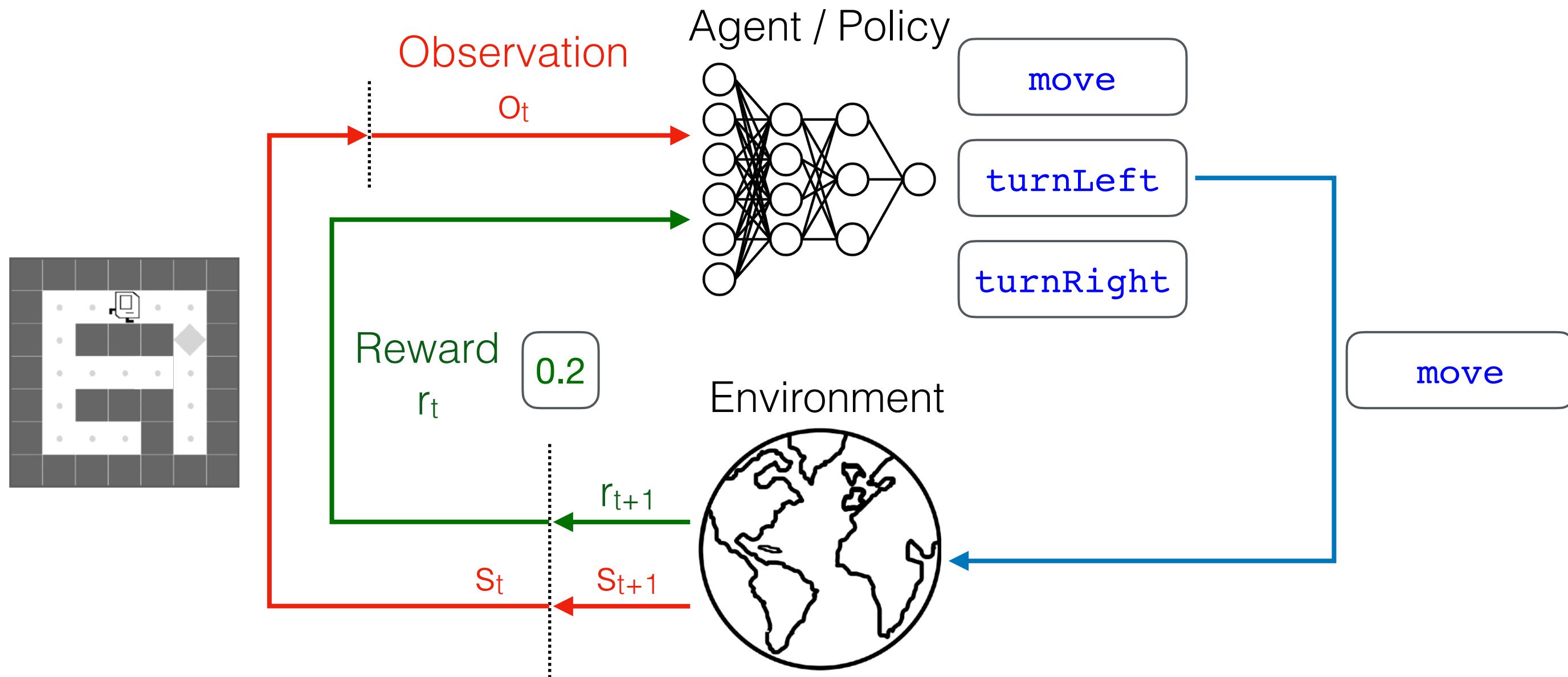


Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Xiong et al. The Microsoft 2017 Conversational Speech Recognition System

Van Den Oord et al. WaveNet: A Generative Model for Raw Audio

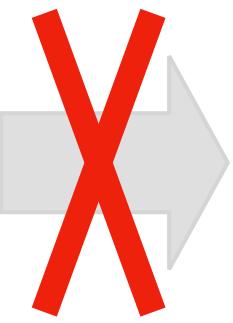
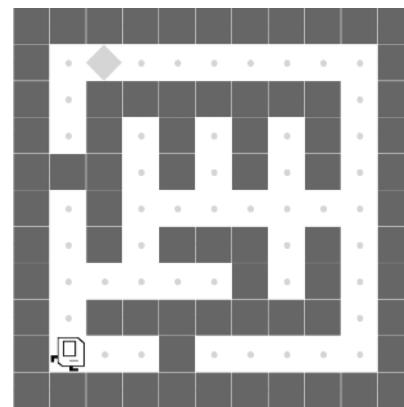
Robot Learning via Deep Reinforcement Learning



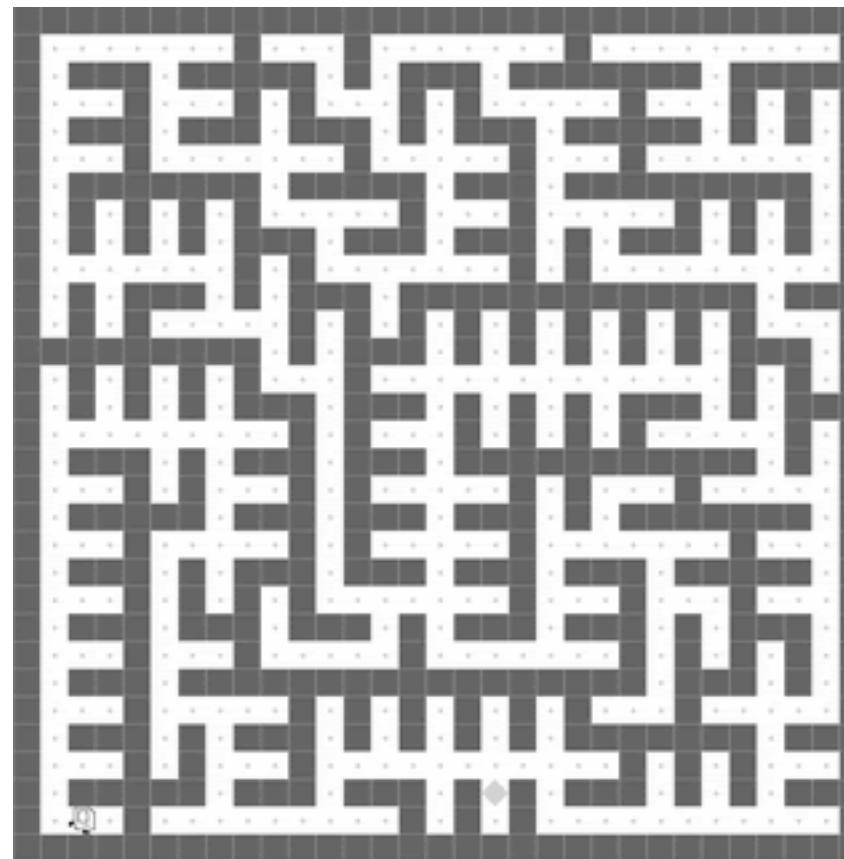
Goal: maximize $\sum_{t=0}^{t=H} \gamma^t R_t(s_t, a_t)$

Robot Learning via Deep Reinforcement Learning - Issues

Generalization



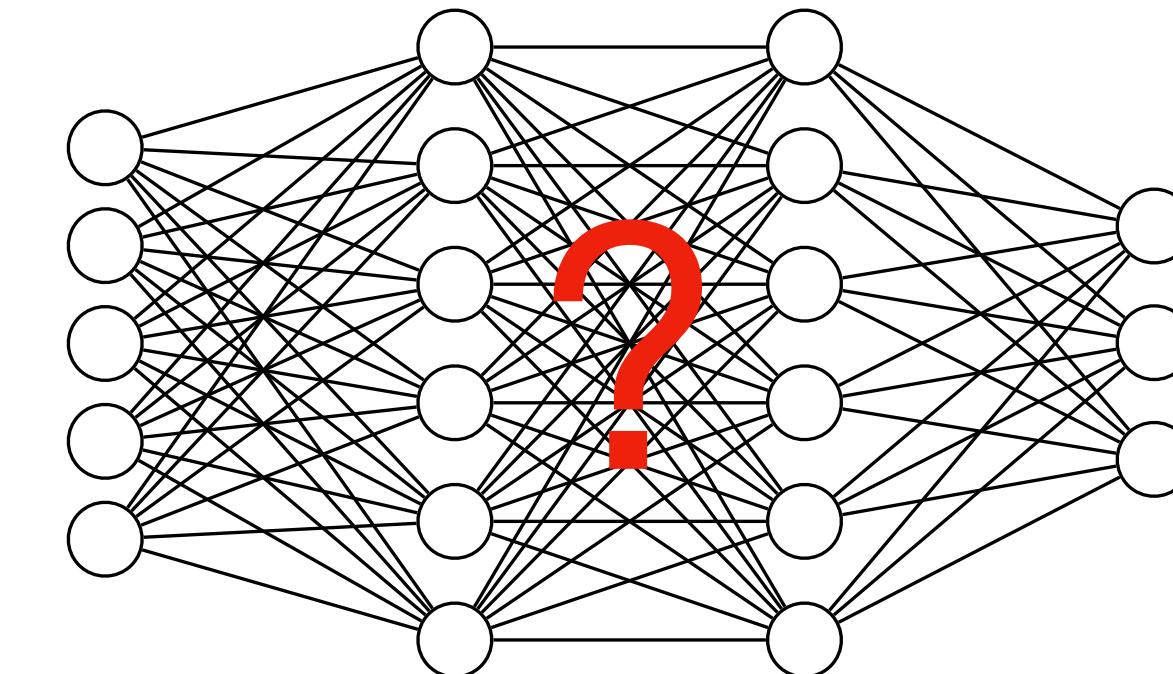
Simple task



Complex task

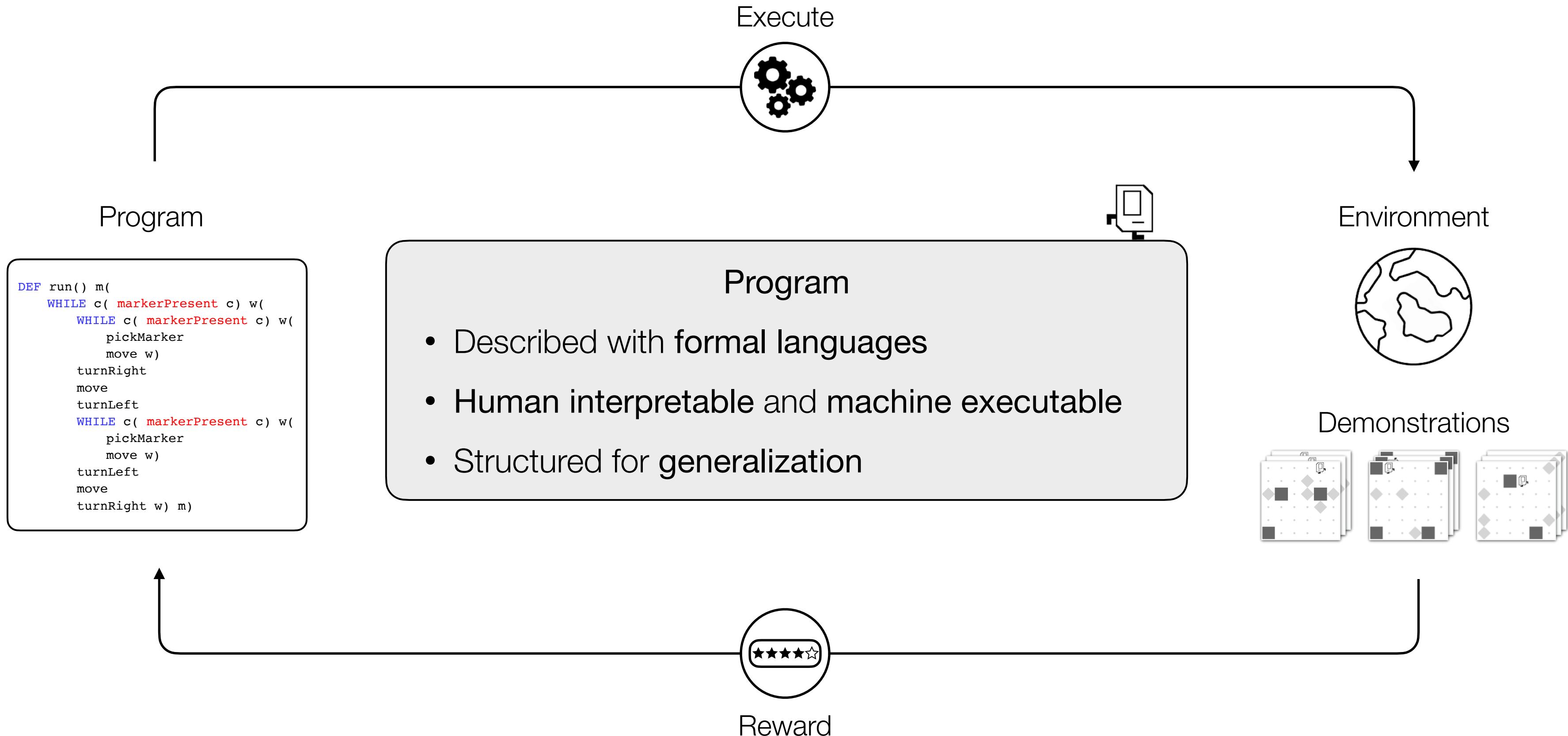
Interpretability

Trust, Safety, and Contestability



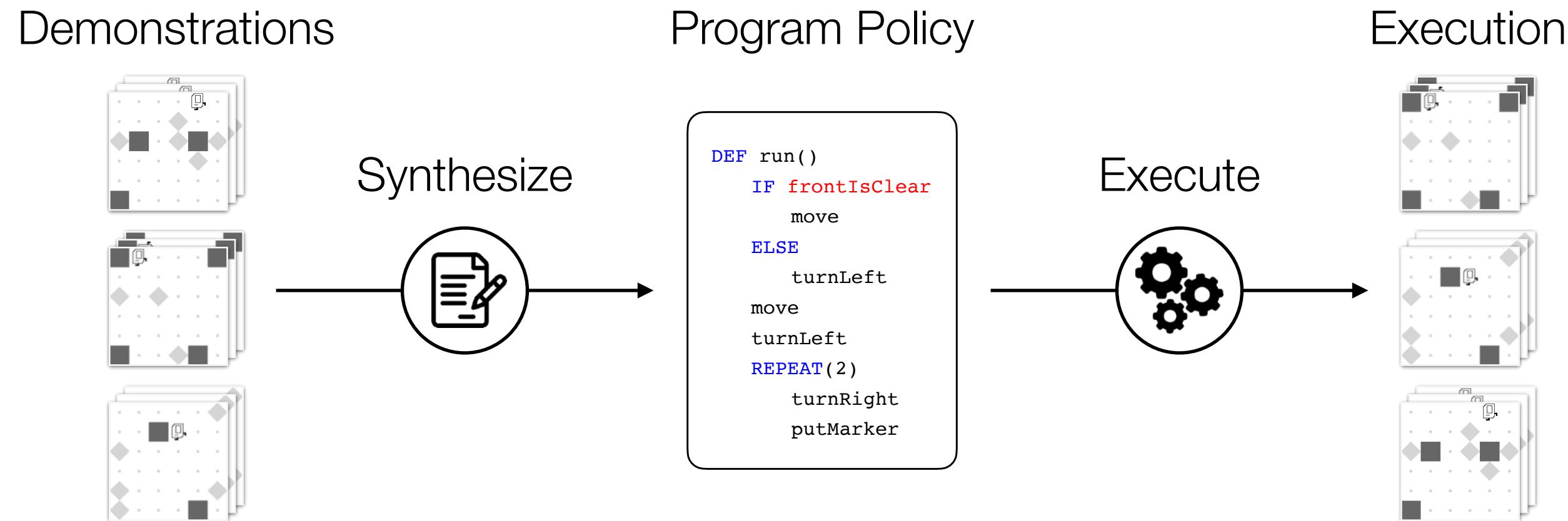
Deep neural network policy

Program as Reinforcement Learning Policies



Neural Program Synthesis from Diverse Demonstration Videos

ICML 2018



Shao-Hua Sun*



Hyeonwoo Noh*

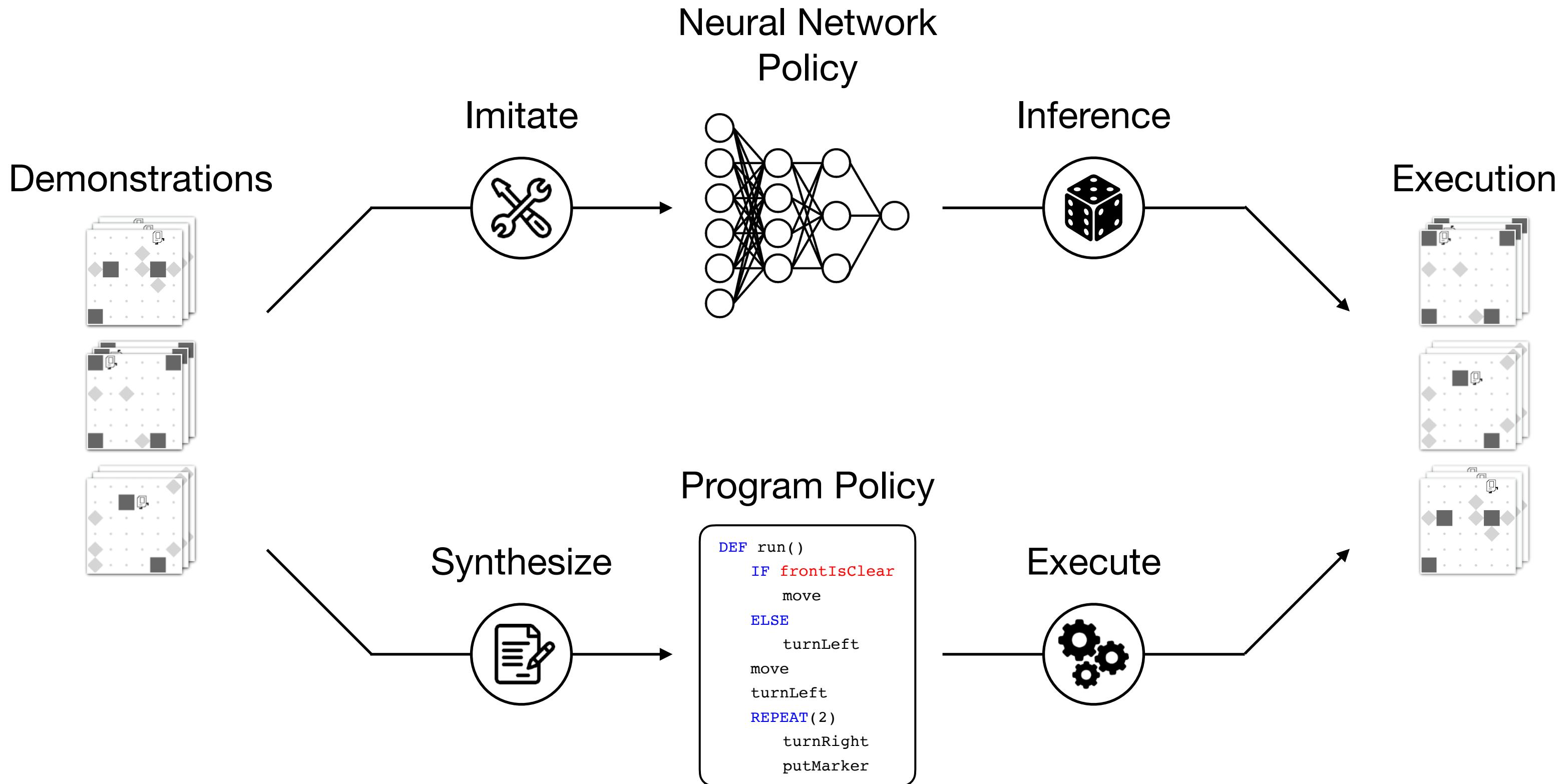


Sriram Somasundaram



Joseph J. Lim

Imitation Learning via Synthesizing Programs

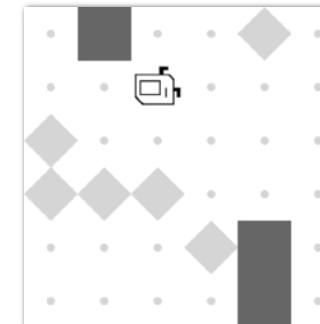
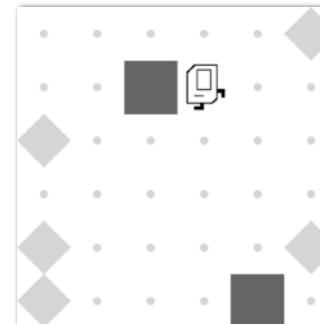
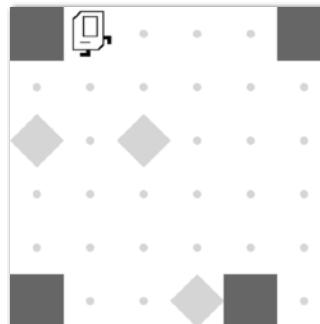
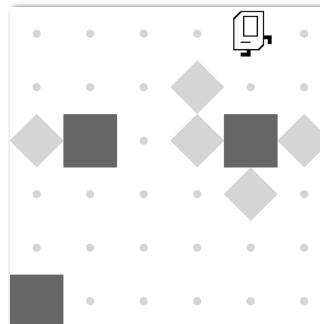


Environments

Karel

Program

```
DEF run()
  IF frontIsClear
    move
  ELSE
    turnLeft
    move
    turnLeft
  REPEAT(2)
    turnRight
    putMarker
```



ViZDoom

Program

```
DEF run()
  WHILE frontIsClear(HellKnight)
    attack
    moveForward
  IF thereIs(Demon)
    moveRight
  ELSE
    moveLeft
    moveBackward
```

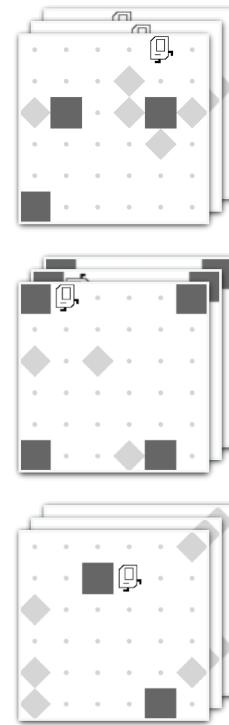


Richard E Pattis. "Karel the robot: a gentle introduction to the art of programming." John Wiley & Sons, Inc., 1981

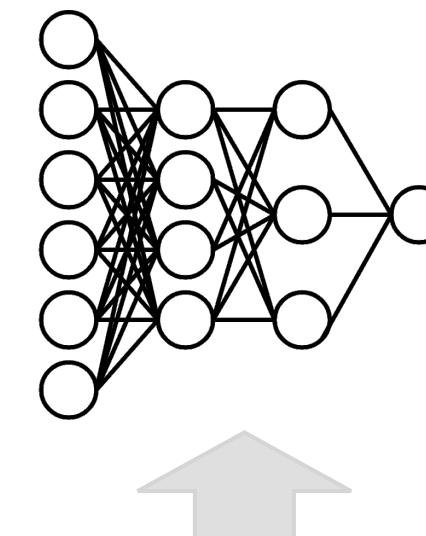
Kempka et al., "Vizdoom: A doom-based ai research platform for visual reinforcement learning." in CIG, 2016

Imitation Learning with Neural Network Policy

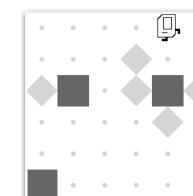
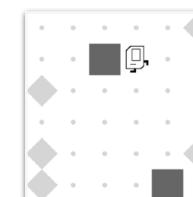
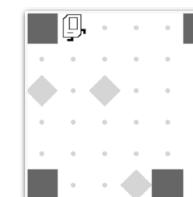
Demonstrations



Neural Network
Policy



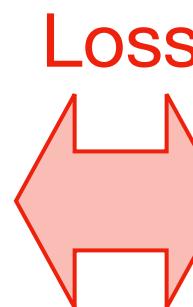
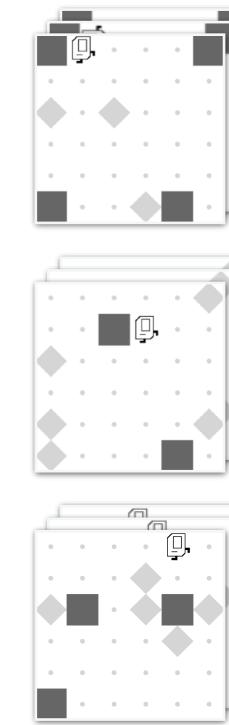
Initial States



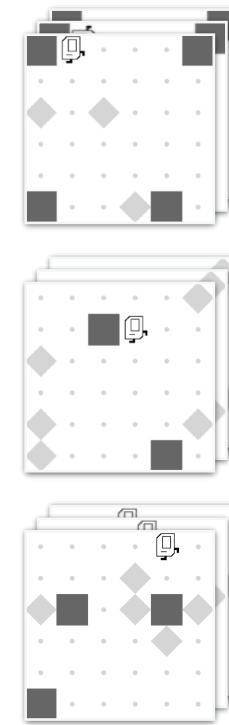
Infer



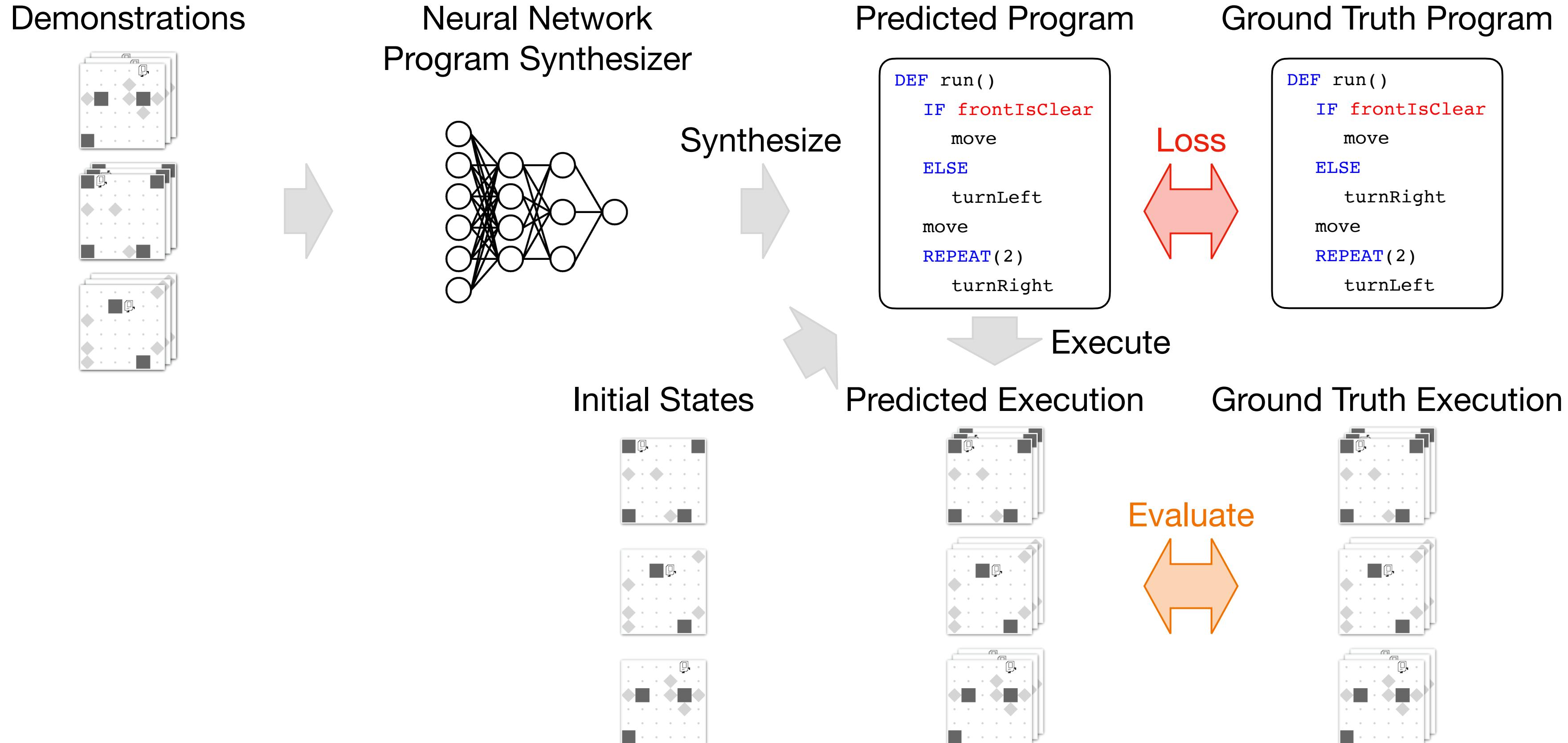
Predicted Execution



Ground Truth Execution

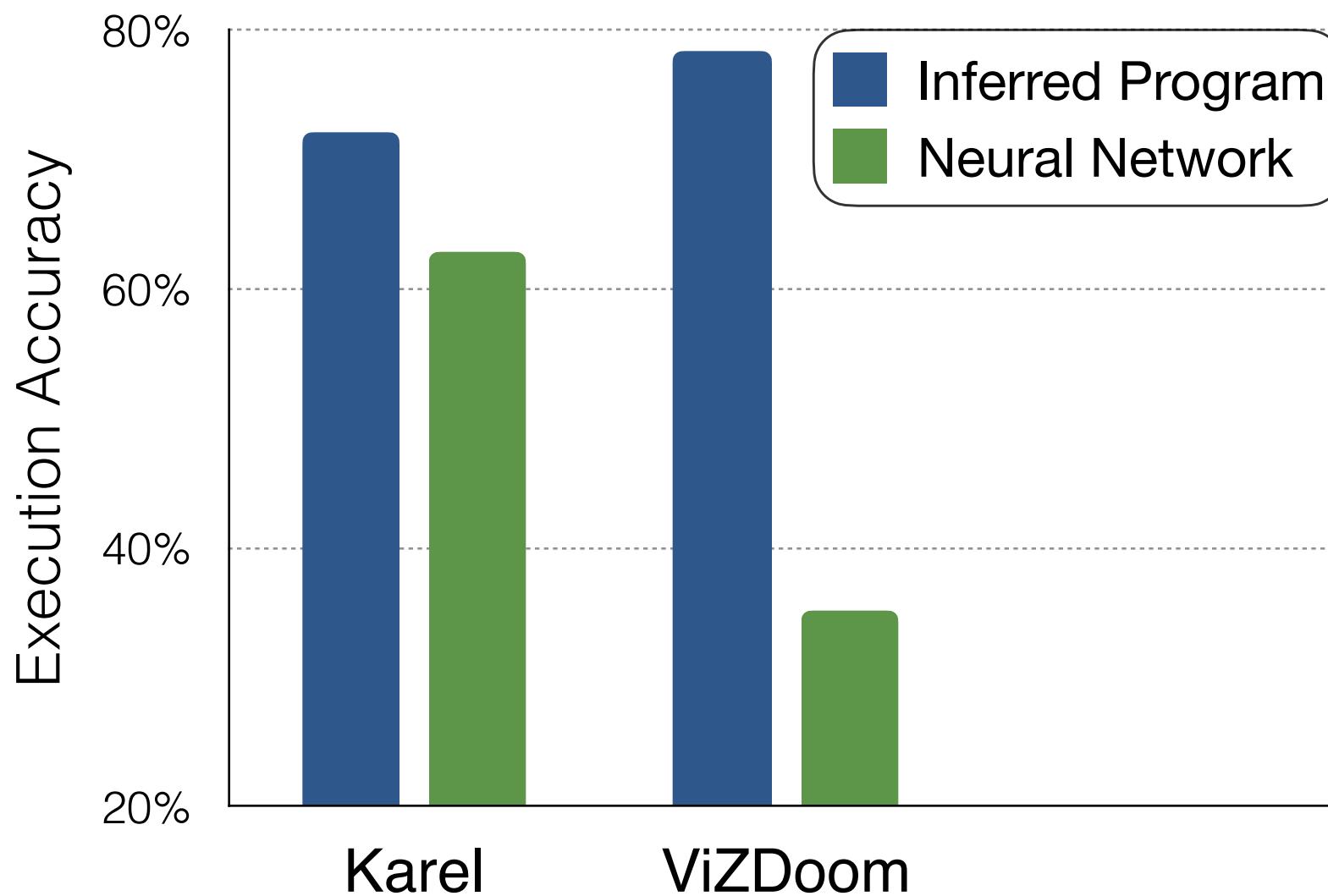


Imitation Learning with Program Policy



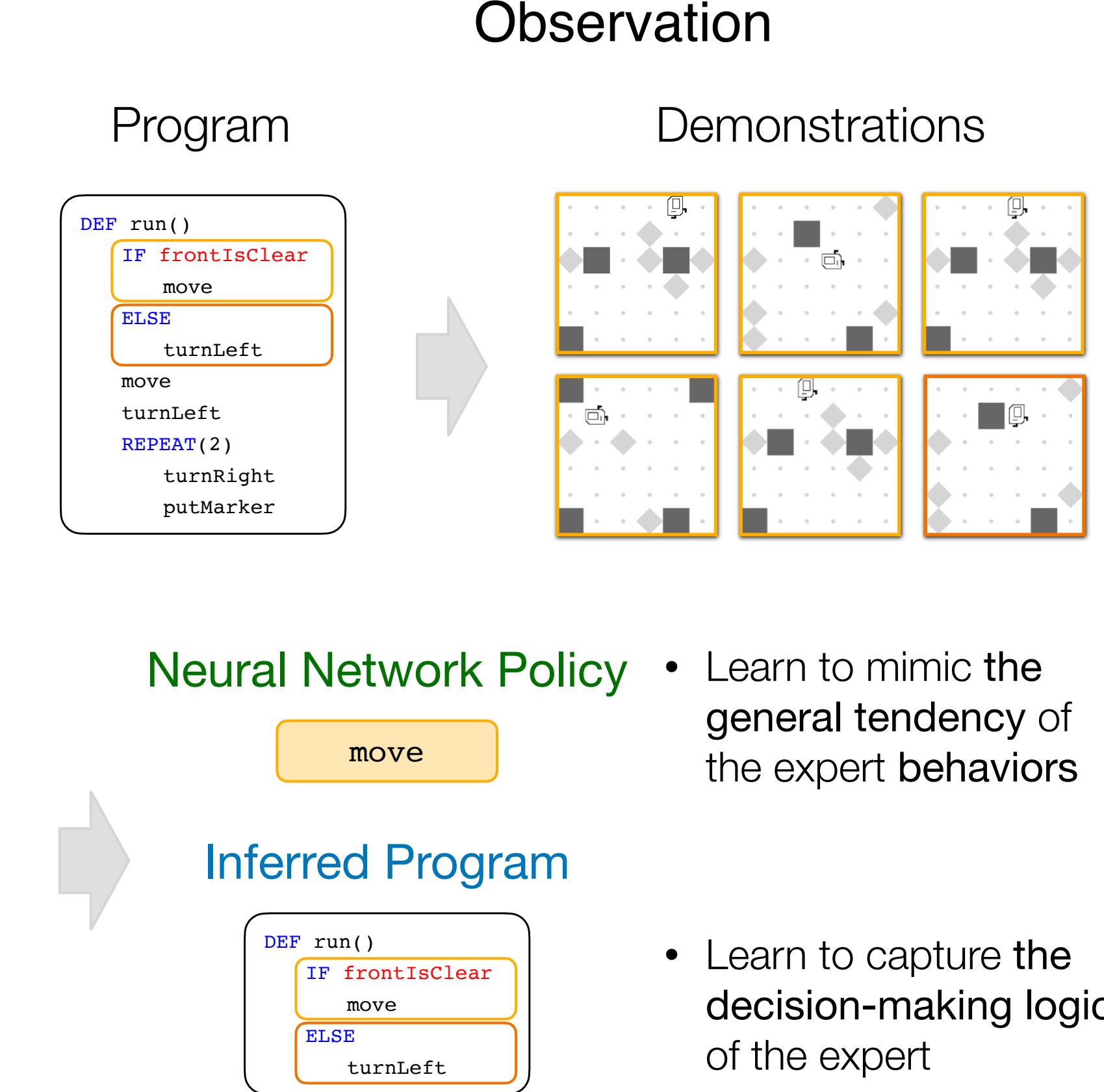
Experimental Results

Quantitative Results



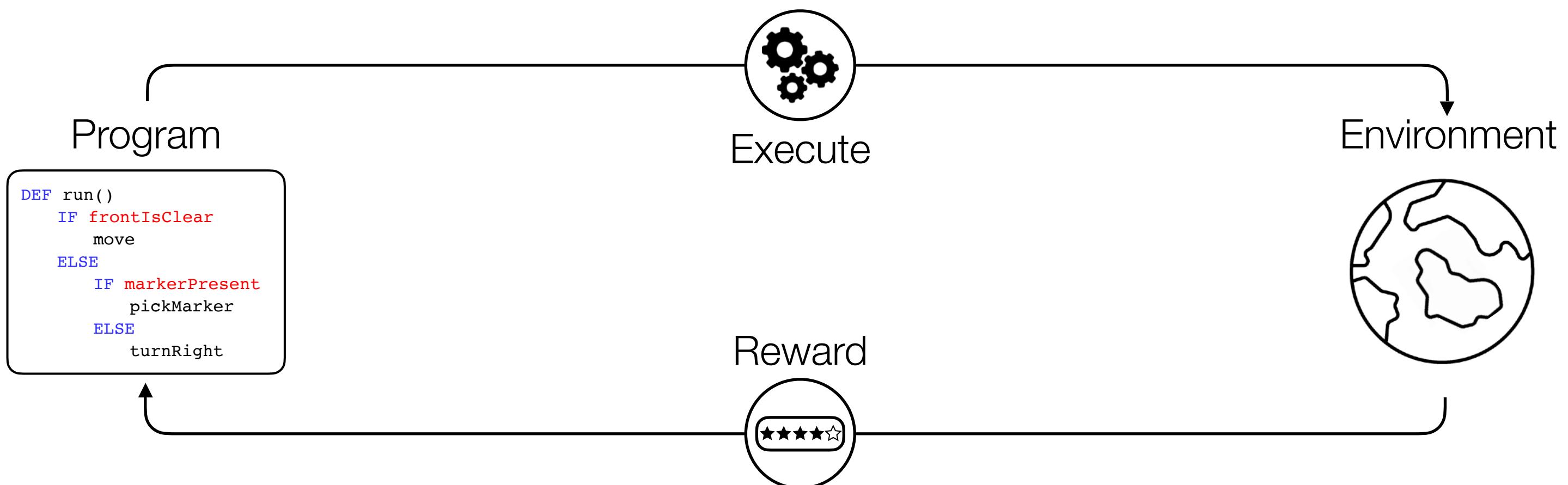
Evaluation: Execute the [inferred program](#) and the [learned neural network policy](#) on a set of unseen initial states and compare them to the [ground truth demonstrations](#)

Observation



Learning to Synthesize Programs as Interpretable and Generalizable Policies

NeurIPS 2021



Dweep Trivedi*



Jesse Zhang*

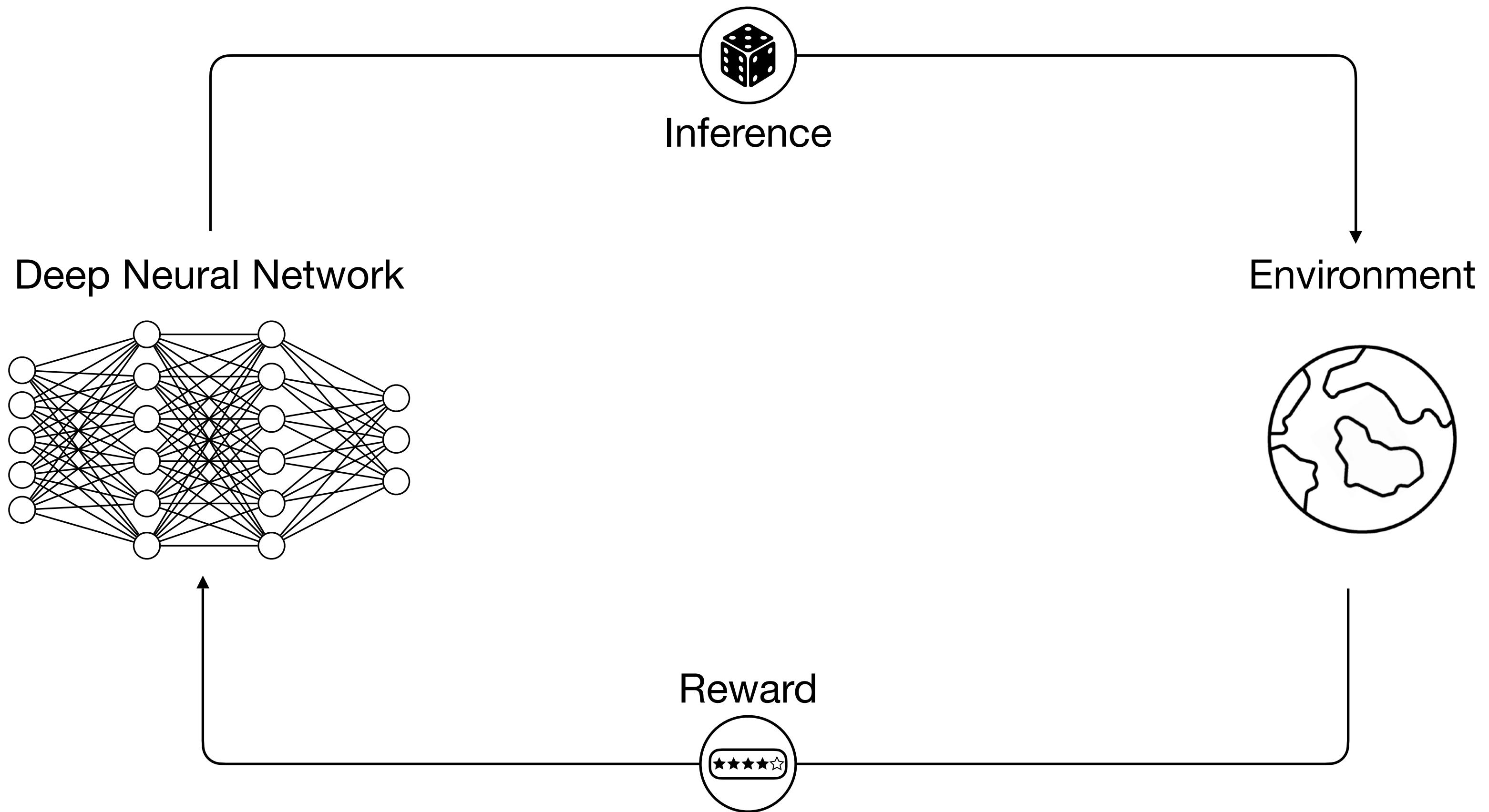


Shao-Hua Sun

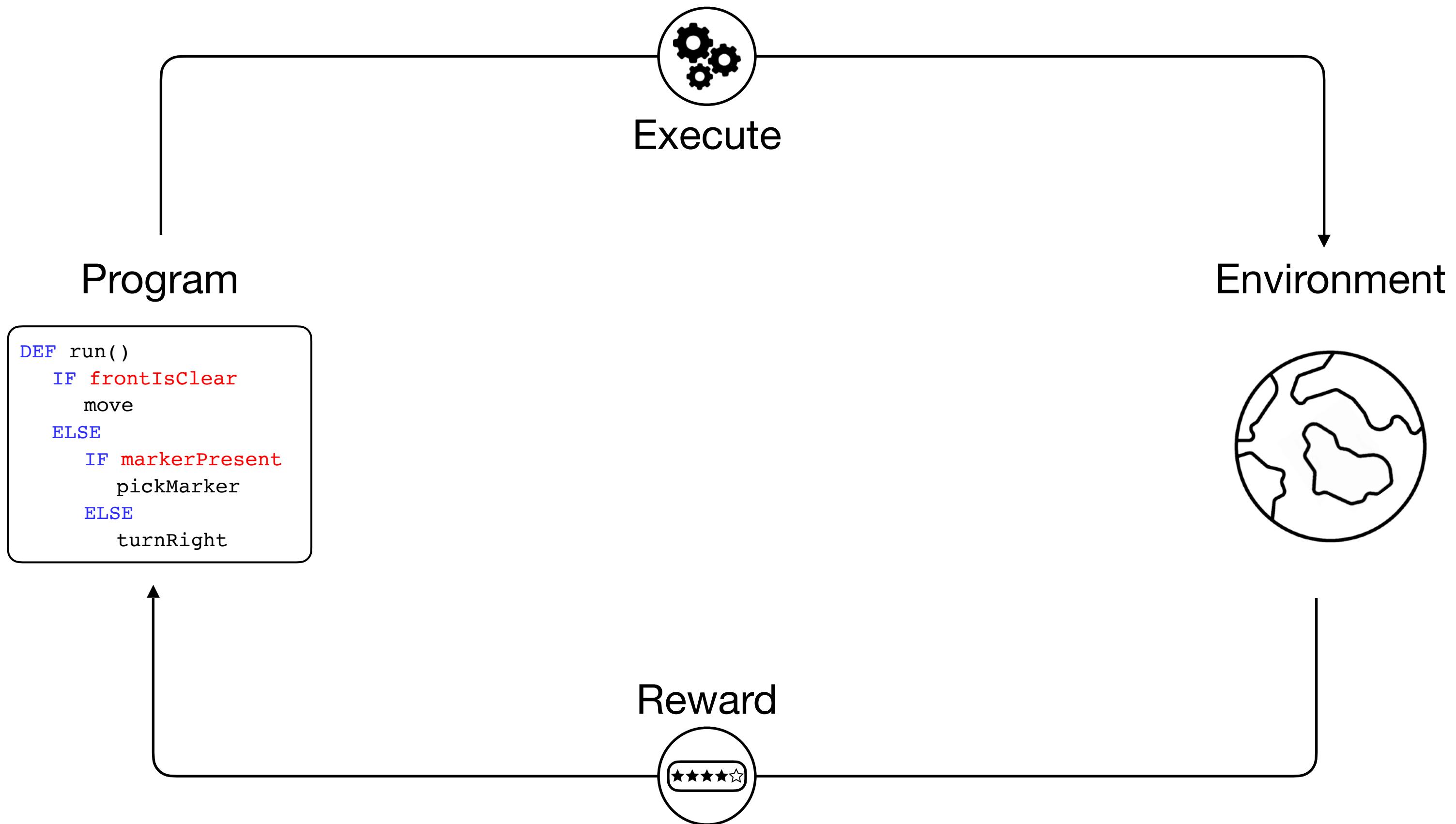


Joseph J. Lim

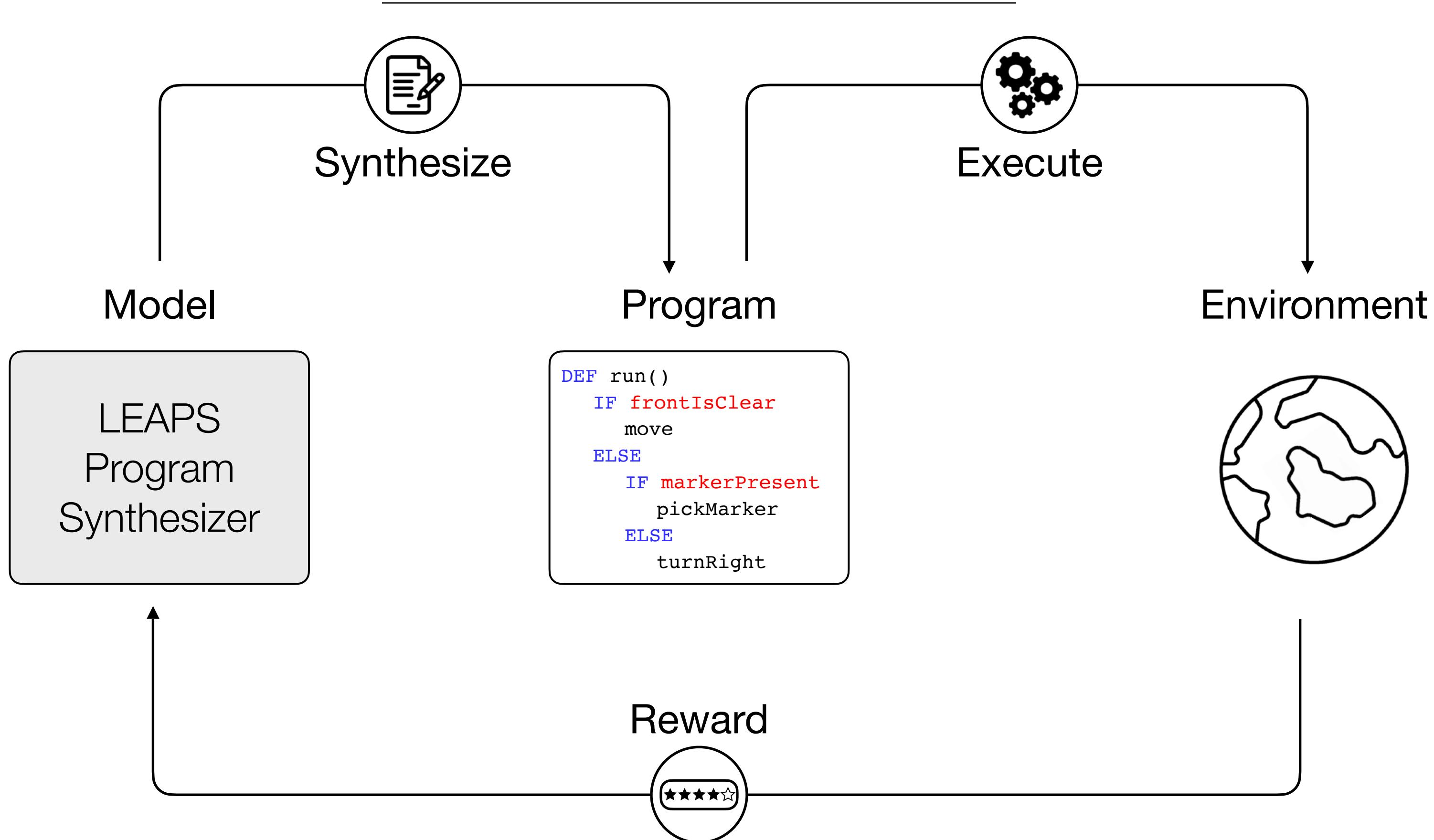
Deep Reinforcement Learning



Reinforcement Learning via Synthesizing Programs



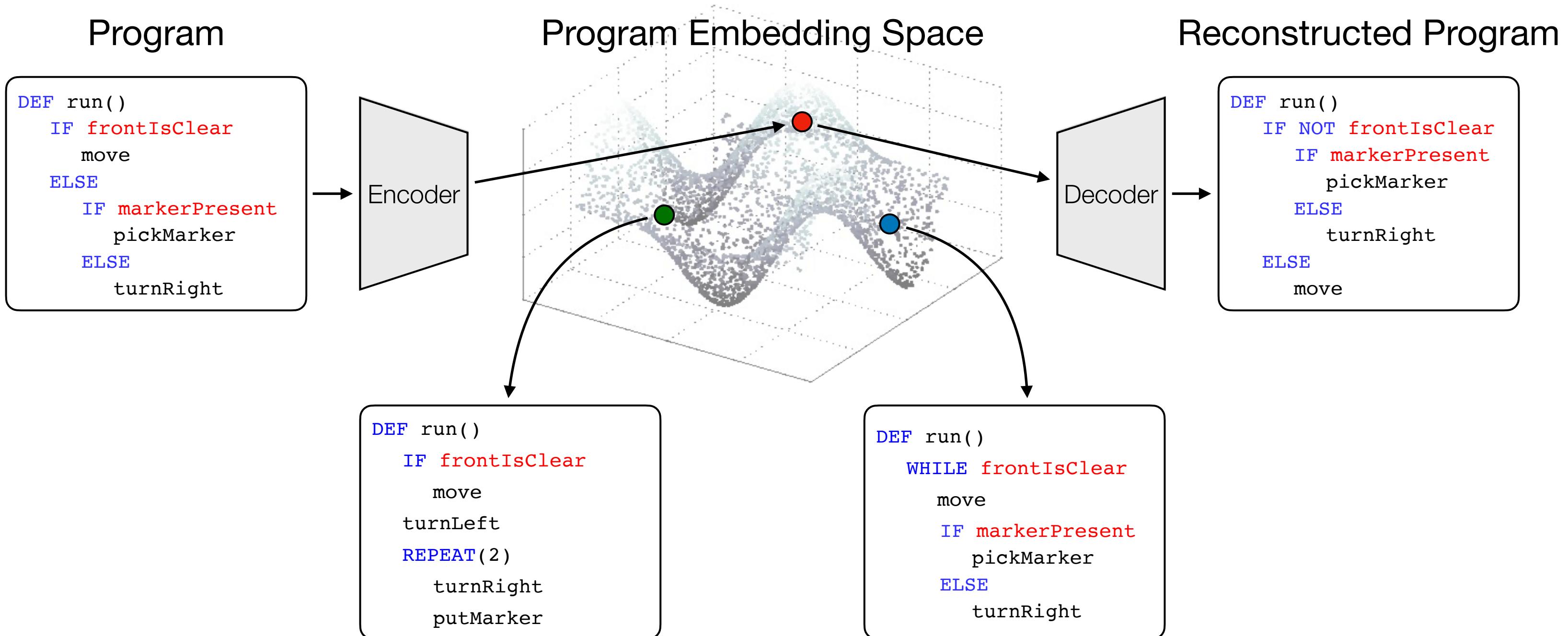
Reinforcement Learning via Synthesizing Programs



LEAPS: Learning Embeddings for Latent Program Synthesis

Stage 1 Learn a program embedding space from randomly generated programs

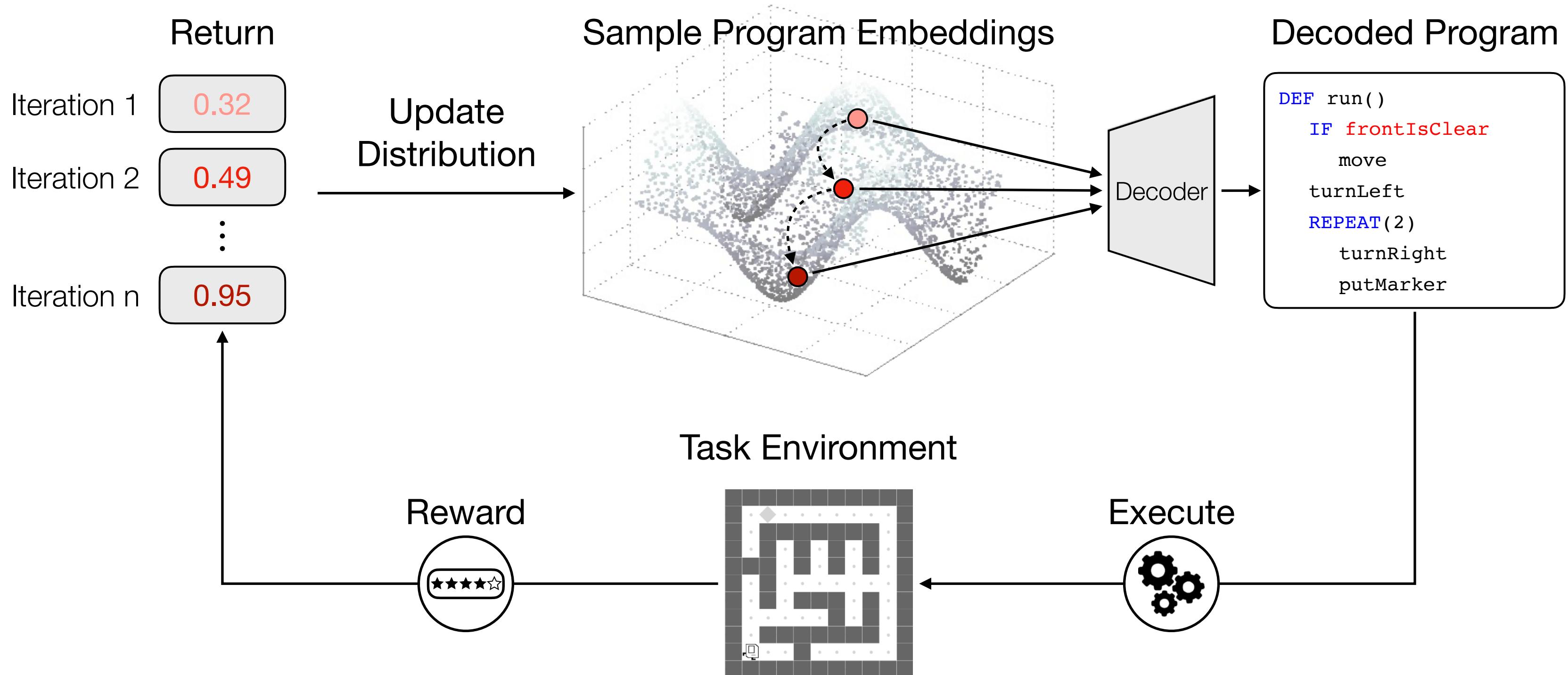
Goal Learn the **grammar** and the **environment dynamics**



LEAPS: Learning Embeddings for Latent Program Synthesis

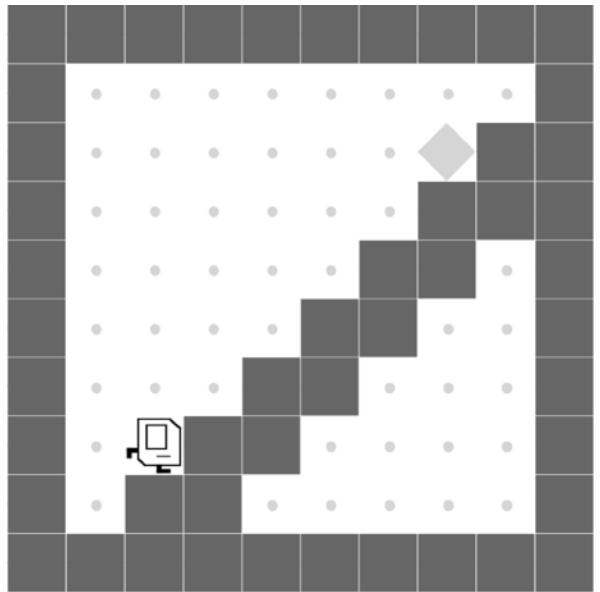
Stage 2 Search for a task-solving program using the cross-entropy method (CEM)

Goal Optimize the **task performance**

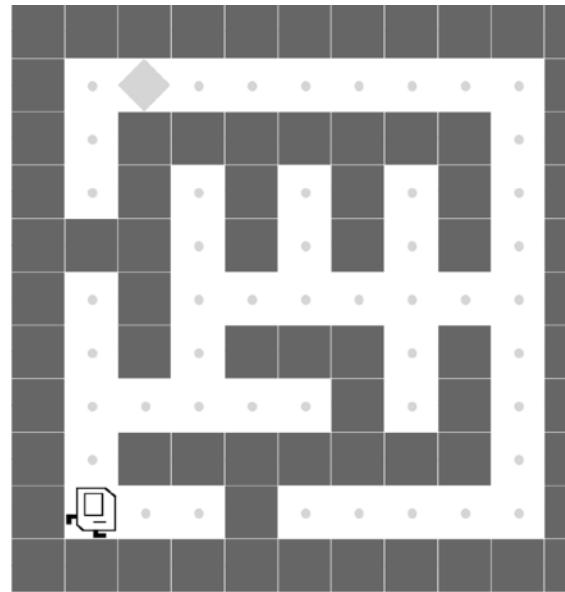


Karel Tasks

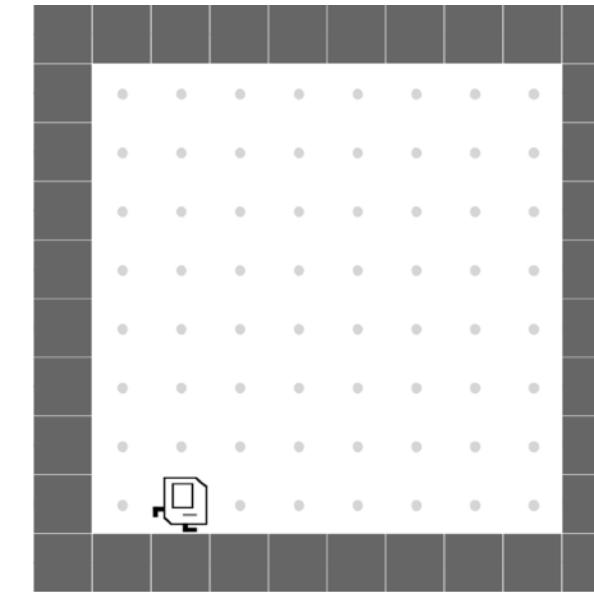
StairClimber



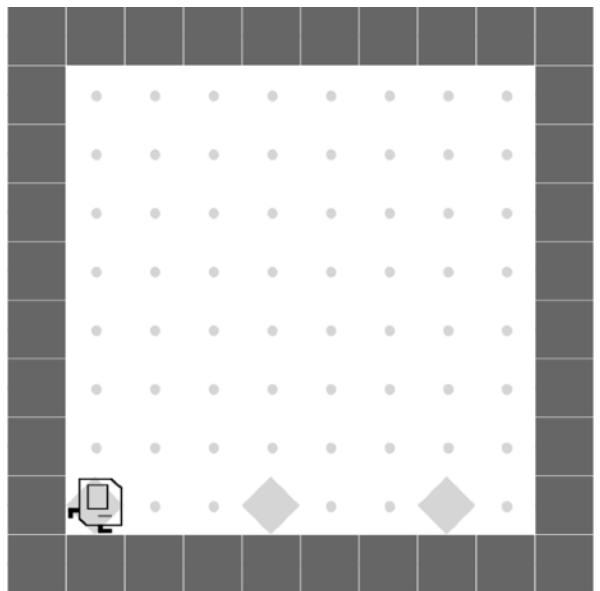
Maze



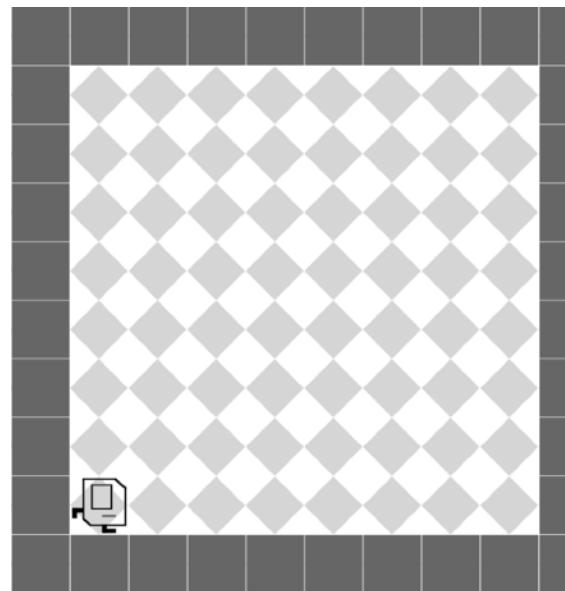
FourCorners



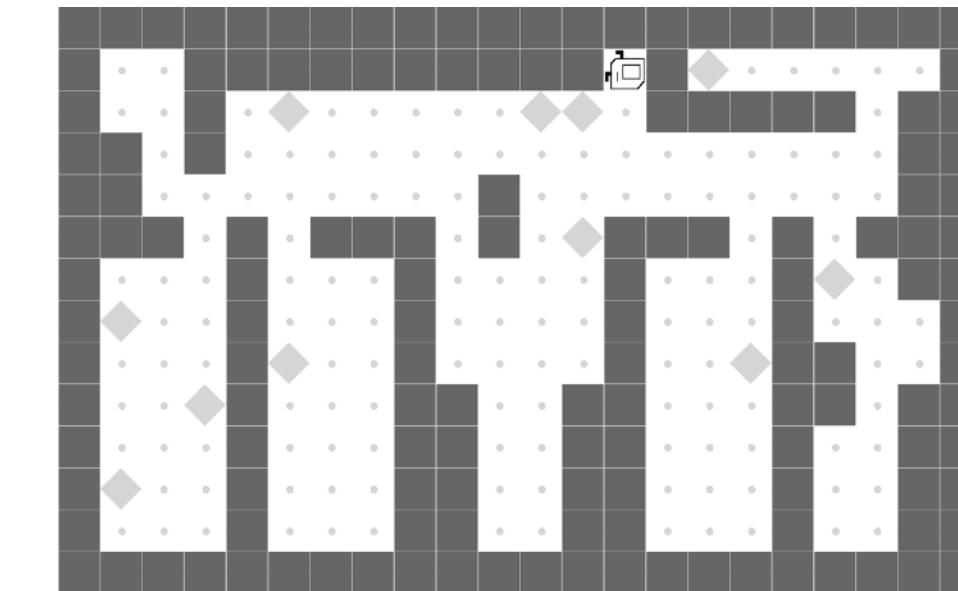
TopOff



Harvester

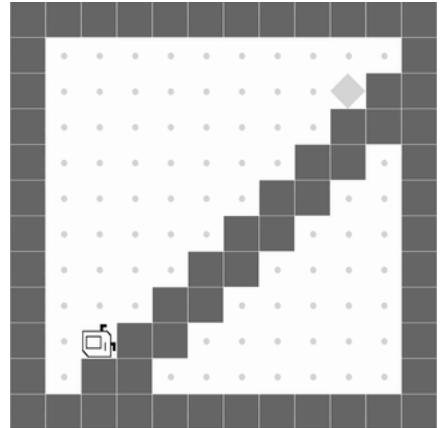


CleanHouse

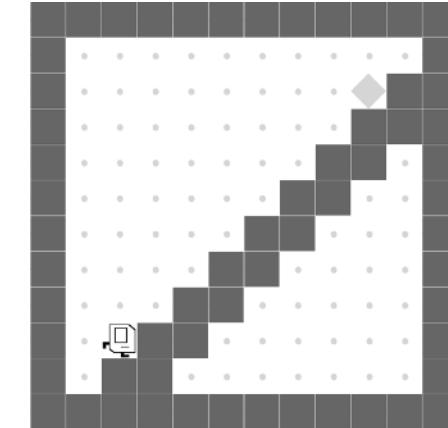


Qualitative Results

StairClimber

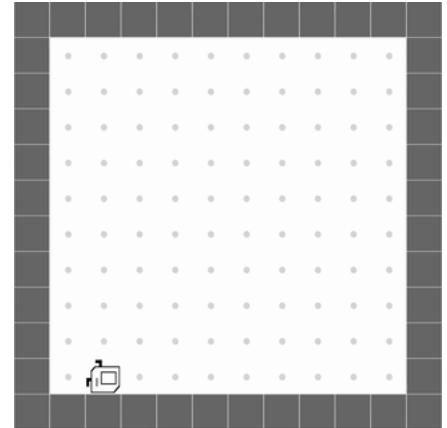


Deep RL

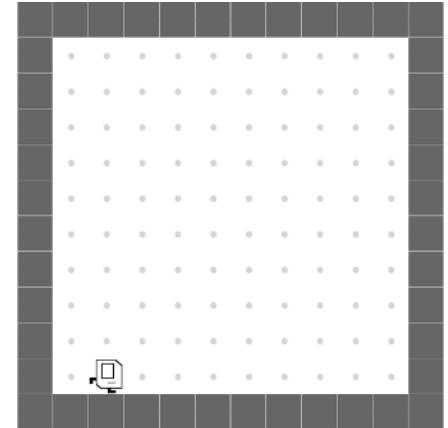


LEAPS

FourCorners

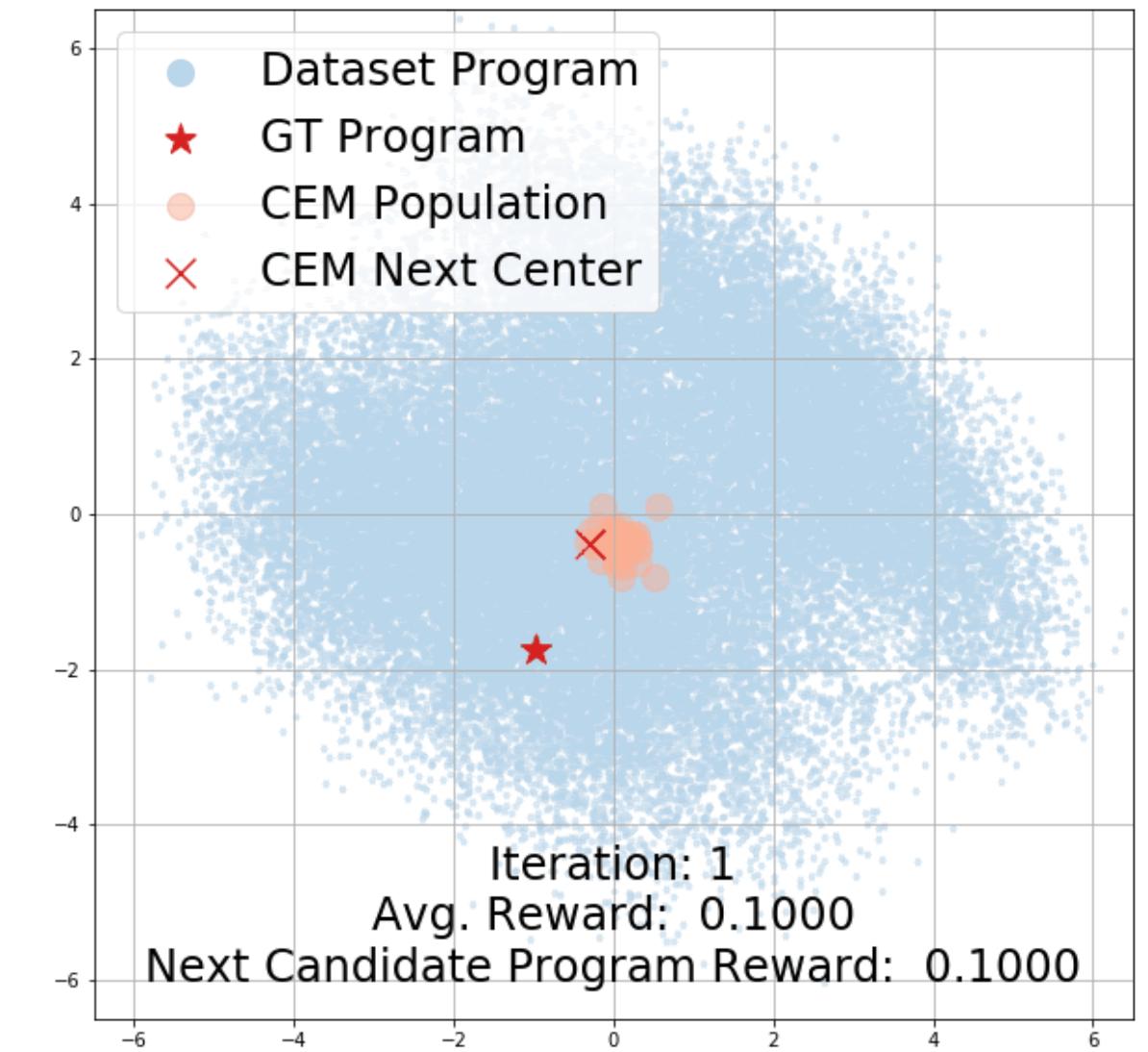


Deep RL

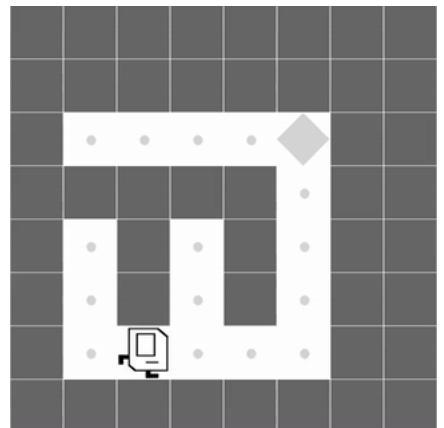


LEAPS

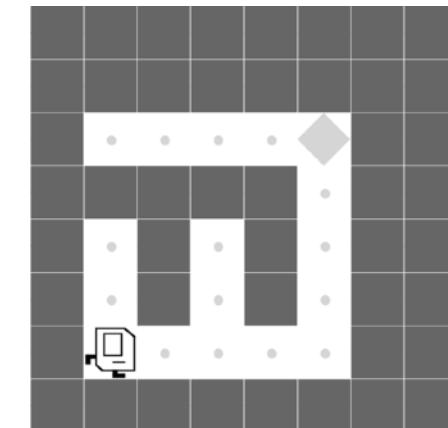
CEM trajectory Visualization



Maze

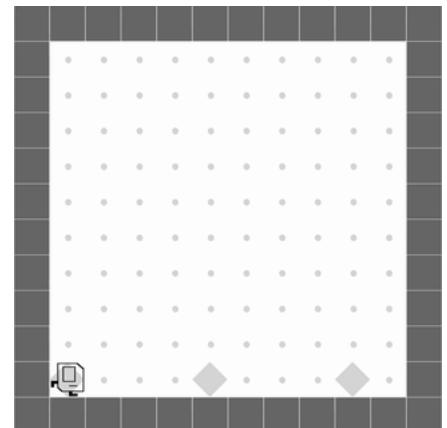


Deep RL

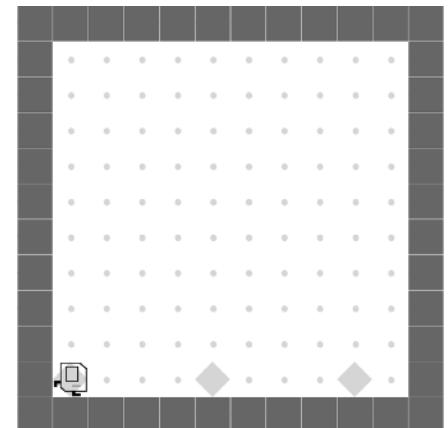


LEAPS

TopOff



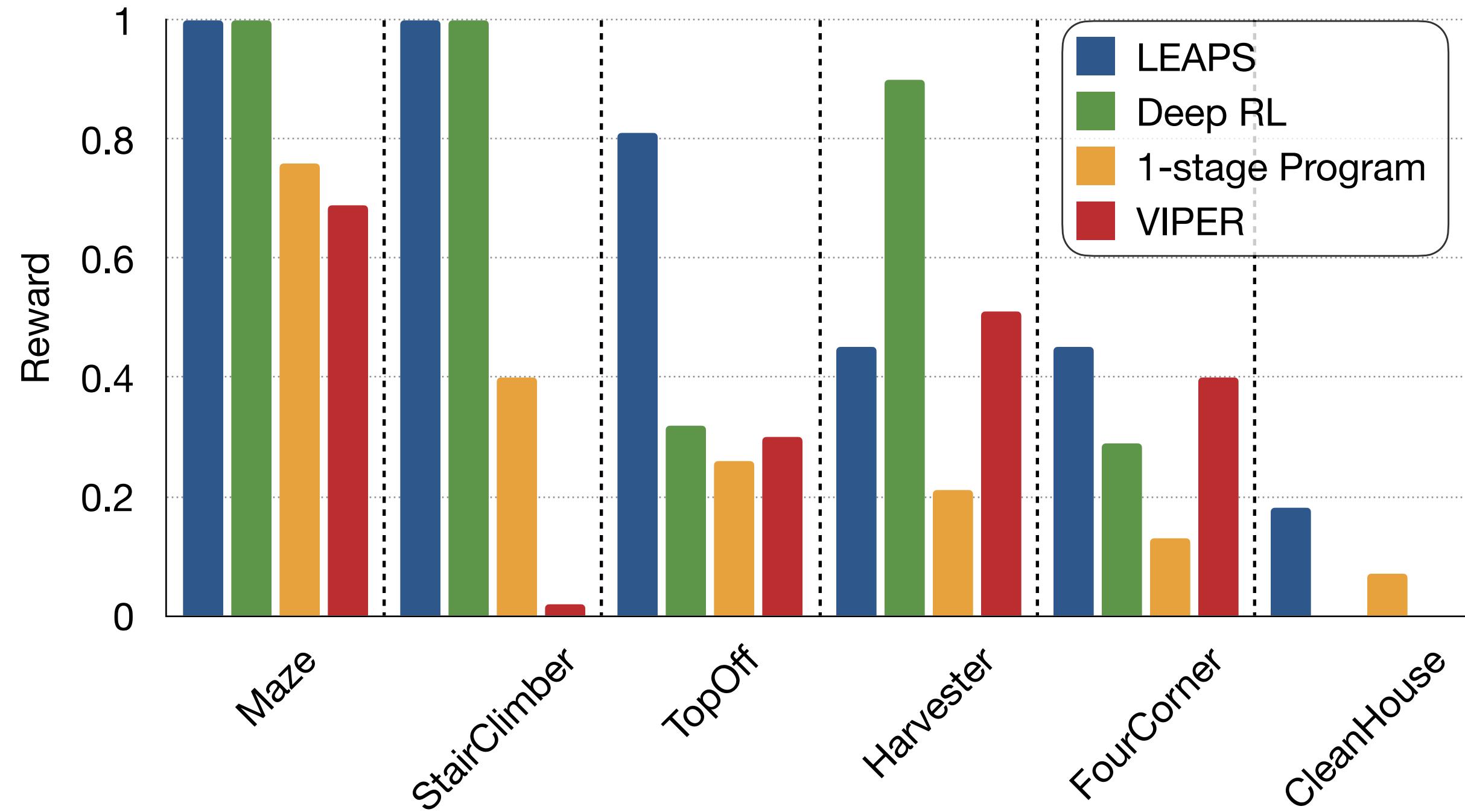
Deep RL



LEAPS

Goal: Search for a StairClimber program
in the learned program embedding space

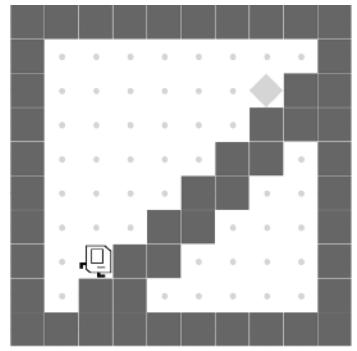
Quantitative Results



LEAPS Zero-shot Generalization

Learning on 8 x 8

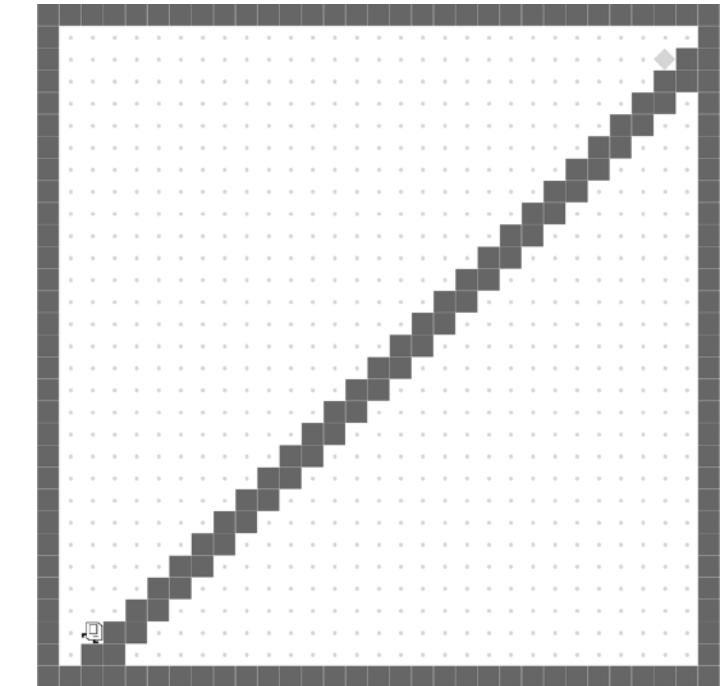
StairClimber



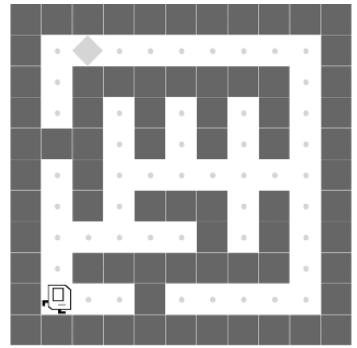
LEAPS
Program
Synthesizer

```
DEF run()
  while noMarkersPresent()
    turnRight
    move
  while rightIsClear()
    turnLeft
```

Evaluation on 100 x 100

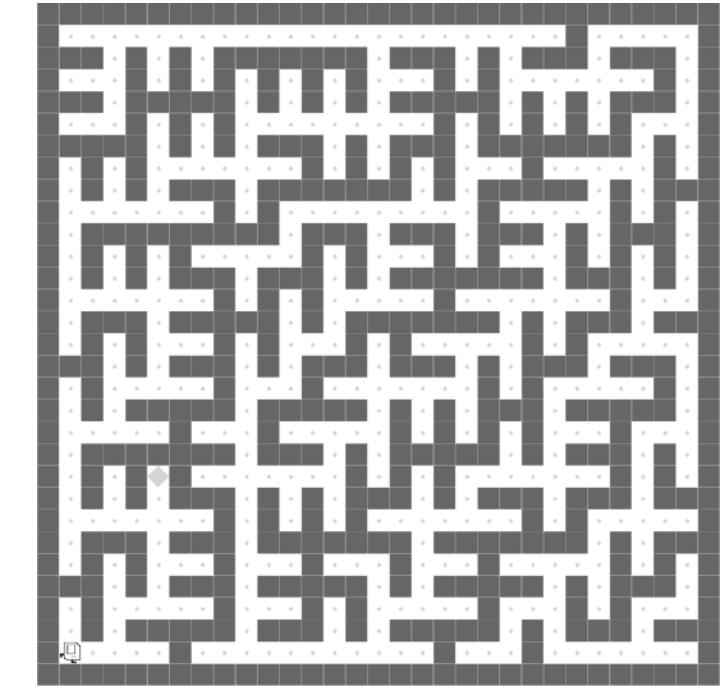


Maze

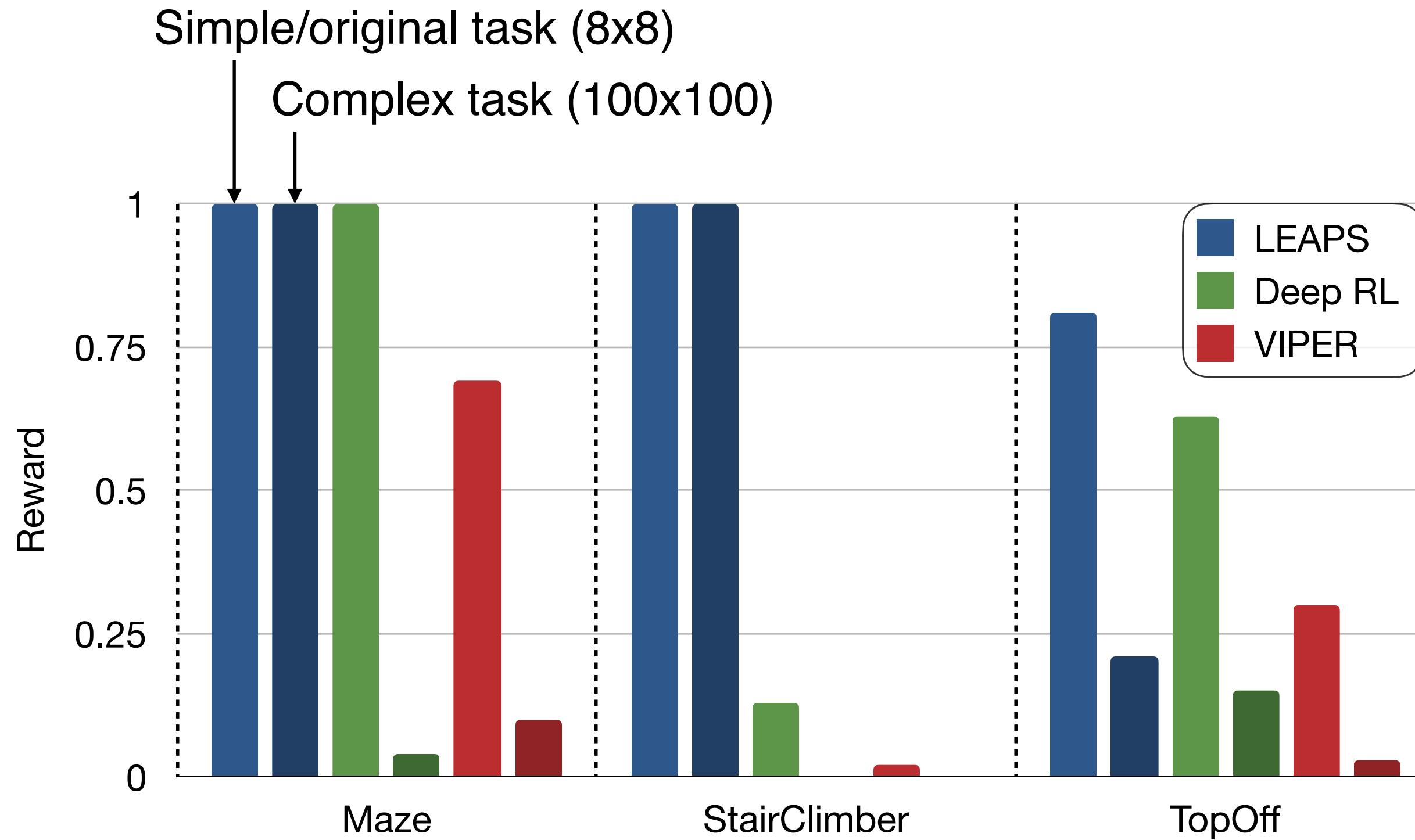


LEAPS
Program
Synthesizer

```
DEF run()
  if frontIsClear()
    turnLeft
  while noMarkersPresent()
    turnRight
    move
```

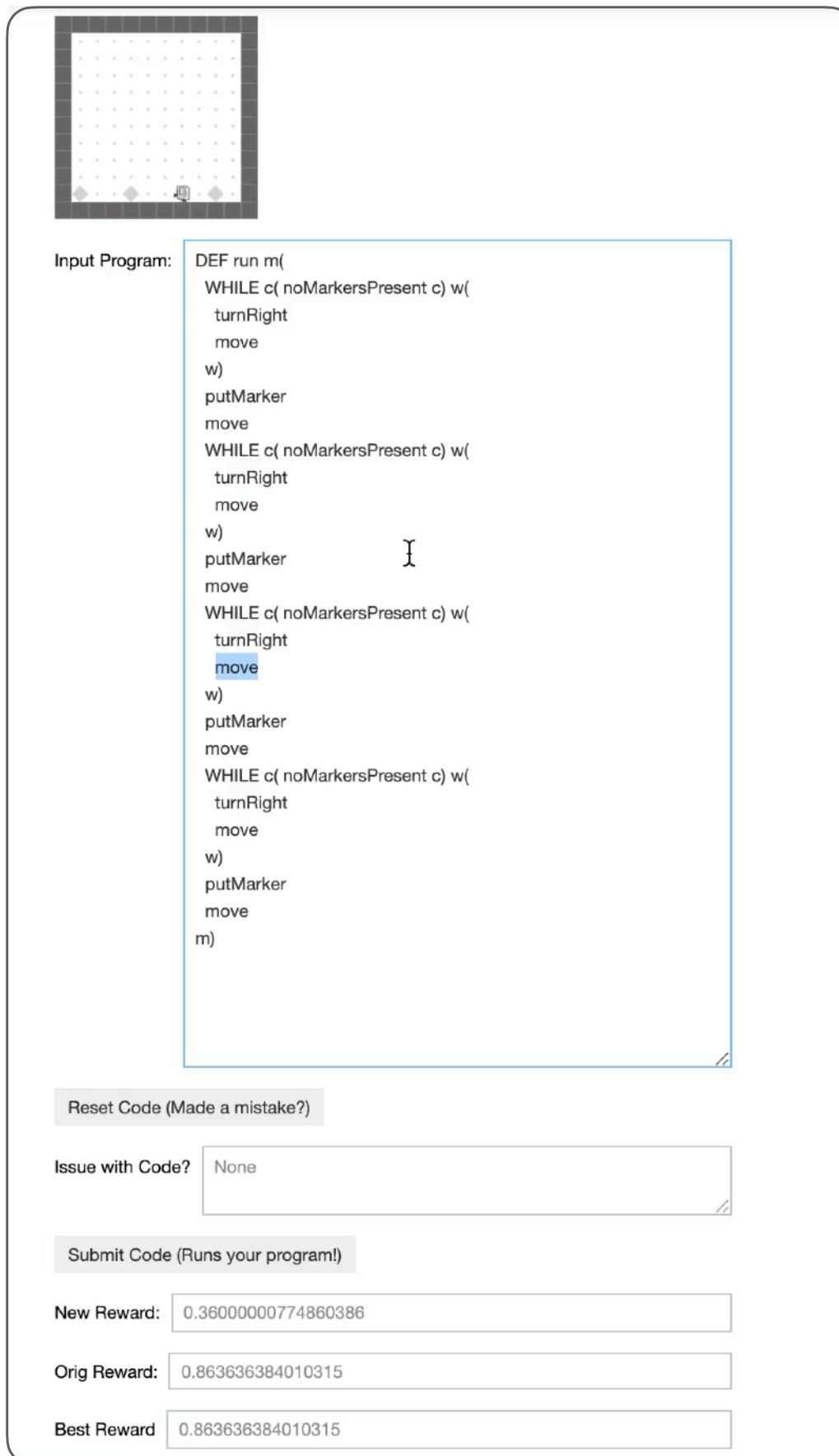


Experimental Results - Zero-shot Generalization

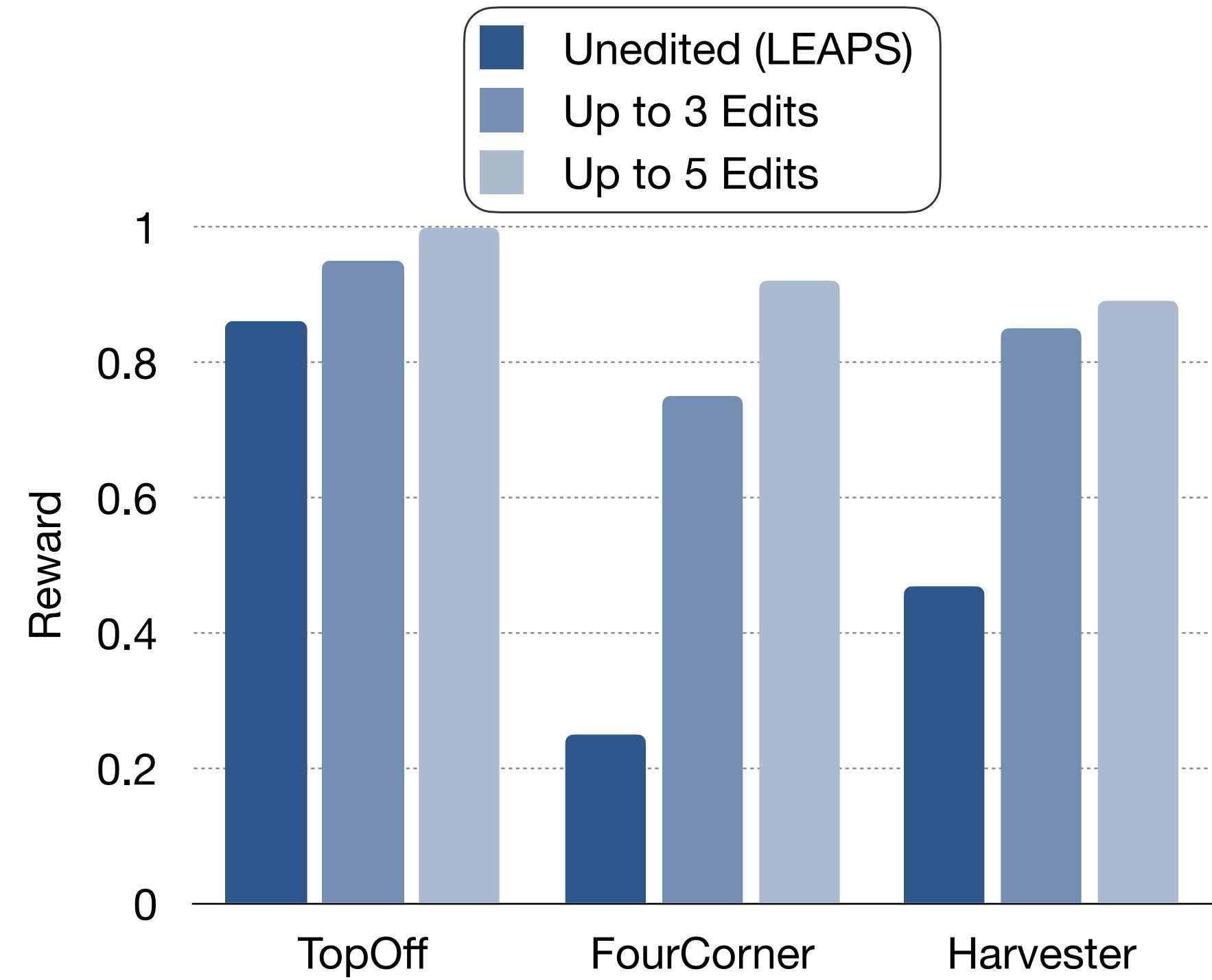


Interpretability & Interactivity

Interactive Debugging Interface

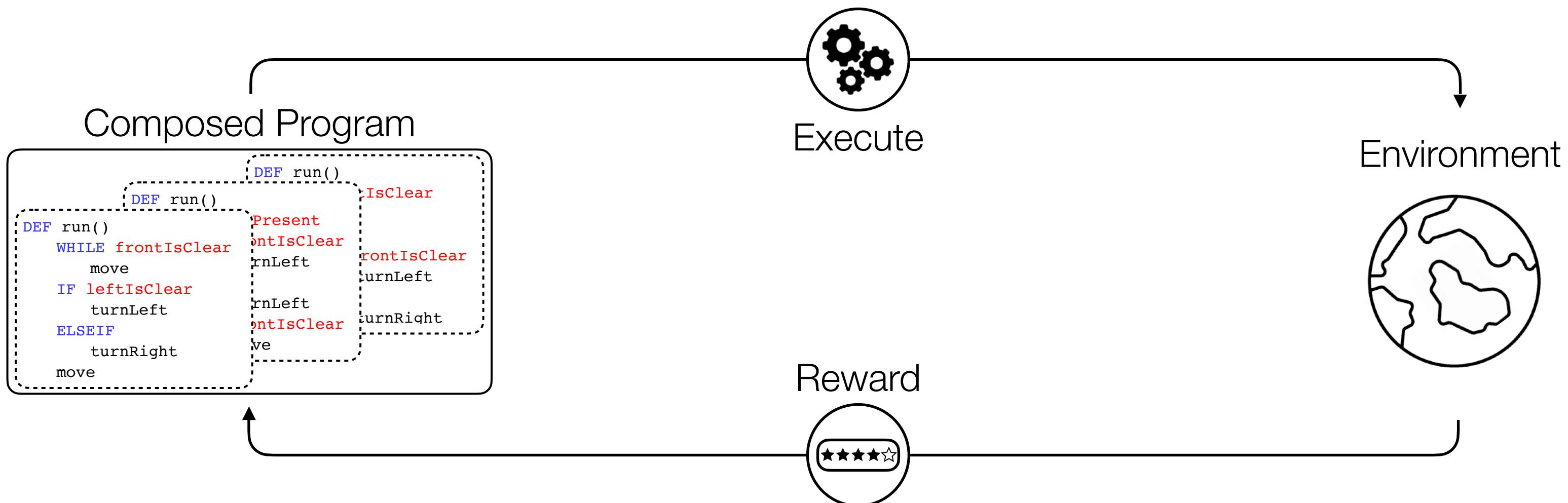


Performance Improvement



Hierarchical Programmatic Reinforcement Learning via Learning to Compose Programs

ICML 2023



Guan-Ting Liu*



En-Pei Hu*



Pu-Jen Cheng



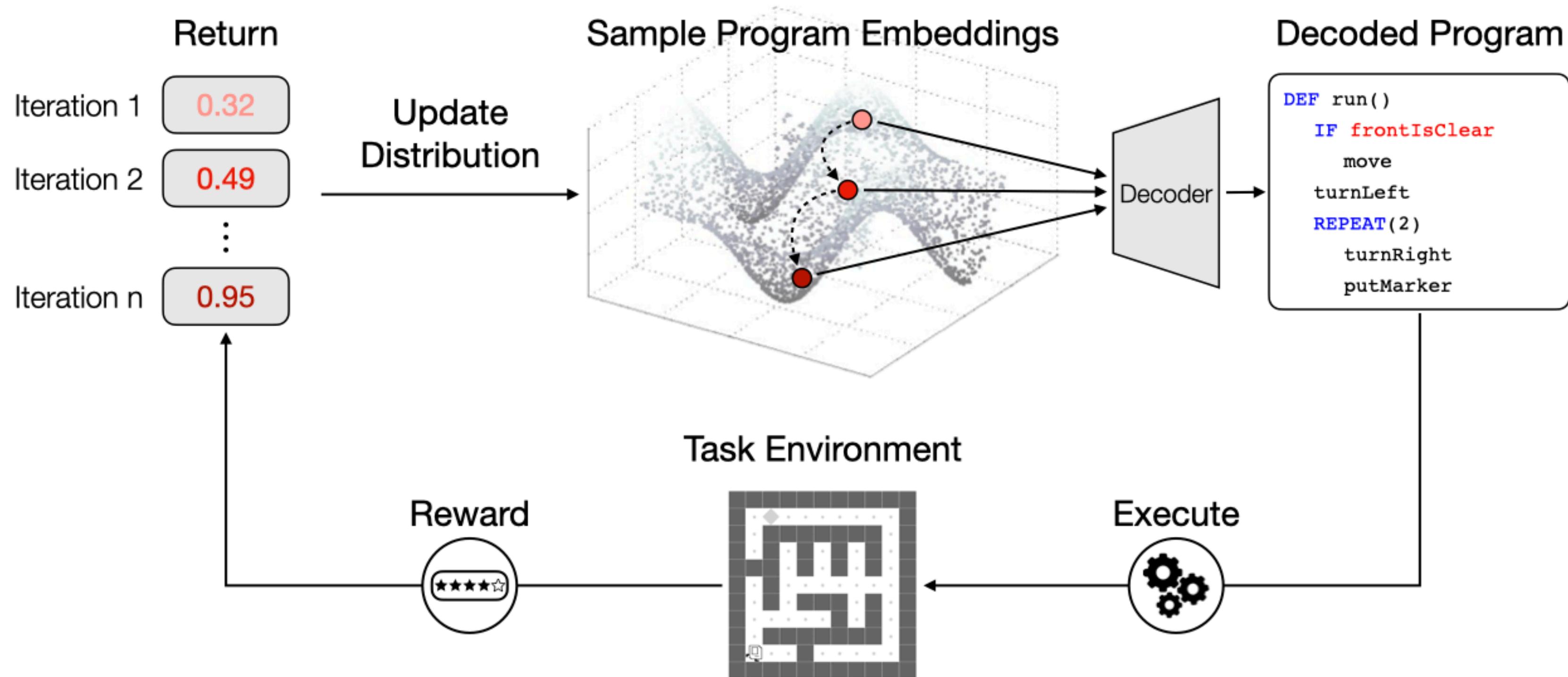
Hung-Yi Lee



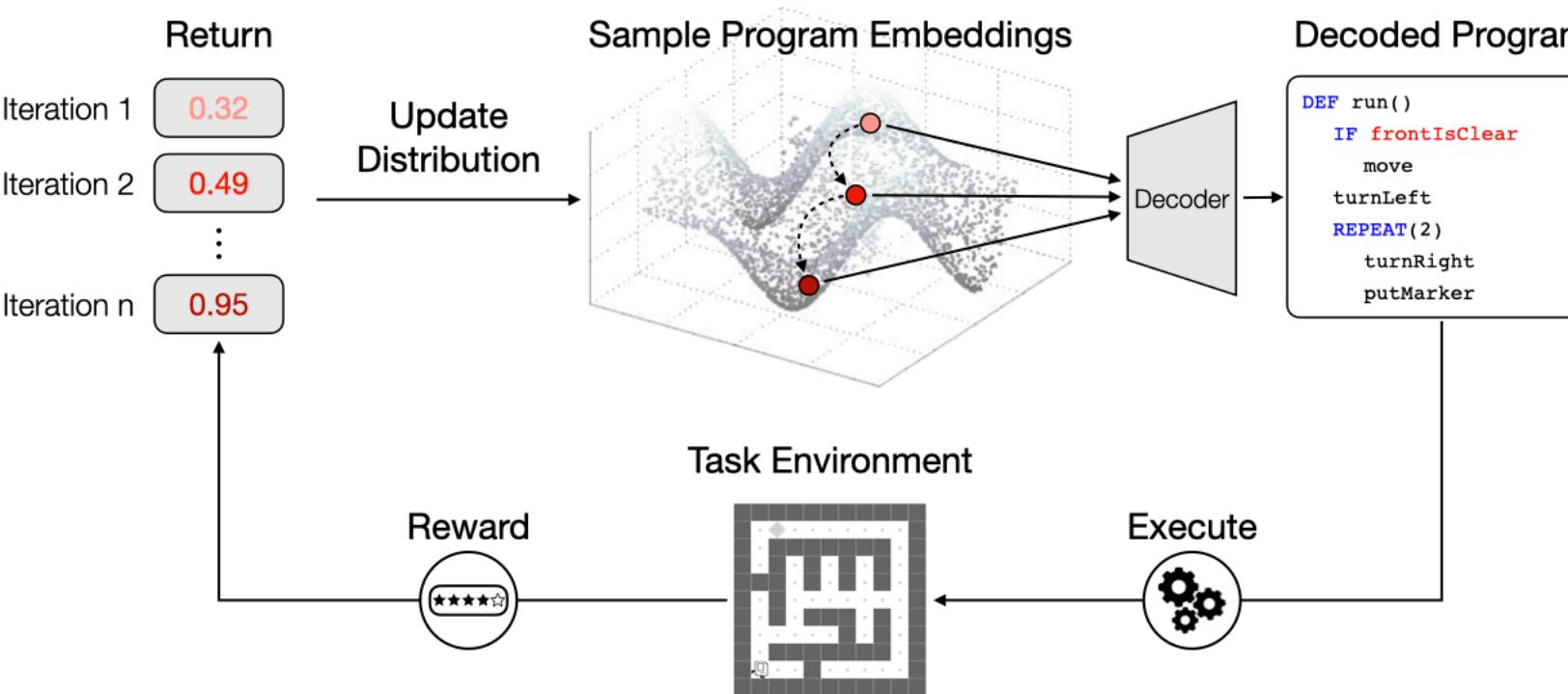
Shao-Hua Sun

LEAPS: Learning Embeddings for Latent Program Synthesis

Stage 2 Searching for a task-solving program using the cross-entropy method



Stage 2 Searching for a task-solving program using the cross-entropy method



Limited program distribution

Search in the program embedding space spanned
by the dataset programs



Cannot synthesize longer or
more complex programs

Poor credit assignment

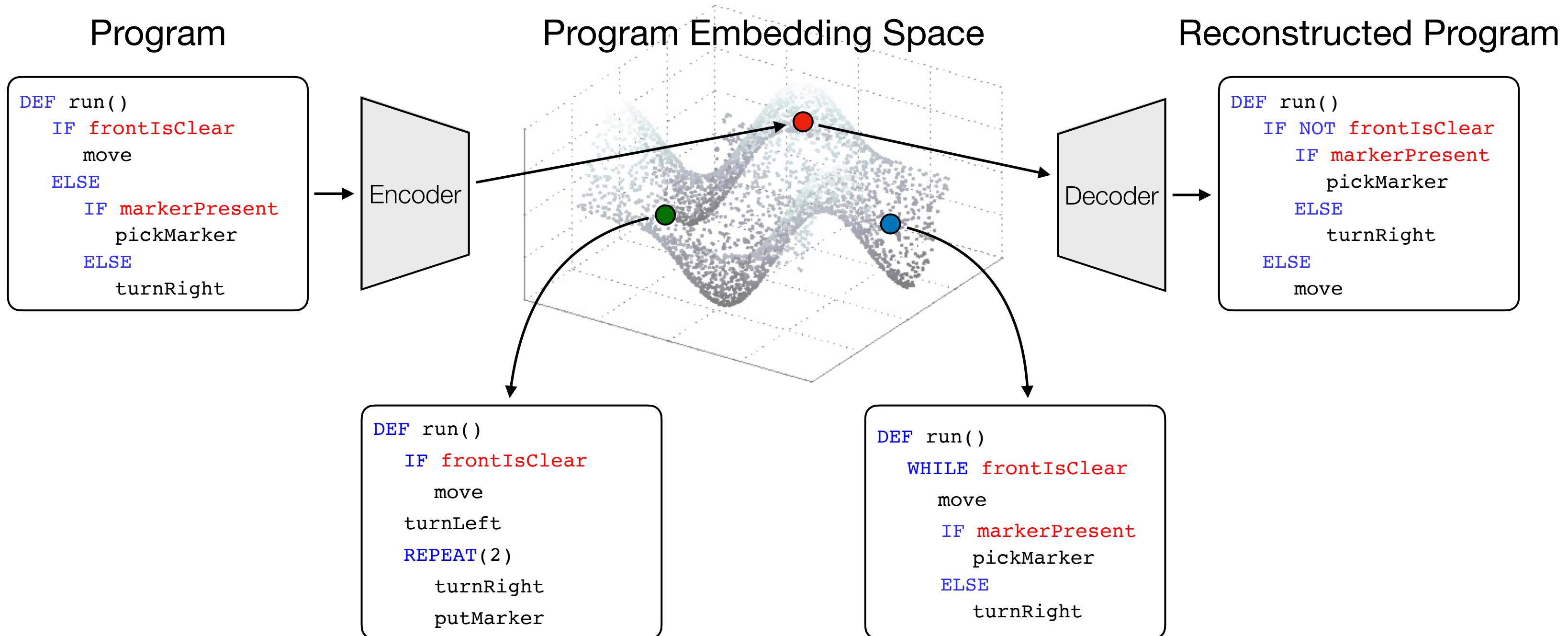
Evaluate each candidate program solely based on
the **cumulative return** of its execution trace



Cannot accurately attribute rewards to
corresponding program parts

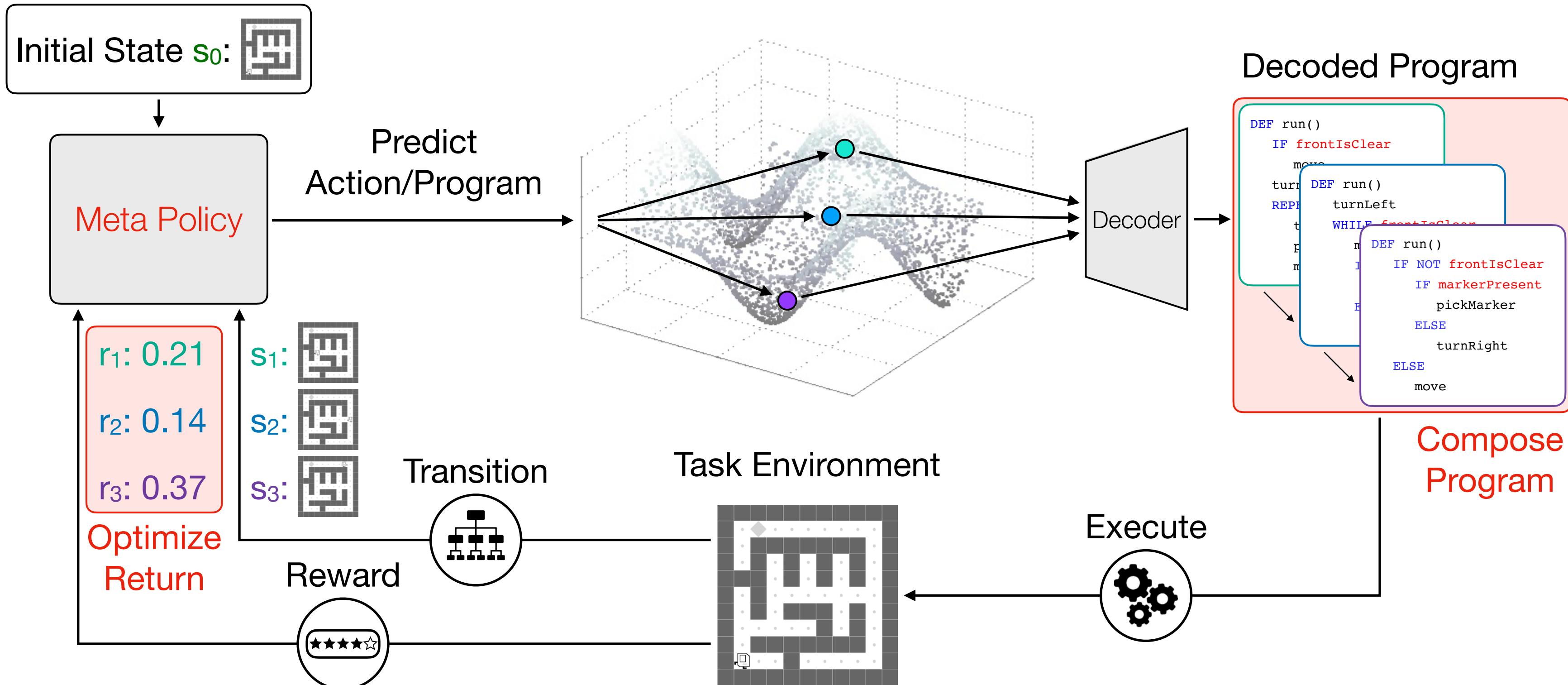
HPRL: Hierarchical Programmatic Reinforcement Learning

Stage 1 Learning a **compressed** program embedding space from randomly generated programs

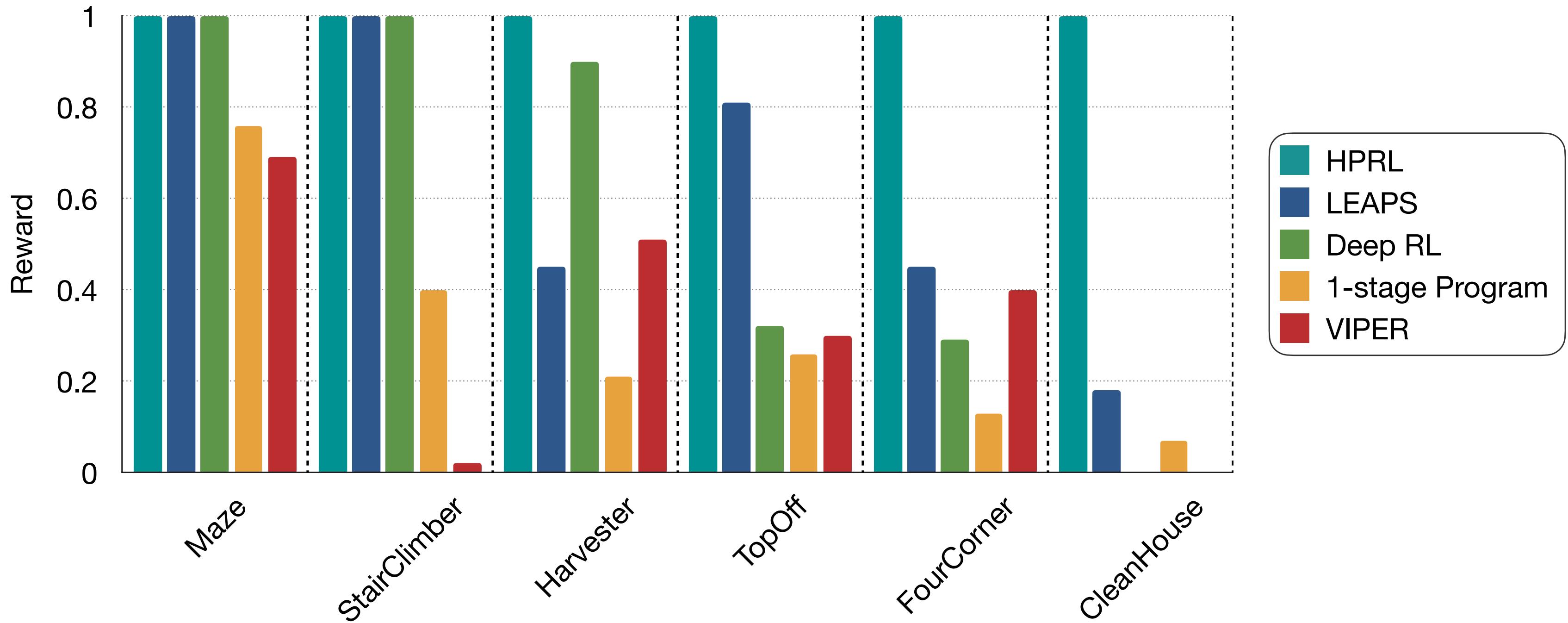


HPRL: Hierarchical Programmatic Reinforcement Learning

Stage 2 Learning a meta policy to produce a series of programs (i.e., predict a series of actions) to yield a composed task-solving program

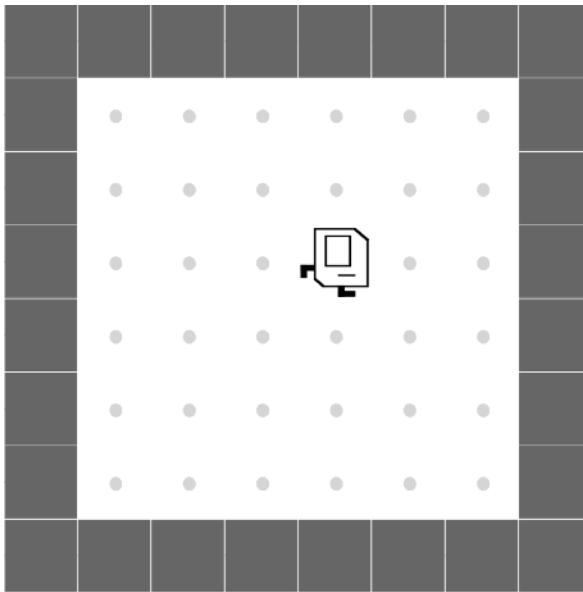


Quantitative Results - Karel Tasks

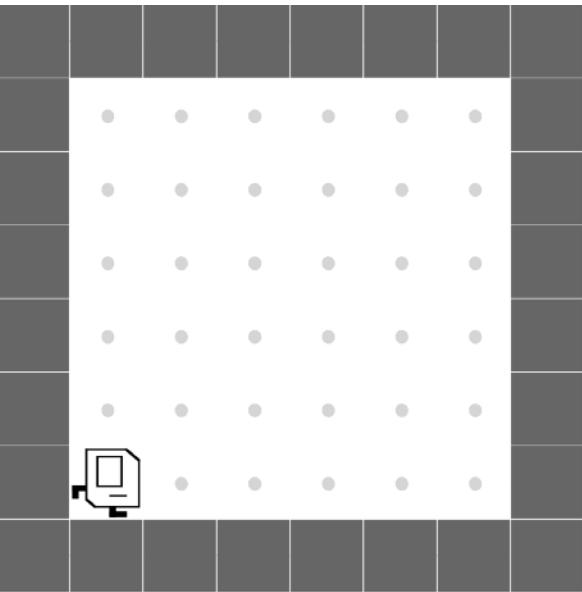


Karel-Hard Tasks

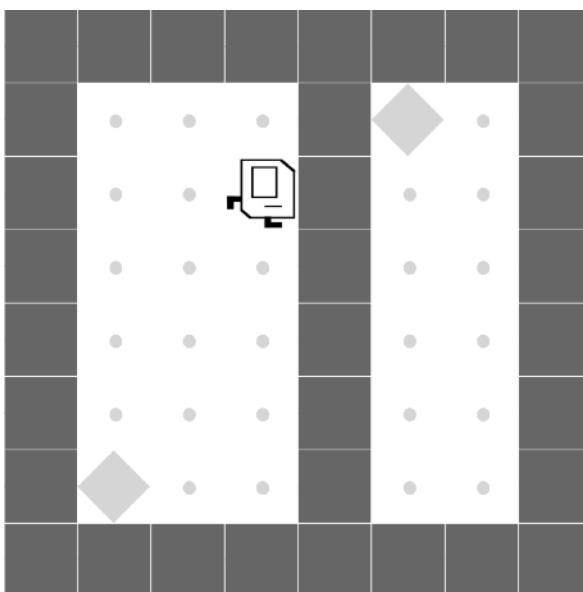
OneStroke



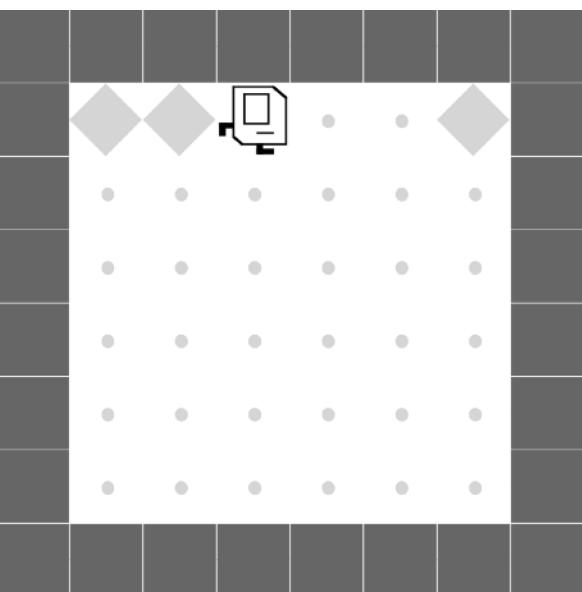
Seeder



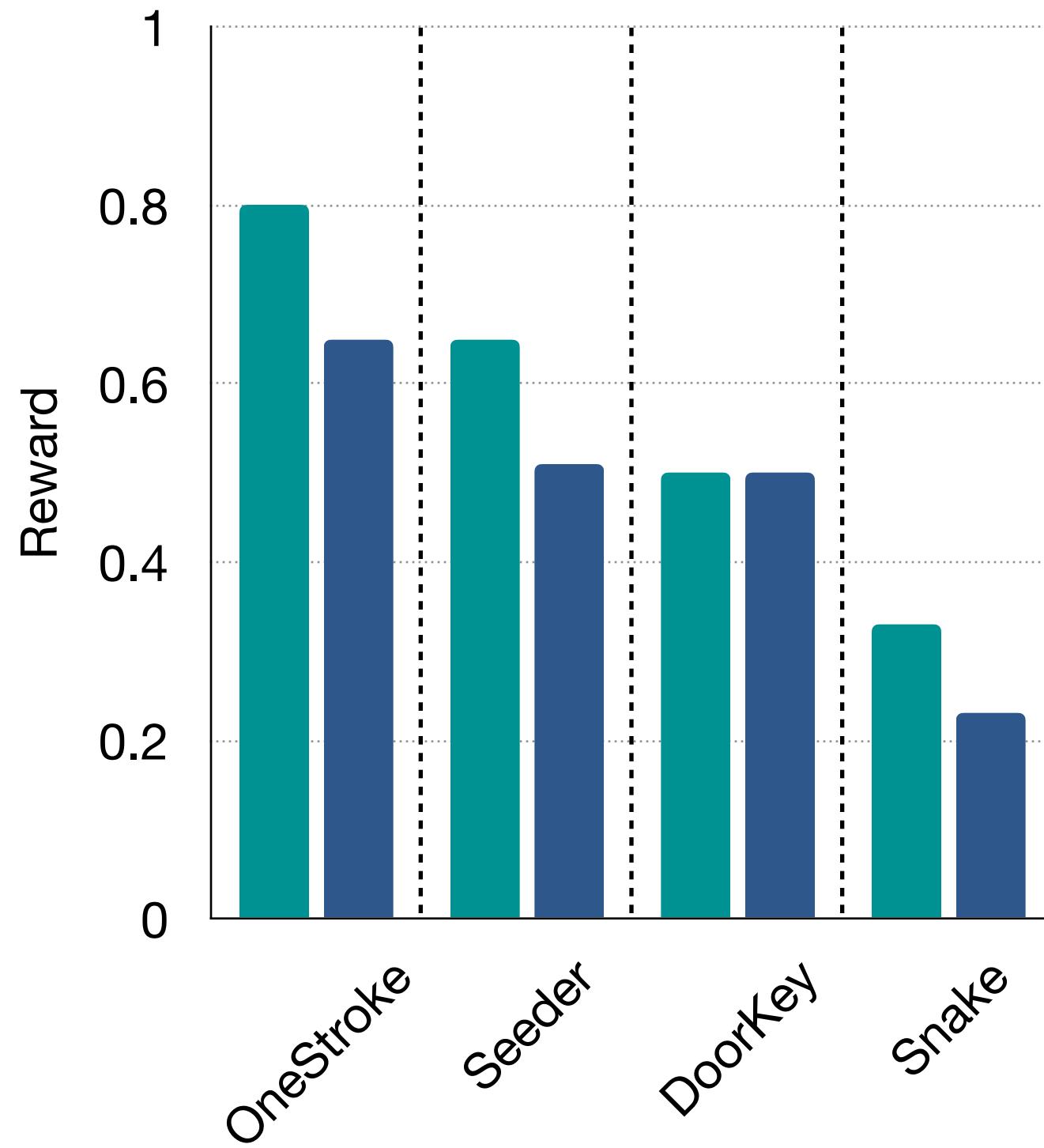
DoorKey



Snake



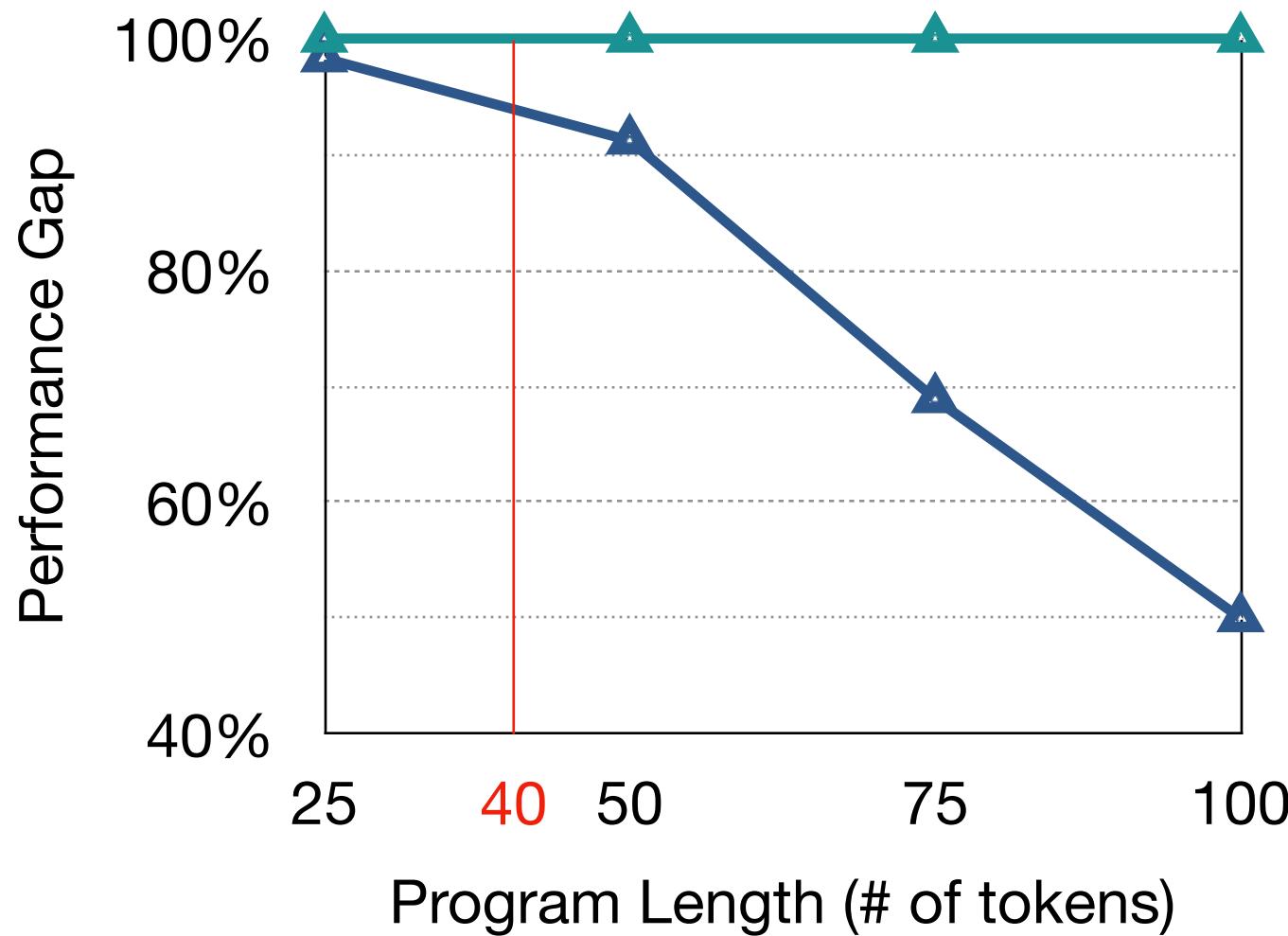
HPRL LEAPS



Additional Experiments

Limited program distribution

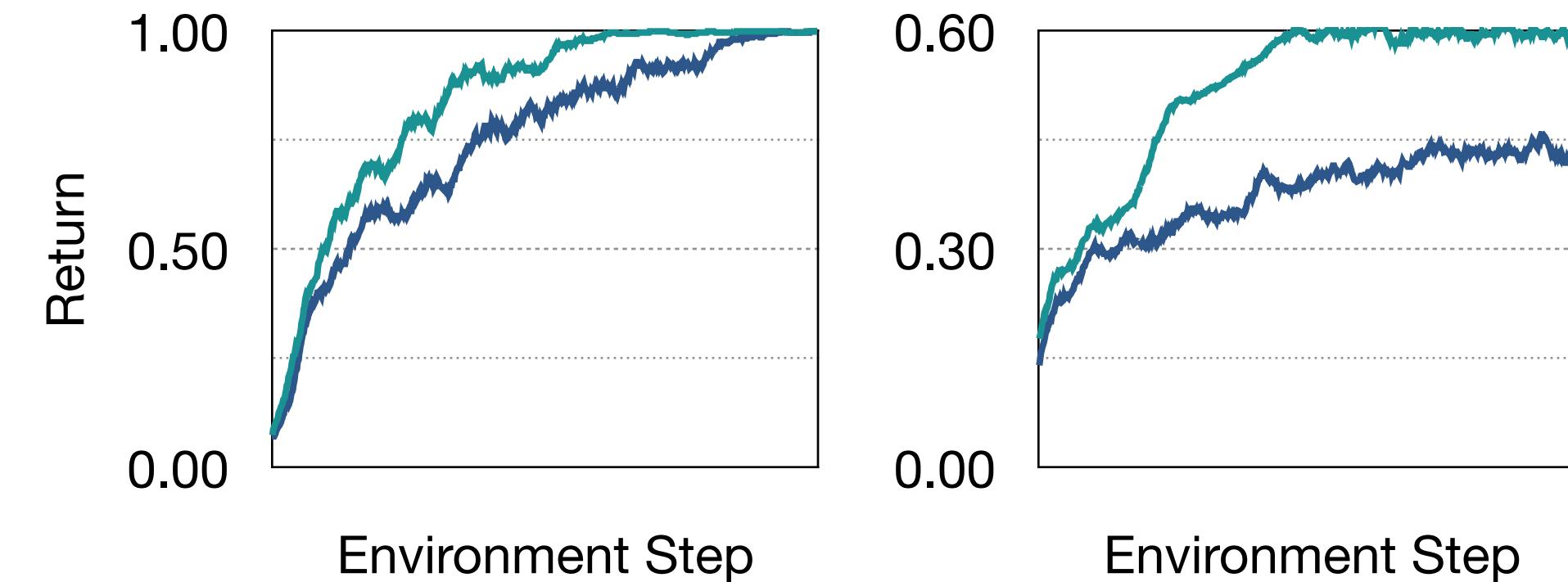
Synthesize out-of-distributionally long programs



Poor credit assignment

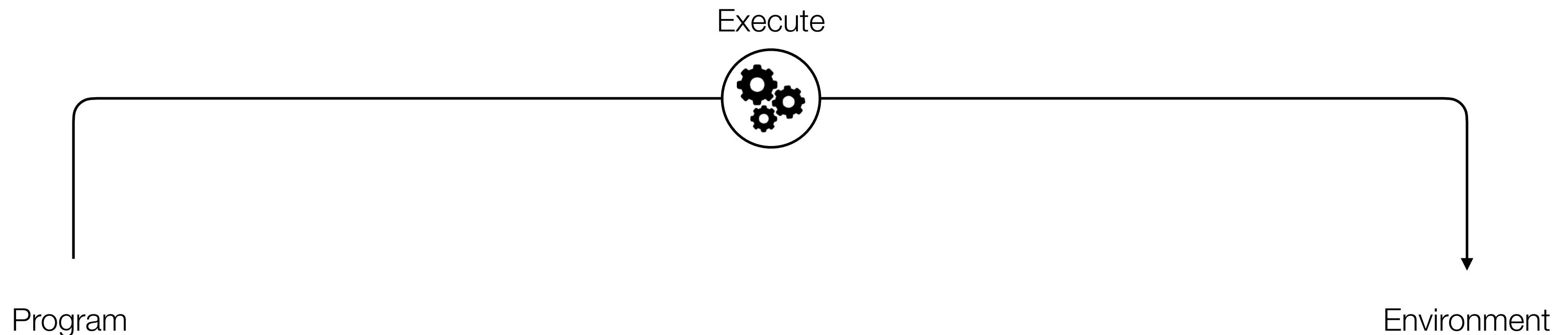
Learning from episodic reward

- Dense: Reward each subprogram based on its execution trace
- Episodic: Reward the entire composed program at the end



- **HPRL** can synthesize programs longer than the dataset programs (< 40 tokens) better than **LEAPS**

- The hierarchical design of **HPRL** allows for better credit assignment with dense rewards, facilitating the learning progress

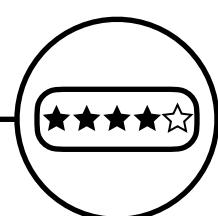
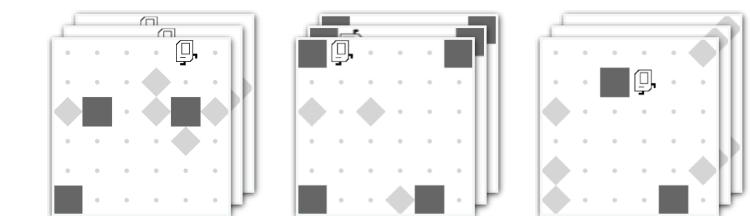


Takeaways

Program Synthesis \times Reinforcement Learning

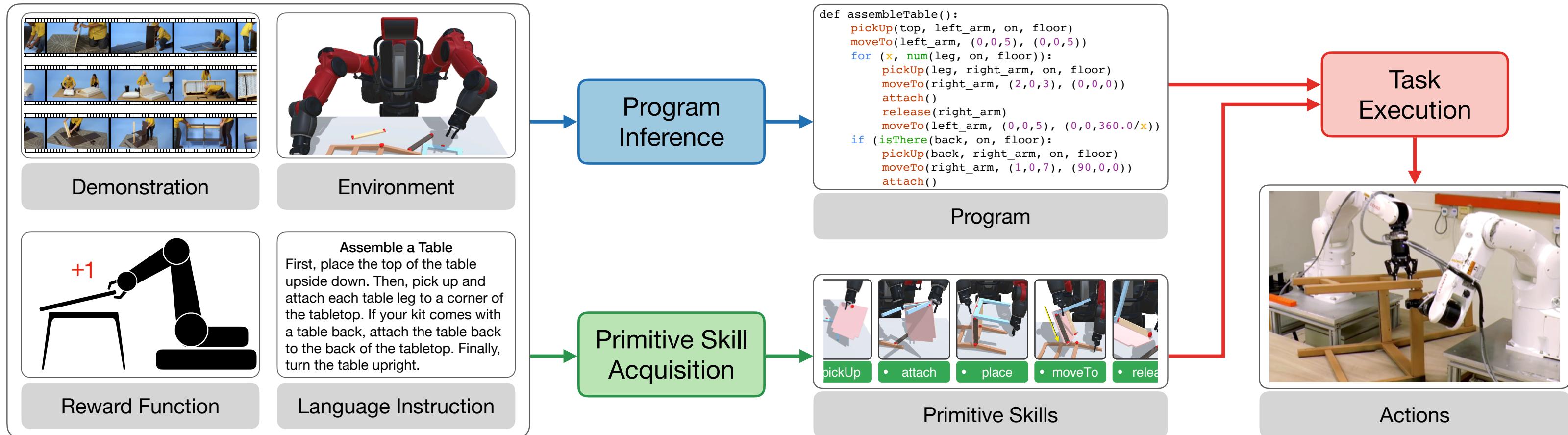
= Interpretable and Generalizable Policies

Demonstrations



Reward

Program-Guided Robot Learning



Key idea

- Represent robot behaviors using programs based on pre-defined and acquired **primitive skills**
- Decouple learning a skill as performing **program inference** and **task execution**

Program

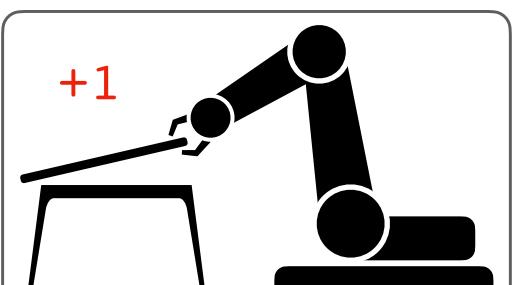
- Human interpretable and machine executable
- Structured for generalization

Program-Guided Robot Learning

Task Specification



Demonstration

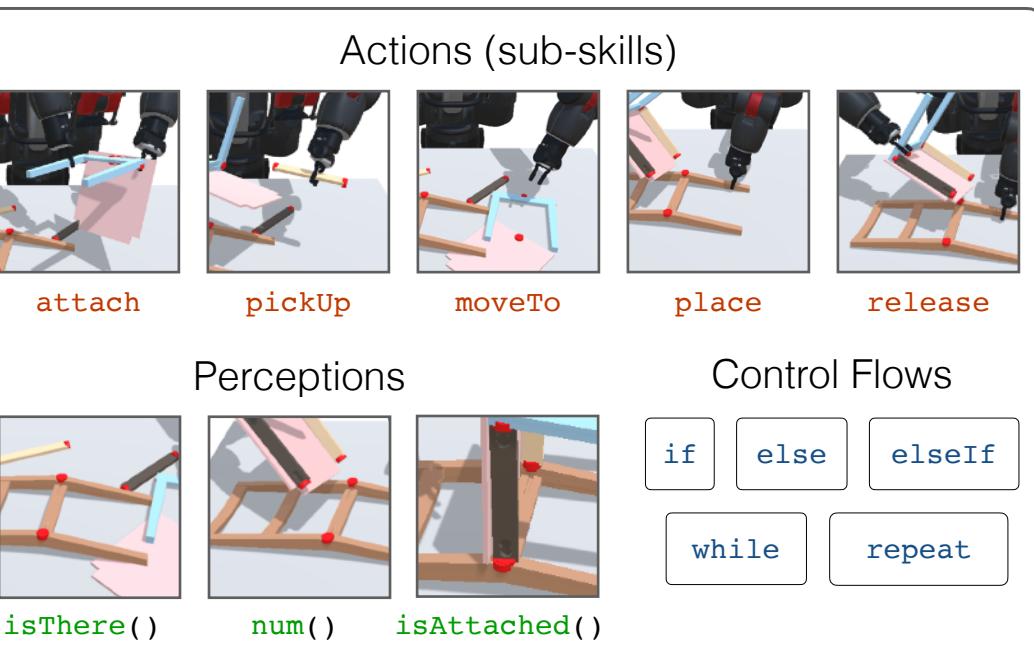


Reward Function

Assemble a Table
First, place the top of the table upside down. Then, pick up and attach each table leg to a corner of the tabletop. If your kit comes with a table back, attach the table back to the back of the tabletop. Finally, turn the table upright.

Language Instruction

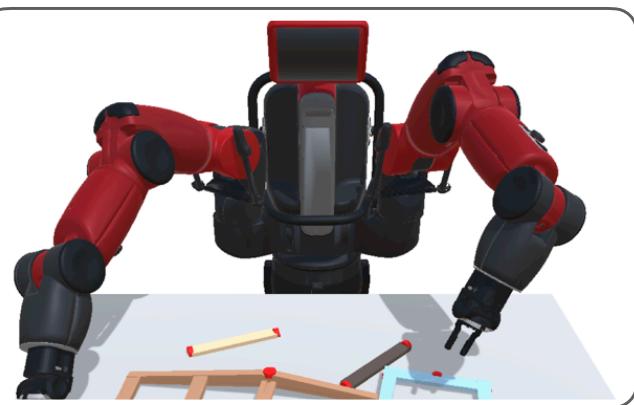
Domain-Specific Language



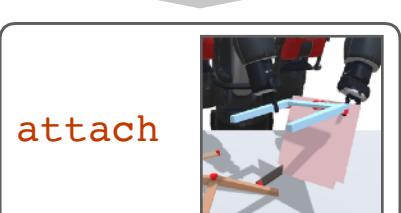
Program

```
def assembleTable():
    pickUp(top, left_arm, on, floor)
    moveTo(left_arm, (0,0,5), (0,0,5))
    for (x, num(leg, on, floor)):
        pickUp(leg, right_arm, on, floor)
        moveTo(right_arm, (2,0,3), (0,0,0))
        attach()
        release(right_arm)
        moveTo(left_arm, (0,0,5), (0,0,360.0/x))
    if (isThere(back, on, floor)):
        pickUp(back, right_arm, on, floor)
        moveTo(right_arm, (1,0,7), (90,0,0))
        attach()
```

Observation



High-Level Plan



Low-Level Execution



Joint u torque

```
[-2.09531783e-19 2.72130735e-05 6.14480786e-22 -3.45474715e-06
 7.42993721e-06 -1.40711141e-04 -3.04253586e-04 -2.07559344e-04
 8.50646247e-05 -3.45474715e-06 7.42993721e-06 -1.40711141e-04
 -3.04253586e-04 -2.07559344e-04 -8.50646247e-05 1.11317030e-04
 -7.03465386e-05 -2.22862221e-05 -1.11317030e-04 7.03465386e-05
 -2.22862221e-05 -2.09531783e-19 2.72130735e-05 6.14480786e-22
 -3.45474715e-06 7.42993721e-06 -1.40711141e-04 -3.04253586e-04
 -2.07559344e-04 8.50646247e-05 -3.45474715e-06 7.42993721e-06
 -1.40711141e-04 -3.04253586e-04 -2.07559344e-04 -8.50646247e-05
 1.11317030e-04 -7.03465386e-05 -2.22862221e-05 -1.11317030e-04]
```

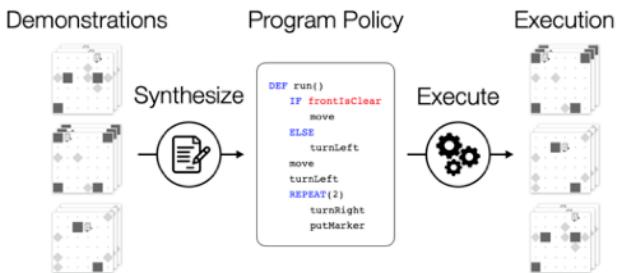
Joint v torque

```
[-2.09531783e-19 2.72130735e-05 6.14480786e-22 -3.45474715e-06
 7.42993721e-06 -1.40711141e-04 -3.04253586e-04 -2.07559344e-04
 8.50646247e-05 -3.45474715e-06 7.42993721e-06 -1.40711141e-04
 -3.04253586e-04 -2.07559344e-04 -8.50646247e-05 1.11317030e-04
 -7.03465386e-05 -2.22862221e-05 -1.11317030e-04 7.03465386e-05
 -2.22862221e-05 -2.09531783e-19 2.72130735e-05 6.14480786e-22
 -3.45474715e-06 7.42993721e-06 -1.40711141e-04 -3.04253586e-04
 -2.07559344e-04 8.50646247e-05 -3.45474715e-06 7.42993721e-06
 -1.40711141e-04 -3.04253586e-04 -2.07559344e-04 -8.50646247e-05
 1.11317030e-04 -7.03465386e-05 -2.22862221e-05 -1.11317030e-04]
```

:

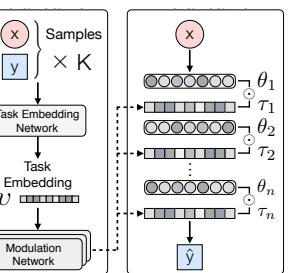
Program Inference

Neural Program Synthesis from Diverse Demonstration Videos



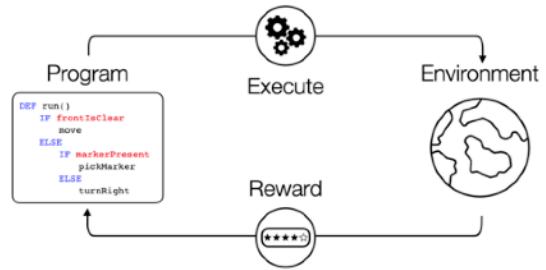
ICML 2018

Multimodal Model-Agnostic Meta-Learning via Task-Aware Modulation



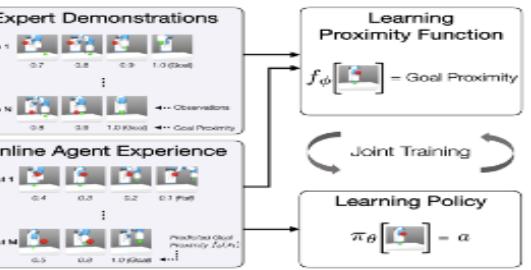
NeurIPS 2019 (Spotlight)

Learning to Synthesize Programs as Interpretable and Generalizable Policies



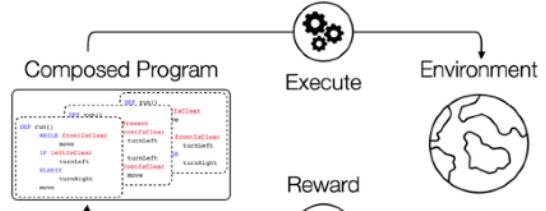
NeurIPS 2021

Generalizable Imitation Learning from Observation via Inferring Goal Proximity



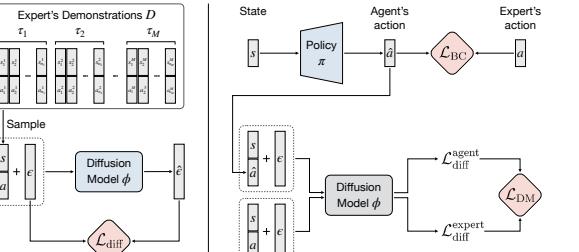
NeurIPS 2021

Hierarchical Programmatic Reinforcement Learning via Learning to Compose Programs



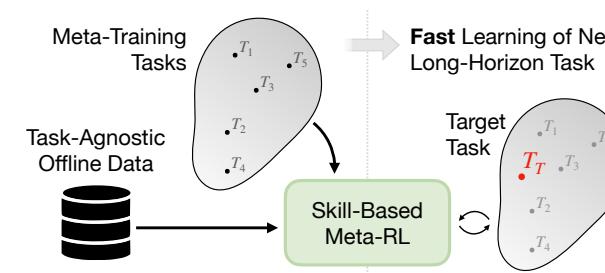
ICML 2023

Diffusion Model-Guided Behavioral Cloning



ICML-W 2023 & Submitted to ICML 2024

Skill-based Meta-Reinforcement Learning

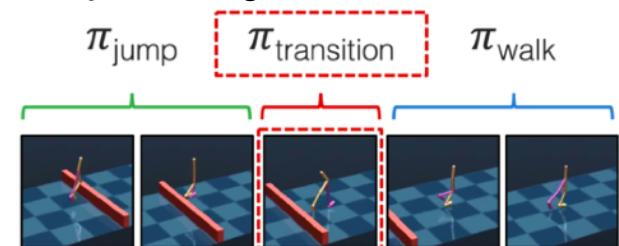


ICLR 2022

Primitive Skill Acquisition

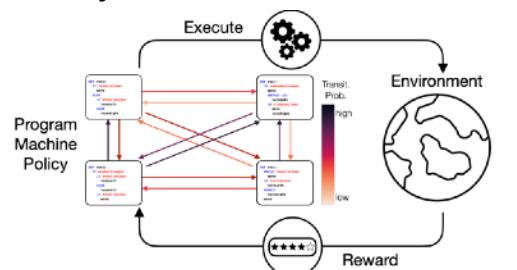
Task Execution

Composing Complex Skills by Learning Transition Policies



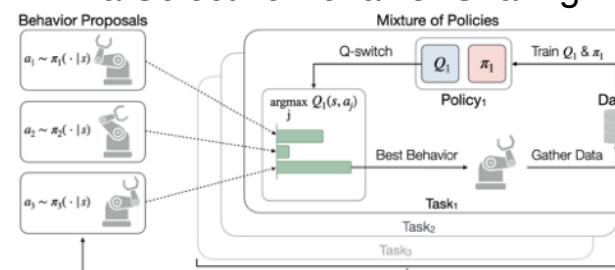
ICLR 2019

Addressing Long-Horizon Tasks by Integrating Program Synthesis and State Machines



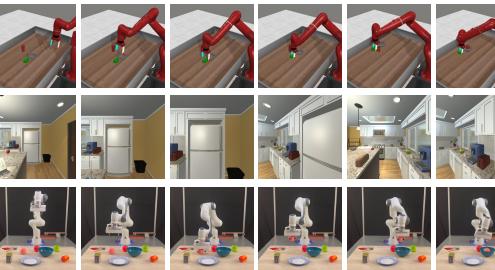
NeurIPS-W 2023 & Submitted to ICML 2024

Efficient Multi-Task Reinforcement Learning via Selective Behavior Sharing



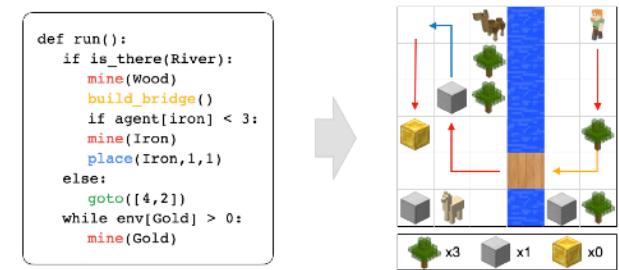
NeurIPS-W 2022 & Submitted to ICML 2024

Learning to Act from Actionless Videos through Dense Correspondences



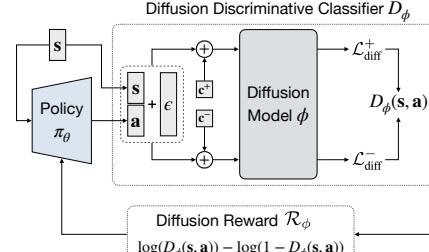
ICLR 2024 (Spotlight)

Program Guided Agent



ICLR 2020 (Spotlight)

Diffusion Rewards Guided Adversarial Imitation Learning



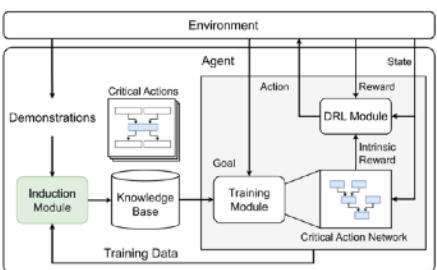
ICLR-W 2024 & Submitted to ICML 2024

Bootstrap Your Own Skills: Learning to Solve New Tasks with Large Language Model Guidance



CoRL 2023 (Oral)

Integrating Planning and Deep Reinforcement Learning via Automatic Induction of Task Substructures



ICLR 2024



Shao-Hua Sun (孫紹華)

Assistant Professor
at National Taiwan University
shaohuas@ntu.edu.tw



Bio

I am an **Assistant Professor** at **National Taiwan University (NTU)** with a joint appointment in **the Department of Electrical Engineering** and **the Graduate Institute of Communication Engineering**. Prior to joining NTU, I recently completed my Ph.D. in Computer Science at the University of Southern California, where I worked in the [Cognitive Learning for Vision and Robotics Lab \(CLVR\)](#). Before that, I received my B.S. degree in Electrical Engineering from NTU. My research interests span **Robot Learning**, **Reinforcement Learning**, **Program Synthesis**, and **Machine Learning**.

Prospective students: I am looking for students interested in machine learning, robot learning, reinforcement learning, and program synthesis. Specifically, I am hiring **M.S.** and **Ph.D. students** admitted to the Data Science and Smart Networking Group at the Graduate Institute of Communication Engineering (電信所丙組/資料科學與智慧網路組) or the Data Science Degree Program (資料科學學位學程) at NTU. Also, I am seeking **undergraduate students**, **research assistants**, and **visitors** with different experience levels. If you are interested in joining my group, please check out [this slide](#) and fill in the Google form.



[Edit profile](#)

Shao-Hua Sun

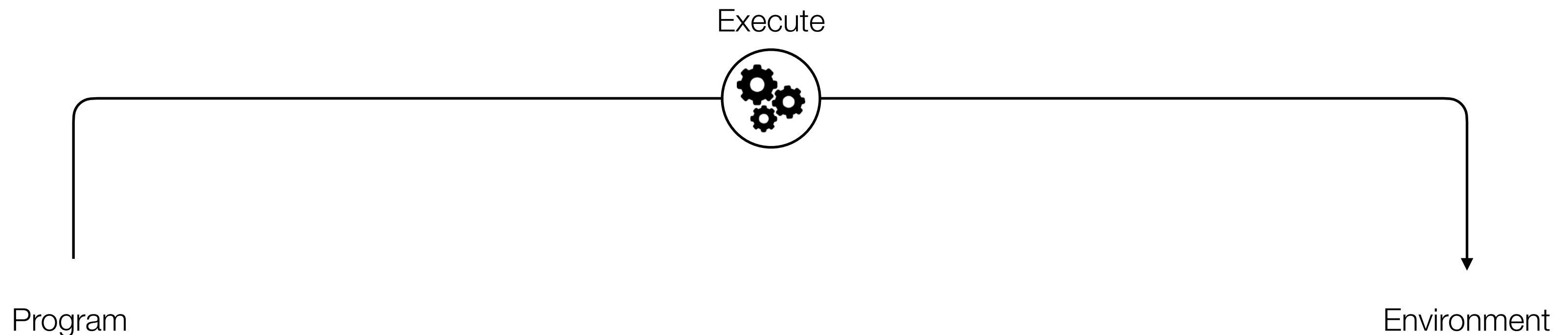
@shaohua0116

Assistant Professor @ National Taiwan University (NTU) | CS Ph.D. [@USC](#) | Robot Learning, Reinforcement Learning, Program Synthesis | 台大電機系助理教授

⌚ Taipei, Taiwan ⚡ shaohua0116.github.io 📅 Joined October 2015

868 Following 2,987 Followers





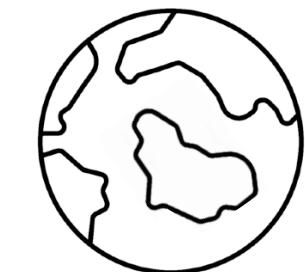
```
DEF run() m(  
  WHILE c( markerPresent c) w(  
    WHILE c( markerPresent c) w(  
      pickMarker  
      move w)  
    turnRight  
    move  
    turnLeft  
  WHILE c( markerPresent c) w(  
    pickMarker  
    move w)  
    turnLeft  
    move  
    turnRight w) m)
```

Thank You

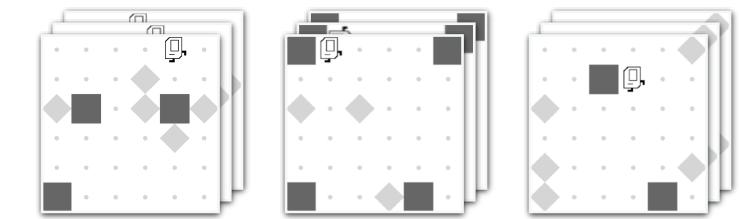


Questions?

Environment



Demonstrations



Reward