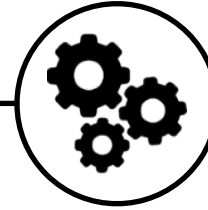


Execute



Learning to Synthesize Programs as Interpretable and Generalizable Reinforcement Learning Policies

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)
```

Environment



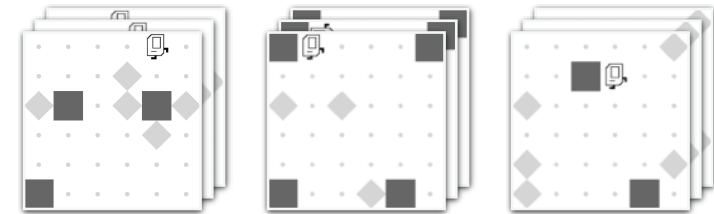
Shao-Hua Sun (孫紹華)

Assistant Professor

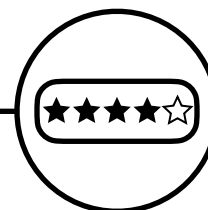
Dept. of Electrical Engineering (EE)

National Taiwan University

Demonstrations



Machine Learning Summer Schools 2024 @ OIST



Reward

Boston Dynamics



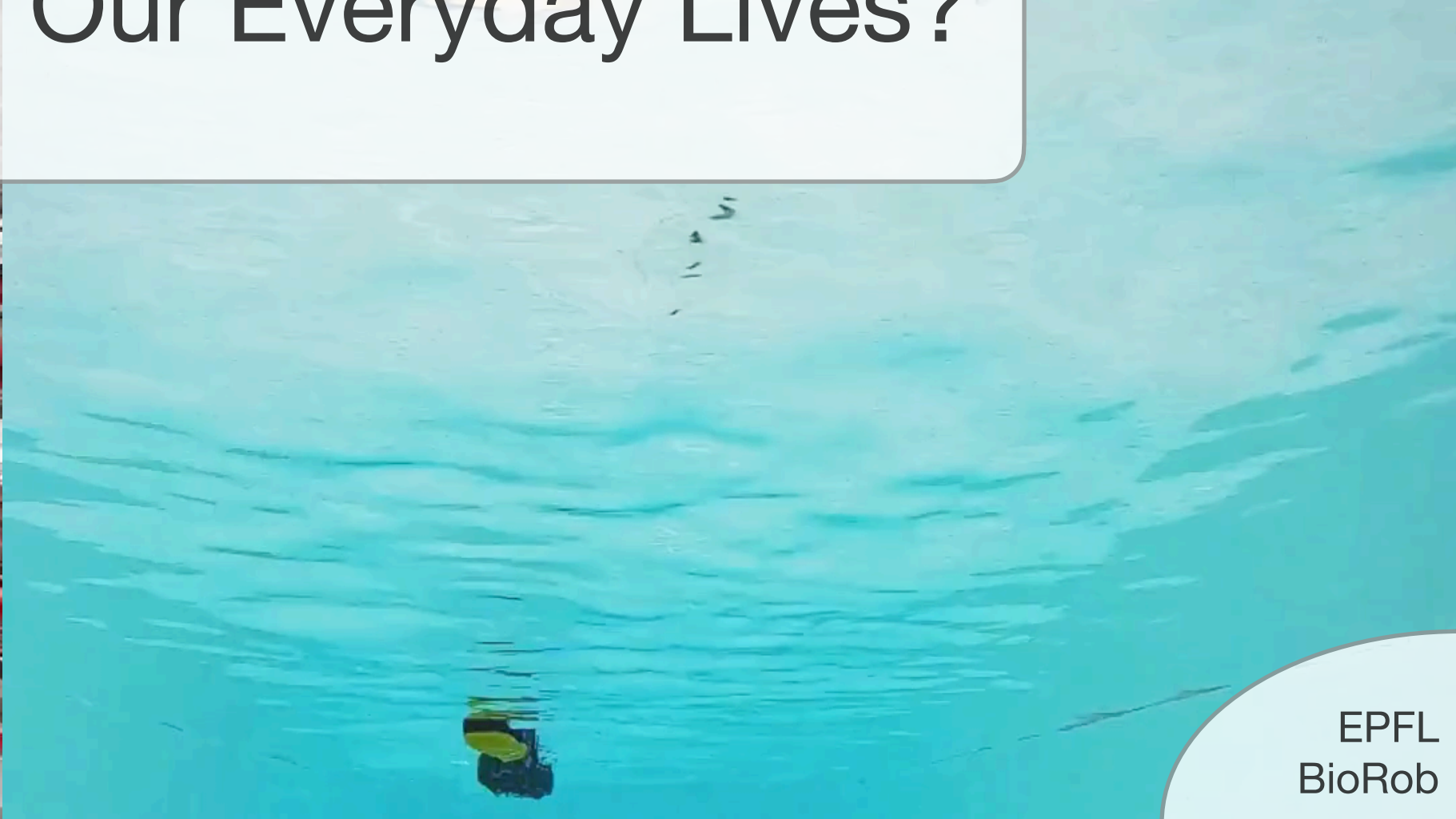
Google Robotics



Why Aren't Robots in Our Everyday Lives?

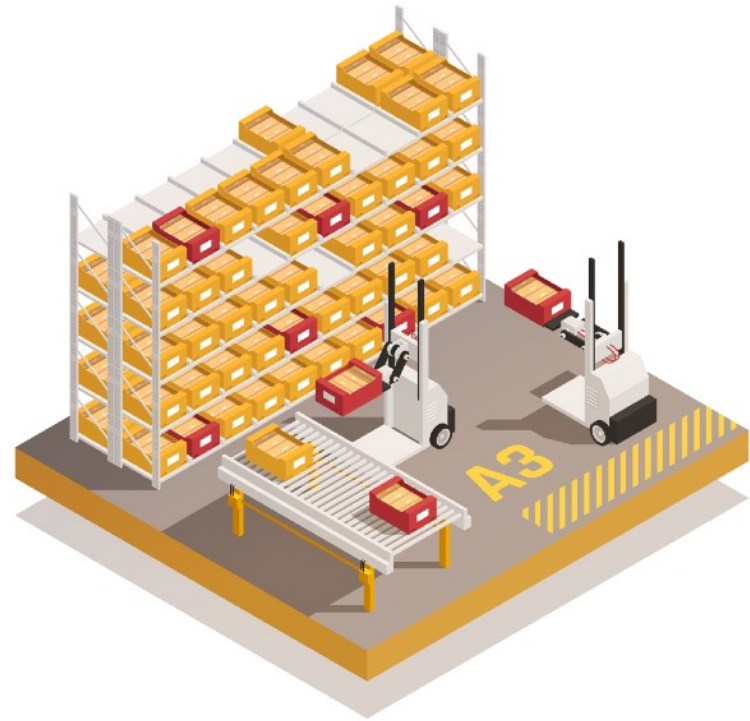


Tesla



EPFL BioRob

Environment



Structured



Unstructured

Object



Known



Unseen

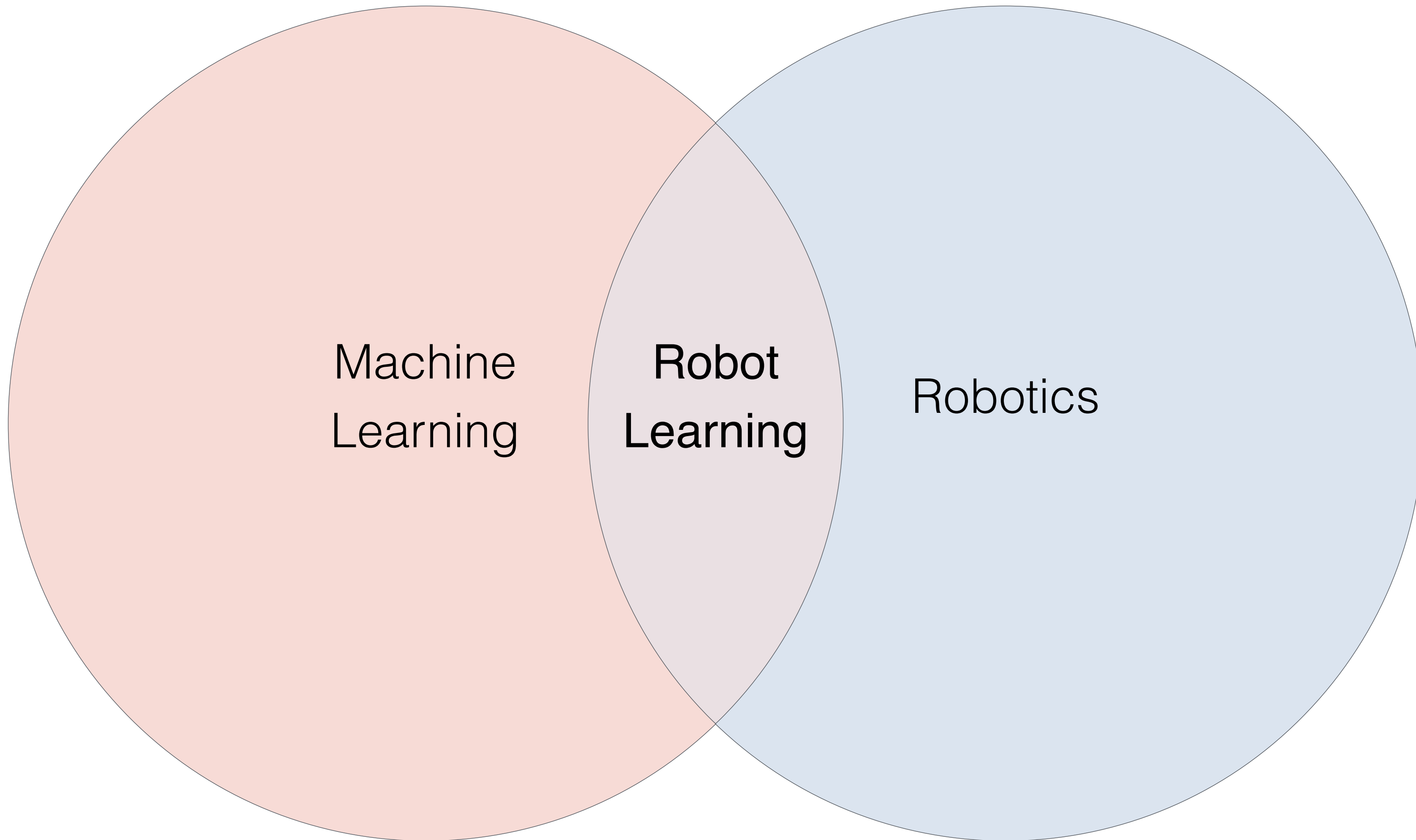
Task



Pre-defined / Pre-programmed



Diverse and Novel



Machine
Learning

**Robot
Learning**

Robotics

Supervised Learning

Image Classification

Model	image size	# parameters	Multi-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.2
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 x 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. Multi-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

Machine Translation

Source	Human	PBMT	GNMT
"The reason Boeing are 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.	6.0	3.0	6.0
"La raison pour laquelle Boeing sont en train de faire, c'est de concentrer davantage de sièges pour rendre leur avion plus compétitive avec nos produits", a déclaré Kevin M. Keniston, chef du confort des passagers de l'Airbus de l'Europe.	6.0	3.0	6.0
"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.	6.0	3.0	6.0
"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 chez l'avionneur européen Airbus.	6.0	3.0	6.0
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."	6.0	3.0	6.0
Interrogé à ce sujet, un responsable de l'administration américaine a répondu : "Les États-Unis n'est pas effectuer une surveillance électronique destiné aux bureaux de la Banque mondiale et du FMI à Washington."	6.0	3.0	6.0
Interrogé à ce sujet, un fonctionnaire de l'administration américaine a répondu: "Les États-Unis n'effectuent pas de surveillance électronique à l'intention des bureaux de la Banque mondiale et du FMI à Washington."	6.0	3.0	6.0
Interrogé sur le sujet, un responsable de l'administration américaine a répondu: "les États-Unis ne mènent pas de surveillance électronique visant les sièges de la Banque mondiale et du FMI à Washington."	6.0	3.0	6.0

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

Question Answering

System	Dev		Test	
	EM	F1	EM	F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - nInet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nInet	-	-	83.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	78.9	86.3	85.8	91.8
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} (SpQ+TrivialQA)	84.2	91.1	85.1	91.8
BERT _{LARGE} (Enb+TrivialQA)	86.2	92.2	87.4	93.2

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

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

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 use-dev				VQA v2 use-test			
	All	Visual	Text	Other	All	Visual	Text	Other
Prior (over correction answer to missing seq [1])	-	-	-	-	25.96	41.26	6.96	1.17
LSTM Language only (blat-exv0 [1])	-	-	-	-	44.36	47.01	51.55	37.37
Deep LSTM QANet (TTT) as reported in [1]	-	-	-	-	54.22	57.96	55.18	49.87
MCN [1] as reported in [1]	-	-	-	-	62.27	74.82	38.28	25.36
LIPMC LRF [1]	-	-	-	-	65.71	82.07	41.06	57.12
Abeno	-	-	-	-	67.59	82.50	44.19	59.97
ENINSE	-	-	-	-	66.77	81.99	46.26	58.26
IDE-USD-DNCC	-	-	-	-	68.09	84.50	45.39	61.01
Proposed model	-	-	-	-	62.07	79.20	59.45	52.63
ResNet VGG7+7, single network	62.07	79.20	59.45	52.63	62.27	79.52	59.71	52.91
Image features from better pre-training, adaptive A, single network	65.32	81.82	64.21	56.05	65.67	82.26	45.96	65.26
ResNet VGG7+7, ensemble	65.36	83.28	61.17	57.09	66.73	83.71	45.71	61.26
Image features from better pre-training, adaptive A, ensemble	69.07	84.69	68.79	62.80	74.24	86.60	64.44	64.12

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

Word Embeddings



On word embeddings - Part 1 by Ruder

Named Entity Recognition

contentSkip to site indexPublicSubscriberLog InSubscriberLog InToday's PaperAdvertisementSupported Org by F.B.I. Agent Peter Strzok PERSON . Who Criticized Trump PERSON in Texts, Is FiredImagePeter Strzok, a top F.B.I. GPE counterintelligence agent who was taken off the special counsel investigation after his disparaging texts about President Trump PERSON were uncovered, was fired. Credit J. Kirspatrick PERSON for The New York TimesBy Adam Goldman GPE and Michael S. SchmidtGPE PERSON 13 CARDINAL 2018WASHINGTON CARDINAL - Peter Strzok PERSON the F.B.I. 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 F.B.I. GPE lawyer, Lisa Page — to PERSON assailing the Russia GPE investigation as an illegitimate "witch hunt." Mr. Strzok PERSON , who rose over 20 years DATE at the F.B.I. 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 F.B.I. 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

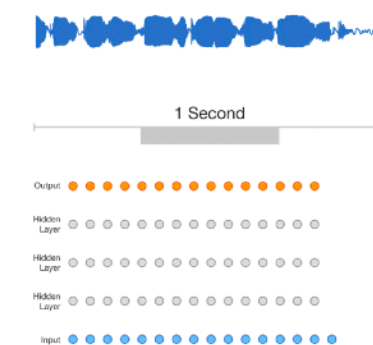
Named Entity Recognition and Classification with SciKit-Learn by Susan Li
Etelevs et al. Named Entity Recognition in Twitter using Images and Text

Speech Recognition

Senone set	Model/combination step	Word Error Rate			
		WER devset ngram-LM	WER test	WER devset LSTM-LMs	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
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	-	-	7.4	5.2
-	+ ngram rescoring	-	-	7.3	5.2
-	+ backchannel penalty	-	-	7.2	5.2
-	-	-	-	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

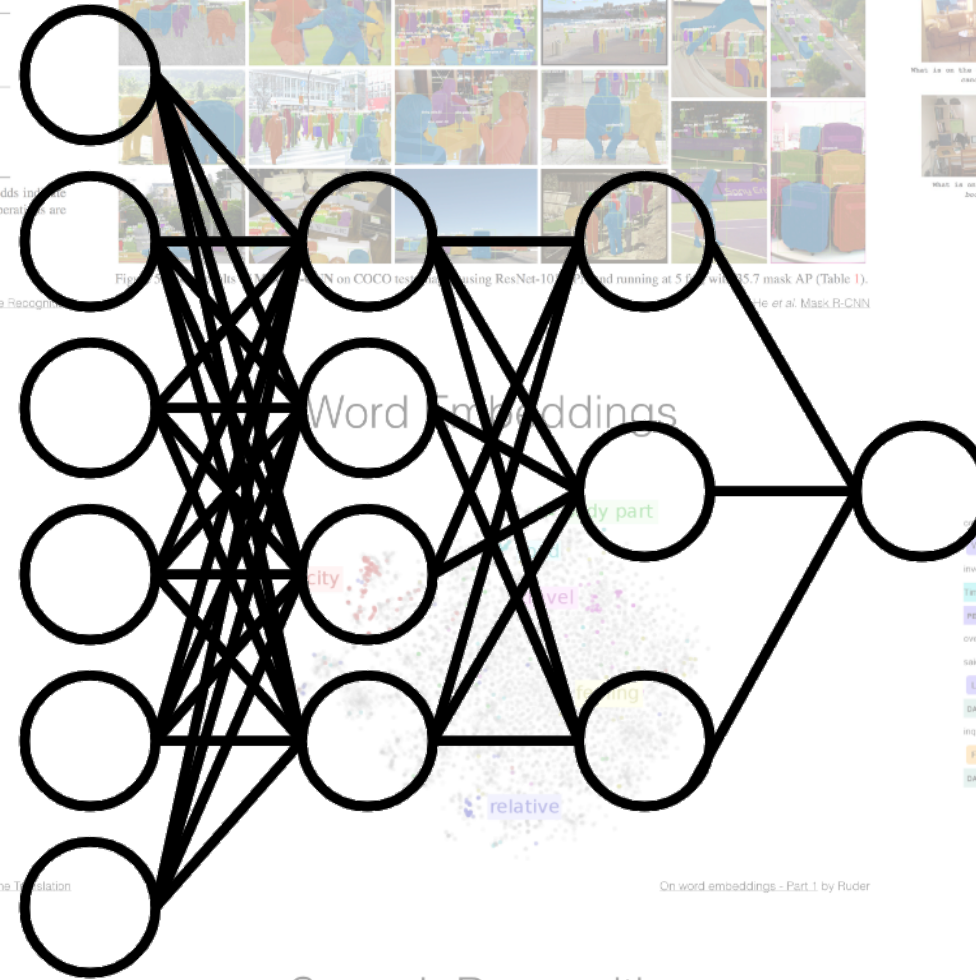
Supervised Learning

Image Classification

Model	image size	# parameters	Multi-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.2
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 x 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. Multi-Adds is 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



Visual Question Answering

Model	VQA v2 test dev			
	All	Visual	Text	Other
MetaNet	43.7	25.96	41.26	51.96
Deep LSTM (Seq2Seq)	44.26	47.01	51.91	57.37
Deep LSTM (Seq2Seq) + LSTM	54.22	51.96	59.18	48.81
Deep LSTM (Seq2Seq) + LSTM + LSTM	62.27	58.21	59.29	51.36
Deep LSTM (Seq2Seq) + LSTM + LSTM + LSTM	65.71	62.07	61.06	55.12
Deep LSTM (Seq2Seq) + LSTM + LSTM + LSTM + LSTM	67.39	62.50	61.19	59.07
Deep LSTM (Seq2Seq) + LSTM + LSTM + LSTM + LSTM + LSTM	66.77	61.90	60.26	56.26
Deep LSTM (Seq2Seq) + LSTM + LSTM + LSTM + LSTM + LSTM + LSTM	66.09	61.50	60.26	55.01

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

Question Answering

System	Dev		Test	
	EM	F1	EM	F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - nlNet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nlNet	-	-	83.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	78.9	86.3	79.5	86.6
R.M. Reader (Single)	81.2	87.9	82.3	88.5
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} (Spl+TriviaQA)	84.2	91.1	85.1	91.8
BERT _{LARGE} (Emb+TriviaQA)	86.2	92.2	87.4	93.2

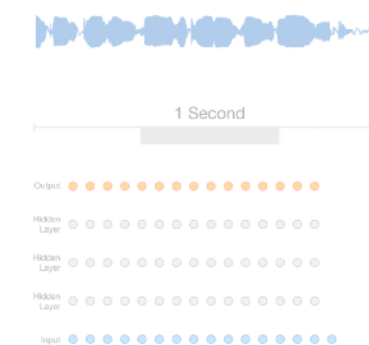
Table 2: SQuAD results. The BERT ensemble is 7x systems which use different pre-training checkpoints and fine-tuning seeds.

Speech Recognition

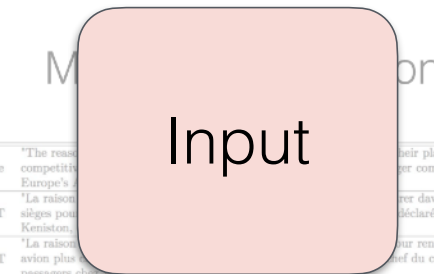
Senone set	Model/combination step	Word Error Rate			
		WER devset	WER test	WER devset	WER test
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.3
27k-puhpum	BLSTM	11.3	8.0	9.2	6.3
9k	BLSTM+ResNet+LACE+CNN-BLSTM	9.6	7.2	7.7	5.4
9k-puhpum	BLSTM+ResNet+LACE	9.7	7.4	7.8	5.4
9k-puhpum	BLSTM+ResNet+LACE+CNN-BLSTM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



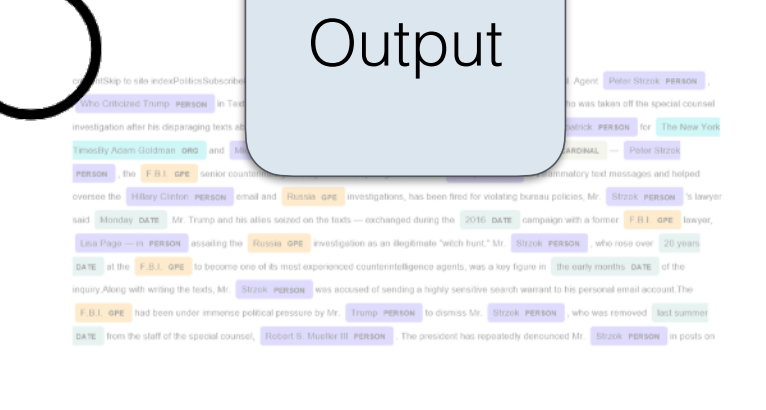
Wu et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

Word Embeddings



On word embeddings - Part 1 by Ruder

Name Entity Recognition



Named Entity Recognition and Classification with Seq2Seq by Susan Li
Elieves et al. Named Entity Recognition in Twitter using Images and Text

Supervised Learning

Image Classification

Model	image size	# parameters	Multi-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.2
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 x 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	32.0 B	81.5	95.8	
SENet [25]	300×300	42.3 M	42.3 B	82.7	96.2
NASNet-A (6 @ 4032)	300×300	23.8 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. Multi-Adds is the number of composite multiply-accumulate operations for a single image. Note that the composite multiply-accumulate operations are the implementation.



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

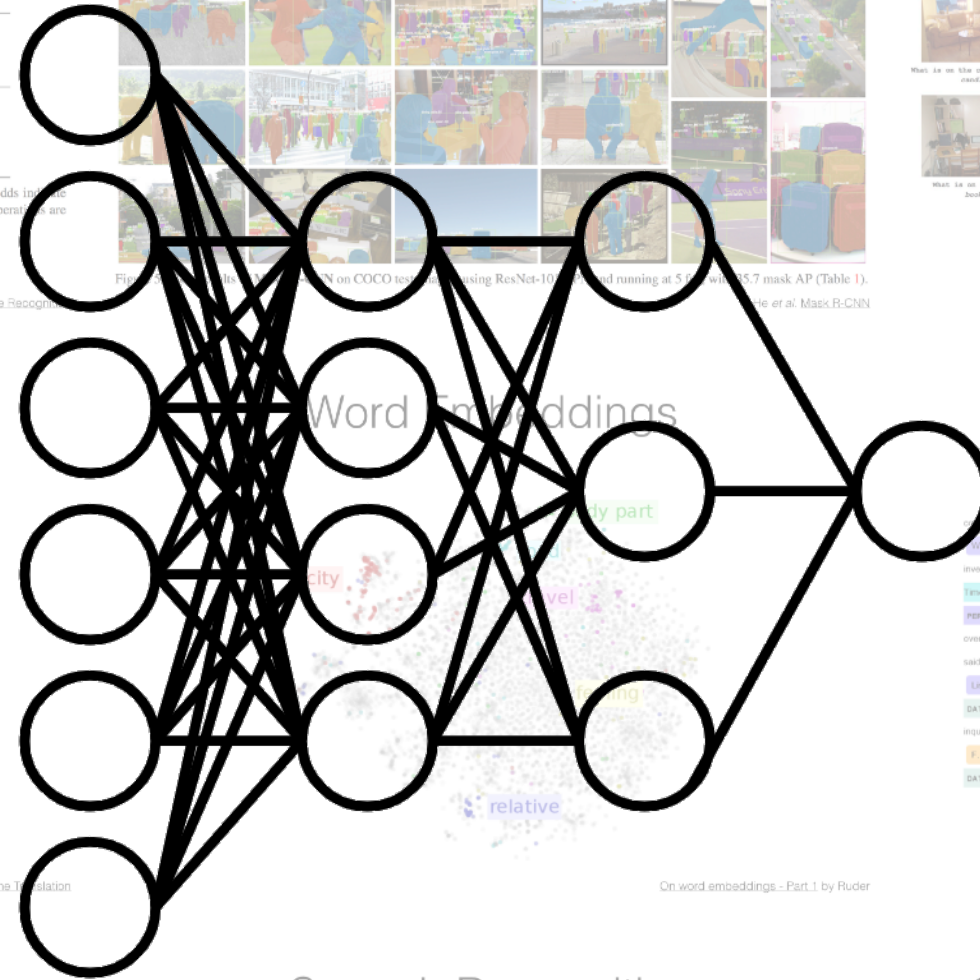
Instance Segmentation



Figure 1. Instance segmentation on COCO test set using ResNet-101 trained running at 5 FPS on a V100 GPU (Table 1).

Figure 2. Architecture for Scalable Image Recognition on COCO test set using ResNet-101 trained running at 5 FPS on a V100 GPU (Table 1).

Figure 3. Architecture for Scalable Image Recognition on COCO test set using ResNet-101 trained running at 5 FPS on a V100 GPU (Table 1).

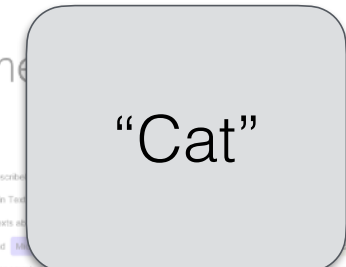


Visual Question Answering

Model	VQA v2 on dev				VQA v2 on test			
	All	Text	Image	Other	All	Text	Image	Other
PixelCNN (residual block + conv64 [1])	—	—	—	—	25.96	41.26	31.96	1.17
LSTM Language only (final conv64 [1])	—	—	—	—	44.26	47.01	51.51	37.37
Deep LSTM (LSTM [1]) as input [61]	—	—	—	—	54.22	57.46	59.18	48.82
MCN [1] (residual block [1])	—	—	—	—	62.27	74.82	78.29	52.36
UPMC LSTM [1]	—	—	—	—	65.71	82.07	81.06	55.12
Allen	—	—	—	—	67.29	82.50	81.19	51.97
ENRNet	—	—	—	—	66.77	81.99	80.26	56.26
IDE-LSTM-DNCC	—	—	—	—	66.69	81.90	80.29	55.62
Proposed model	—	—	—	—	67.02	81.29	80.40	52.63
ResNet (resnet-101, single network)	62.02	78.29	78.40	52.63	62.27	78.52	78.71	52.96
Image features from ResNet (resnet-101, single network)	65.52	81.82	84.23	56.03	65.67	82.26	82.96	56.26
ResNet (resnet-101, ensemble)	66.26	82.18	83.17	57.10	66.17	81.71	82.19	57.26
Deep Fusion from ResNet (resnet-101, ensemble)	67.07	84.49	86.79	62.80	74.24	86.68	86.44	63.12

Class

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



Question Answering

System	Dev		Test	
	EM	F1	EM	F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - nlNet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nlNet	-	-	83.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	78.9	86.3	79.5	86.6
R.M. Reader (Single)	81.2	87.9	82.3	88.5
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} (Emb+TriviaQA)	86.2	92.2	87.4	93.2

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

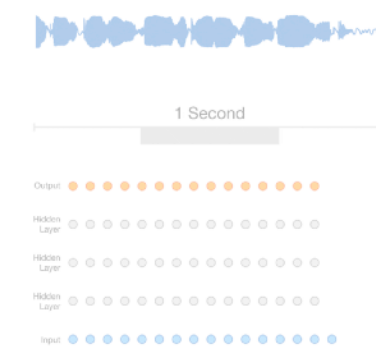
Devin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Speech Recognition

Senone set	Model/combination step	Word Error Rate			
		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
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

Supervised Learning

Image Classification

Model	image size	# parameters	Multi-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.2
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 x 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. Multi-Adds is 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.

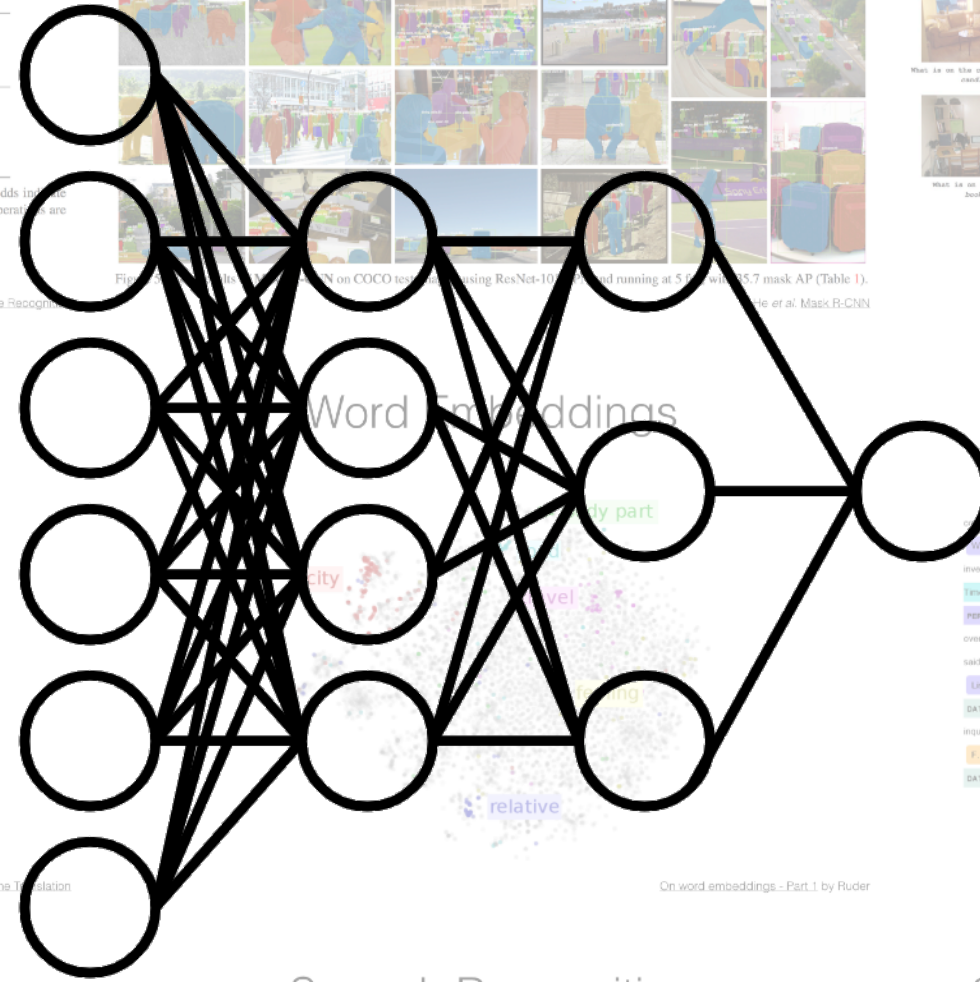
English sentence

France is never cold in September

System	Dev EM	Dev F1	Test EM	Test F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - nInet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nInet	-	-	83.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	78.9	86.3	79.5	86.6
R.M. Reader (Single)	81.2	87.9	82.3	88.5
R.M. Reader (Ensemble)	80.8	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} (Emb+TriviaQA)	86.2	92.2	87.4	93.2

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

Instance Segmentation



Speech Recognition

Senone set	Model/combination step	Word Error Rate			
		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
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

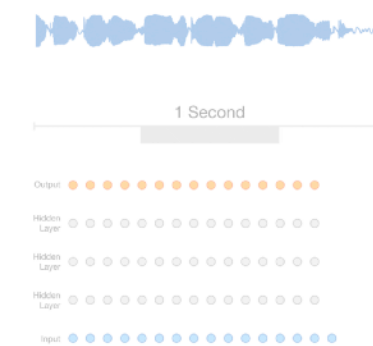
Visual Question Answering

Model	VQA v2 dev set				VQA v2 test set			
	All	Visual	Text	Other	All	Visual	Text	Other
MetaNet	-	-	-	-	25.96	41.26	31.96	13.17
Deep LSTM (single) [11]	-	-	-	-	44.26	47.01	51.51	37.37
Deep LSTM (ensemble) [11]	-	-	-	-	54.02	57.46	59.18	48.81
RCNN [1]	-	-	-	-	62.27	74.82	78.28	52.86
Deep LSTM [11]	-	-	-	-	65.71	82.07	81.06	55.12
Allen	-	-	-	-	67.39	82.50	81.19	59.87
ENRNet	-	-	-	-	68.77	83.96	84.52	58.26
Deep LSTM+ENRNet	-	-	-	-	68.89	84.50	85.19	59.81
Proposed model	-	-	-	-	62.02	78.29	78.40	52.61
ResNet (single) [11]	-	-	-	-	62.27	78.52	78.71	52.96
Image features from ResNet (ensemble) [11]	-	-	-	-	65.52	81.82	84.23	60.87
ResNet (ensemble) [11]	-	-	-	-	66.26	82.18	83.17	61.29
Deep LSTM from ResNet (ensemble) [11]	-	-	-	-	69.07	84.48	84.89	64.80
Deep LSTM from ResNet (ensemble) [11]	-	-	-	-	74.24	86.68	86.84	64.12

French sentence

la france est jamais froid en septembre

Speech Synthesis (text-to-speech)



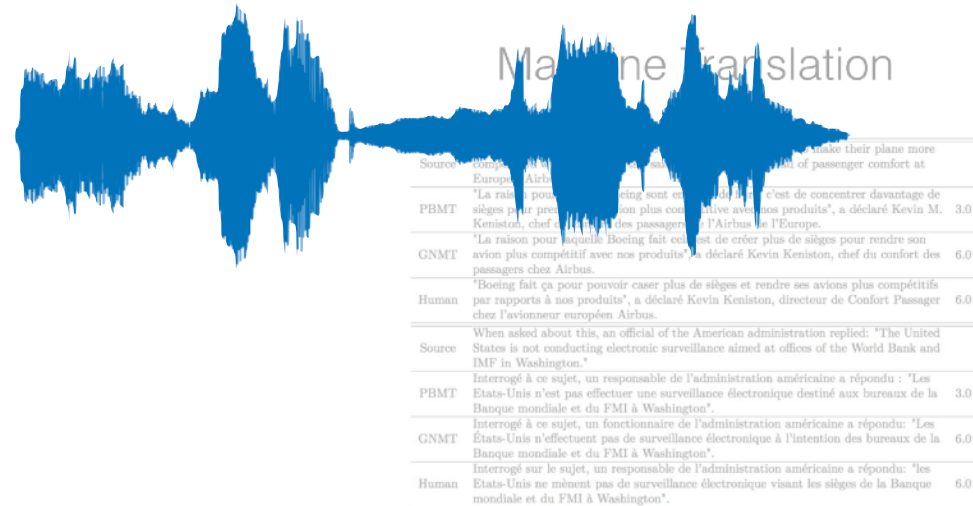
Supervised Learning

Image Classification

Model	image size	# parameters	Multi-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.2
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 x 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. Multi-Adds is 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.

Waveform



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

Instance Segmentation

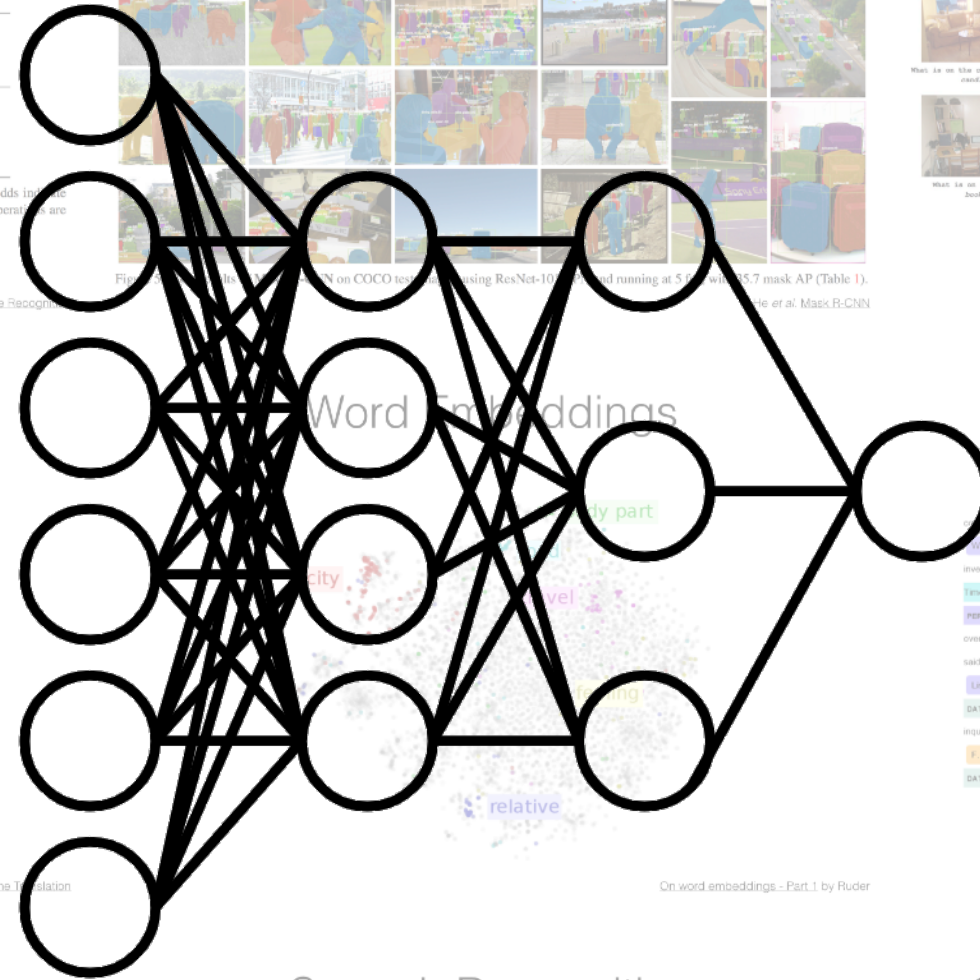


Figure 1. Performance of Mask R-CNN on COCO test set using ResNet-101 backbone and running at 5 FPS on a GTX 1080 Ti GPU. Mask AP (Table 1).

Visual Question Answering



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

Text

This is a supervised learning method

Question Answering

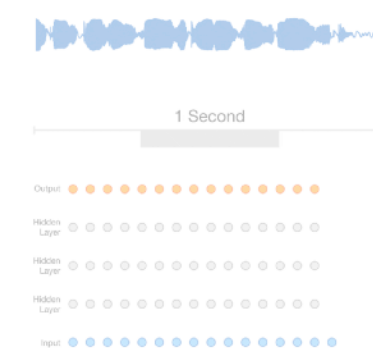
System	Dev		Test	
	EM	F1	EM	F1
Leaderboard (Oct 8th, 2018)				
Human	-	-	82.3	91.2
#1 Ensemble - nlnet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nlnet	-	-	83.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Published				
BIDAF+ELMo (Single)	-	85.8	-	-
R.M. Reader (Single)	78.9	86.3	79.5	86.6
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} (Spl+TriviaQA)	84.2	91.1	85.1	91.8
BERT _{LARGE} (Emb+TriviaQA)	86.2	92.2	87.4	93.2

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

Speech Recognition

Senone set	Model/combination step	Word Error Rate			
		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
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

Speech Synthesis (text-to-speech)



Supervised Learning

Image Classification

Model	Image size	# parameters	Multi-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224x224	11.2 M	1.94 B	74.8	92.2
NASNet-A (5 @ 1538)	299x299	10.9 M	2.35 B	78.6	94.2
Inception V3 [60]	299x299	23.8 M	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. Multi-Adds is 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. MaskRCNN on COCO test set using ResNet-101 trained running at 5 FPS on a V100 GPU. mAP: 35.7 mask AP (Table 1).

Visual Question Answering

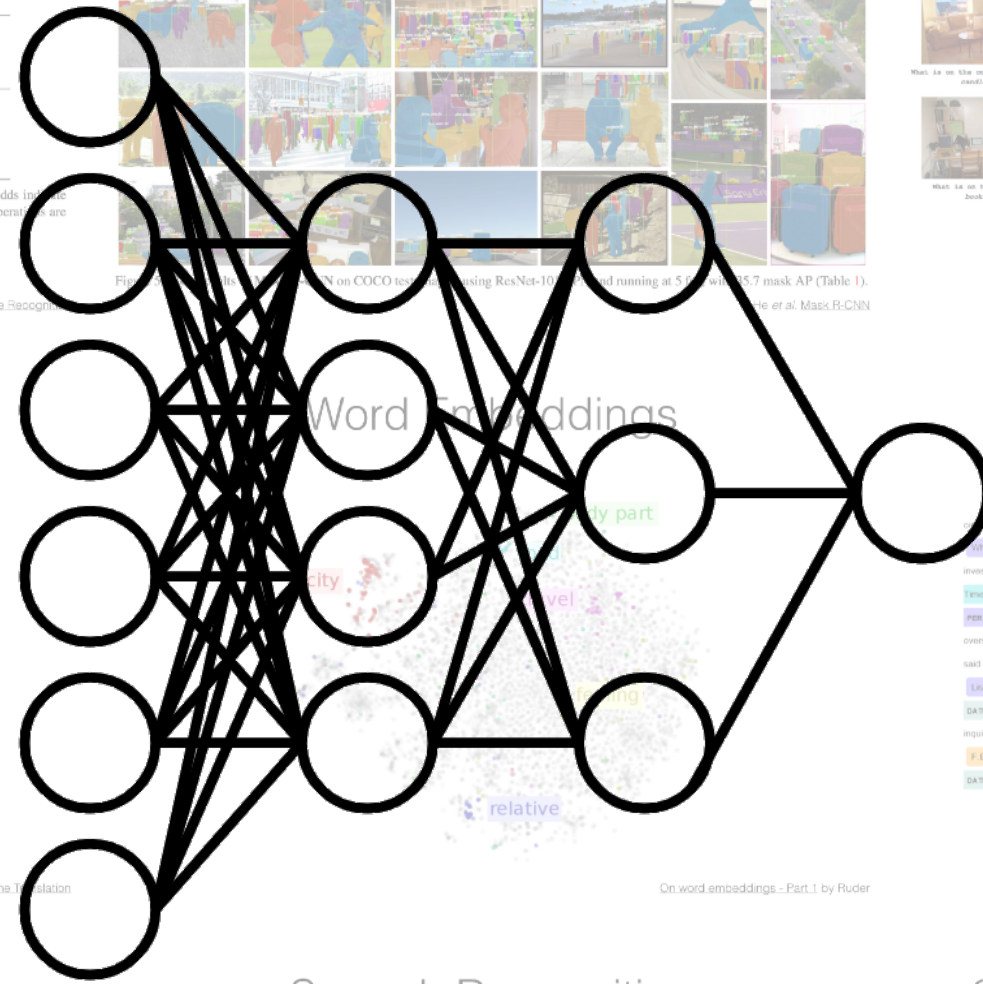
Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA v2 Leaderboard [1].

Question Answering

Source	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 3.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT
"The reason for their plane more comfortable at Europe's airports is that they have a larger cabin width and more legroom." 6.0	Human	GNMT	PBMT

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

Speech Recognition



Speech Recognition

System set	Model/combination step	WER devset	WER test	WER ngram-LM	WER LSTM-LMs
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.0	9.3	6.4
9k-puhpum	BLSTM+ResNet+LACE	9.6	7.2	8.0	6.4
9k-puhpum	BLSTM+ResNet+LACE+CNN-BLSTM	9.7	7.4	7.8	5.4
27k	BLSTM+ResNet+LACE	10.0	7.5	8.0	5.8
	+ backchannel penalty			7.2	5.1

Word Error Rate

Speech Synthesis (text-to-speech)

Named Entity Recognition and Classification with Self-Leaky by Susan Li

Question Answering

System	Dev EM	Dev F1	Test EM	Test F1
#1 Ensemble - nlnet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nlnet	-	-	83.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Ours	80.8	88.5	86.2	92.2
BERTBASE (SQuAD)	80.8	88.5	86.2	92.2
BERTLARGE (SQuAD)	80.8	88.5	86.2	92.2
BERTLARGE (EM+T5v1v2Q)	86.2	92.2	87.4	93.2

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

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

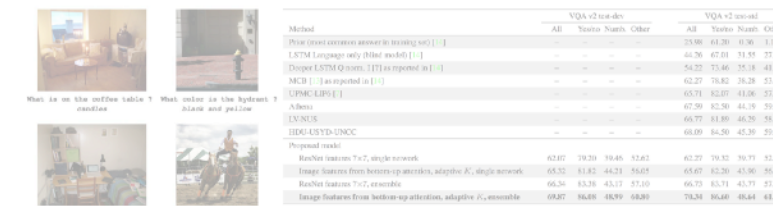
Model	image size	# parameters	Multi-Adds	Top 1 Acc. (%)	Top 5 Acc. (%)
Inception V2 [29]	224×224	11.2 M	1.94 B	74.8	92.2
NASNet-A (5 @ 1538)	299×299	10.9 M	2.35 B	78.6	94.2
Inception V3 [60]	299×299	23.8 M	5.72 B	78.8	94.4
Xception [9]	299×299	22.8 M	8.38 B	79.0	94.5
Inception ResNet V2 [58]	299×299	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 x 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. Multi-Adds is 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 sizes for [25] evaluated from open-source implementation.

Instance Segmentation



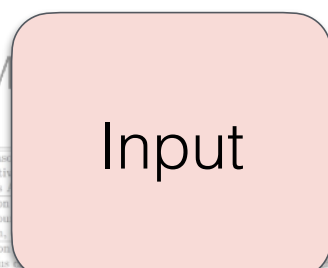
Visual Question Answering



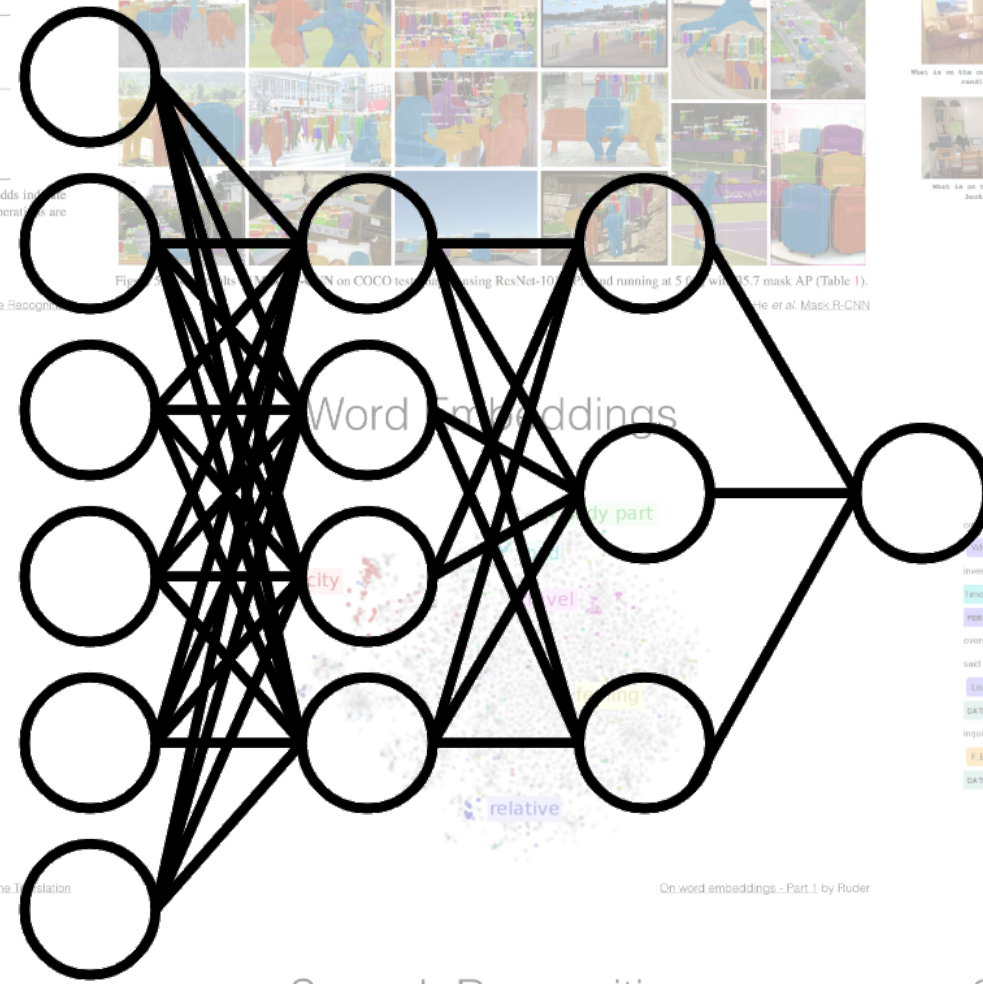
Action

- move
- turnLeft
- turnRight

State



Word Embeddings



Output

Question Answering

System	Dev EM	Dev F1	Test EM	Test F1
Leaderboard (Oct 8th, 2018)				
#1 Ensemble - nlnet	-	-	86.0	91.7
#2 Ensemble - QANet	-	-	84.5	90.5
#1 Single - nlnet	-	-	83.5	90.1
#2 Single - QANet	-	-	82.5	89.3
Ours	86.8	88.5	-	-
BERT _{Large} (E8B+HybridQA)	86.2	92.2	87.4	94.2

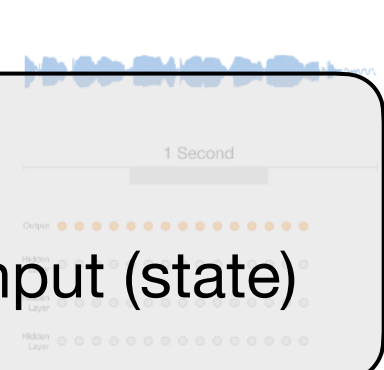
Table 2: SQuAD results. The BERT ensemble is 7x systems which use different pre-training checkpoints and fine-tuning seeds.

Speech Recognition

System set	Model/combination step	WER devset	WER test	WER ngram-LM	WER LSTM-LMs
9k	BLSTM	11.5	8.3	9.2	6.3
27k	BLSTM	11.4	8.6	9.3	6.3
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
	+ backchannel penalty			7.2	5.1

Xiong et al. The Microsoft 2017 Conversational Speech Recognition System

Speech Synthesis (text-to-speech)

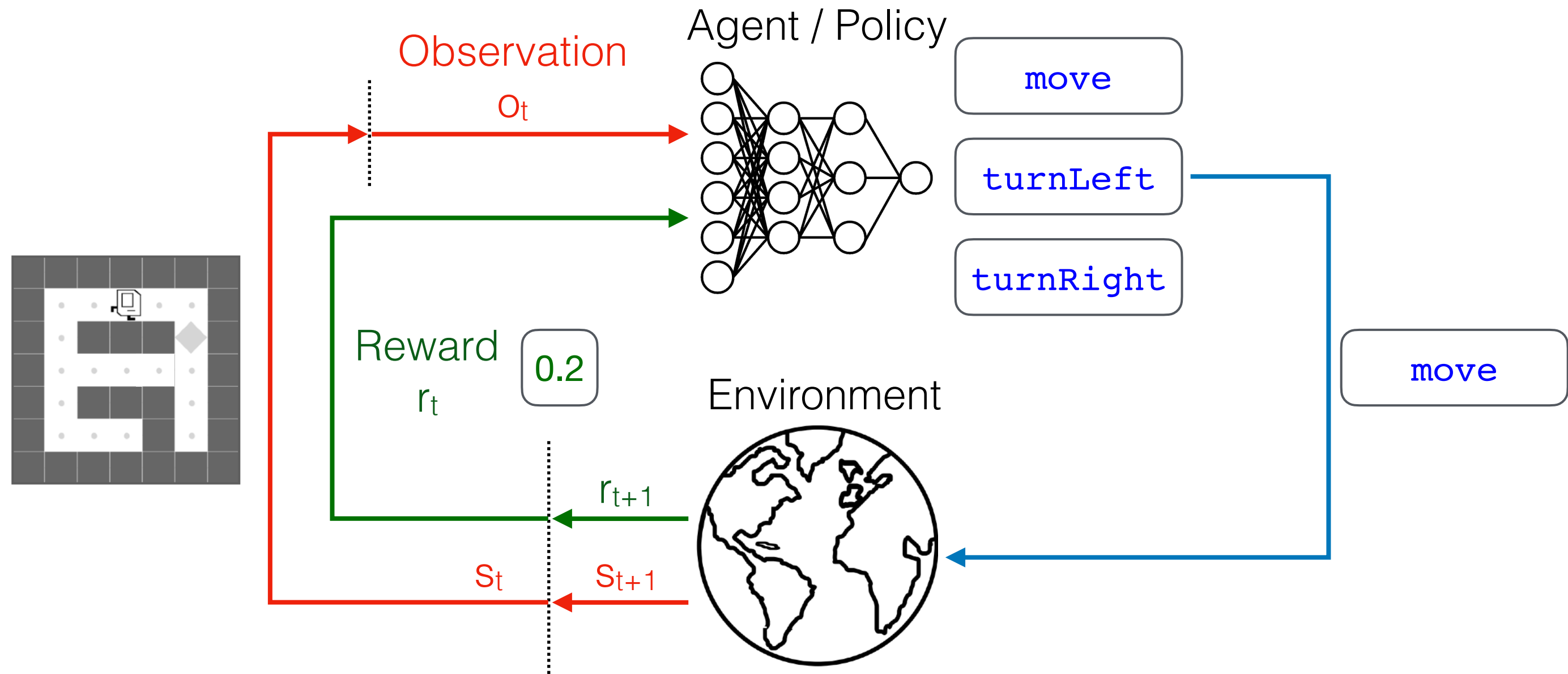


Van Den Oord et al. WaveNet: A Generative Model for Raw Audio

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

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

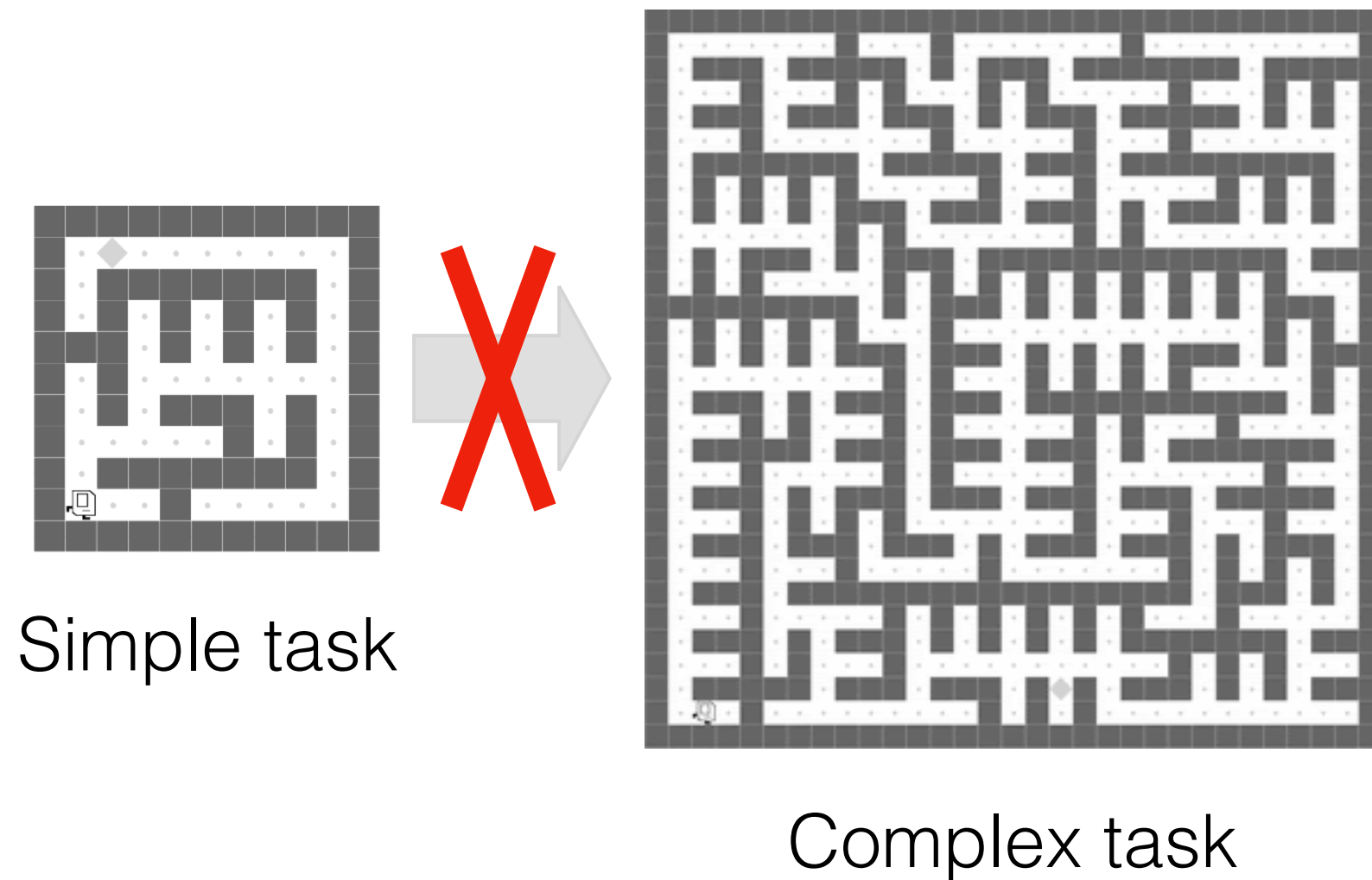
Robot Learning via Deep Reinforcement Learning



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

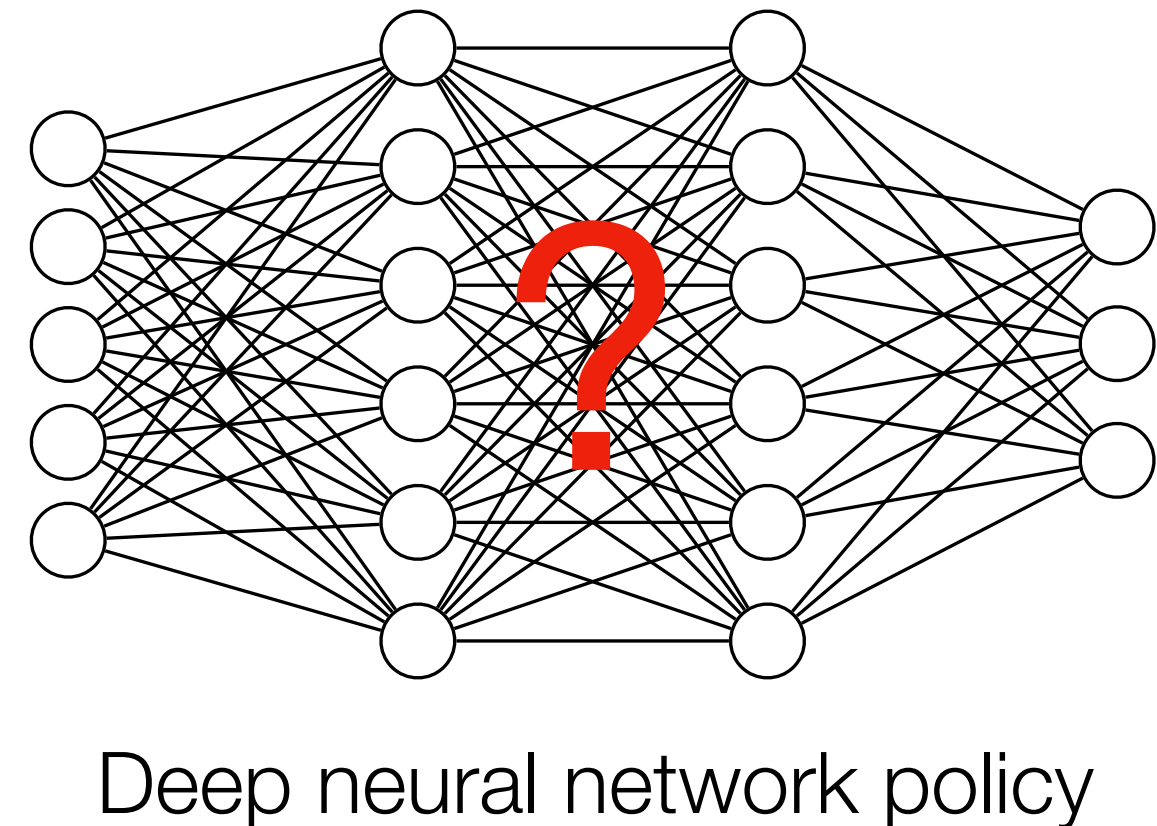
Robot Learning via Deep Reinforcement Learning - Issues

Generalization

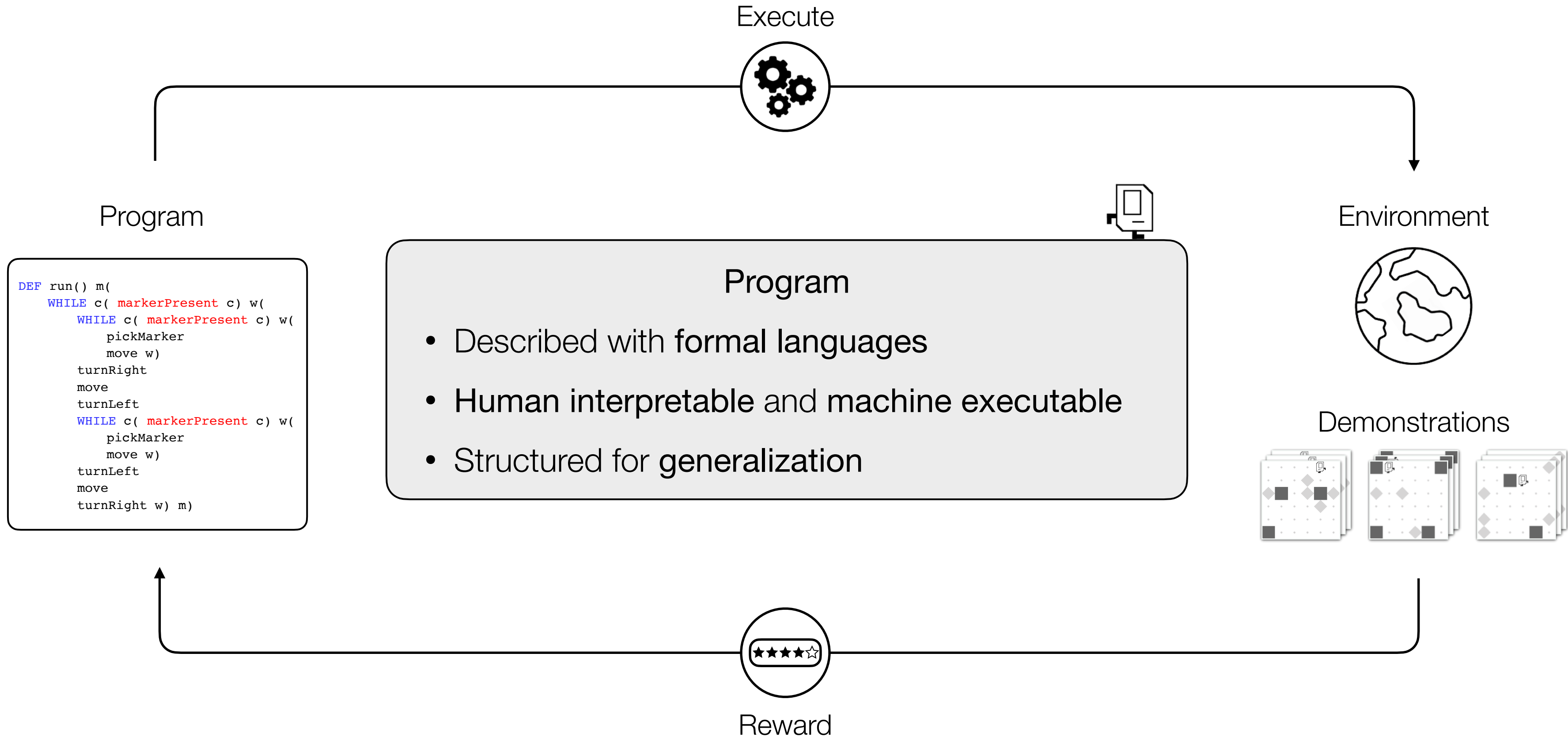


Interpretability

Trust, Safety, and Contestability



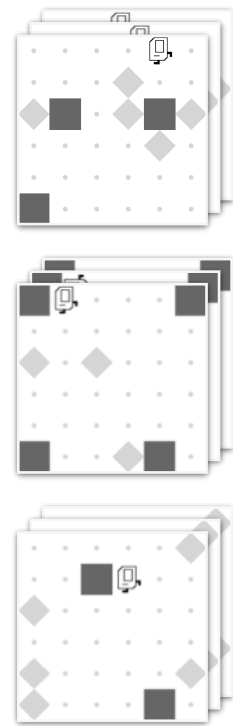
Program as Reinforcement Learning Policies



Neural Program Synthesis from Diverse Demonstration Videos

ICML 2018

Demonstrations



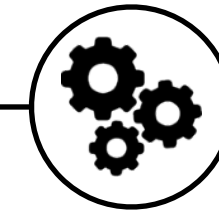
Synthesize



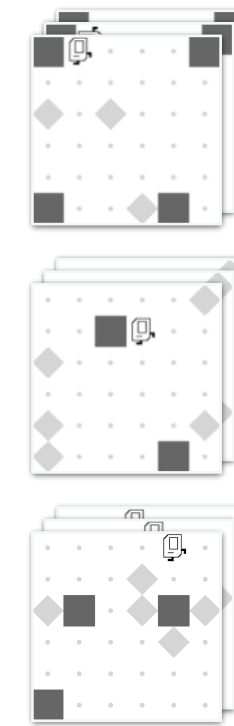
Program Policy

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

Execute



Execution



Shao-Hua Sun*



Hyeonwoo Noh*

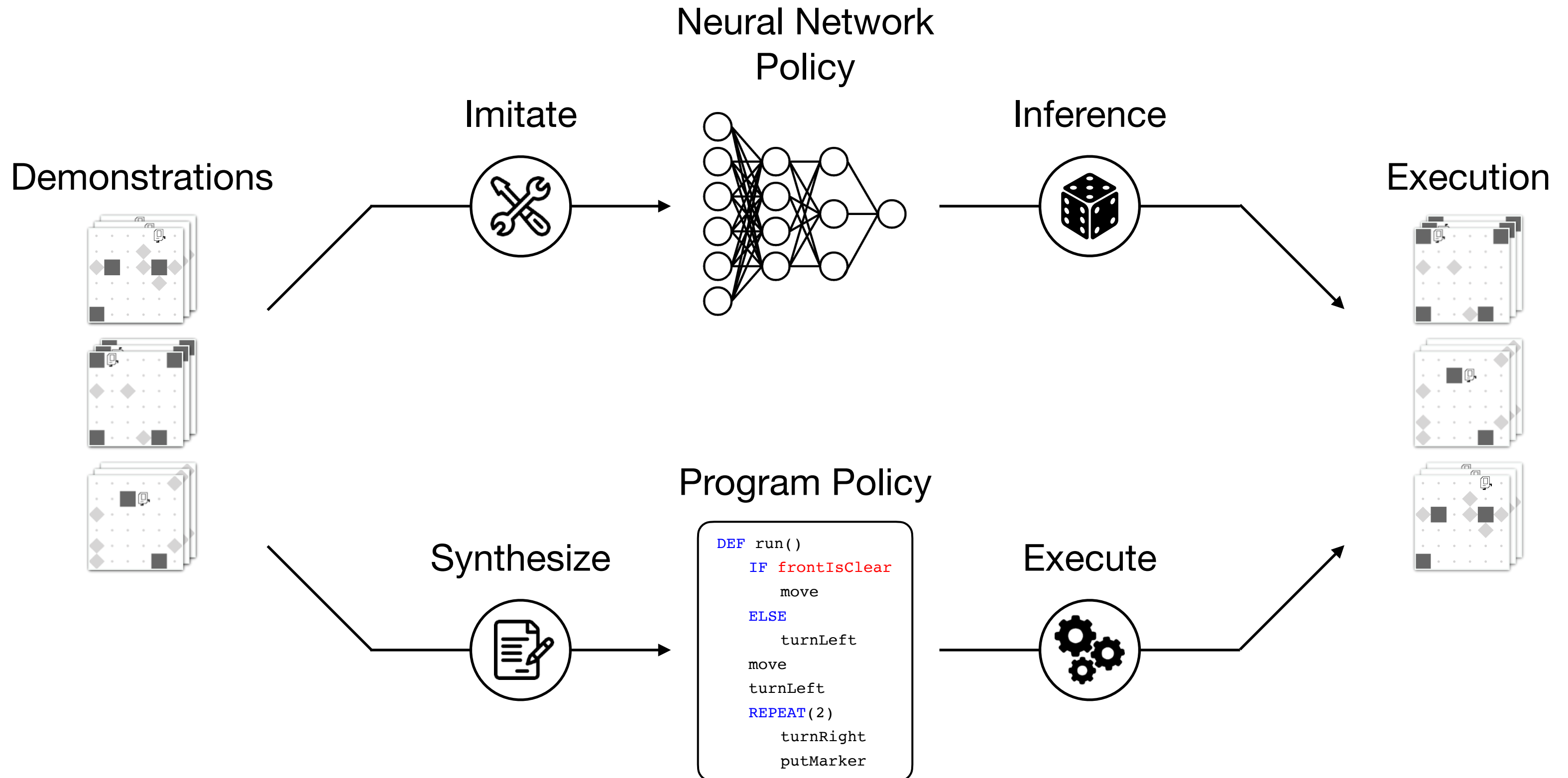


Sriram Somasundaram



Joseph J. Lim

Imitation Learning via Synthesizing Programs



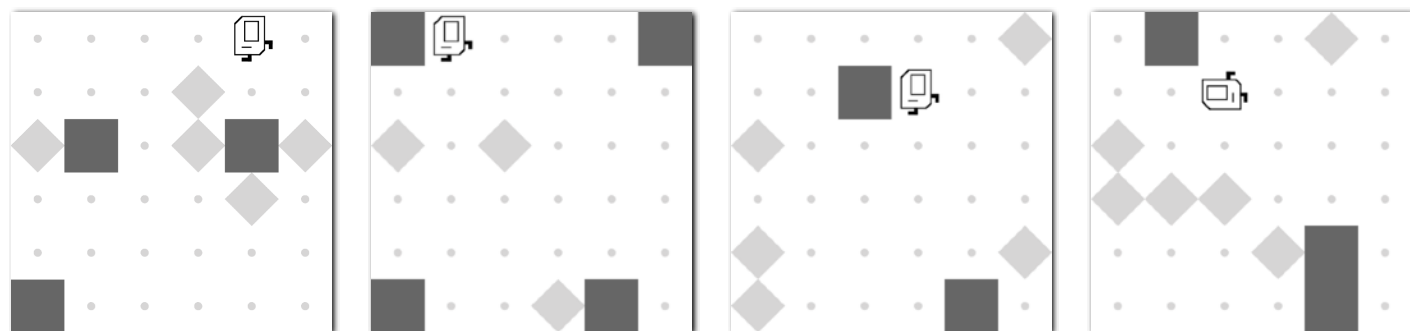
Environments

Karel

Program

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

Demonstrations



ViZDoom

Program

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

Demonstrations

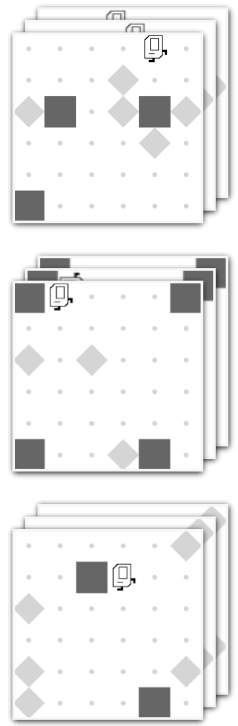


Richard E Pattis. "Karel the robot: a gentle introduction to the art of programming." John Wiley & Sons, Inc., 1981

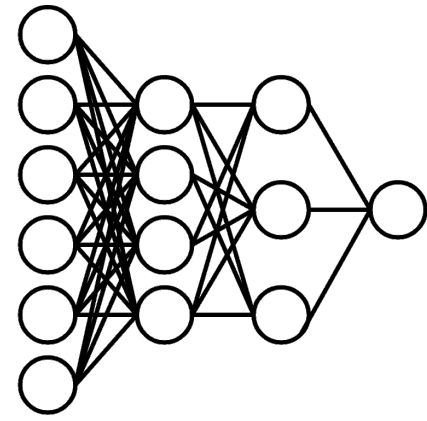
Kempka et al., "Vizdoom: A doom-based ai research platform for visual reinforcement learning." in CIG, 2016

Imitation Learning with Neural Network Policy

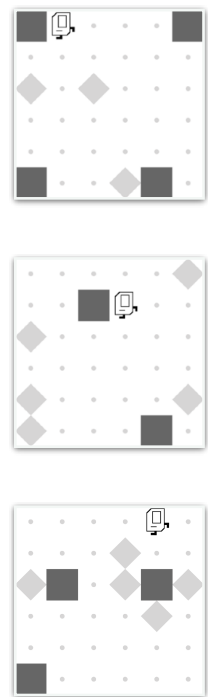
Demonstrations



Neural Network
Policy



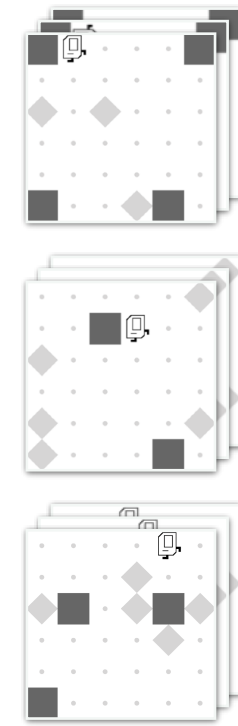
Initial States



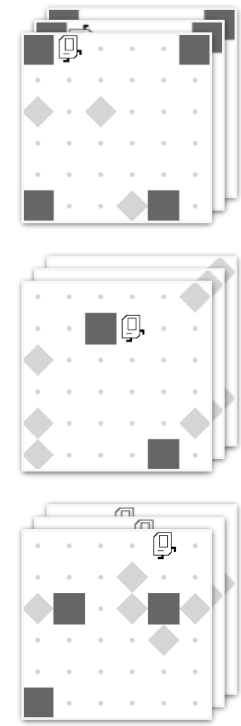
Infer



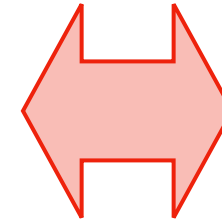
Predicted Execution



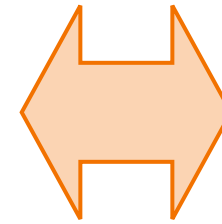
Ground Truth Execution



Loss

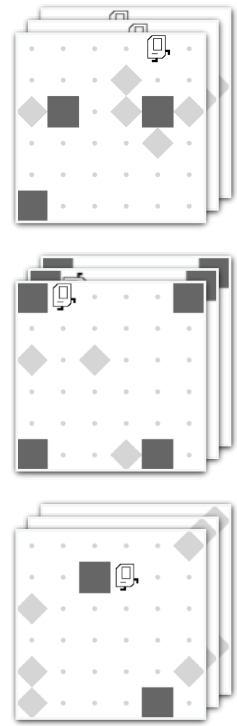


Evaluate

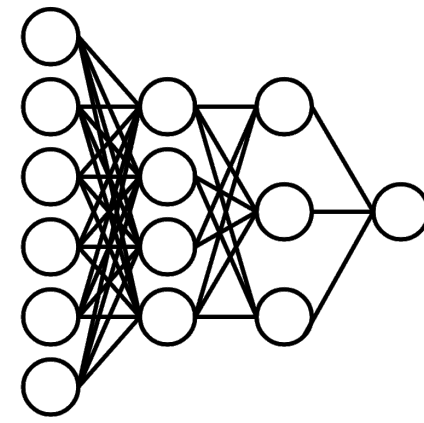


Imitation Learning with Program Policy

Demonstrations



Neural Network
Program Synthesizer



Synthesize

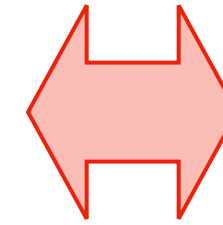
Predicted Program

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

Ground Truth Program

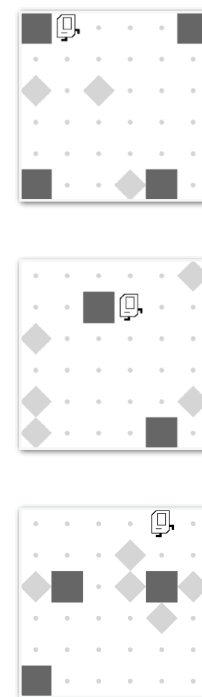
```
DEF run()  
  IF frontIsClear  
    move  
  ELSE  
    turnRight  
    move  
  REPEAT(2)  
    turnLeft
```

Loss

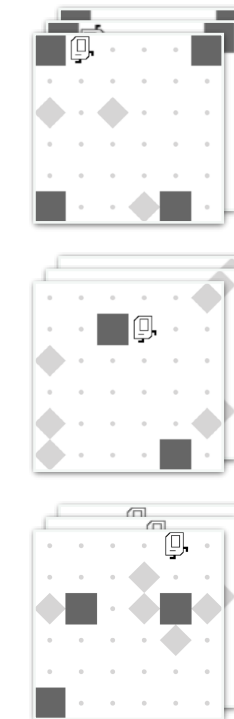


Execute

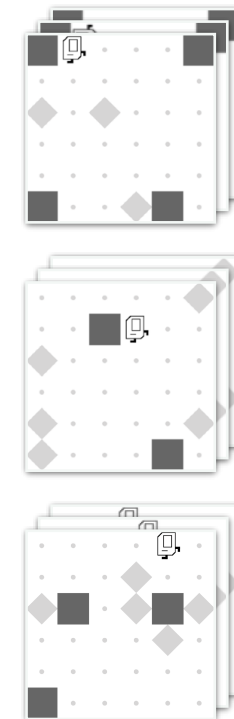
Initial States



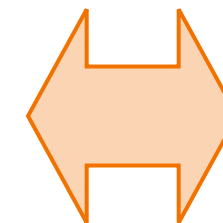
Predicted Execution



Ground Truth Execution

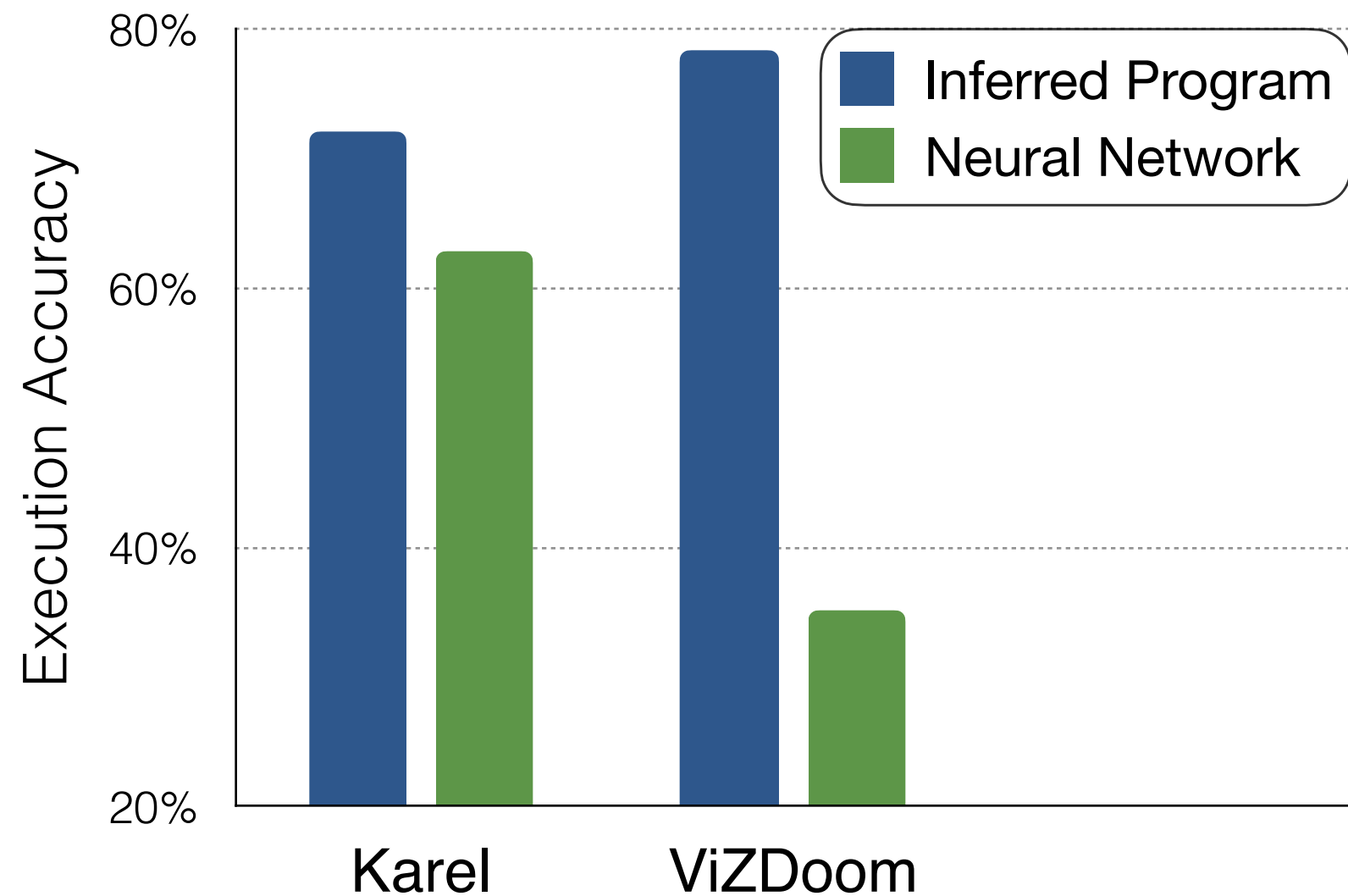


Evaluate



Experimental Results

Quantitative Results



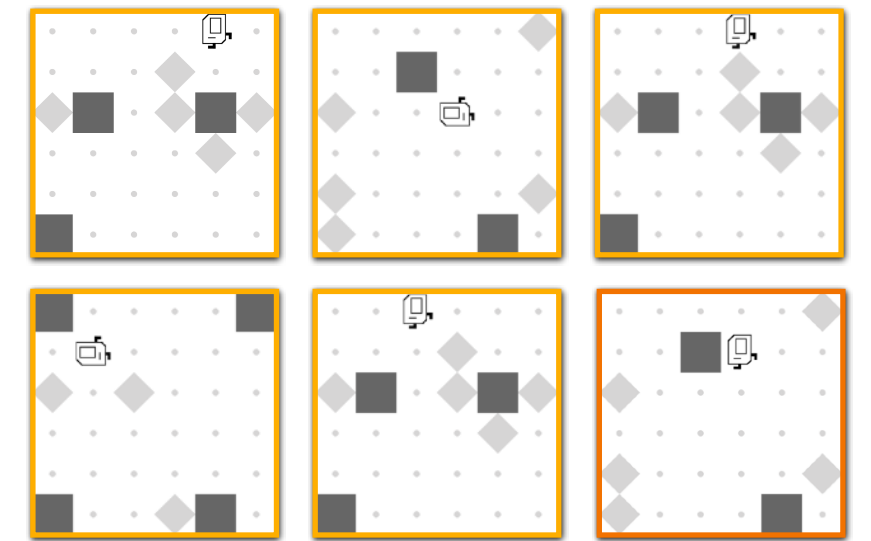
Evaluation: Execute the **inferred program** and the **learned neural network policy** on a set of unseen initial states and compare them to the **ground truth demonstrations**

Observation

Program

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

Demonstrations



Neural Network Policy

move

- Learn to mimic the general tendency of the expert behaviors

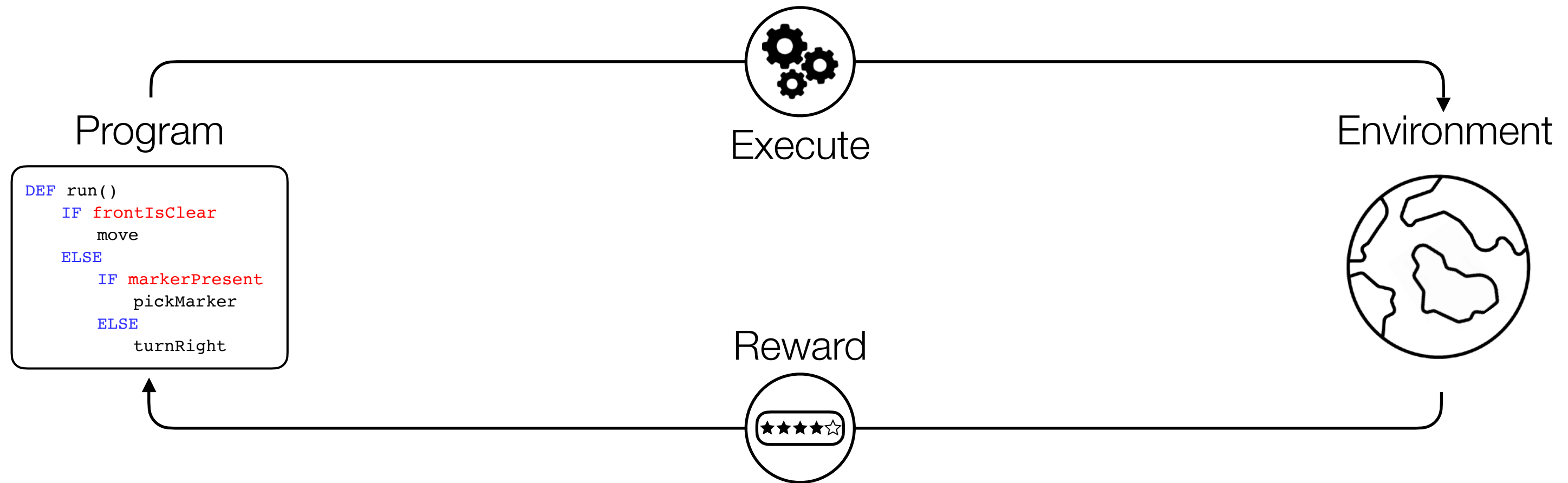
Inferred Program

```
DEF run()  
  IF frontIsClear  
    move  
  ELSE  
    turnLeft
```

- Learn to capture the decision-making logics of the expert

Learning to Synthesize Programs as Interpretable and Generalizable Policies

NeurIPS 2021



Dweep Trivedi*



Jesse Zhang*

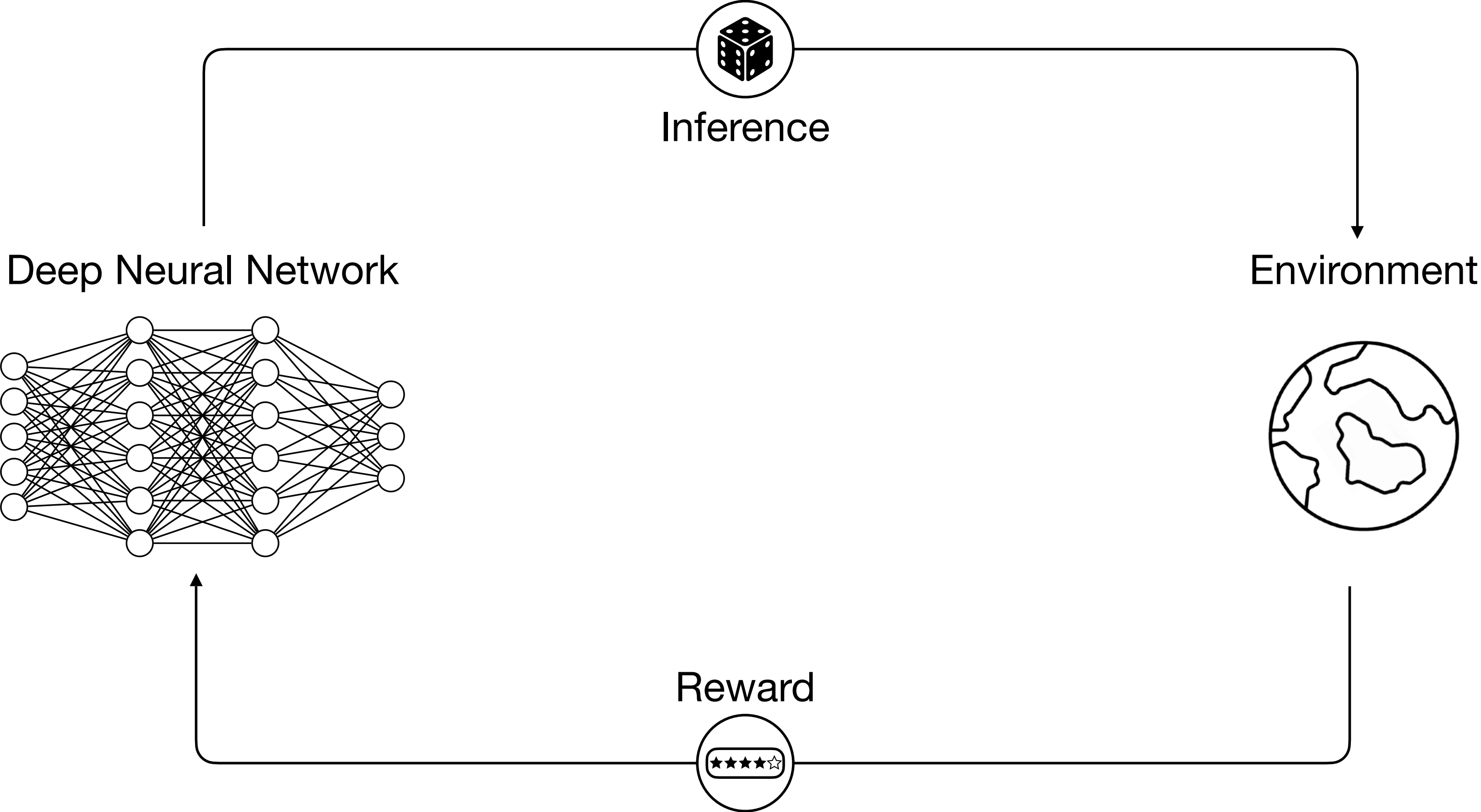


Shao-Hua Sun

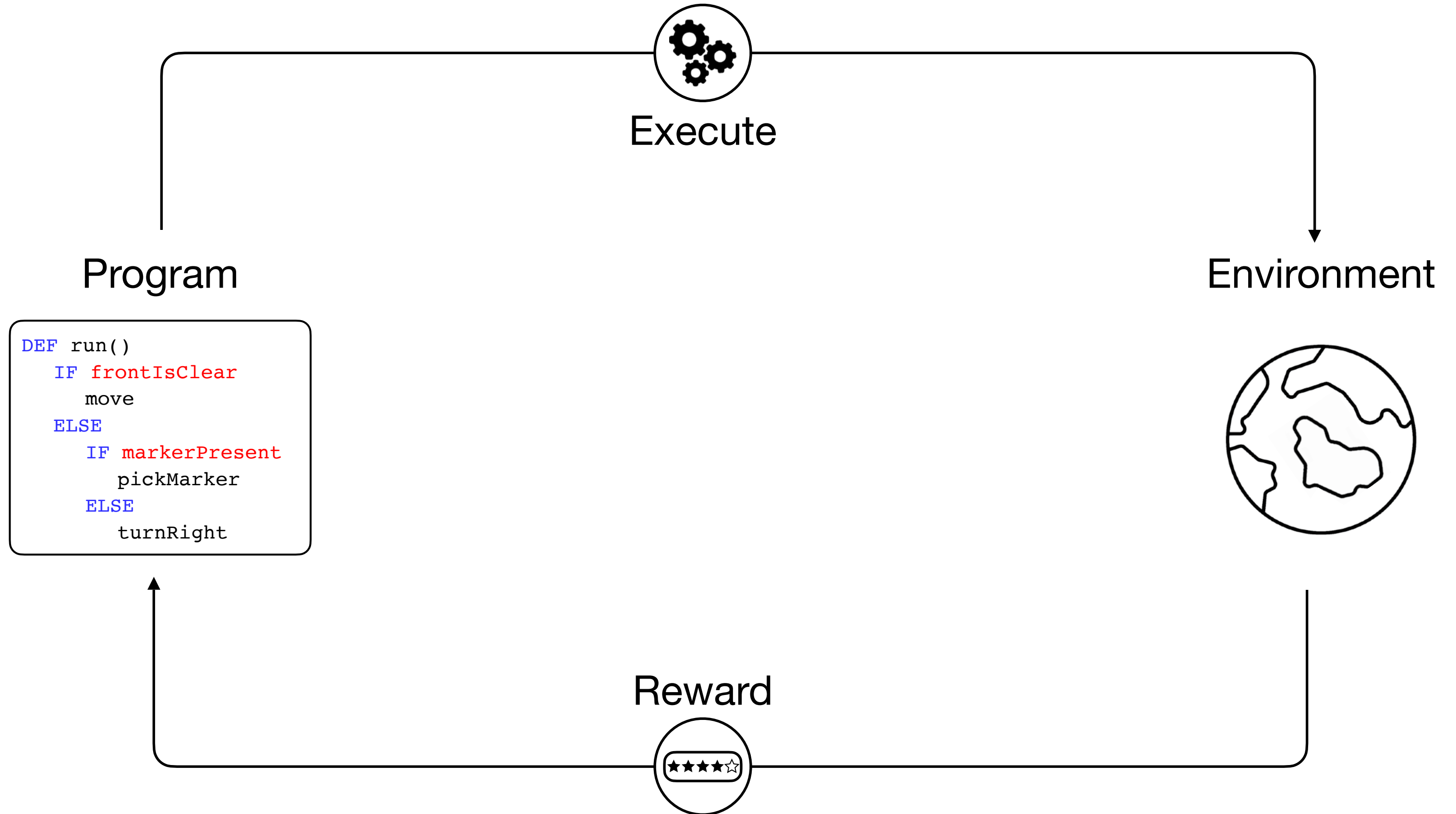


Joseph J. Lim

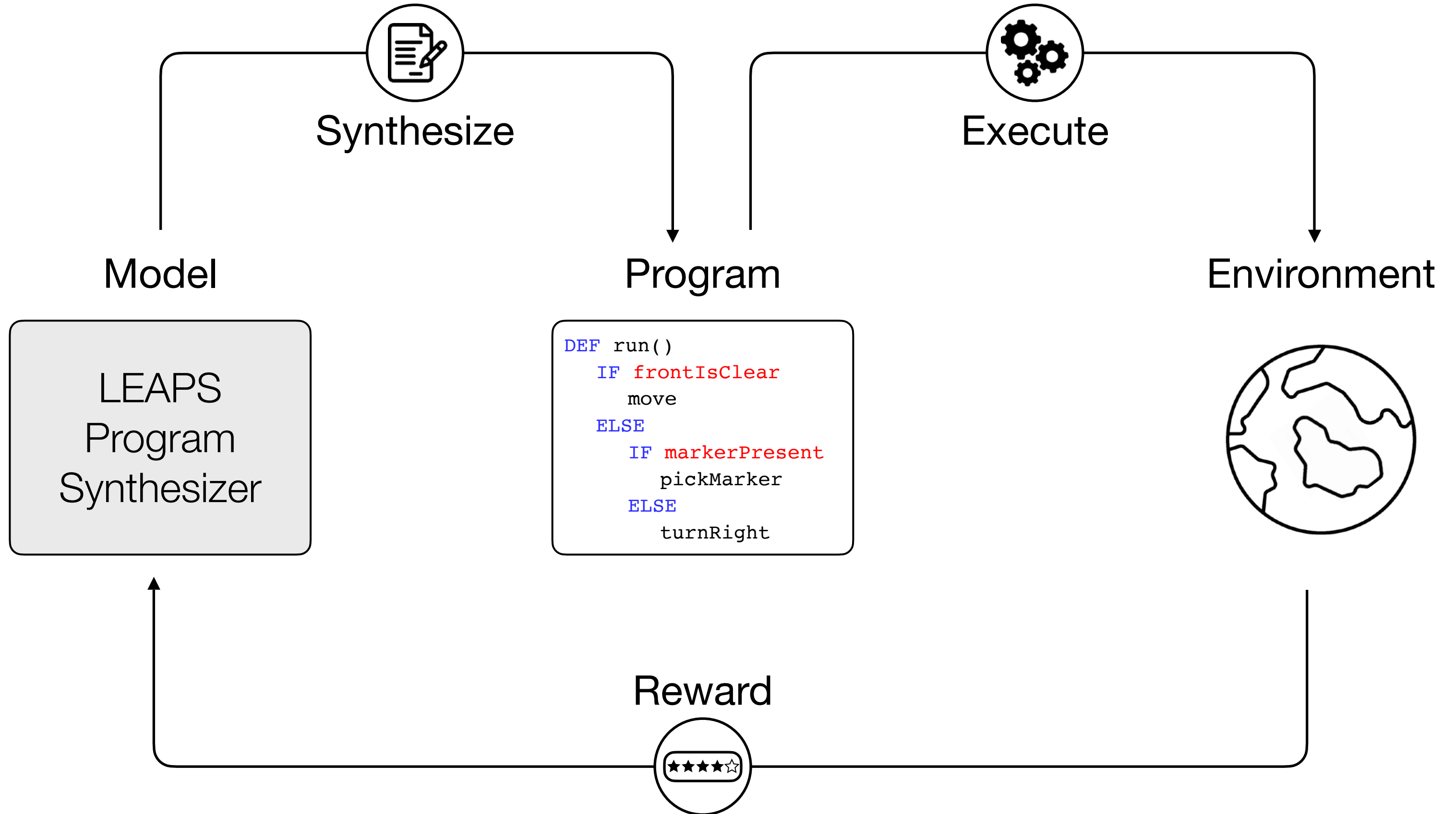
Deep Reinforcement Learning



Reinforcement Learning via Synthesizing Programs



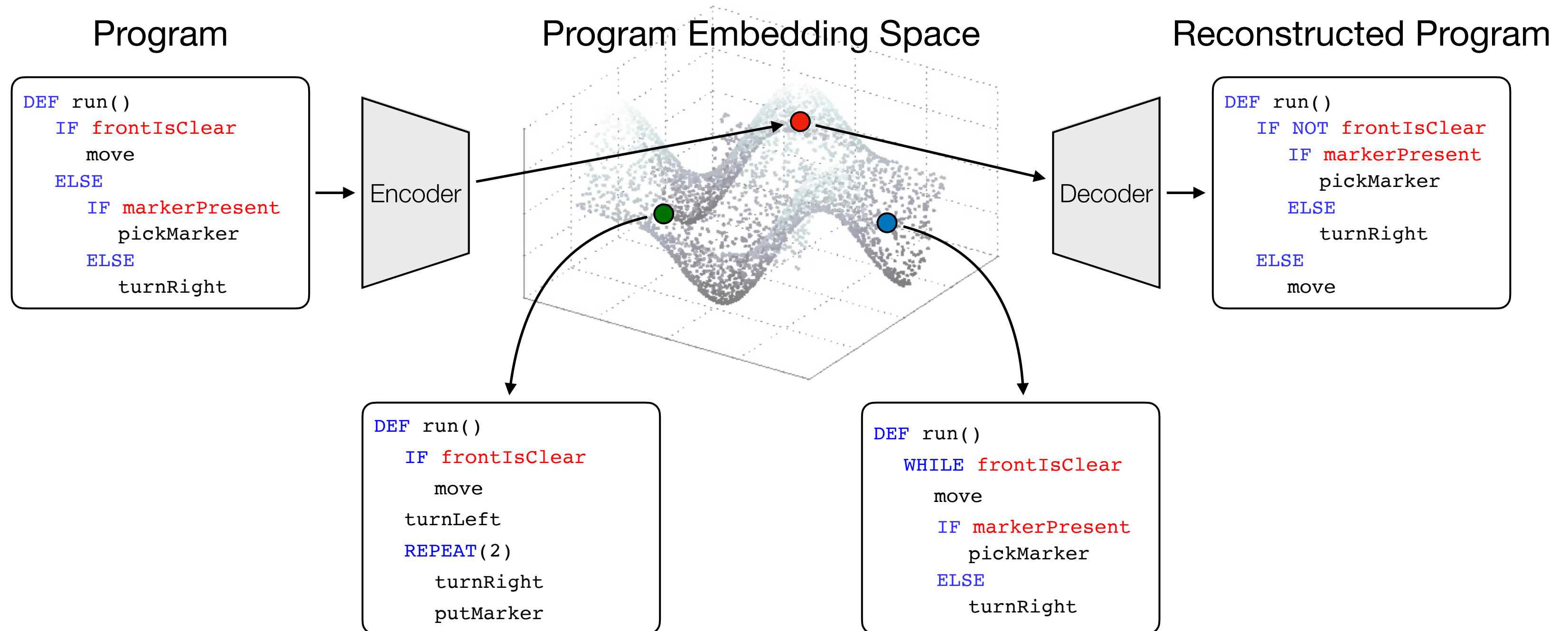
Reinforcement Learning via Synthesizing Programs



LEAPS: Learning Embeddings for Latent Program Synthesis

Stage 1 Learn a program embedding space from randomly generated programs

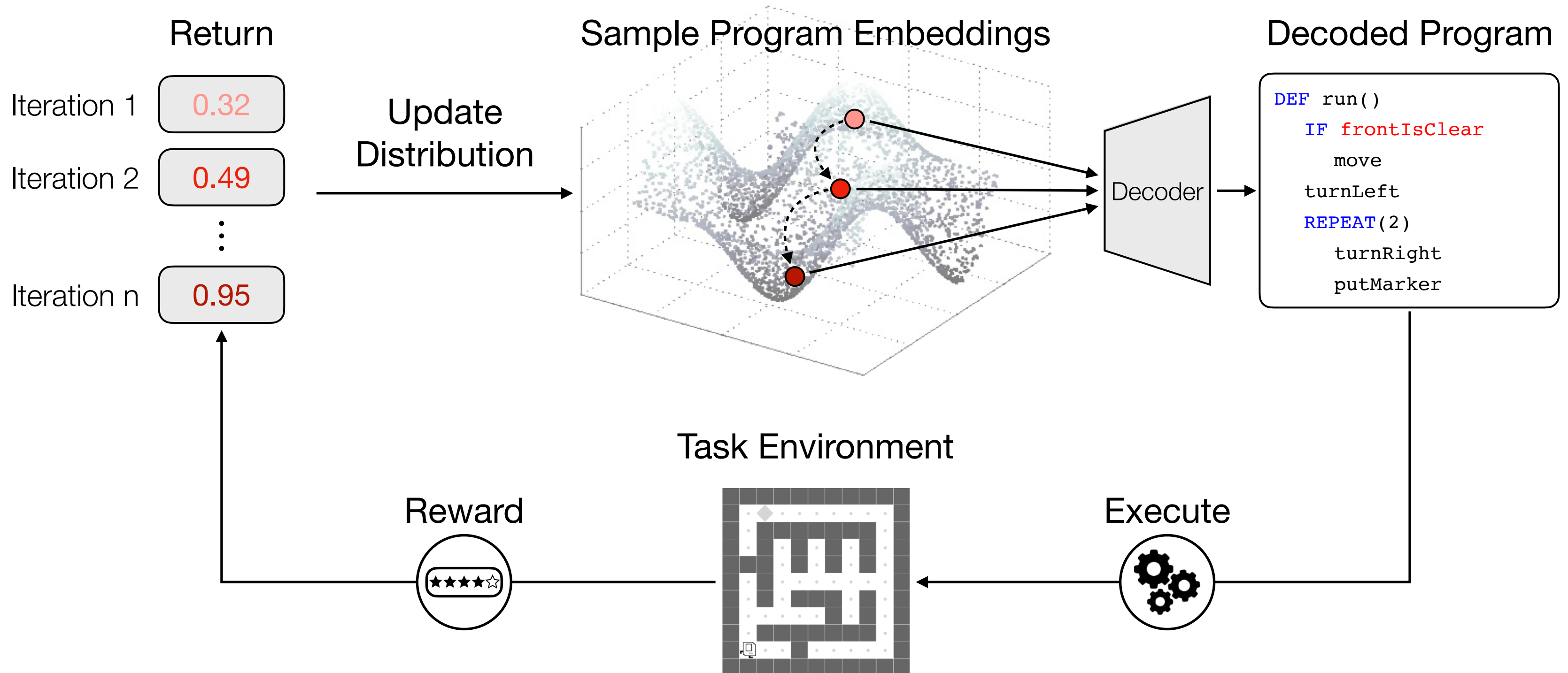
Goal Learn the **grammar** and the **environment dynamics**



LEAPS: Learning Embeddings for Latent Program Synthesis

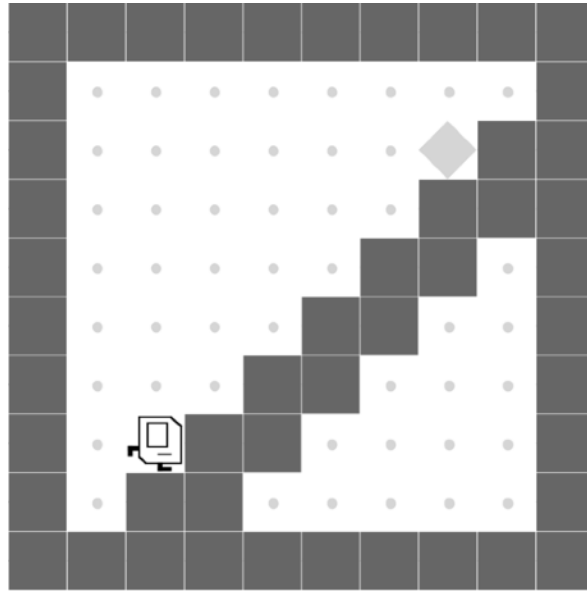
Stage 2 Search for a task-solving program using the cross-entropy method (CEM)

Goal Optimize the **task performance**

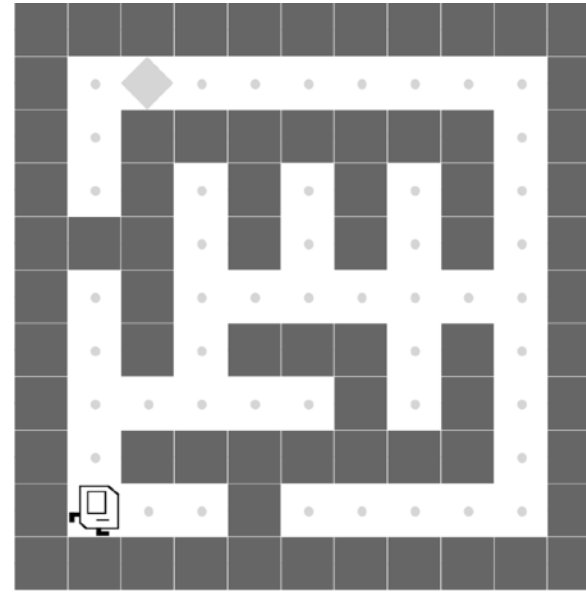


Karel Tasks

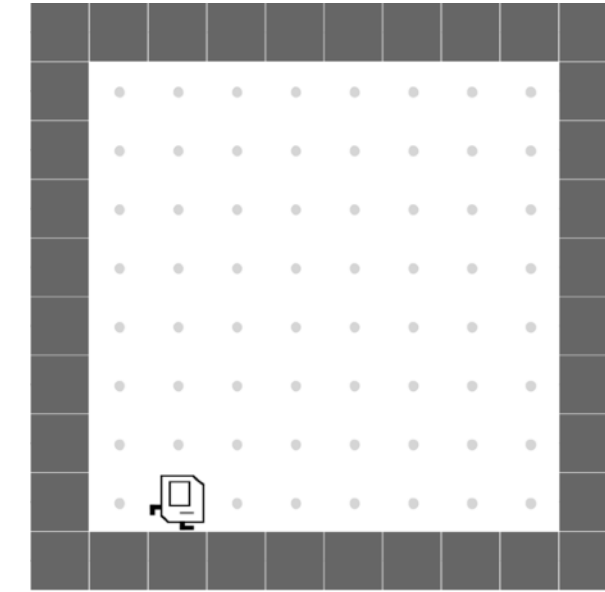
StairClimber



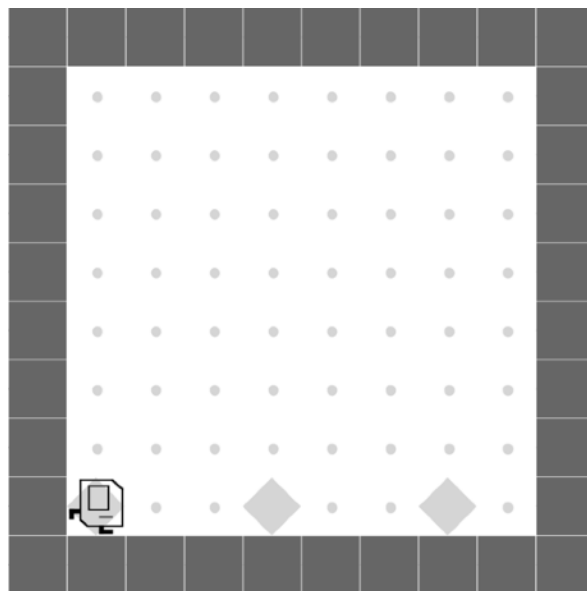
Maze



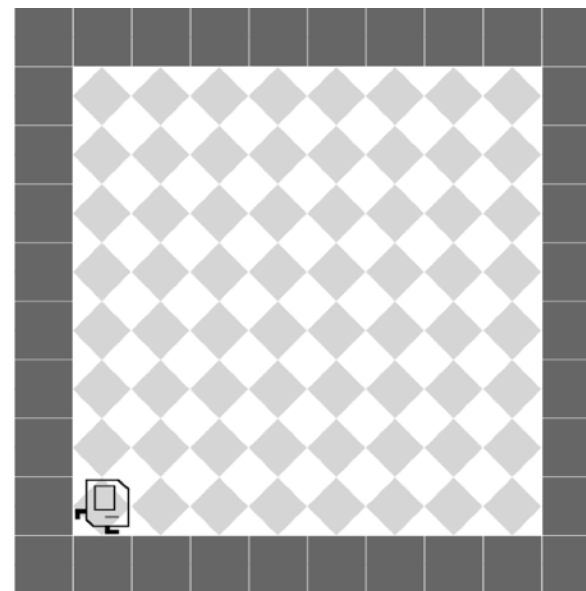
FourCorners



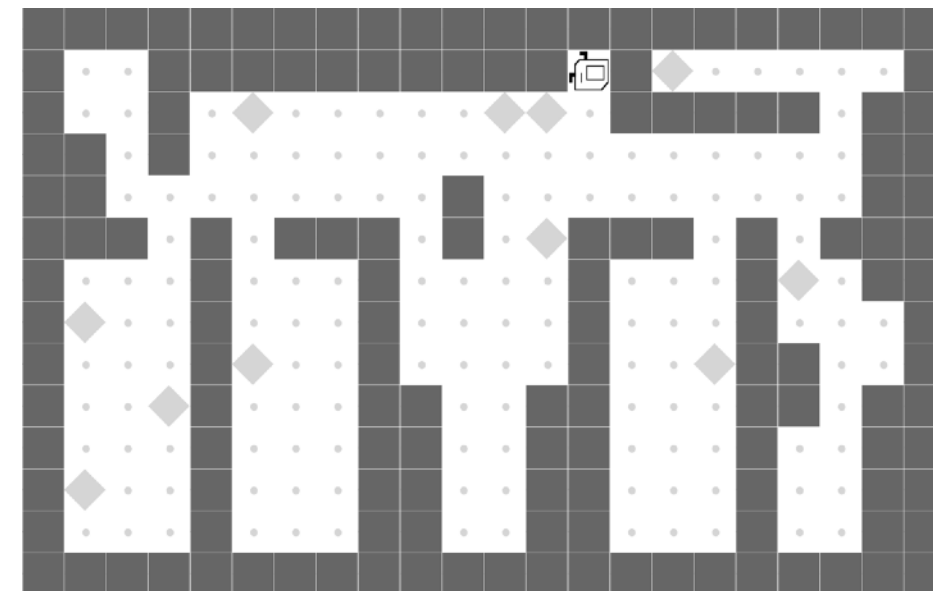
TopOff



Harvester

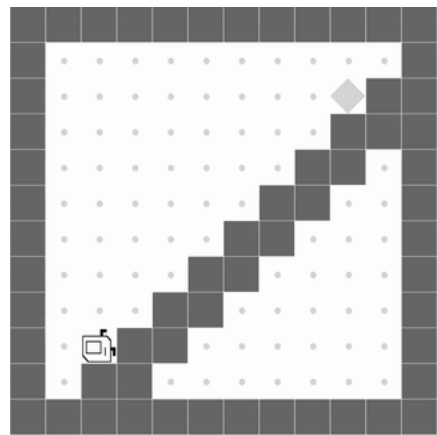


CleanHouse

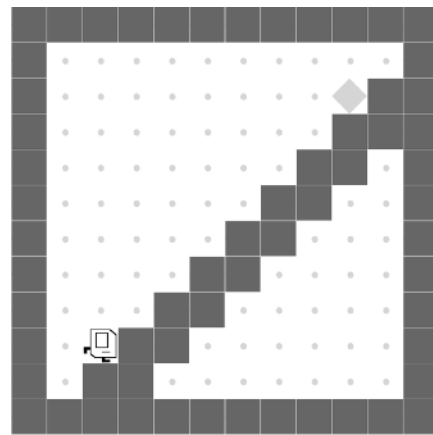


Qualitative Results

StairClimber

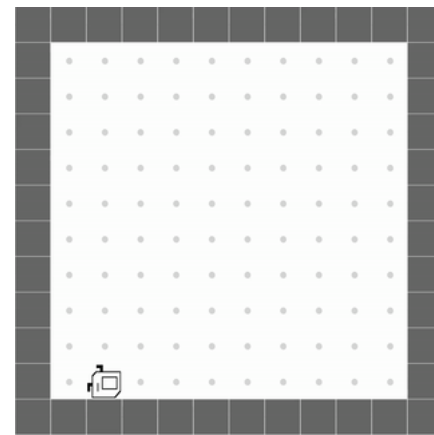


Deep RL

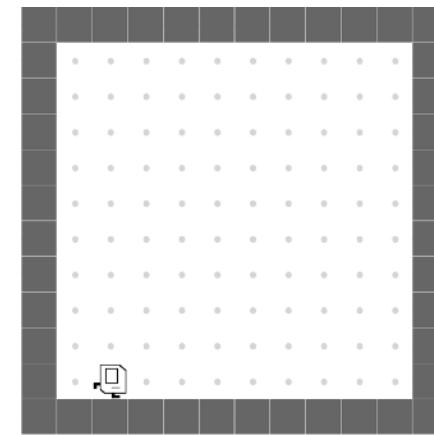


LEAPS

FourCorners

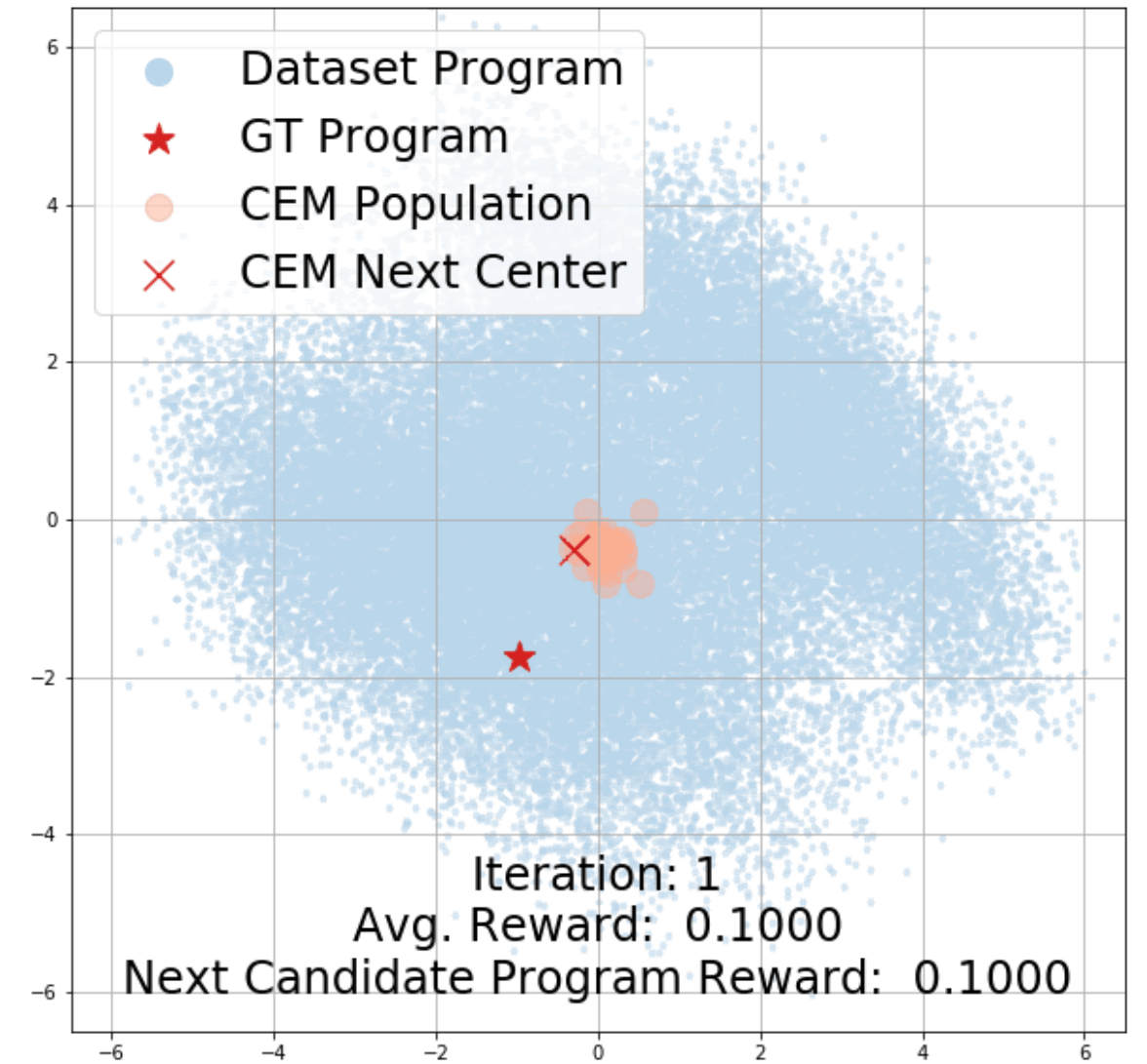


Deep RL

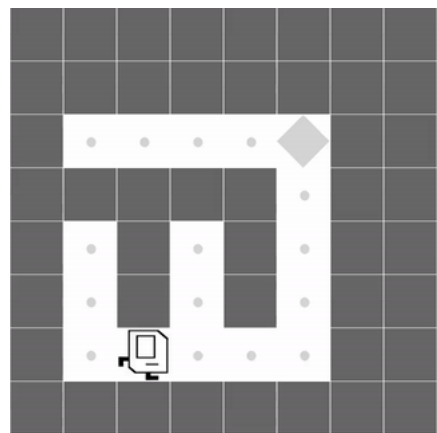


LEAPS

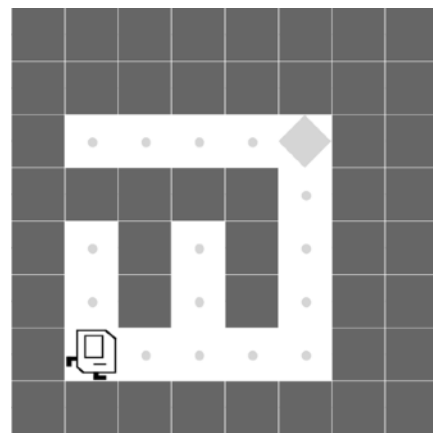
CEM trajectory Visualization



Maze

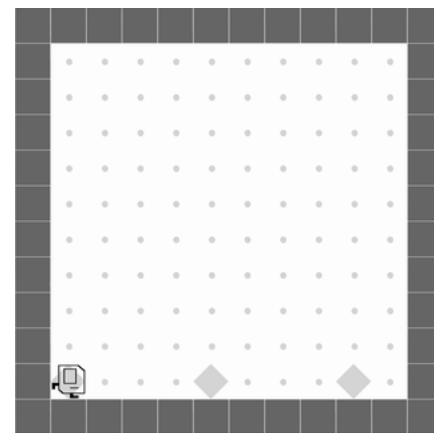


Deep RL

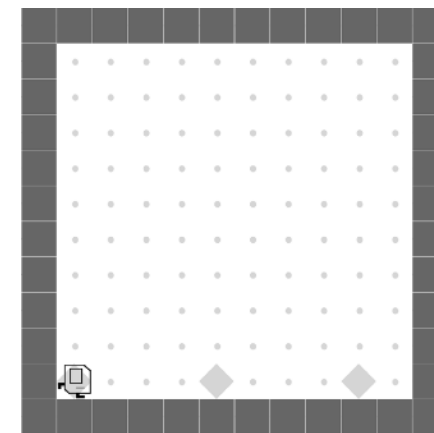


LEAPS

TopOff



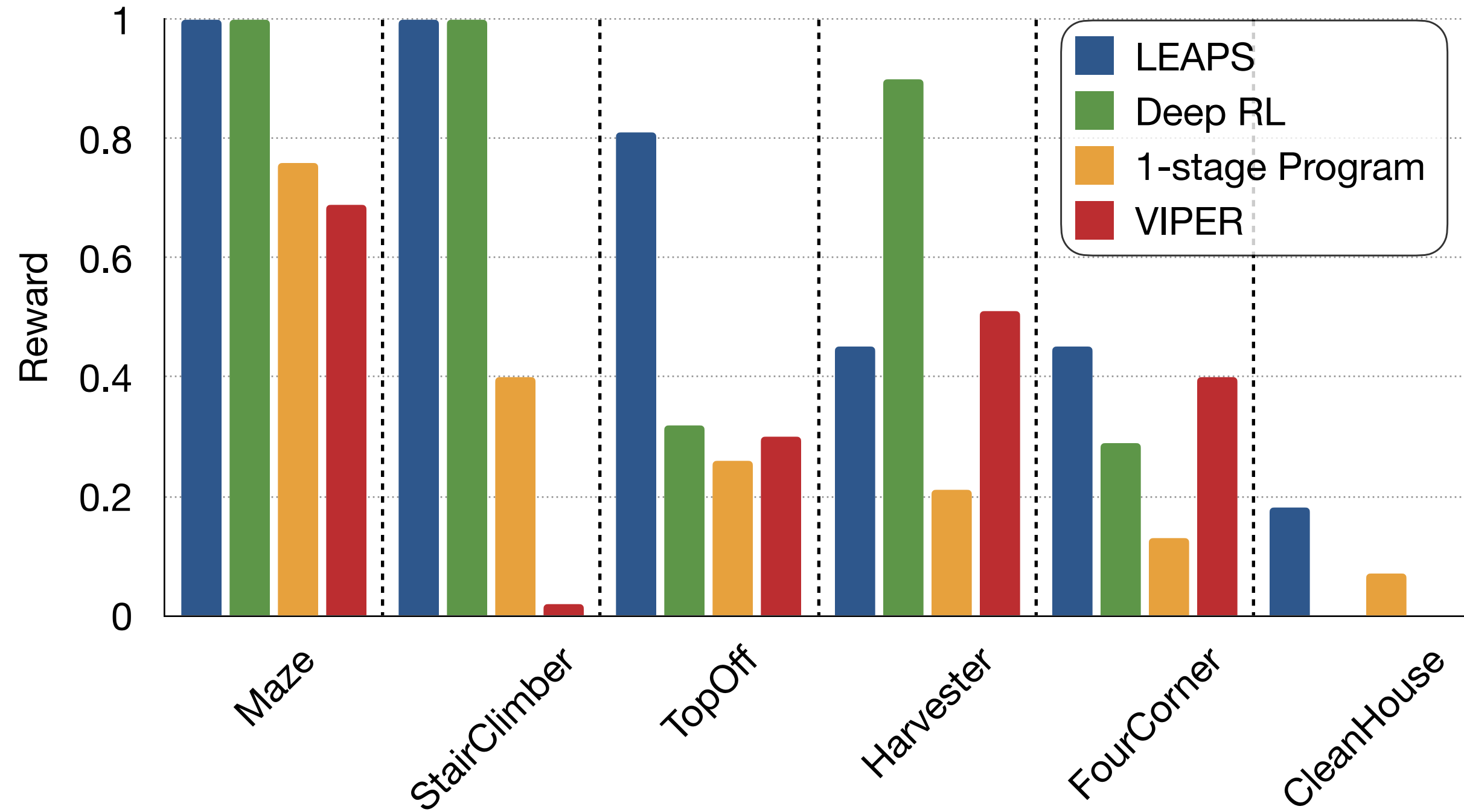
Deep RL



LEAPS

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

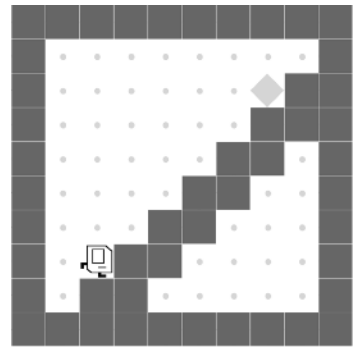
Quantitative Results



LEAPS Zero-shot Generalization

Learning on 8 x 8

StairClimber



LEAPS
Program
Synthesizer

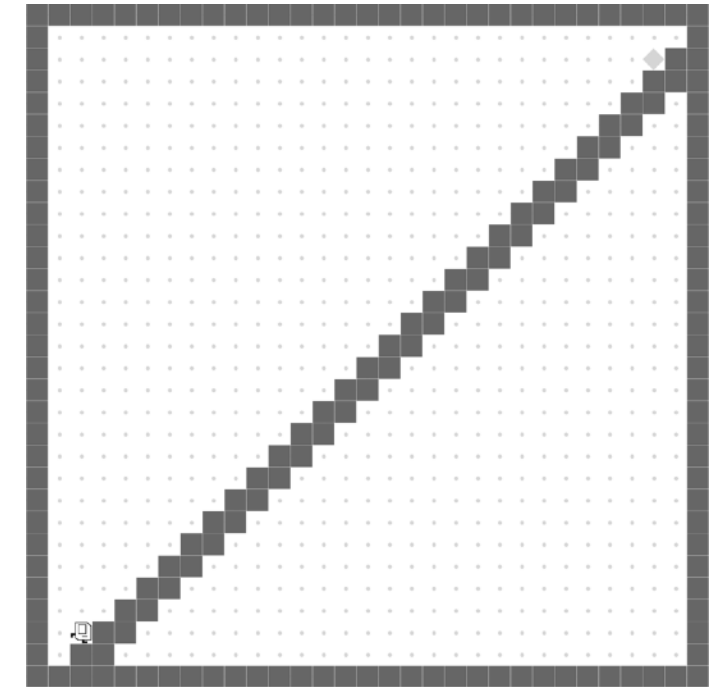


LEAPS Program Policy

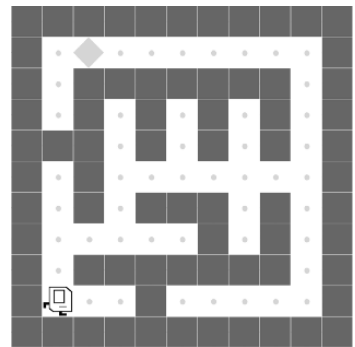
```
DEF run()  
  while noMarkersPresent()  
    turnRight  
    move  
  while rightIsClear()  
    turnLeft
```



Evaluation on 100 x 100



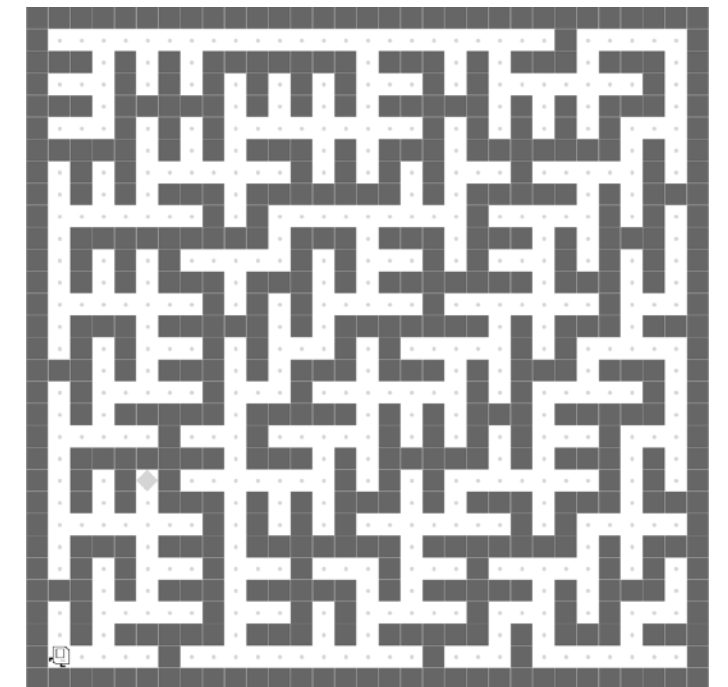
Maze



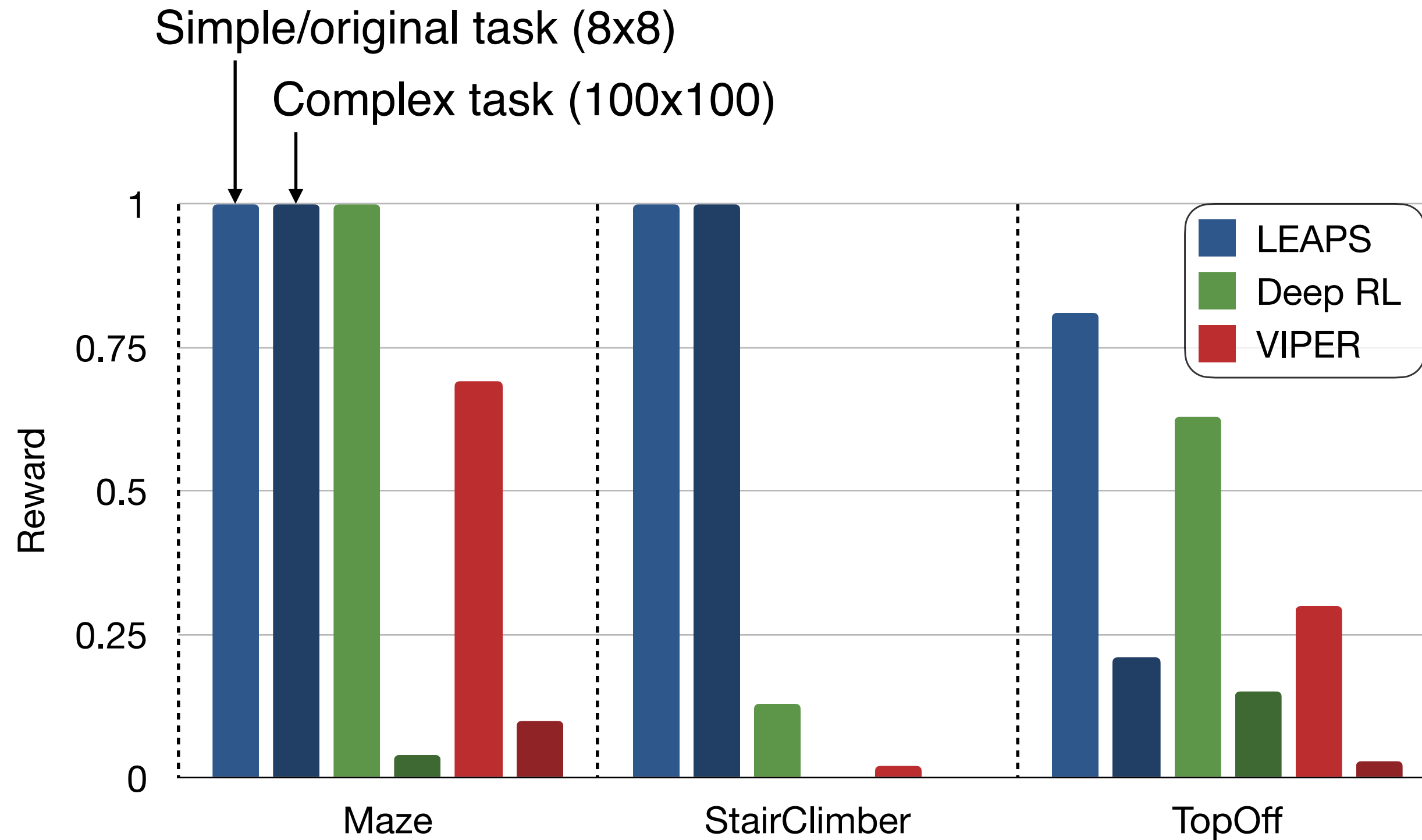
LEAPS
Program
Synthesizer



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



Experimental Results - Zero-shot Generalization



Interpretability & Interactability

Interactive Debugging Interface

```
Input Program: DEF run m(  
  WHILE c( noMarkersPresent c) w(  
    turnRight  
    move  
  w)  
  putMarker  
  move  
  WHILE c( noMarkersPresent c) w(  
    turnRight  
    move  
  w)  
  putMarker  
  move  
  WHILE c( noMarkersPresent c) w(  
    turnRight  
    move  
  w)  
  putMarker  
  move  
  WHILE c( noMarkersPresent c) w(  
    turnRight  
    move  
  w)  
  putMarker  
  move  
  m)
```

Reset Code (Made a mistake?)

Issue with Code? None

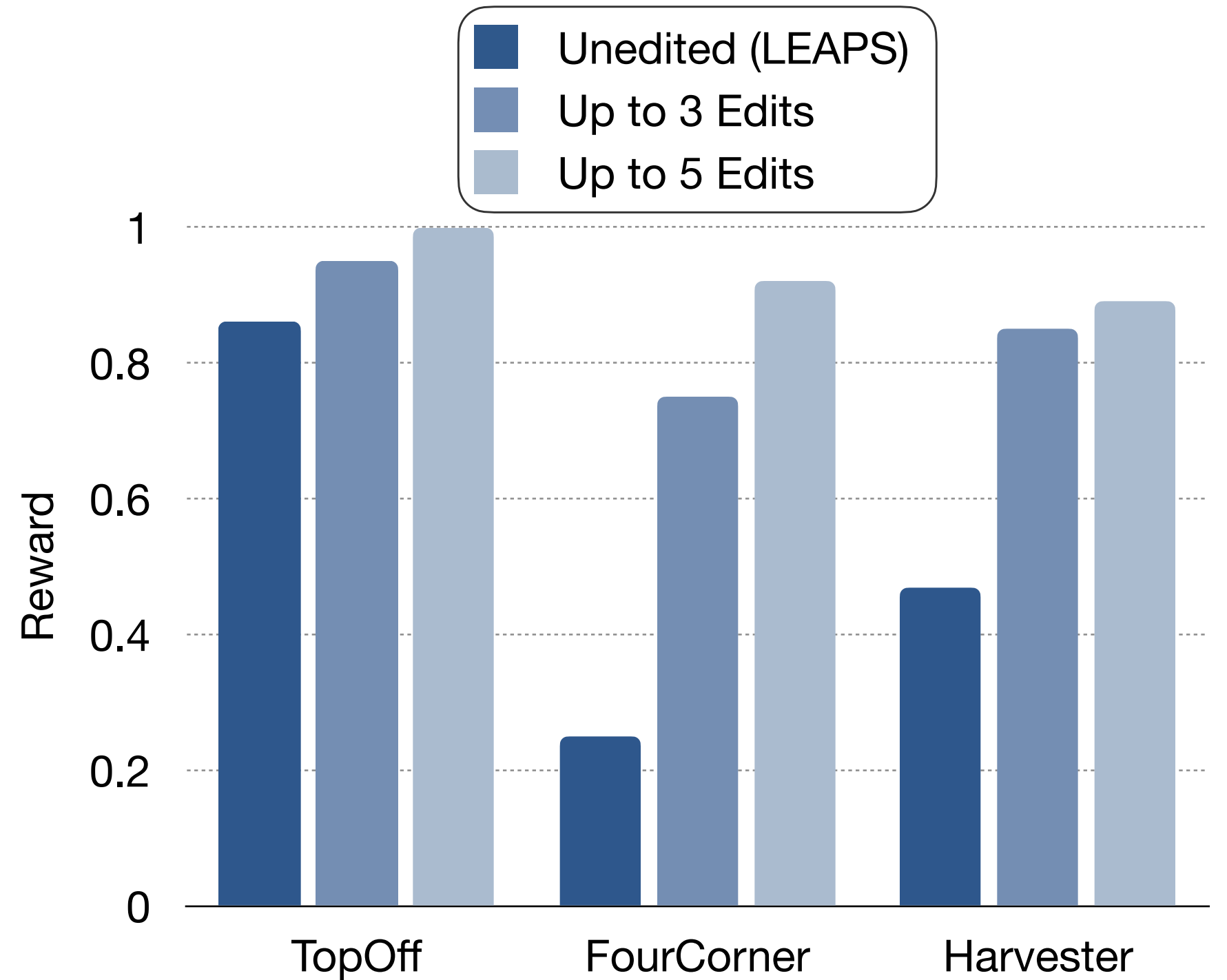
Submit Code (Runs your program!)

New Reward: 0.36000000774860386

Orig Reward: 0.863636384010315

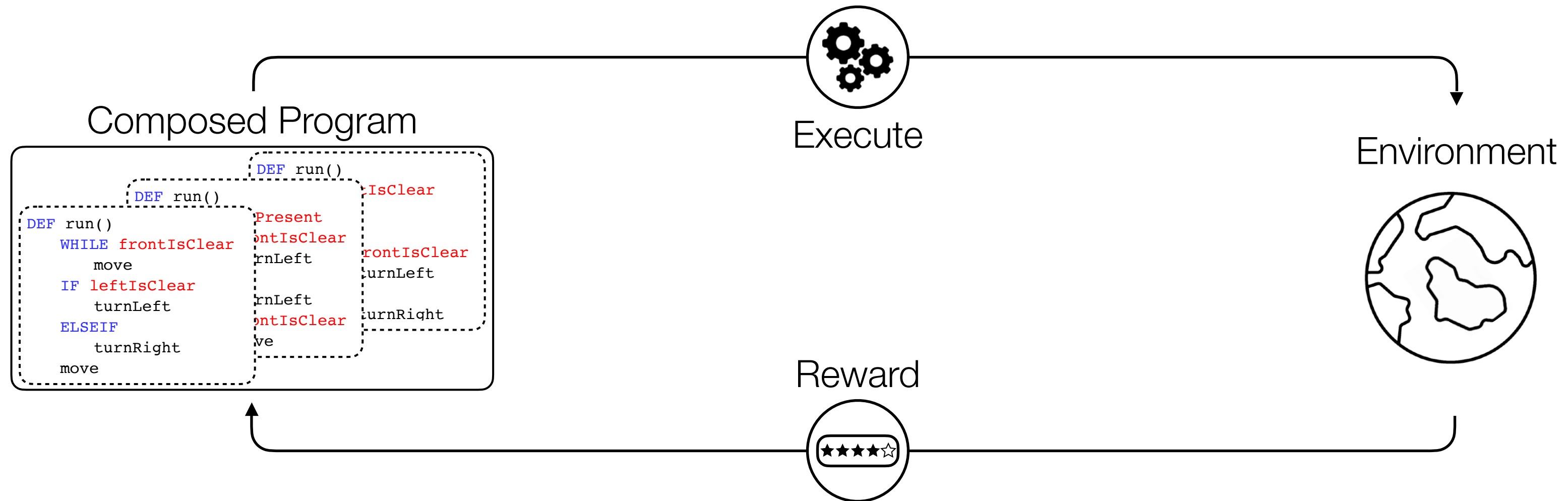
Best Reward: 0.863636384010315

Performance Improvement



Hierarchical Programmatic Reinforcement Learning via Learning to Compose Programs

ICML 2023



Guan-Ting Liu*



En-Pei Hu*



Pu-Jen Cheng



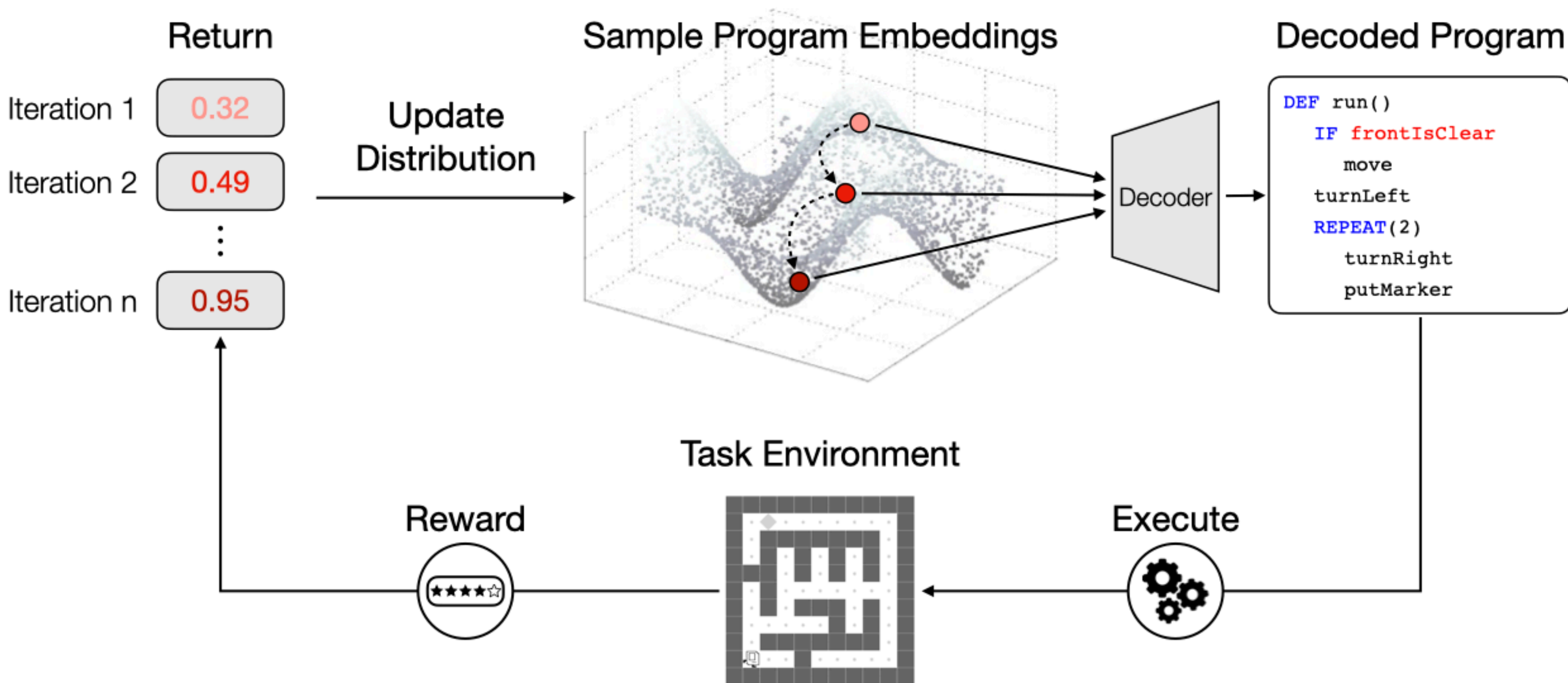
Hung-Yi Lee



Shao-Hua Sun

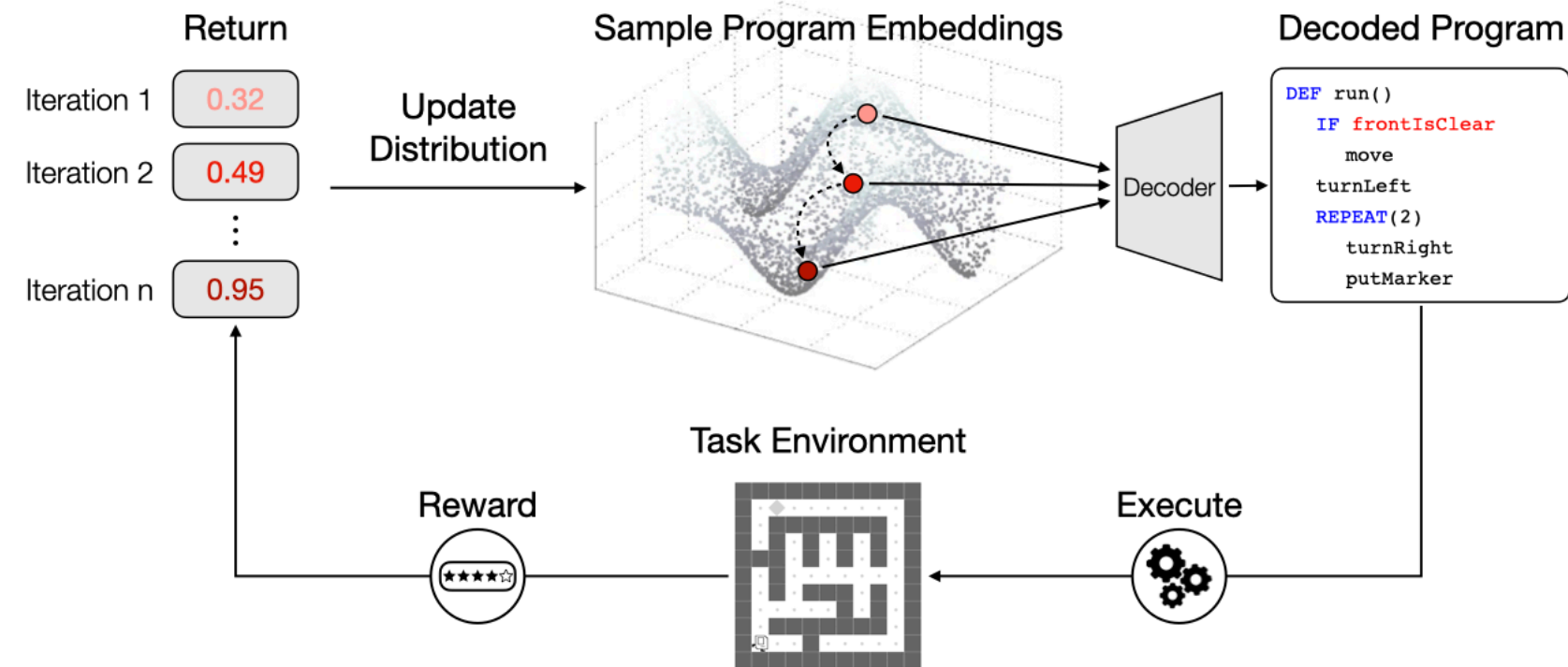
LEAPS: Learning Embeddings for Latent Program Synthesis

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



LEAPS: Learning Embeddings for Latent Program Synthesis

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

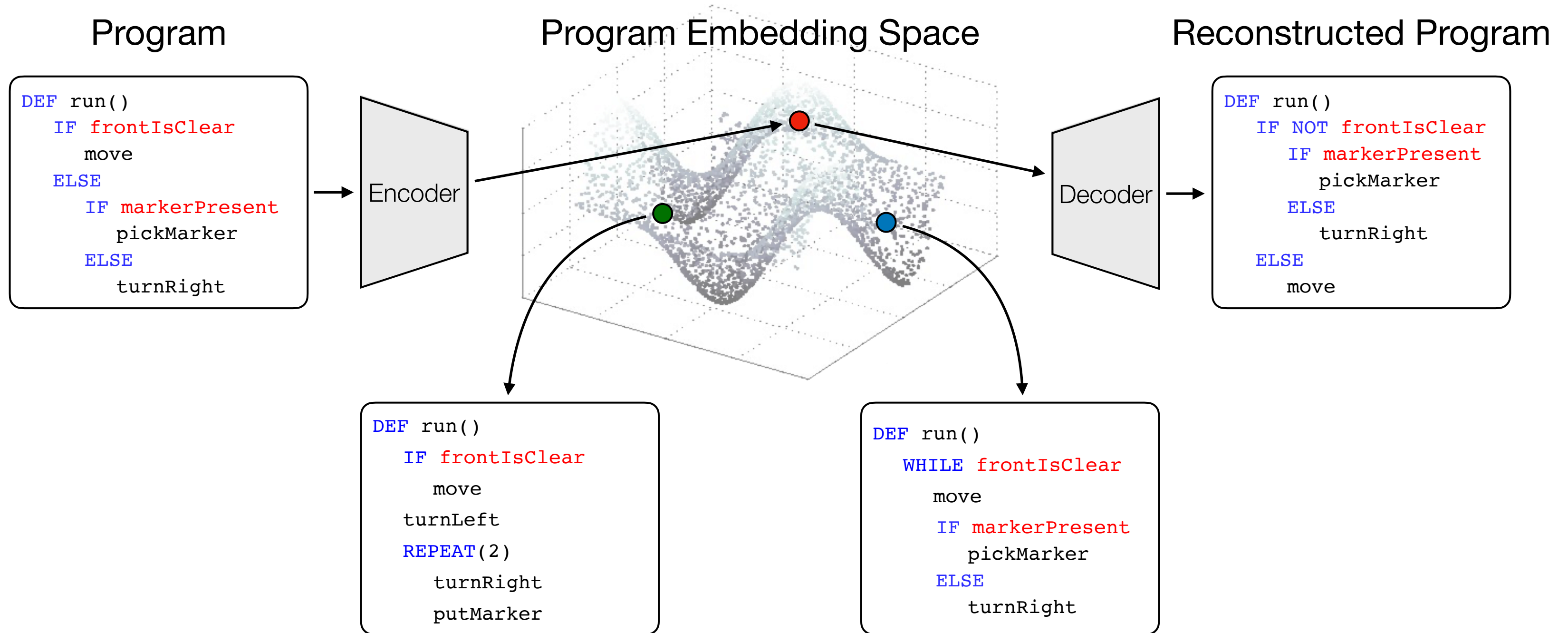
Evaluate each candidate program solely based on the cumulative return of its execution trace



Cannot accurately attribute rewards to corresponding program parts

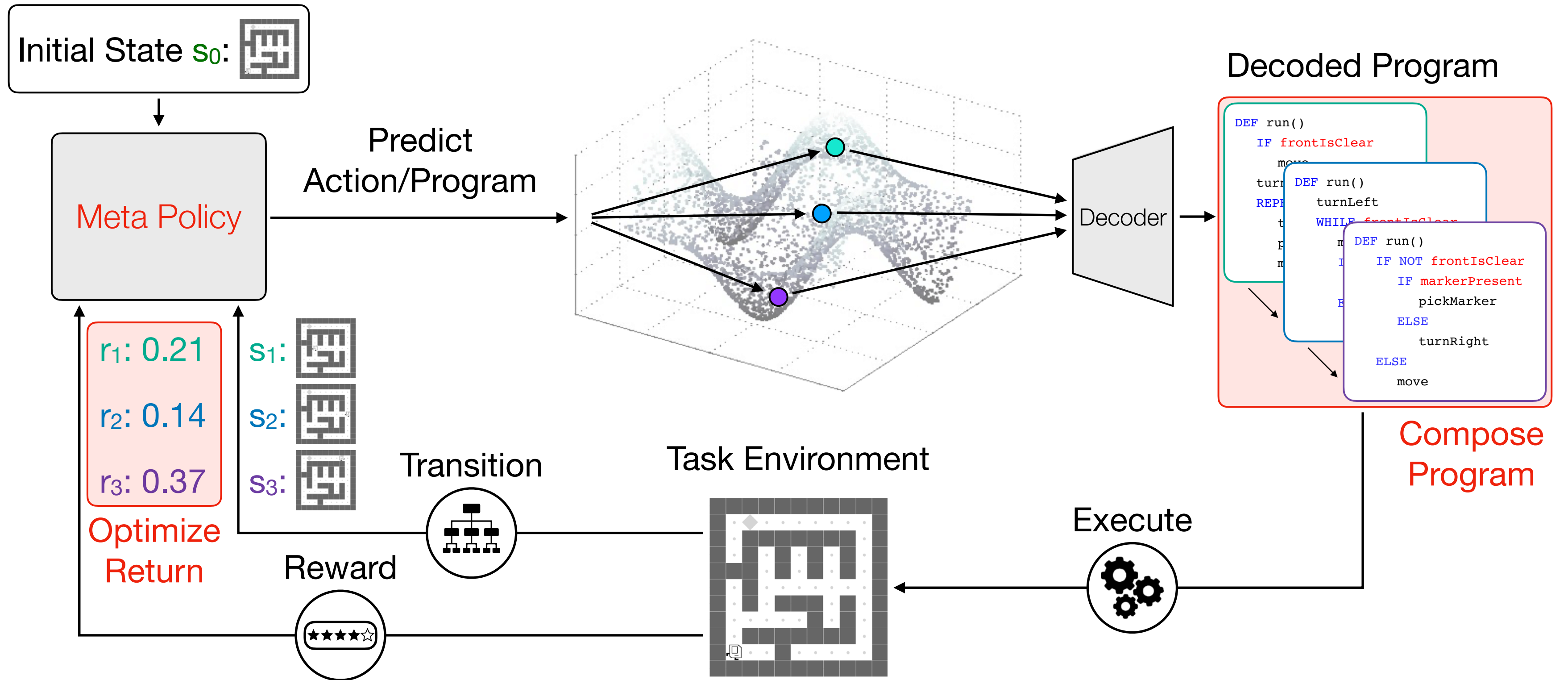
HPRL: Hierarchical Programmatic Reinforcement Learning

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

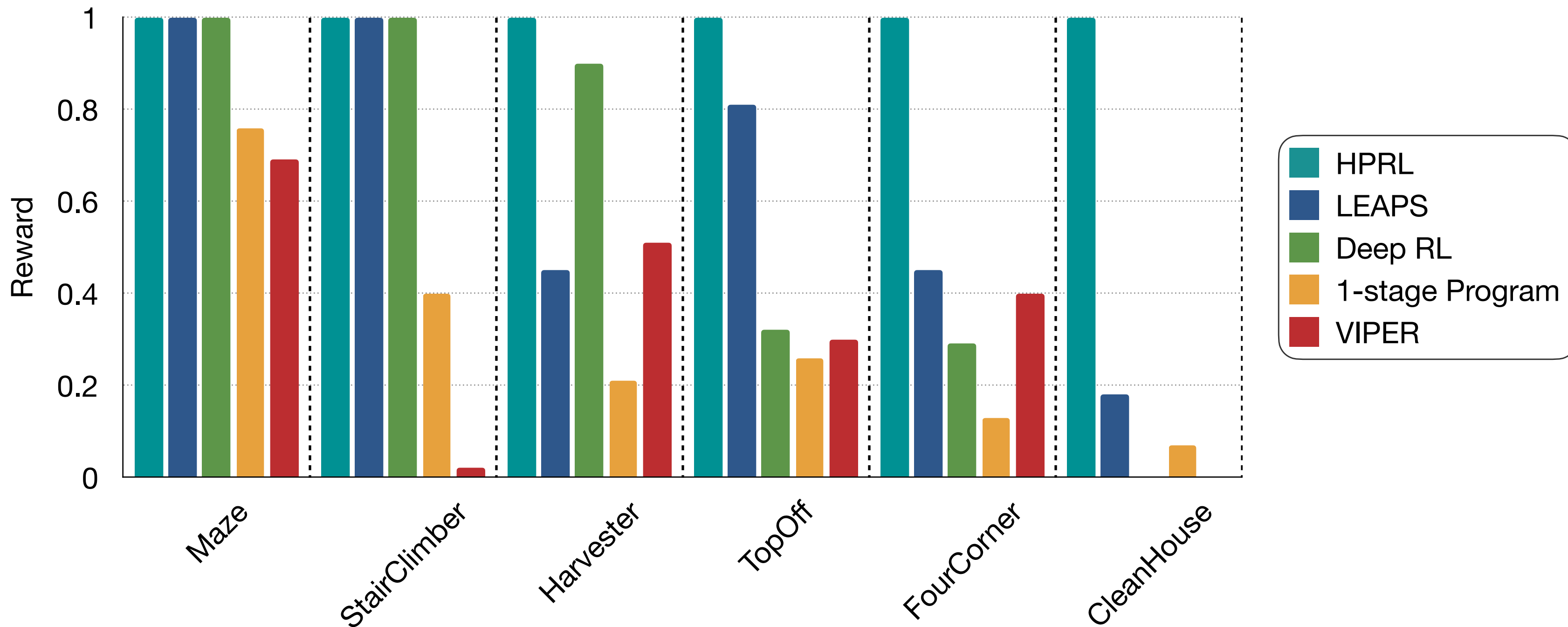


HPRL: Hierarchical Programmatic Reinforcement Learning

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

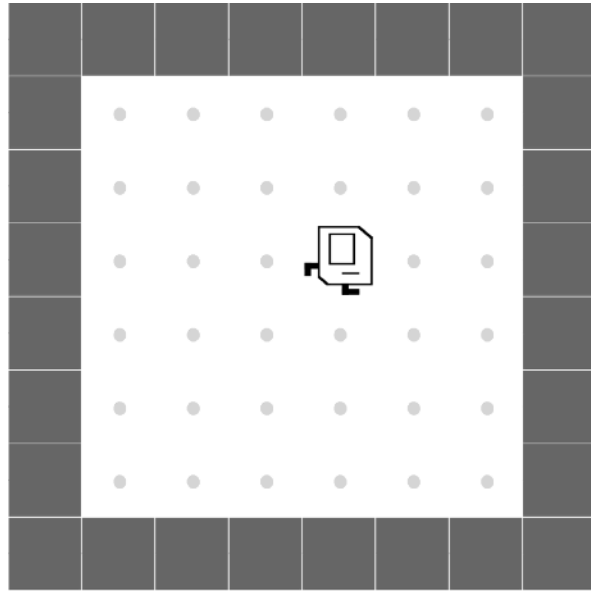


Quantitative Results - Karel Tasks

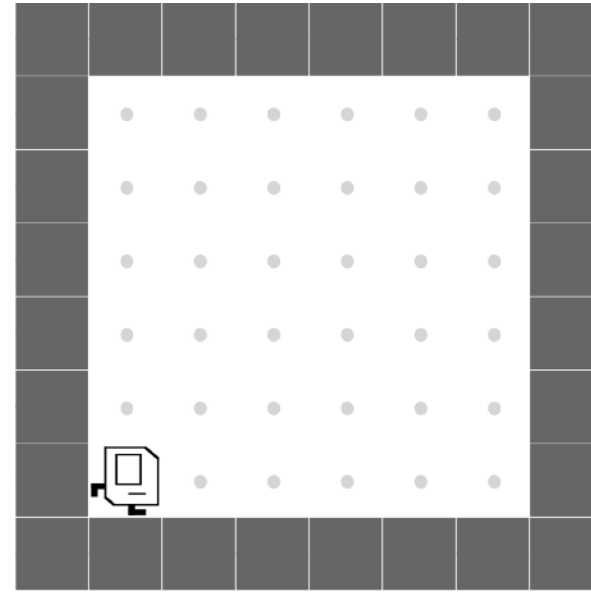


Karel-Hard Tasks

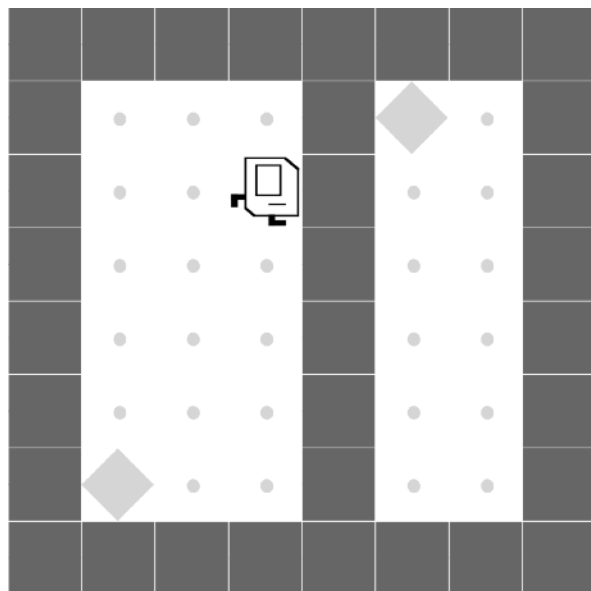
OneStroke



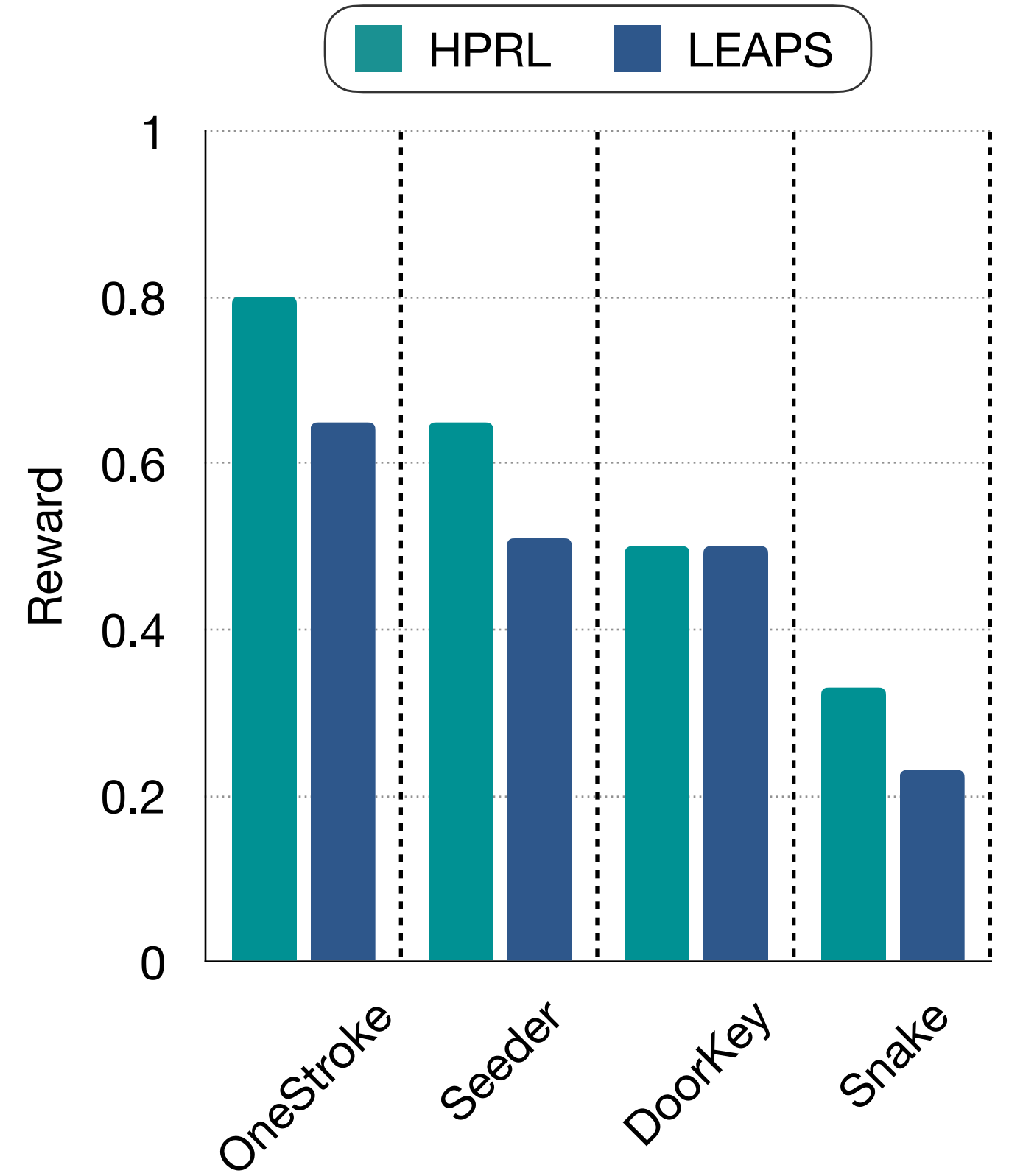
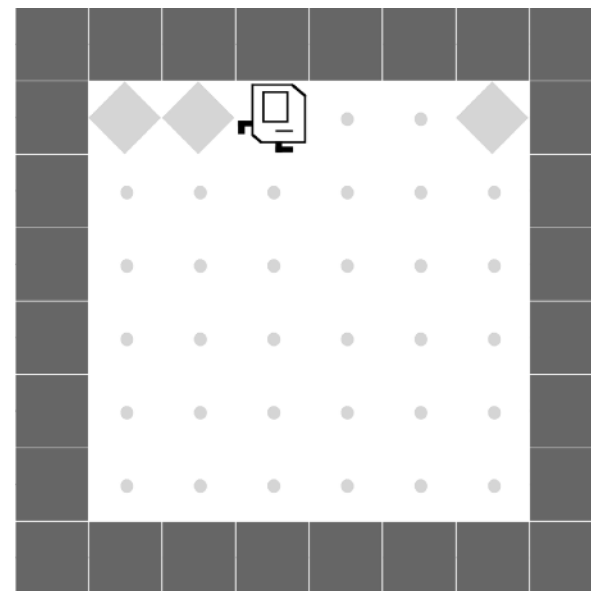
Seeder



DoorKey



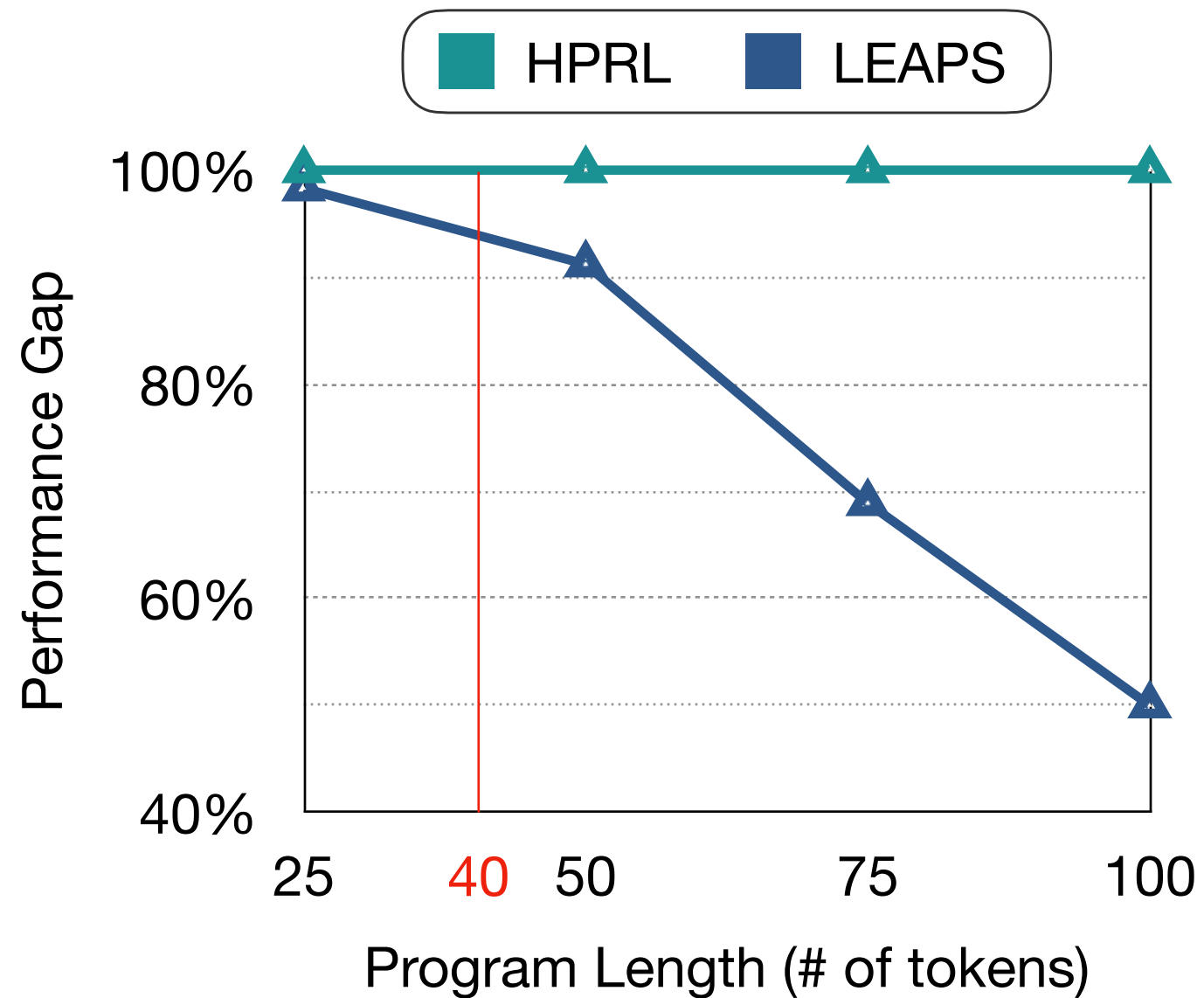
Snake



Additional Experiments

Limited program distribution

Synthesize out-of-distributionally long programs

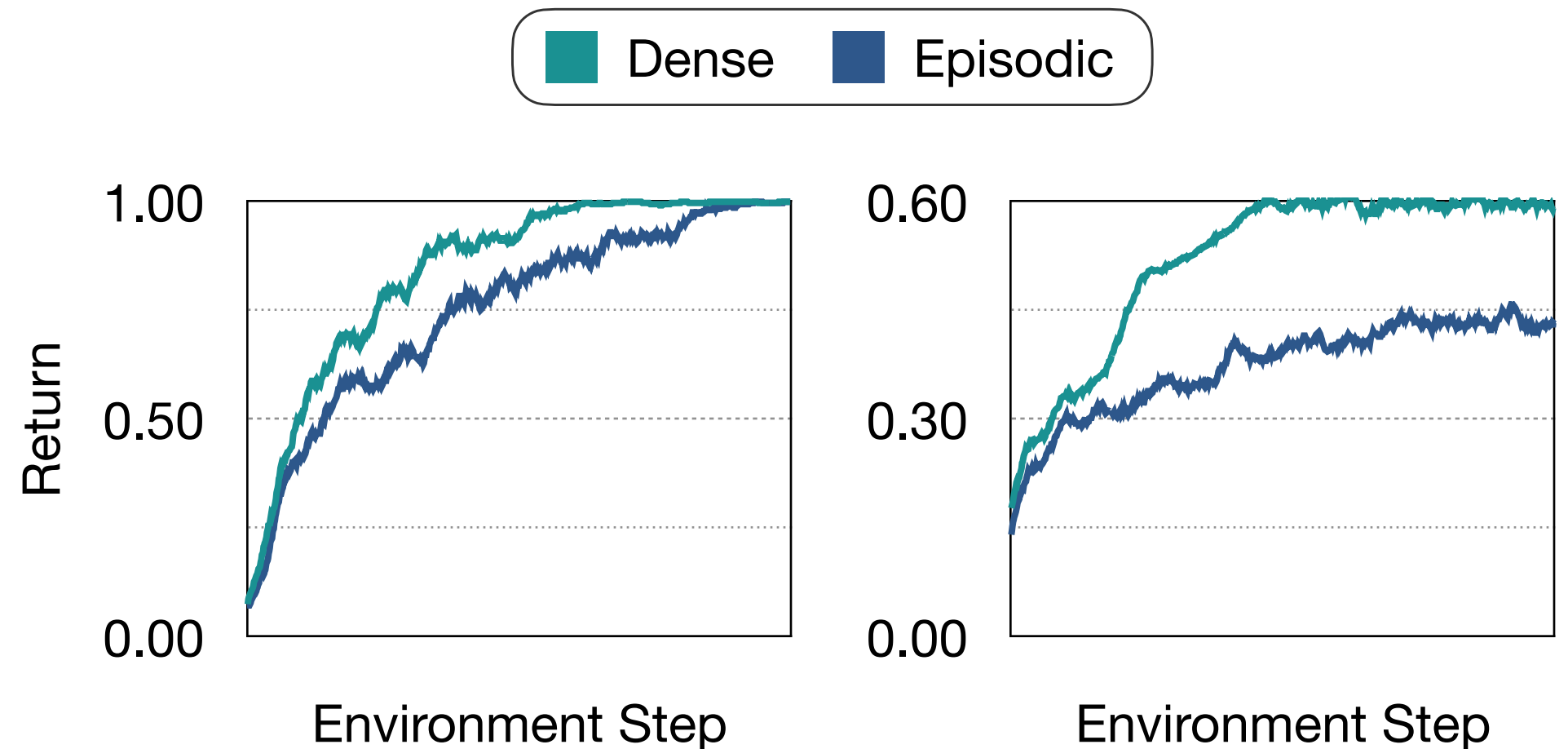


- **HPRL** can synthesize programs longer than the dataset programs (< 40 tokens) better than **LEAPS**

Poor credit assignment

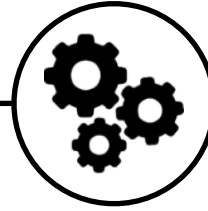
Learning from episodic reward

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



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

Execute



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)

```

Environment

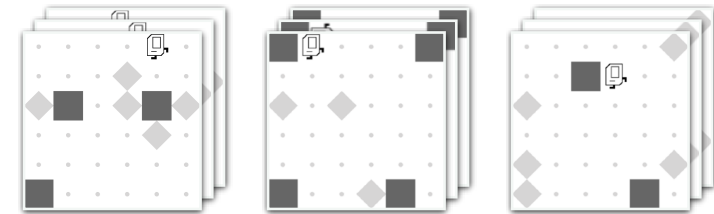


Takeaways

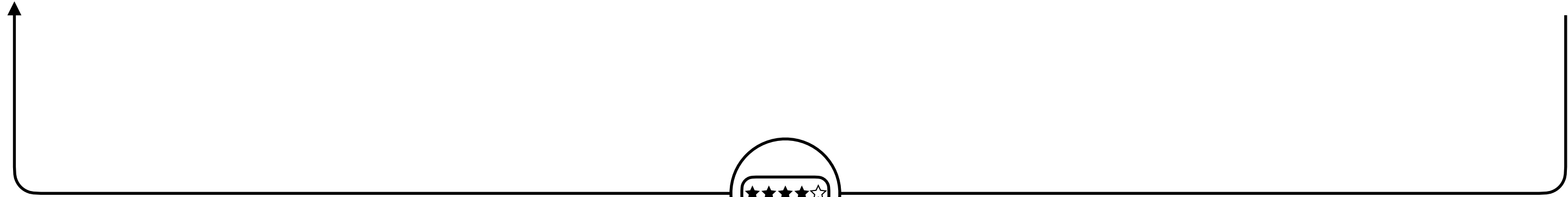
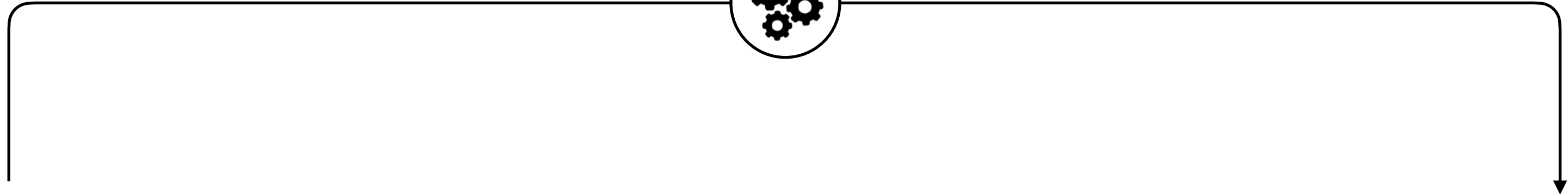
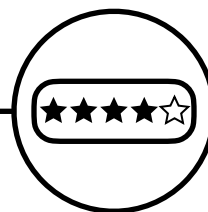
Program Synthesis × Reinforcement Learning

= Interpretable and Generalizable Policies

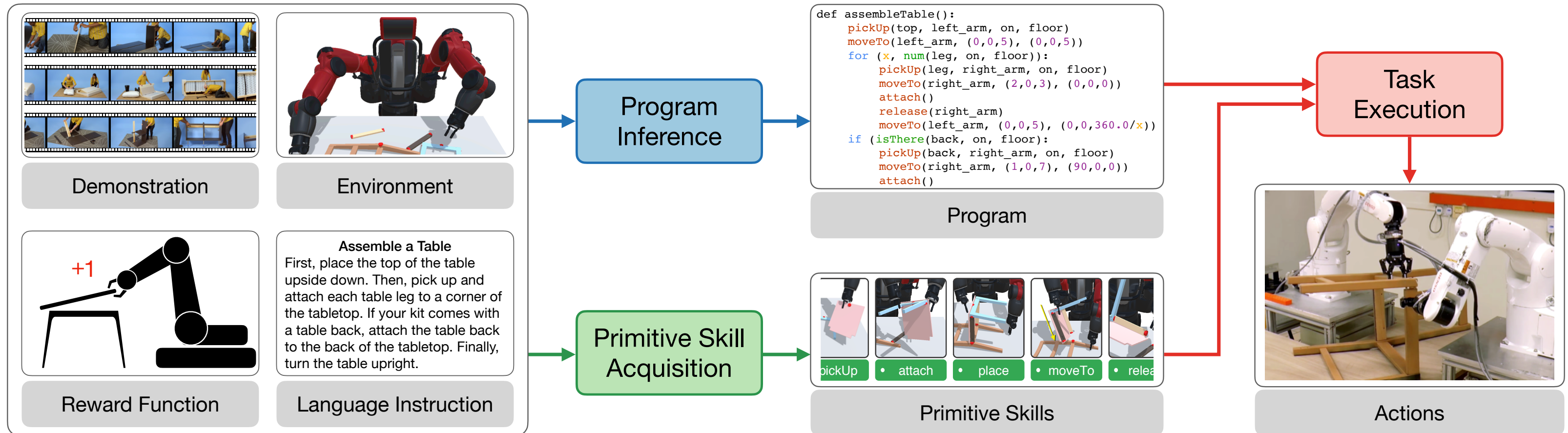
Demonstrations



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

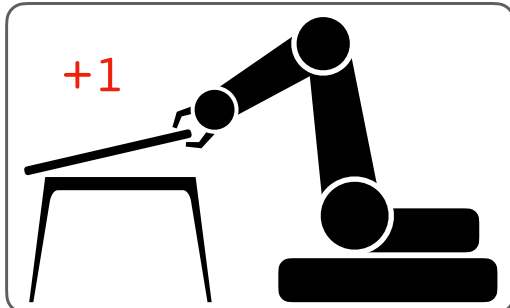
- Described with formal languages
- Human interpretable and machine executable
- Structured for **generalization**

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

Actions (sub-skills)

attach pickUp moveTo place release

Perceptions

isThere() num() isAttached()

Control Flows

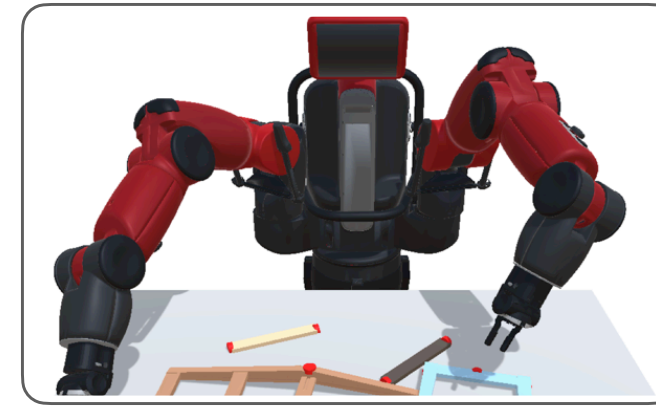
if else elif

while repeat

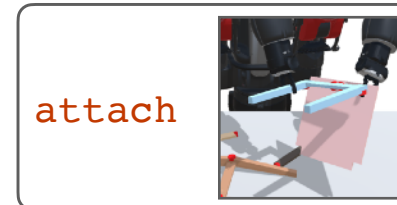
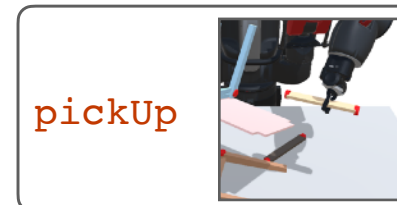
Program

```
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()
```

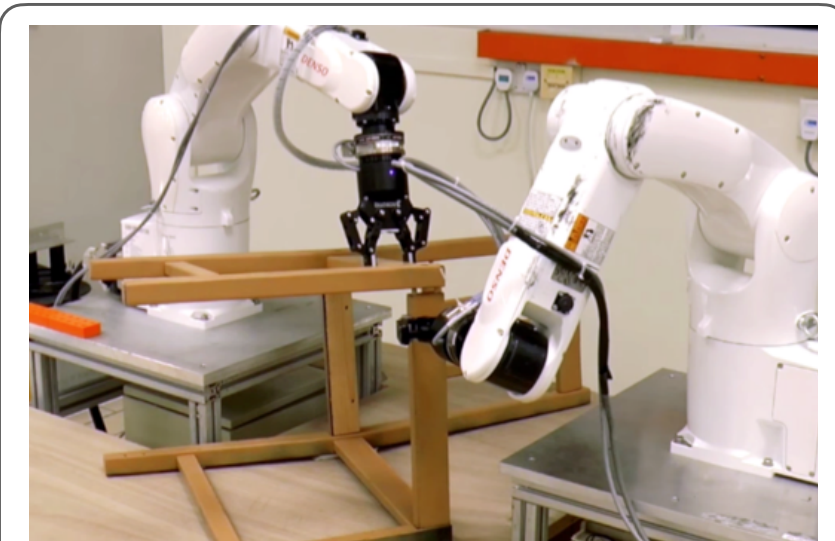
Observation



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]
```

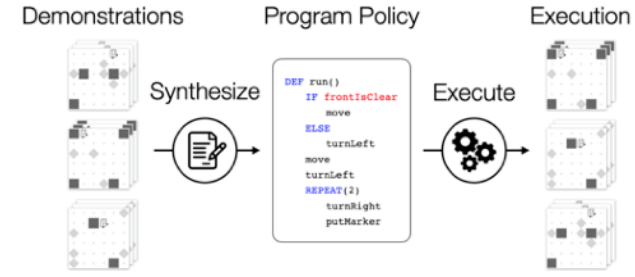
Joint v torque

```
[-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

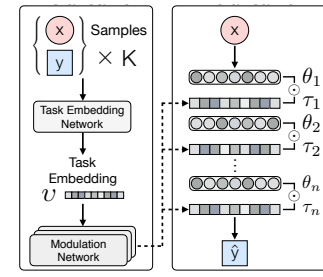
Neural Program Synthesis from Diverse Demonstration Videos



ICML 2018

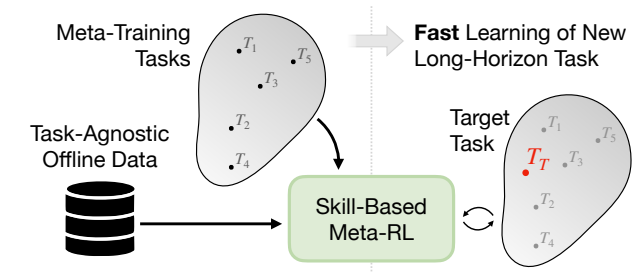
Primitive Skill Acquisition

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



NeurIPS 2019 (Spotlight)

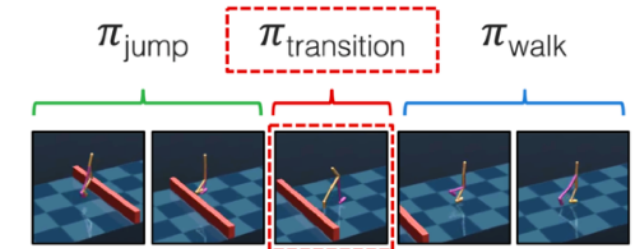
Skill-based Meta-Reinforcement Learning



ICLR 2022

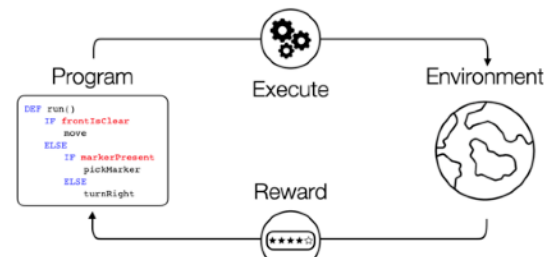
Task Execution

Composing Complex Skills by Learning Transition Policies



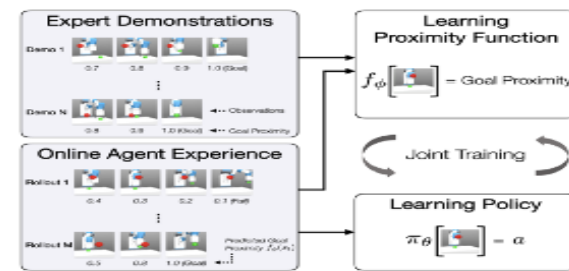
ICLR 2019

Learning to Synthesize Programs as Interpretable and Generalizable Policies



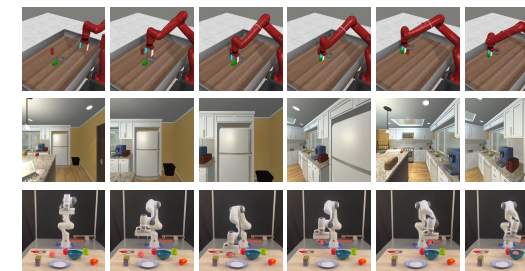
NeurIPS 2021

Generalizable Imitation Learning from Observation via Inferring Goal Proximity



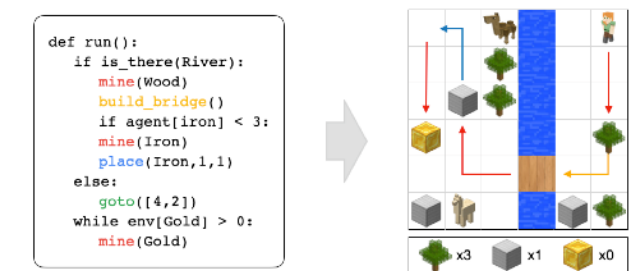
NeurIPS 2021

Learning to Act from Actionless Videos through Dense Correspondences



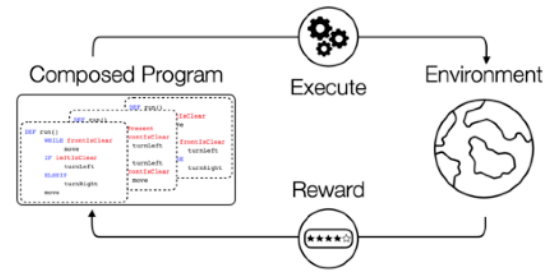
ICLR 2024 (Spotlight)

Program Guided Agent



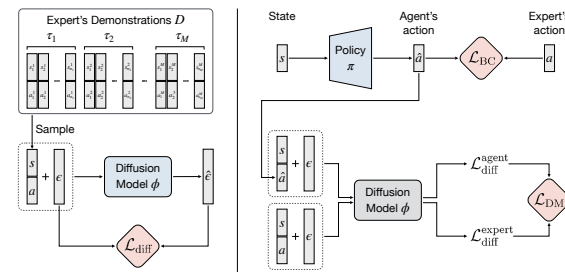
ICLR 2020 (Spotlight)

Hierarchical Programmatic Reinforcement Learning via Learning to Compose Programs



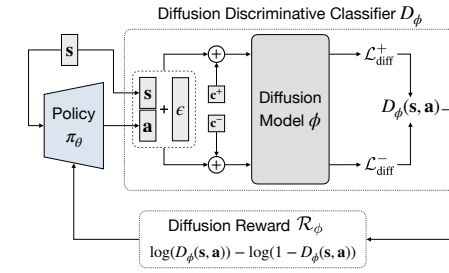
ICML 2023

Diffusion Model-Guided Behavioral Cloning



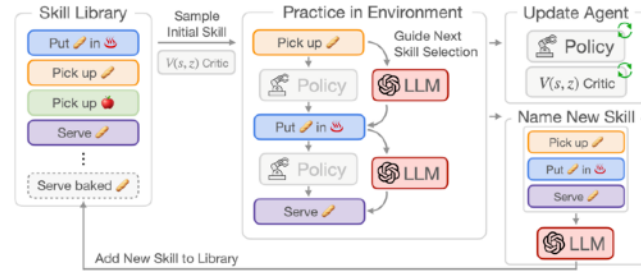
ICML-W 2023 & Submitted to ICML 2024

Diffusion Rewards Guided Adversarial Imitation Learning



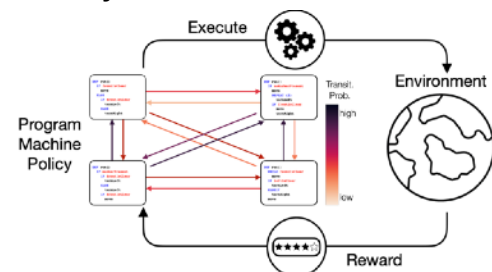
ICLR-W 2024 & Submitted to ICML 2024

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



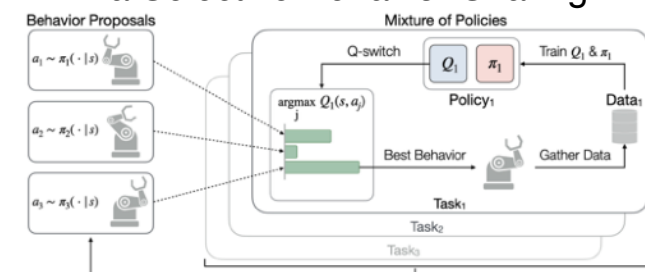
CoRL 2023 (Oral)

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



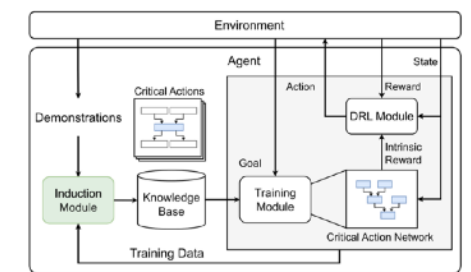
NeurIPS-W 2023 & Submitted to ICML 2024

Efficient Multi-Task Reinforcement Learning via Selective Behavior Sharing



NeurIPS-W 2022 & Submitted to ICML 2024

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

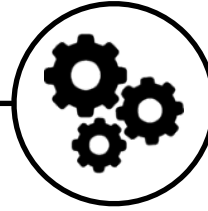
Assistant Professor @ National Taiwan University (NTU) | CS Ph.D. @USC | Robot Learning, Reinforcement Learning, Program Synthesis | 台大電機系助理教授

📍 Taipei, Taiwan 🔗 shaohua0116.github.io 📅 Joined October 2015

868 Following 2,987 Followers



Execute



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)

```

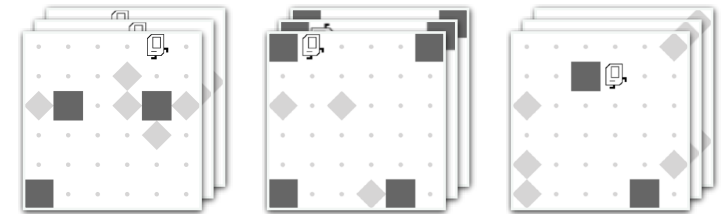
Environment



Thank You

Questions?

Demonstrations



Reward

