Execute

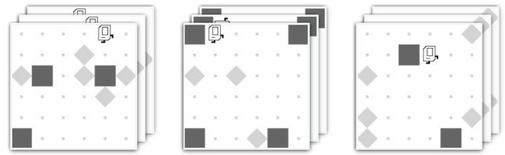
Learning to Synthesize **Programs** as Interpretable and Generalizable Reinforcement Learning **Policies**

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

Demonstrations



Shao-Hua Sun (孫紹華)

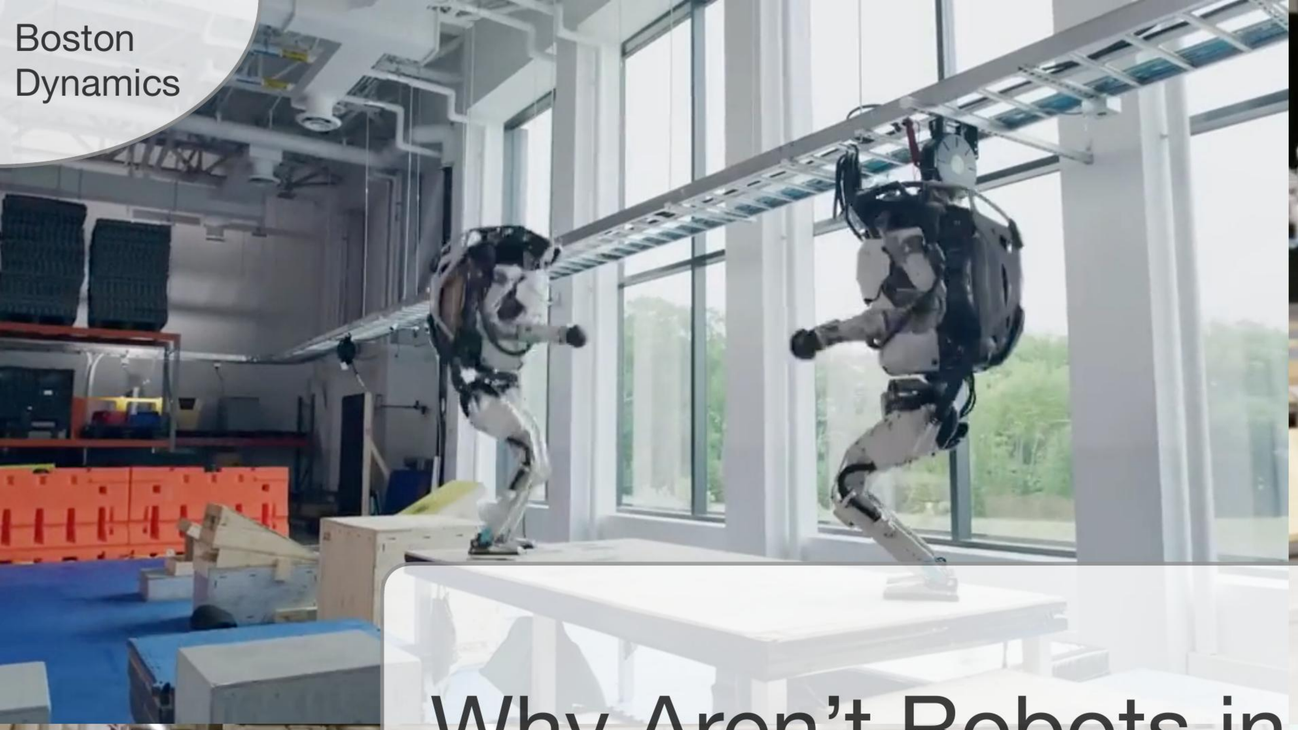
Assistant Professor

Dept. of Electrical Engineering (EE)

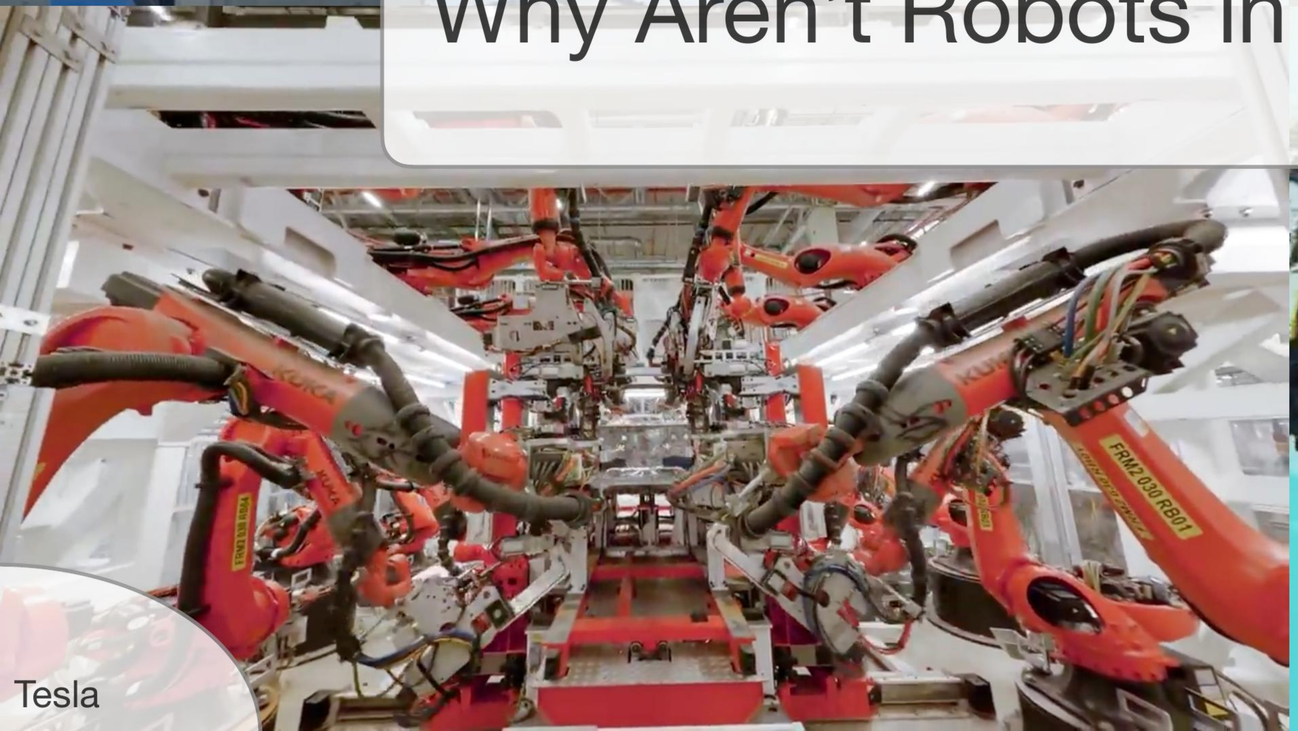
National Taiwan University

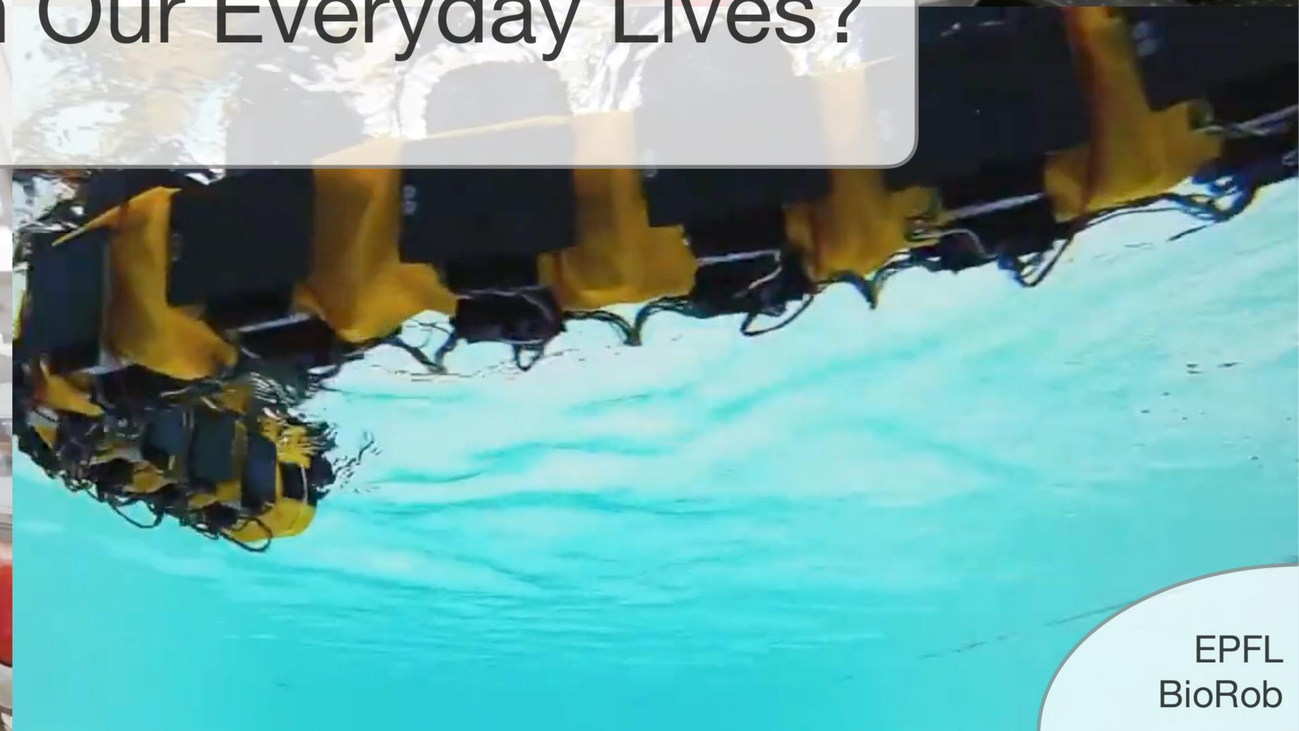
@ TAAI AI Forum 人工智慧論壇 2023

Reward

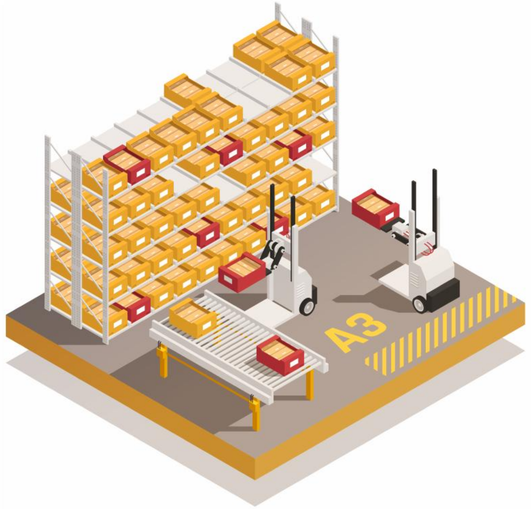


Why Aren’t Robots in Our Everyday Lives?





Environment



Structured

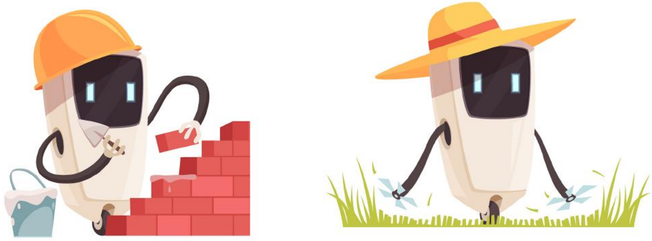
Unstructured

Object

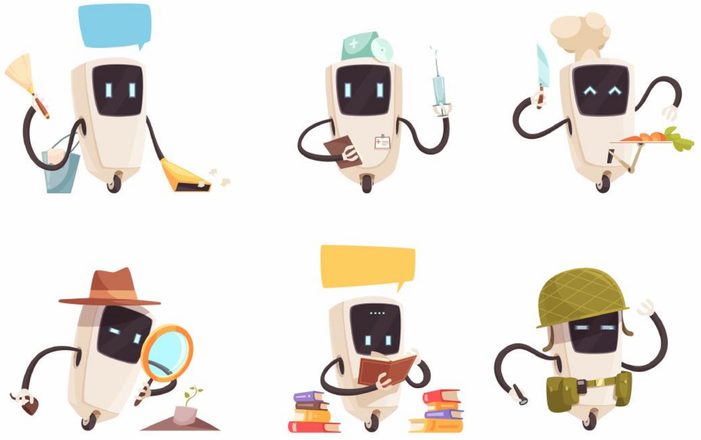
Known

Unseen

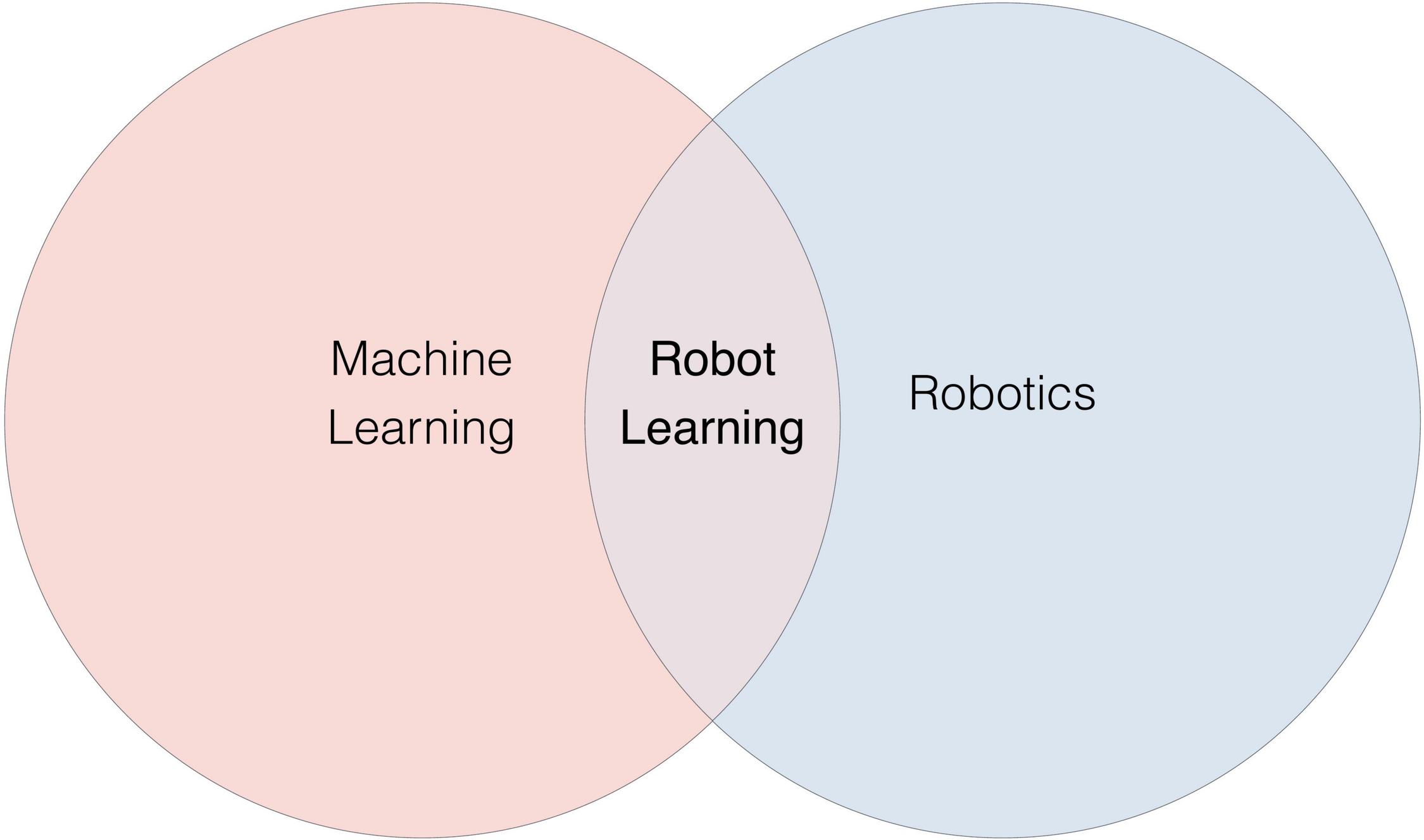
Task



Pre-defined / Pre-programmed



Diverse and Novel


Machine
Learning
Robot
Learning
Robotics

Supervised Learning

Image Classification

Model	image size	# parameters	Mult-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. Mult-Adds indicate the number of composite multiply-accumulate operations for a single image. Note that the composite multiple-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 All	Yes/no	Num.	Other	VQA v2 test-std All	Yes/no	Num.	Other
Prior (most common answer in training set) [ ]	–	–	–	–	25.28	61.20	0.36	1.17
LSTM Language only (blind model) [ ]	–	–	–	–	44.26	67.01	31.55	27.37
Deeper LSTM Q norm. I [7] as reported in [ ]	–	–	–	–	54.22	73.46	35.18	41.83
MCB [ ] as reported in [ ]	–	–	–	–	62.27	78.82	38.28	53.36
UPMC-LIP6 [ ]	–	–	–	–	65.71	82.07	41.06	57.12
Athena	–	–	–	–	67.59	82.50	44.19	59.97
LV-NUS	–	–	–	–	68.77	83.89	46.20	58.70
HDU-USYD-UNCC	–	–	–	–	68.19	84.50	45.39	59.01
Proposed model								
ResNet features 7×7, single network	62.07	79.20	39.06	52.62	62.27	79.32	39.37	52.59
Image features from bottom-up attention, adaptive K, single network	65.32	81.82	44.21	56.05	65.67	82.20	43.90	56.26
ResNet features 7×7, ensemble	66.34	83.36	43.17	57.10	66.73	83.71	43.77	57.20
Image features from bottom-up attention, adaptive K, ensemble	**69.87**	**86.09**	**48.99**	**60.80**	**70.34**	**86.60**	**48.64**	**61.15**

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

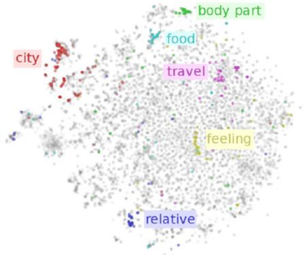
Teney *et al.* Tips and Tricks for Visual Question Answering: Learnings from the 2017 Challenge

Machine Translation

Source	"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.	
PBMT	"La raison pour laquelle Boeing sont en train de faire, c'est de concentrer davantage de sièges pour prendre leur avion plus compétitive avec nos produits", a déclaré Kevin M. Keniston, chef du confort des passagers de l'Airbus de l'Europe.	3.0
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.	6.0
Human	"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapports à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez l'avionneur européen Airbus.	6.0
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".	3.0
GNMT	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
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".	6.0

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

Named Entity Recognition

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

- Input Question:
  `Where do water droplets collide with ice crystals to form precipitation?`
- Input Paragraph:
  `... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. ...`
- Output Answer:
  `within a cloud`

System	Dev EM	Dev F1	Test EM	Test 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$_{LARGE}$ (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 systems which use different pre-training checkpoints and fine-tuning seeds.

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

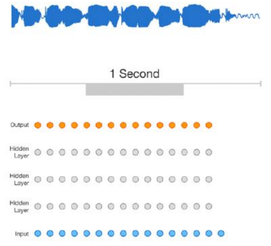
Speech Recognition

Word Error Rate

Senone set	Model/combination step	WER devset	WER test	WER devset	WER test
		ngram-LM		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			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

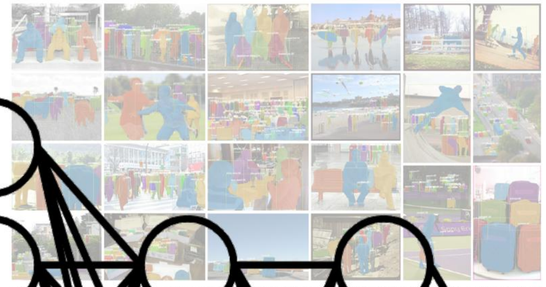
Image Classification

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	92.2
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. Mult-Adds indicates the number of composite multiply-accumulate operations for a single image. Note that the composite multiple-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 Recog

Instance Segmentation



…N on COCO te… using ResNet-10… …l running at 5… …7 mask AP (Table 1).

He *et al.* Mask R-CNN

Visual Question Answering

Method	VQA v2 test-dev All	Yes/no	Numb.	Other	VQA v2 test-std All	Yes/no	Numb.	Other
Prior (most common answer in training set) [ ]	–	–	–	–	25.28	61.20	0.36	1.17
LSTM Language only (blind model) [ ]	–	–	–	–	44.26	67.01	31.55	27.37
Deeper LSTM Q norm. I [7] as reported in [ ]	–	–	–	–	54.22	73.46	35.18	41.83
MCB [ ] as reported in [ ]	–	–	–	–	62.27	78.82	38.28	53.36
UPMC-LIP6 [ ]	–	–	–	–	65.71	82.07	41.06	57.12
Athena	–	–	–	–	67.59	82.50	44.19	59.07
LV-NUS	–	–	–	–	68.77	84.09	46.29	58.96
HDU-USYD-UNCC	–	–	–	–	68.79	84.50	45.59	59.01
Proposed model								
ResNet features 7×7, single network	62.07	79.23	39.96	52.62	62.27	79.32	39.77	52.59
Image features from bottom-up attention, adaptive K, single network	65.32	81.82	44.21	56.05	65.67	82.20	43.90	56.26
ResNet features 7×7, ensemble	66.94	83.38	43.43	57.10	66.73	83.71	43.37	57.20
Image features from bottom-up attention, adaptive K, ensemble	**69.87**	**86.09**	**48.99**	**60.80**	**70.34**	**86.60**	**48.64**	**61.15**

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

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

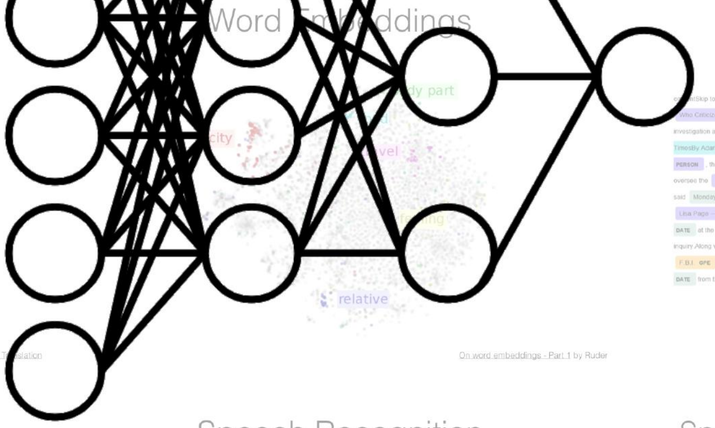
M… …on

Source	"The roas… …eir plane more competiti… …er comfort at Europe's …
PBMT	"La raisc… …rer davantage de sièges pou… …déclaré Kevin M. Keniston, …
GNMT	"La raisc… …ur rendre son avion plus … …ort du confort des passagers …
Human	"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapports à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager 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 États-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 …slation

Input

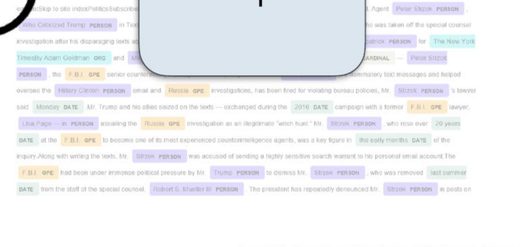
Word Em…ddings



On word embeddings - Part 1 by Ruder

Output

Name… …nition



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

- Input Question:
  `Where do water droplets collide with ice crystals to form precipitation?`
- Input Paragraph:
  `... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. ...`
- Output Answer:
  `within a cloud`

System	Dev EM	Dev F1	Test EM	Test 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$_{LARGE}$ (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 systems which use different pre-training checkpoints and fine-tuning seeds.

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

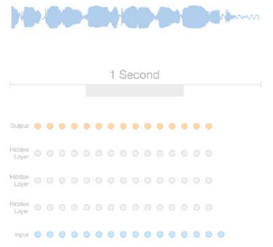
Speech Recognition

Word Error Rate

Senone set	Model/combination step	WER devset ngram-LM	WER test ngram-LM	WER devset LSTM-LMs	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			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

Instance Segmentation

Visual Question Answering

Image

Class

“Cat”

Word Embeddings

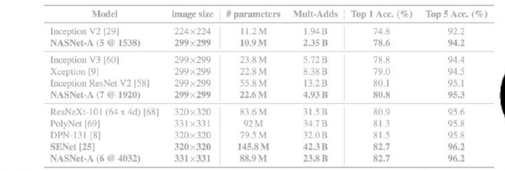
Question Answering

Speech Recognition

Speech Synthesis (text-to-speech)

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	92.2
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. Mult-Adds indicates the number of composite multiply-accumulate operations for a single image. Note that the composite multiple-accumulate operations are calculated for single image are reported in the table. Model size for [25] calculated from open-source implementation.

Zoph *et al.* Learning Transferable Architectures for Scalable Image Recog

Instance Segmentation

...N on COCO test-... using ResNet-10... ...l running at 5 ... ...5.7 mask AP (Table 1).

...ky *et al.* Mask R-CNN

Visual Question Answering

Method	VQA v2 test-dev All	Yes/no	Numb.	Other	VQA v2 test-std All	Yes/no	Numb.	Other
Prior (most common answer in training set) [ ]	-	-	-	-	25.28	61.20	0.36	1.17
LSTM Language only (blind model) [ ]	-	-	-	-	44.26	67.01	31.55	27.37
Deeper LSTM Q norm. I [7] as reported in [ ]	-	-	-	-	54.22	73.46	35.18	41.83
MCB [ ] as reported in [ ]	-	-	-	-	62.27	78.82	38.28	53.36
UPMC-LIP6 [ ]	-	-	-	-	65.71	82.07	41.06	57.12
Athena	-	-	-	-	67.59	82.50	44.19	59.07
LV-NUS	-	-	-	-	68.77	84.89	46.20	58.90
HDU-USYD-UNCC	-	-	-	-	68.59	84.50	47.35	59.01
Proposed model								
ResNet features 7×7, single network	62.07	79.23	39.98	52.62	62.27	79.32	39.77	52.59
Image features from bottom-up attention, adaptive K, single network	65.32	81.82	44.21	56.05	65.67	82.20	43.90	56.26
ResNet features 7×7, ensemble	66.94	83.38	43.43	57.10	66.73	83.71	43.37	57.20
Image features from bottom-up attention, adaptive K, ensemble	69.87	86.09	48.99	60.80	70.34	86.60	48.64	61.15

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

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

English sentence

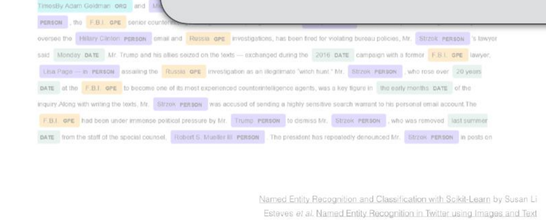
France is never cold in September

French sentence

la france est jamais froid en septembre

Word Embeddings

Name



Named Entity Recognition and Classification with Scikit-Learn by Susan Li

Esteves *et al.* Named Entity Recognition in Twitter using Images and Text

...on

Human	...their plane more ...re comfort at
	...er davantage de ...déclaré Kevin M. 3.0
	...ur rendre son ...f du confort des 6.0
Human	"Boeing fait ça pour pouvoir caser plus de sièges et rendre ses avions plus compétitifs par rapports à nos produits", a déclaré Kevin Keniston, directeur de Confort Passager chez l'avionneur européen Airbus. 6.0
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". 3.0
GNMT	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
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". 6.0

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

...on Word Embeddings - Part 1 by Ruder

Question Answering

- Input Question:
  Where do water droplets collide with ice crystals to form precipitation?
- Input Paragraph:
  ... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. ...
- Output Answer:
  within a cloud

System	Dev EM	Dev F1	Test EM	Test 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$_{LARGE}$ (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 systems which use different pre-training checkpoints and fine-tuning seeds.

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

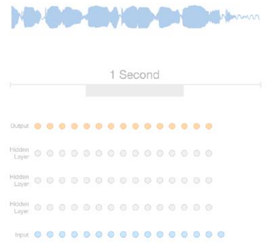
Speech Recognition

Word Error Rate

Senone set	Model/combination step	WER devset ngram-LM	WER test ngram-LM	WER devset LSTM-LMs	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			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

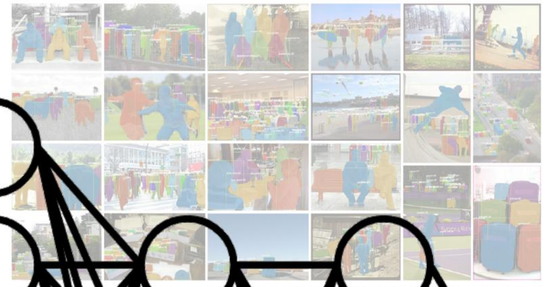
Image Classification

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	92.2
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. Mult-Adds indicates the number of composite multiply-accumulate operations for a single image. Note that the composite multiple-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 Recog

Instance Segmentation



…N on COCO test… using ResNet-10… …running at 5… …7 mask AP (Table 1).

…et al. Mask R-CNN

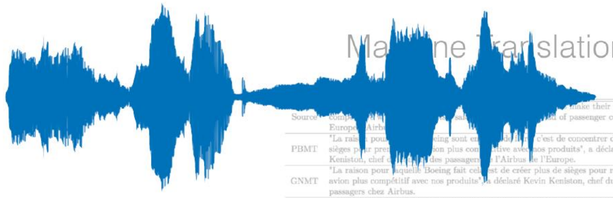
Visual Question Answering

Method	VQA v2 test-dev All	Yes/no	Numb.	Other	VQA v2 test-std All	Yes/no	Numb.	Other
Prior (most common answer in training set) [11]	-	-	-	-	25.28	61.20	0.36	1.17
LSTM Language only (blind model) [11]	-	-	-	-	44.26	67.01	31.55	27.37
Deeper LSTM Q norm. I [37] as reported in [11]	-	-	-	-	54.22	73.46	35.18	41.83
MCB [11] as reported in [11]	-	-	-	-	62.27	78.82	38.28	53.36
UPMC-LIP6 [11]	-	-	-	-	65.71	82.07	41.06	57.12
Athena	-	-	-	-	65.79	82.50	44.19	59.07
LV-NUS	-	-	-	-	66.77	82.80	46.20	58.58
HDU-USYD-UNCC	-	-	-	-	66.79	84.50	45.39	59.01
Proposed model								
ResNet features 7×7, single network	62.07	79.23	39.96	52.62	62.27	79.32	39.77	52.39
(single) features from bottom-up attention, adaptive K, single network	65.32	81.82	44.21	56.05	65.67	82.20	43.90	56.26
ResNet features 7×7, ensemble	66.94	83.38	43.43	57.10	66.73	83.71	43.37	57.20
(image features from bottom-up attention, adaptive K, ensemble	69.87	86.28	48.99	61.86	70.34	86.60	48.64	61.15

Table 3. Comparison of our best model with competing methods. Excerpt from the official VQA *v2 Leaderboard* [11].

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

Waveform



Text

This is a supervised learning method

Machine Translation

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 États-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

Word Embeddings

On word embeddings - Part 1 by Ruder

Name

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

- Input Question:
  `Where do water droplets collide with ice crystals to form precipitation?`
- Input Paragraph:
  `... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. ...`
- Output Answer:
  `within a cloud`

System	Dev EM	Dev F1	Test EM	Test 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$_{LARGE}$ (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 systems which use different pre-training checkpoints and fine-tuning seeds.

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

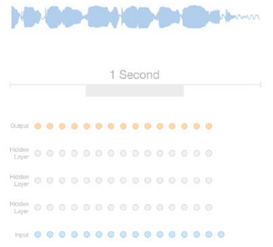
Speech Recognition

Word Error Rate

Senone set	Model/combination step	WER devset ngram-LM	WER test ngram-LM	WER devset LSTM-LMs	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			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

Instance Segmentation

Visual Question Answering

Input

Output

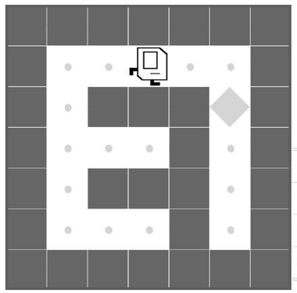
Question Answering

Speech Recognition

Speech Synthesis (text-to-speech)

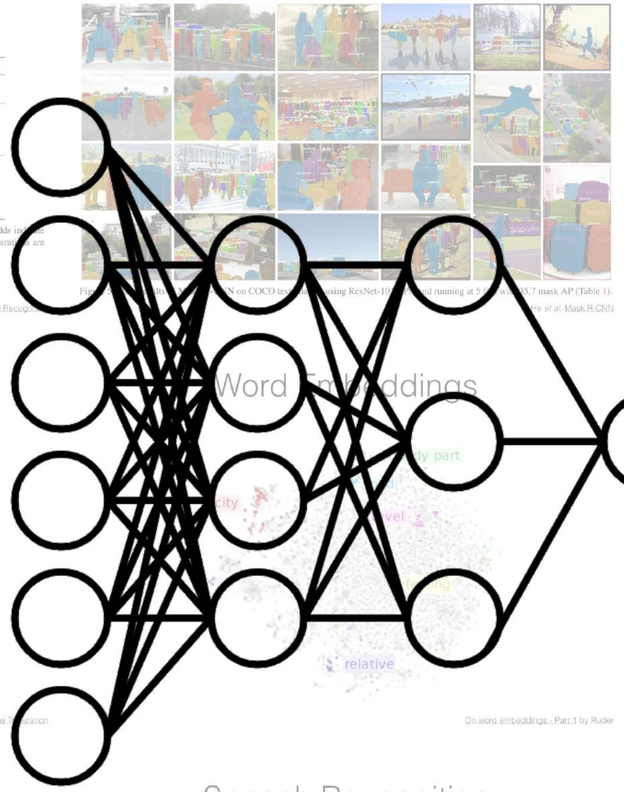
- The model has no control over the **dataset** it learns from
- Ground truth **output** can be specified given **input**

Robot Learning via Supervised Learning



State

Input



Action

Output

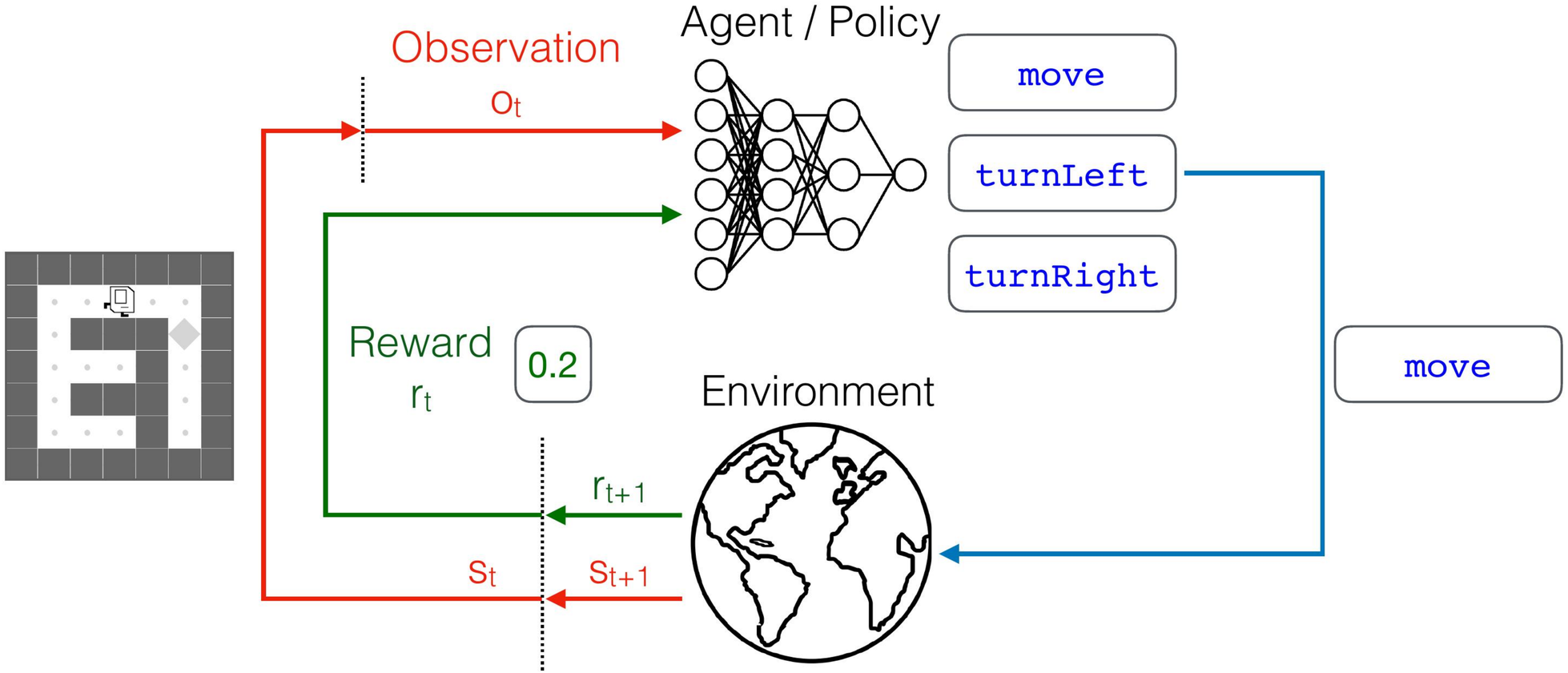
move

turnLeft

turnRight

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

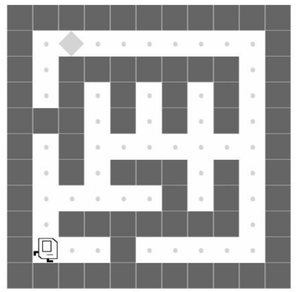
Robot Learning via Deep Reinforcement Learning



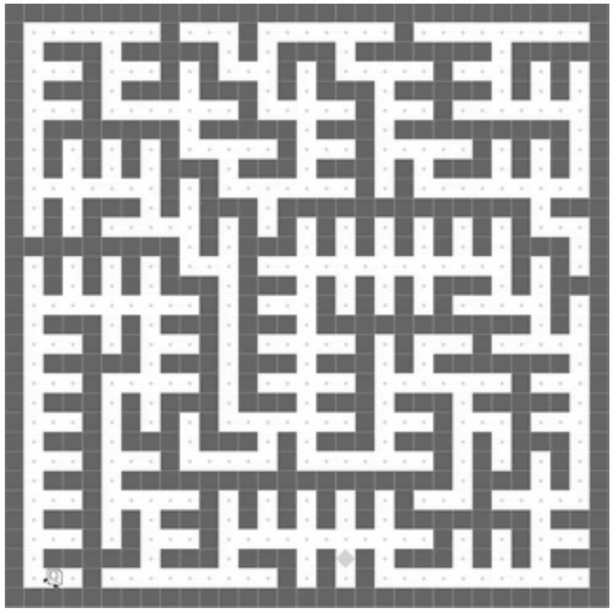
Goal: maximize $\sum_{t=0}^{t=H} \gamma^t R_t(s_t, a_t)$

Robot Learning via Deep Reinforcement Learning - Issues

Generalization



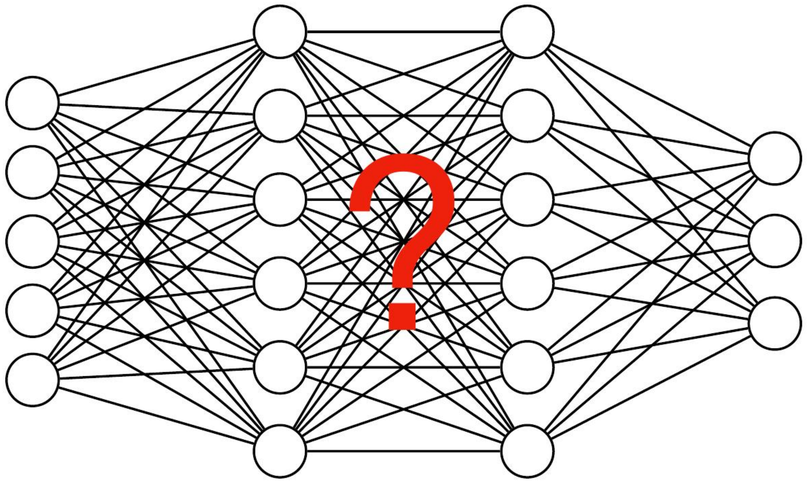
Simple task



Complex task

Interpretability

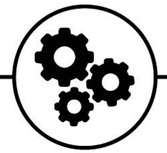
Trust, Safety, and Contestability



Deep neural network policy

Program as Reinforcement Learning Policies

Execute



Program

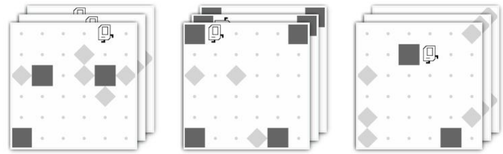
```
DEF run() m(
    WHILE c( markerPresent c) w(
        WHILE c( markerPresent c) w(
            pickMarker
            move w)
        turnRight
        move
        turnLeft
        WHILE c( markerPresent c) w(
            pickMarker
            move w)
        turnLeft
        move
        turnRight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

Program

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

Environment

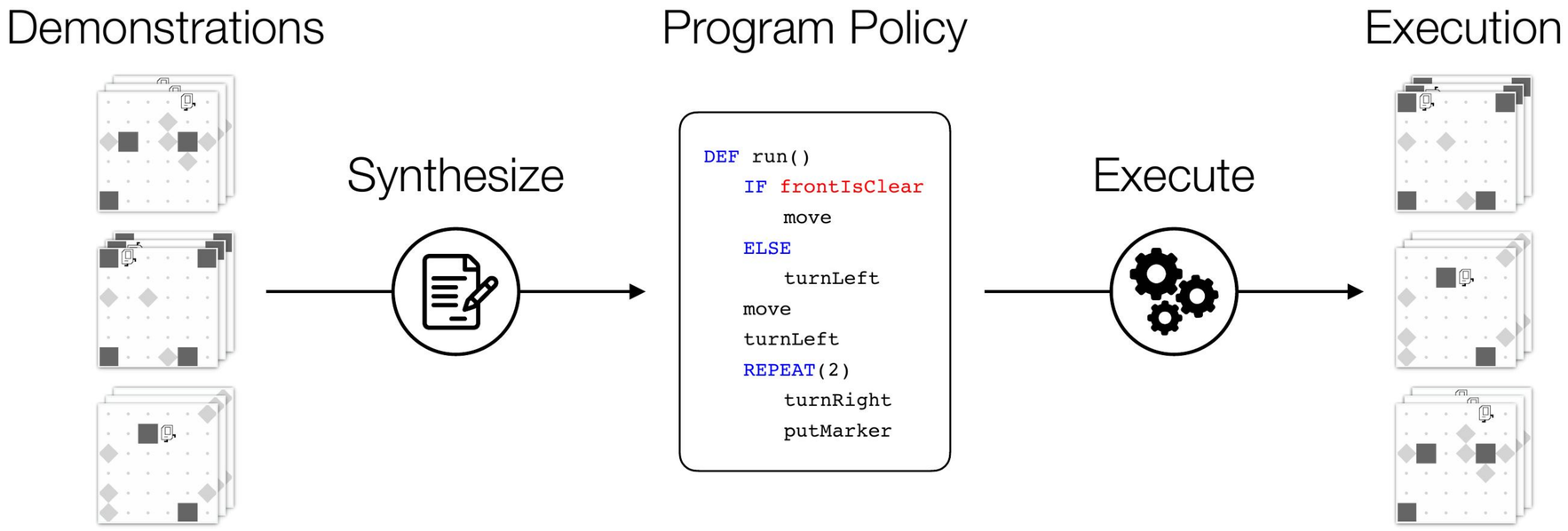
Demonstrations



Reward

Neural Program Synthesis from Diverse Demonstration Videos

ICML 2018



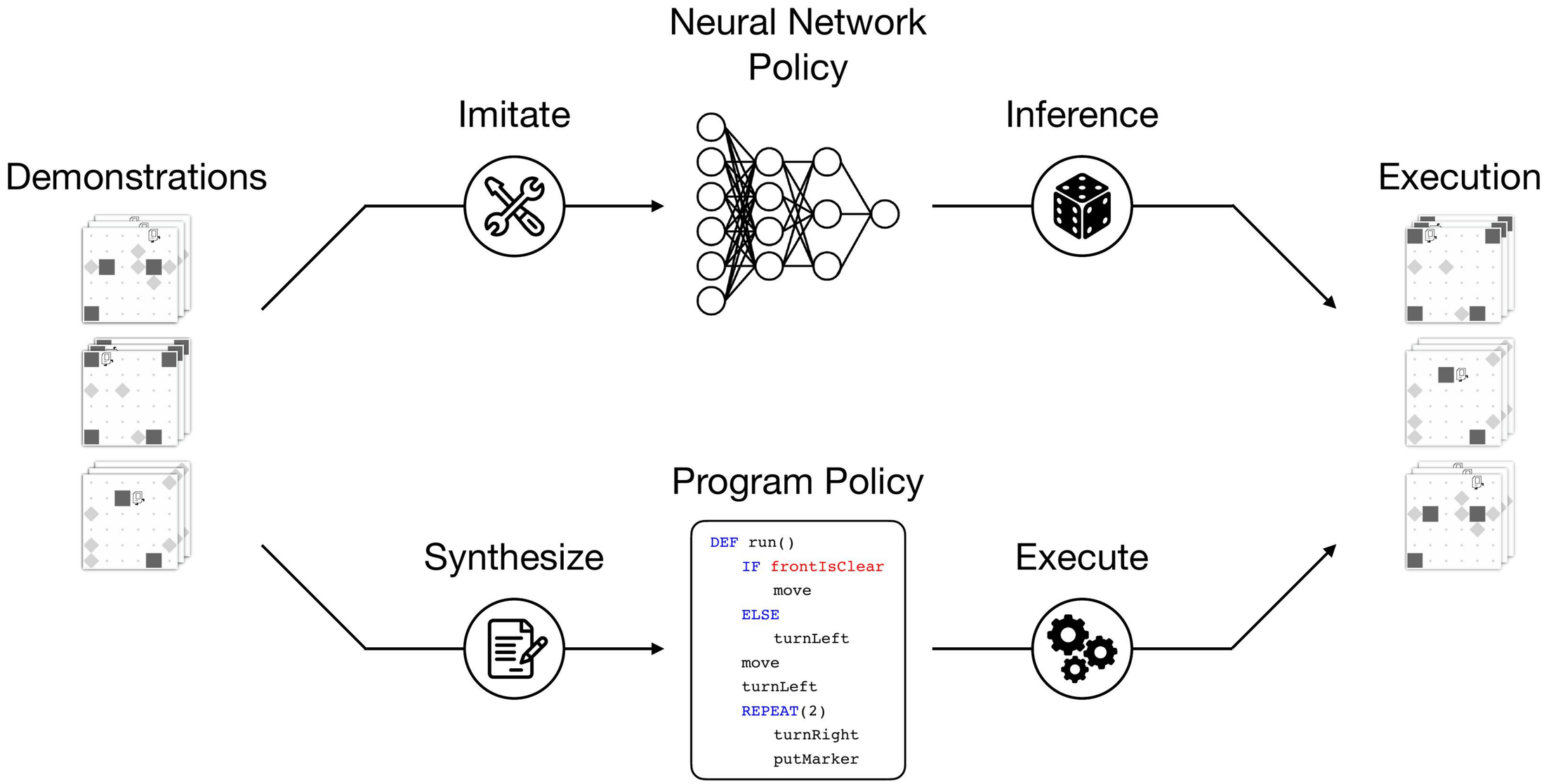

Shao-Hua Sun


Hyeonwoo Noh


Sriram Somasundaram


Joseph J. Lim

Imitation Learning via Synthesizing Programs



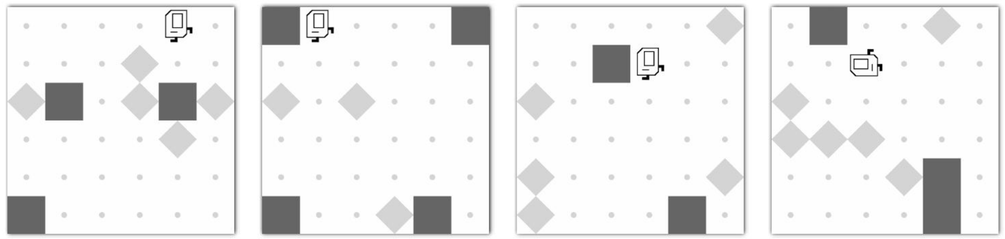
Environments

Karel

Program

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

Demonstrations



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

ViZDoom

Program

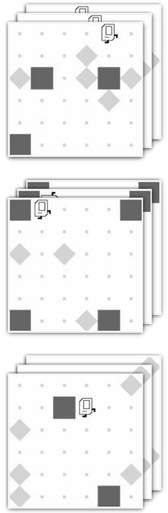
```
DEF run()
   WHILE frontIsClear(HellKnight)
      attack
      moveForward
   IF thereIs(Demon)
      moveRight
   ELSE
      moveLeft
      moveBackward
```

Demonstrations

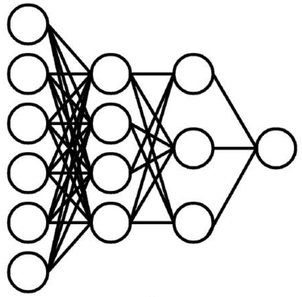
Kempka et al., “Vizdoom: A doom-based ai research platform for visual reinforcement learning.” in CIG, 2016

Imitation Learning with Neural Network Policy

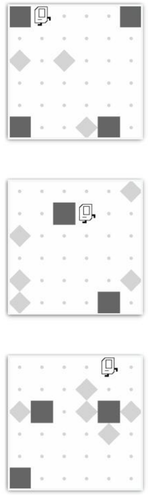
Demonstrations



Neural Network Policy

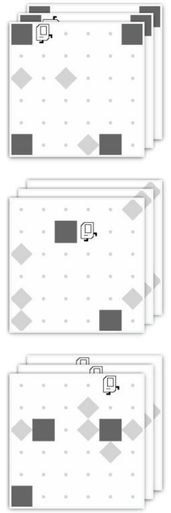


Initial States

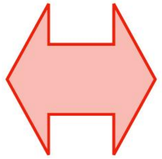


Infer

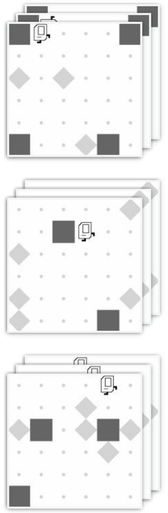
Predicted Execution



Loss

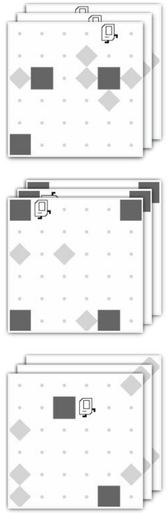


Ground Truth Execution

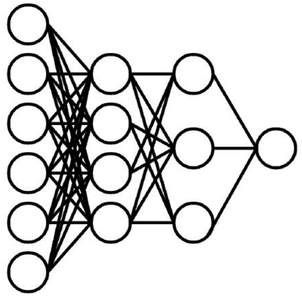


Imitation Learning with Program Policy

Demonstrations



Neural Network
Program Synthesizer



Synthesize

Predicted Program

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

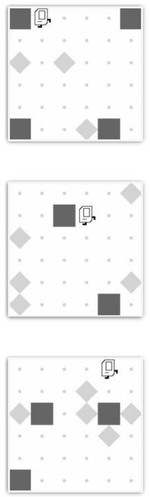
Loss

Ground Truth Program

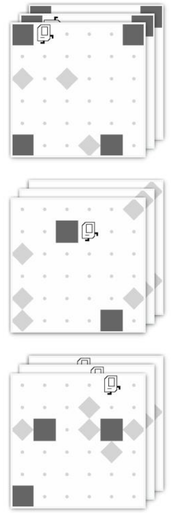
```
DEF run()
   IF frontIsClear
      move
   ELSE
      turnRight
   move
   REPEAT(2)
      turnLeft
```

Execute

Initial States



Predicted Execution

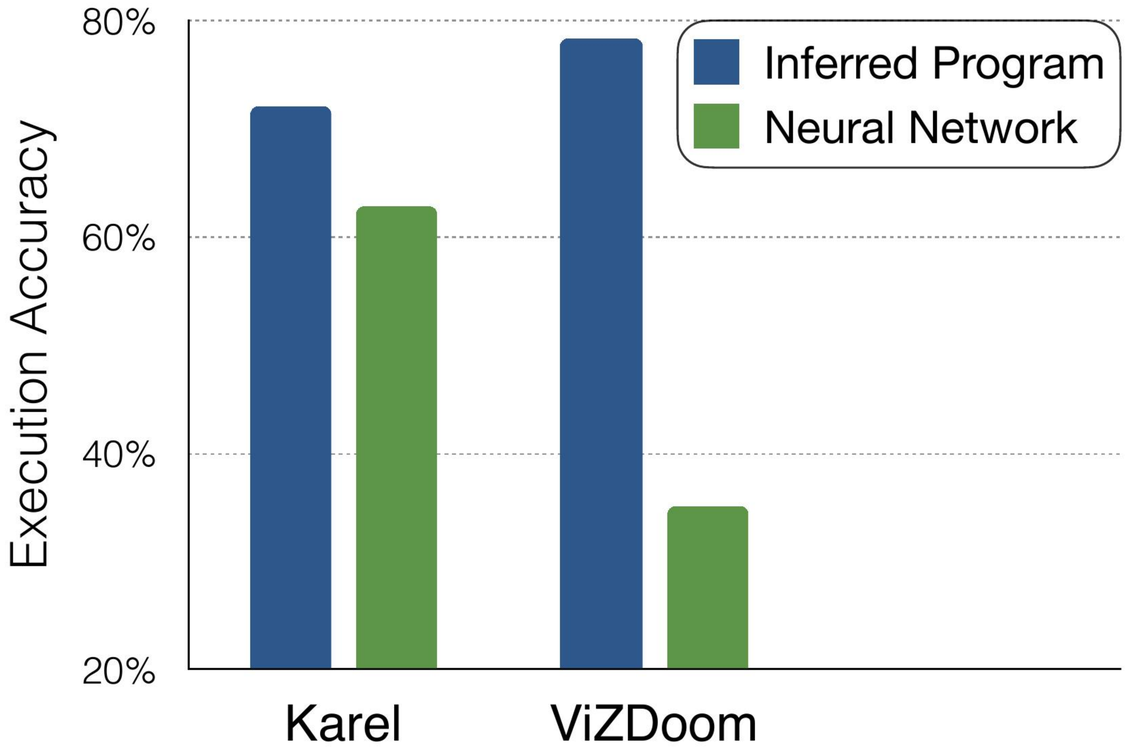


Evaluate

Ground Truth Execution

Experimental Results

Quantitative Results



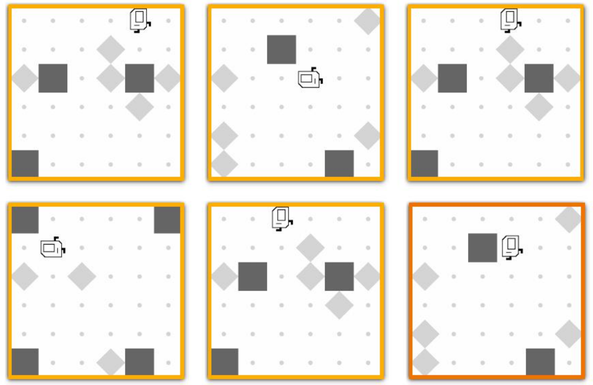
Evaluation: Execute the inferred program and the learned neural network policy on a set of unseen initial states and compare them to the ground truth demonstrations

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**

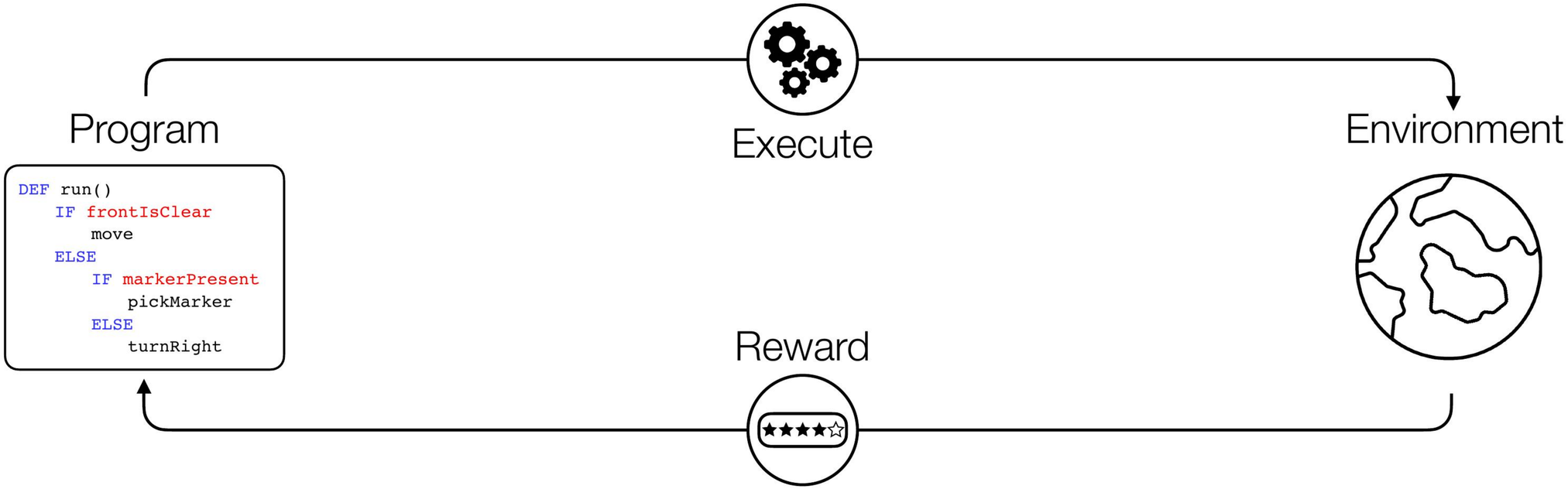
Inferred Program

```
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



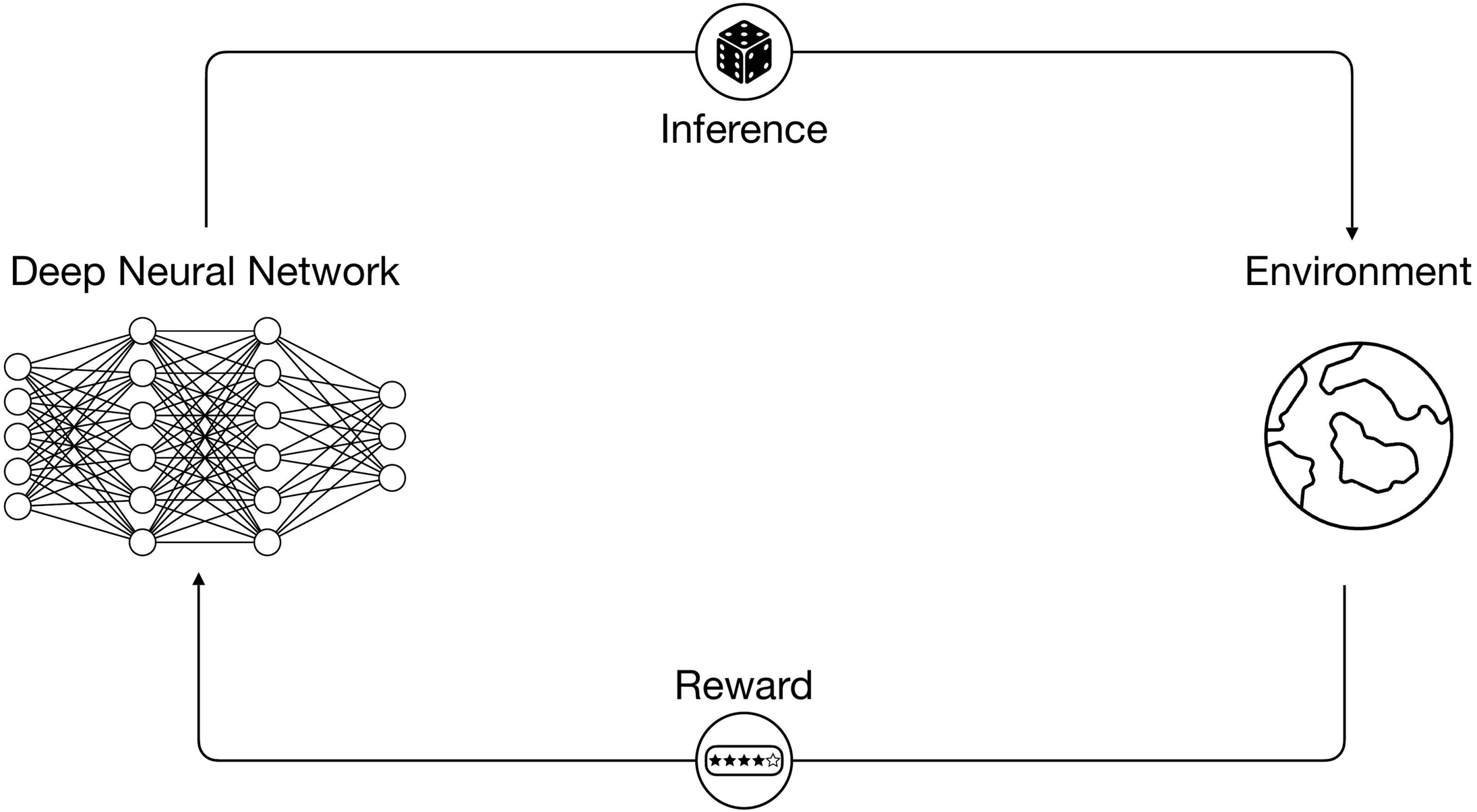

Dweep Trivedi


Jesse Zhang

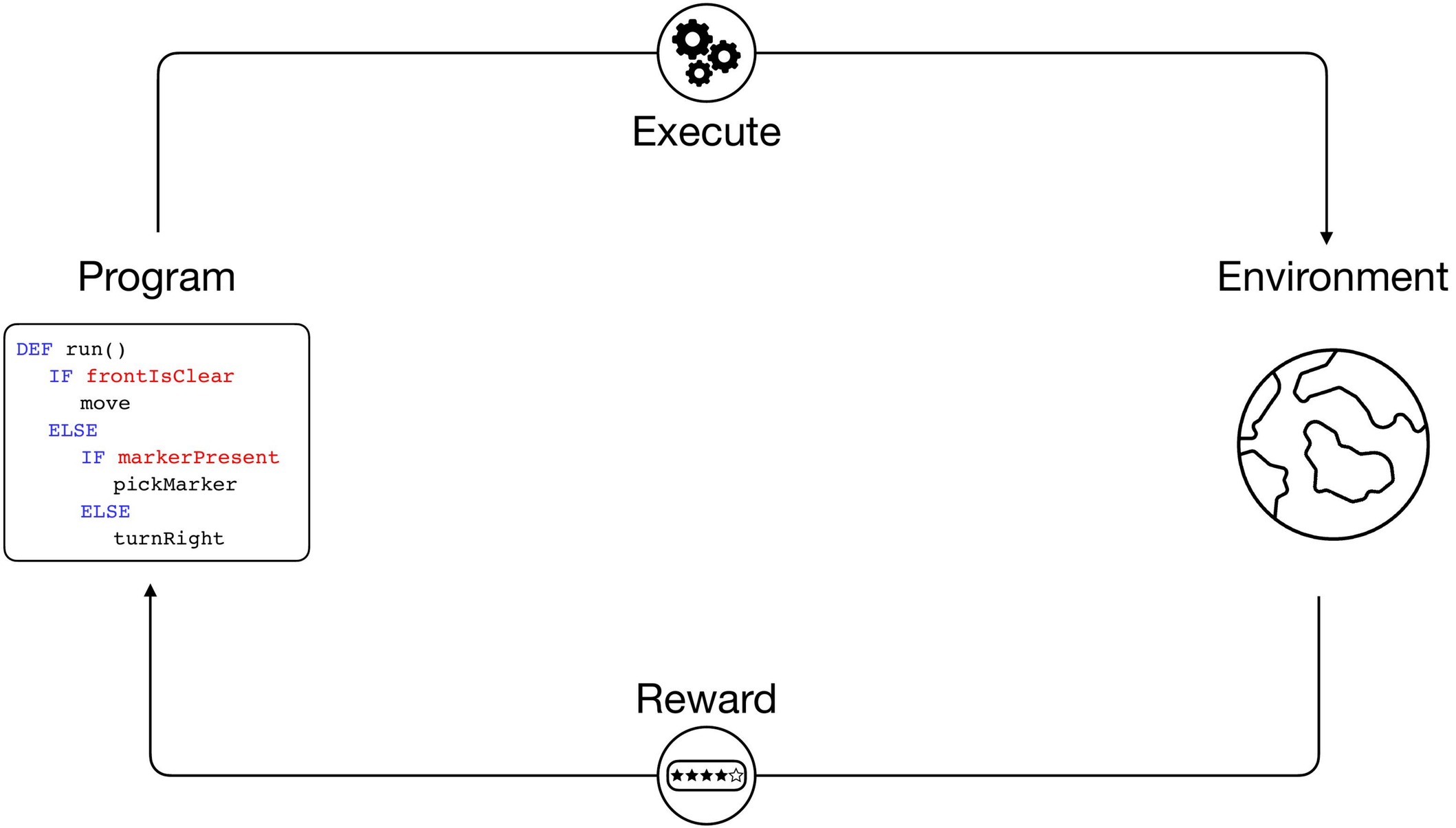

Shao-Hua Sun


Joseph J. Lim

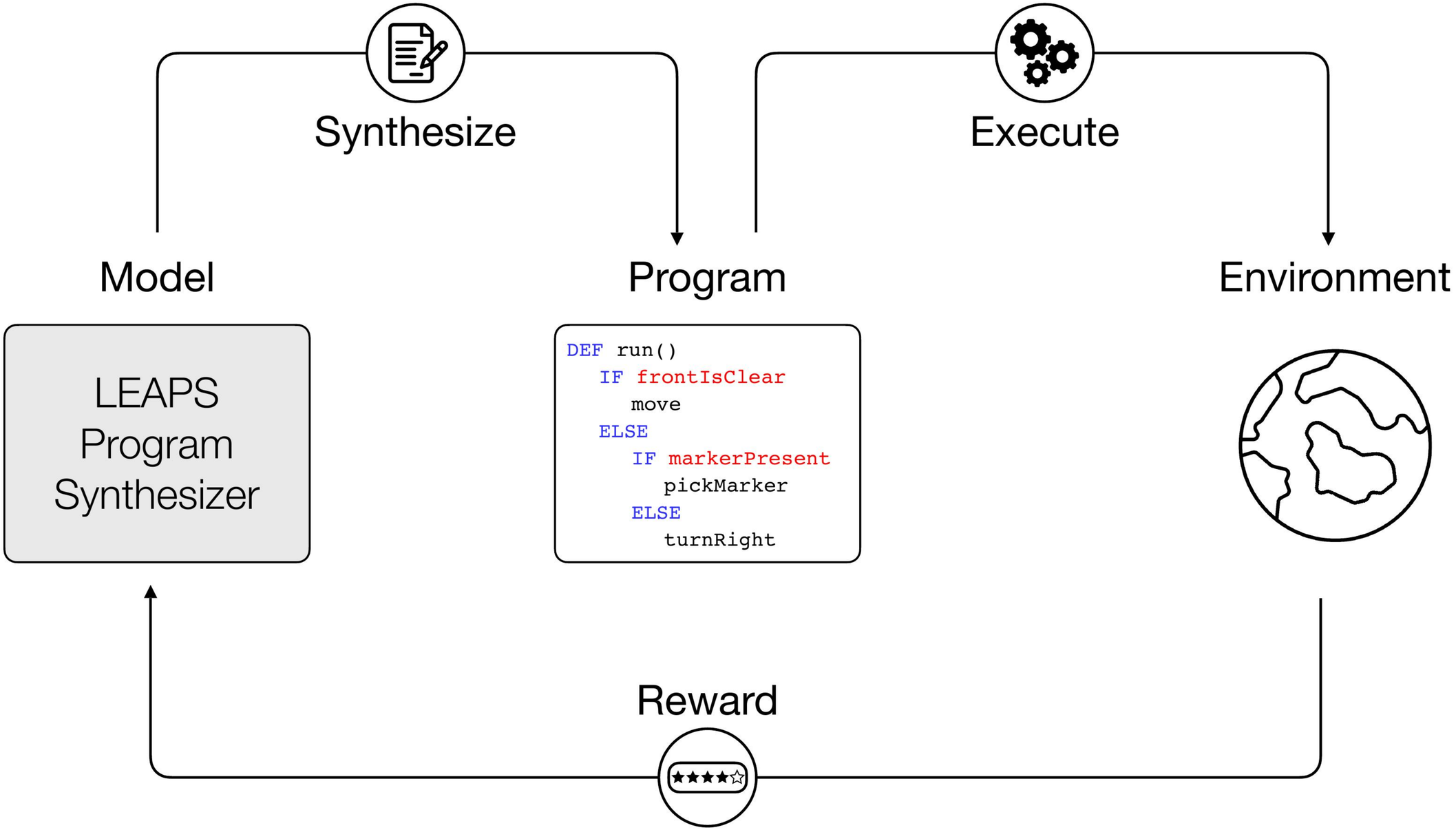
Deep Reinforcement Learning



Reinforcement Learning via Synthesizing Programs



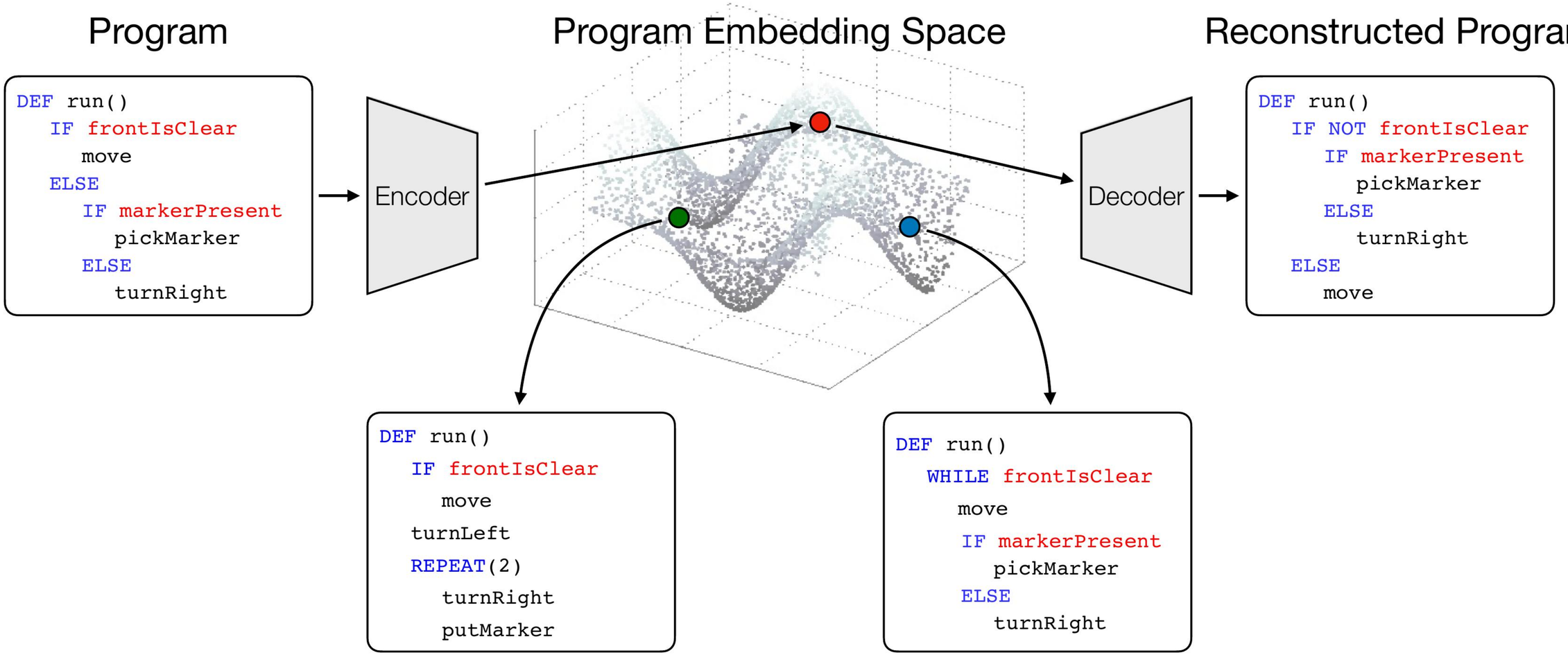
Reinforcement Learning via Synthesizing Programs



LEAPS: Learning Embeddings for IAtent Program Synthesis

Stage 1 Learn a **program embedding space** from randomly generated programs

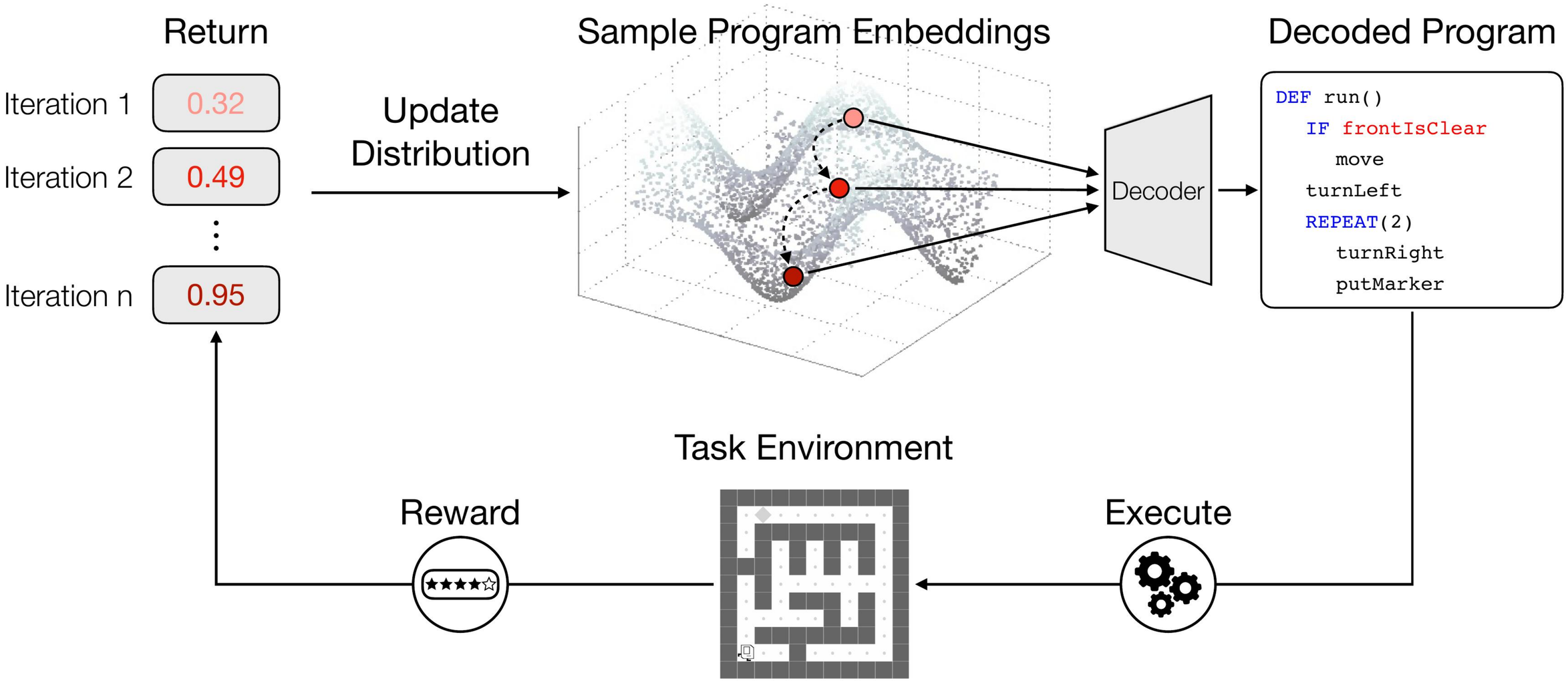
Goal Learn the grammar and the environment dynamics



LEAPS: Learning Embeddings for lAtent Program Synthesis

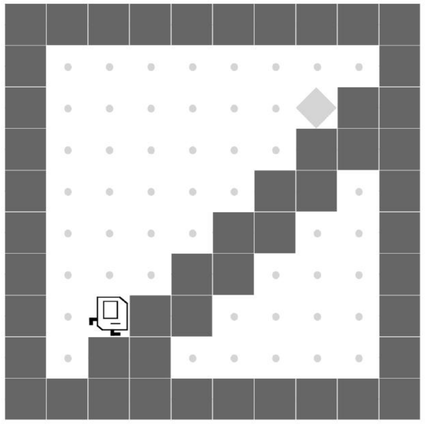
Stage 2 Search for a **task-solving program** using the **cross-entropy method (CEM)**

Goal Optimize the task performance

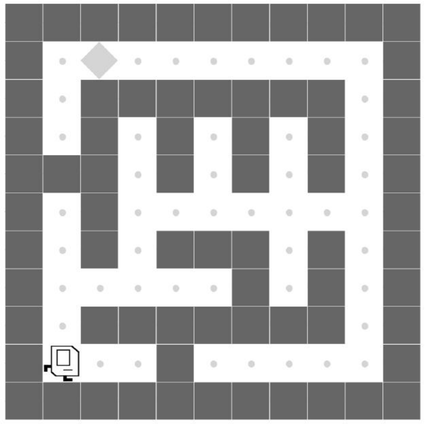


Karel Tasks

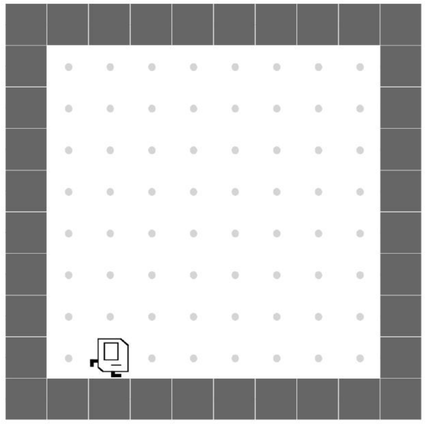
StairClimber



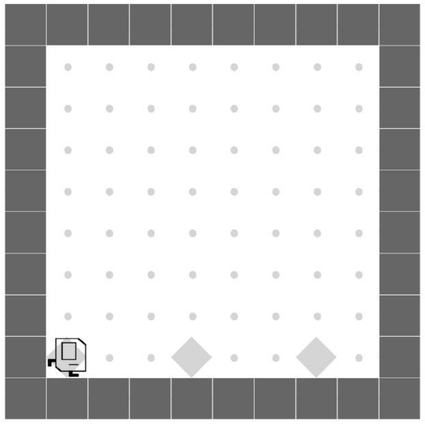
Maze



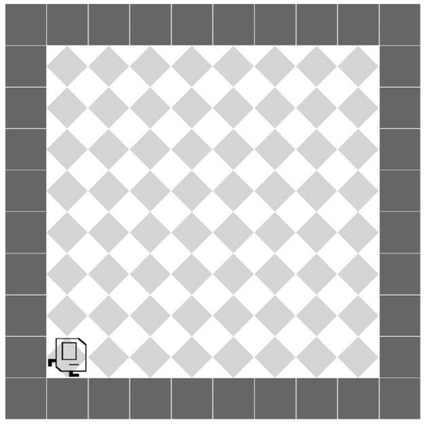
FourCorners



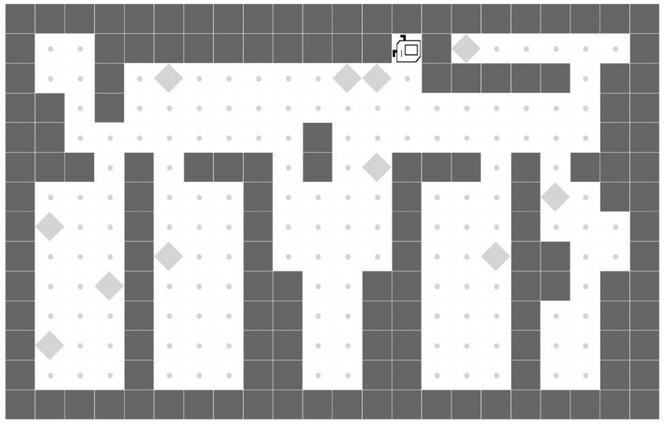
TopOff



Harvester

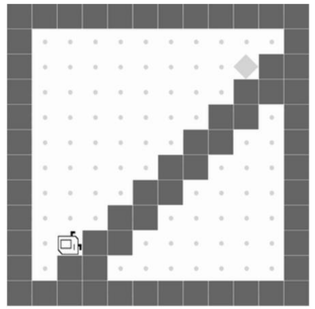


CleanHouse

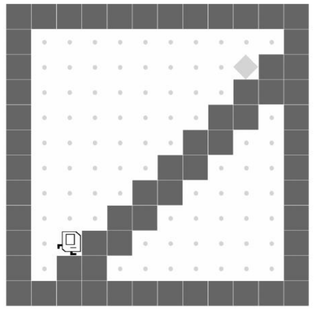


Qualitative Results

StairClimber

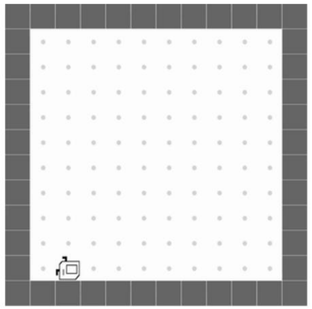


Deep RL

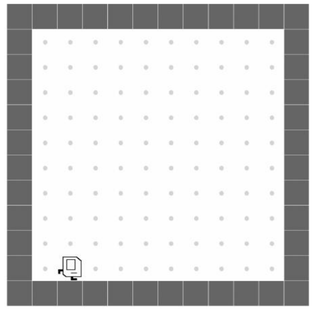


LEAPS

FourCorners

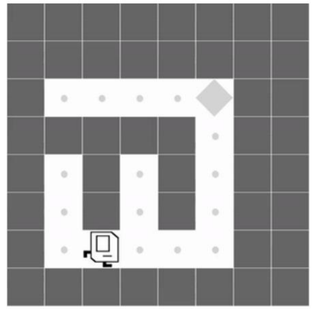


Deep RL

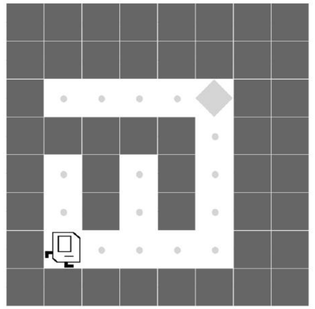


LEAPS

Maze

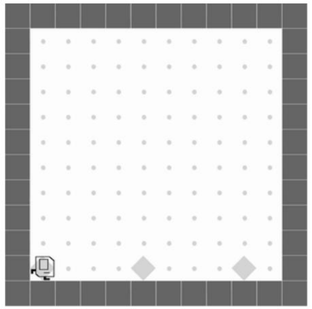


Deep RL

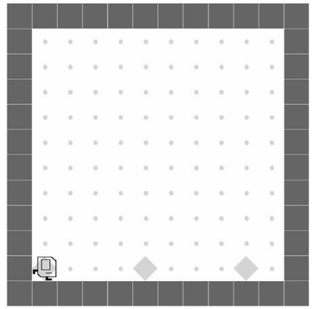


LEAPS

TopOff

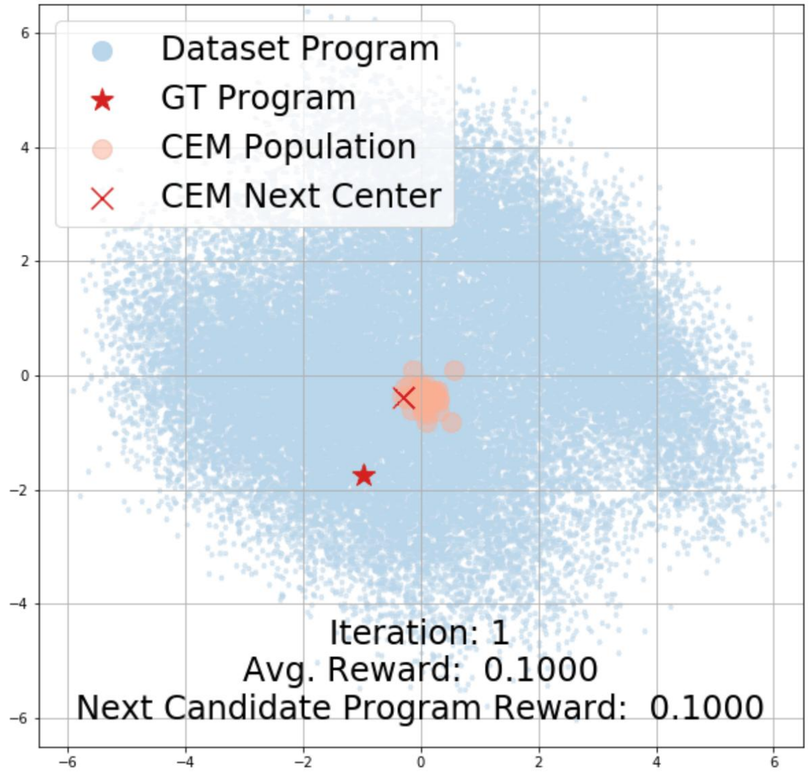


Deep RL



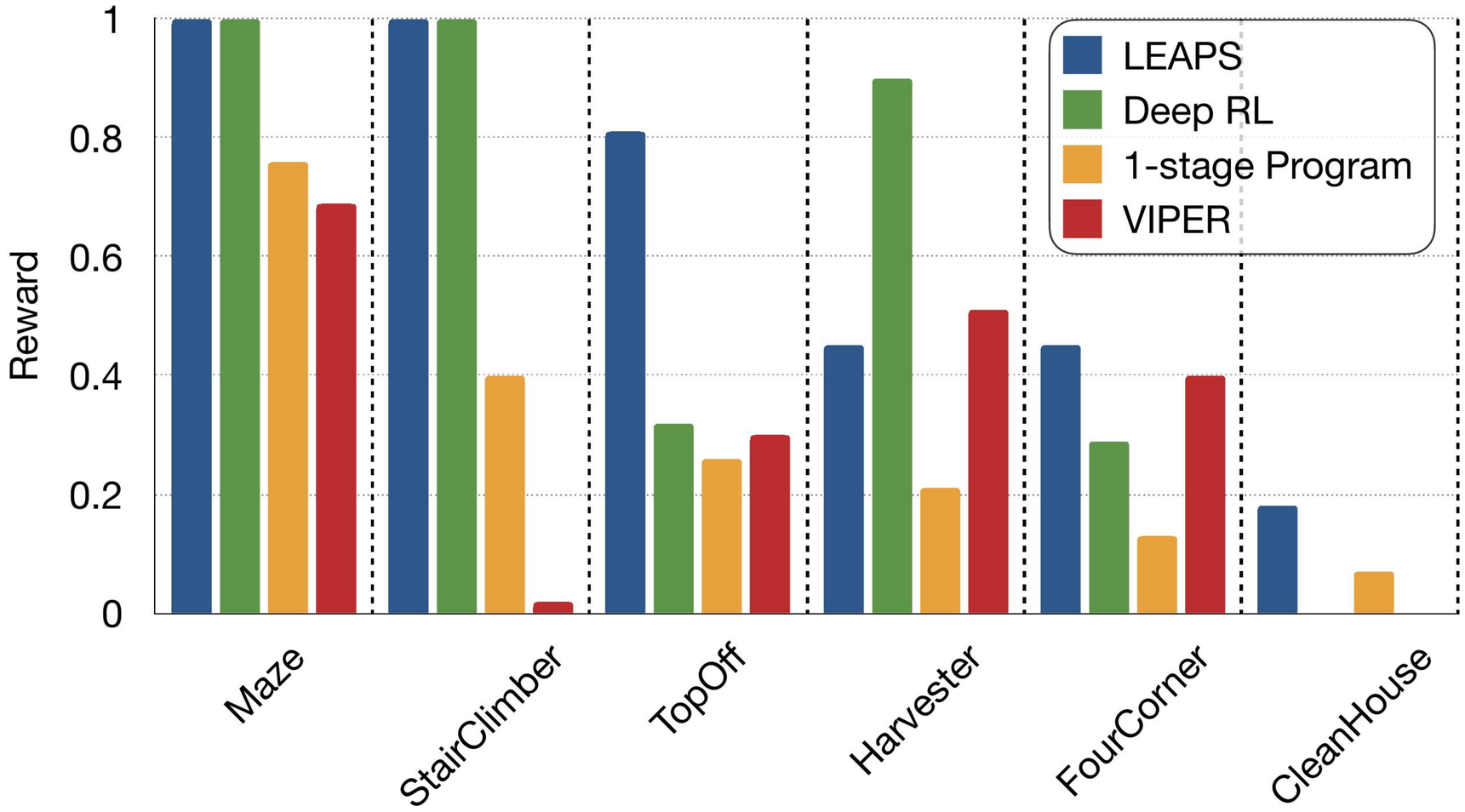
LEAPS

CEM trajectory Visualization



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

Quantitative Results



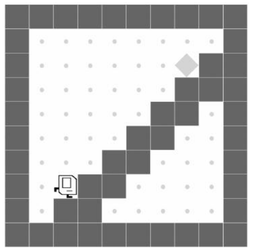
LEAPS Zero-shot Generalization

Learning on 8 x 8

LEAPS Program Policy

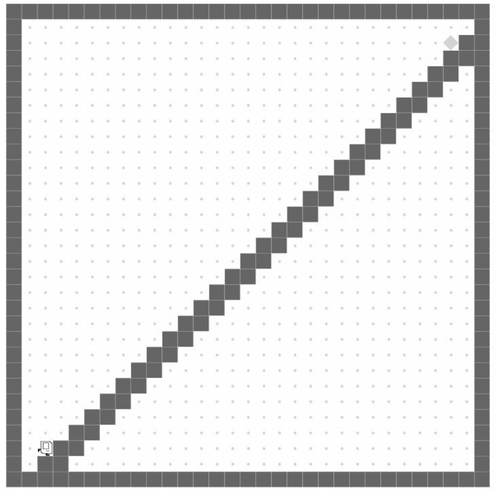
Evaluation on 100 x 100

StairClimber

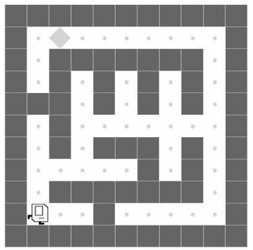


LEAPS Program Synthesizer

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

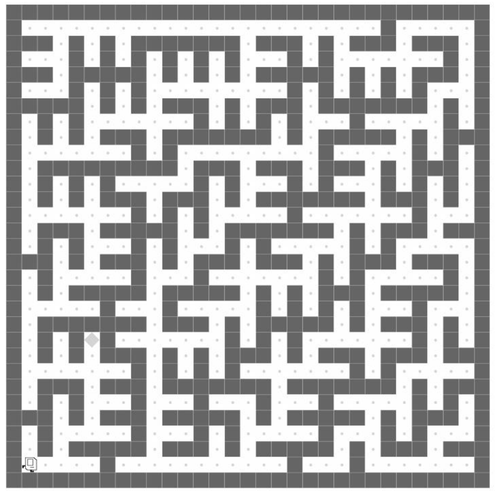


Maze

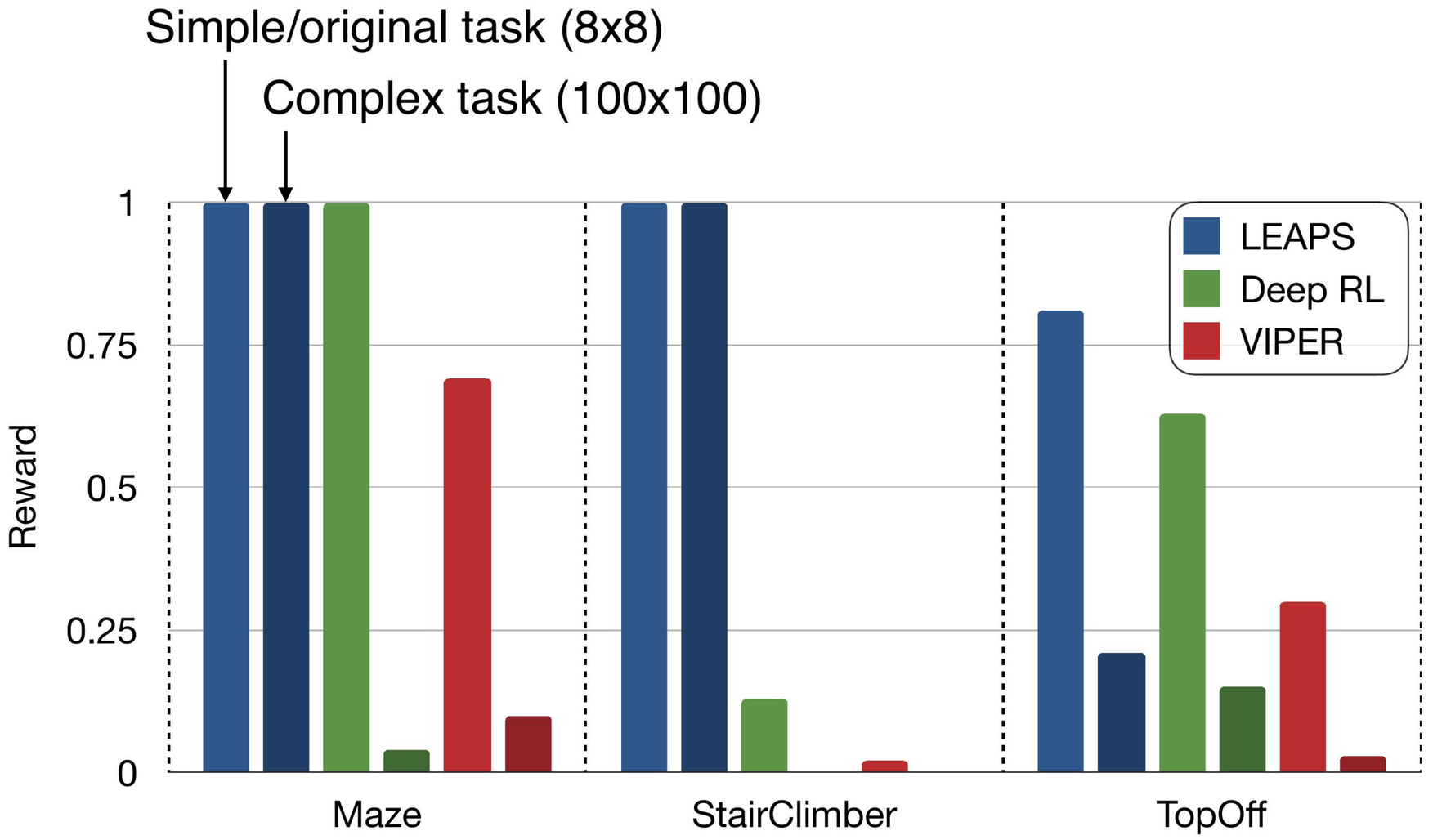


LEAPS Program Synthesizer

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

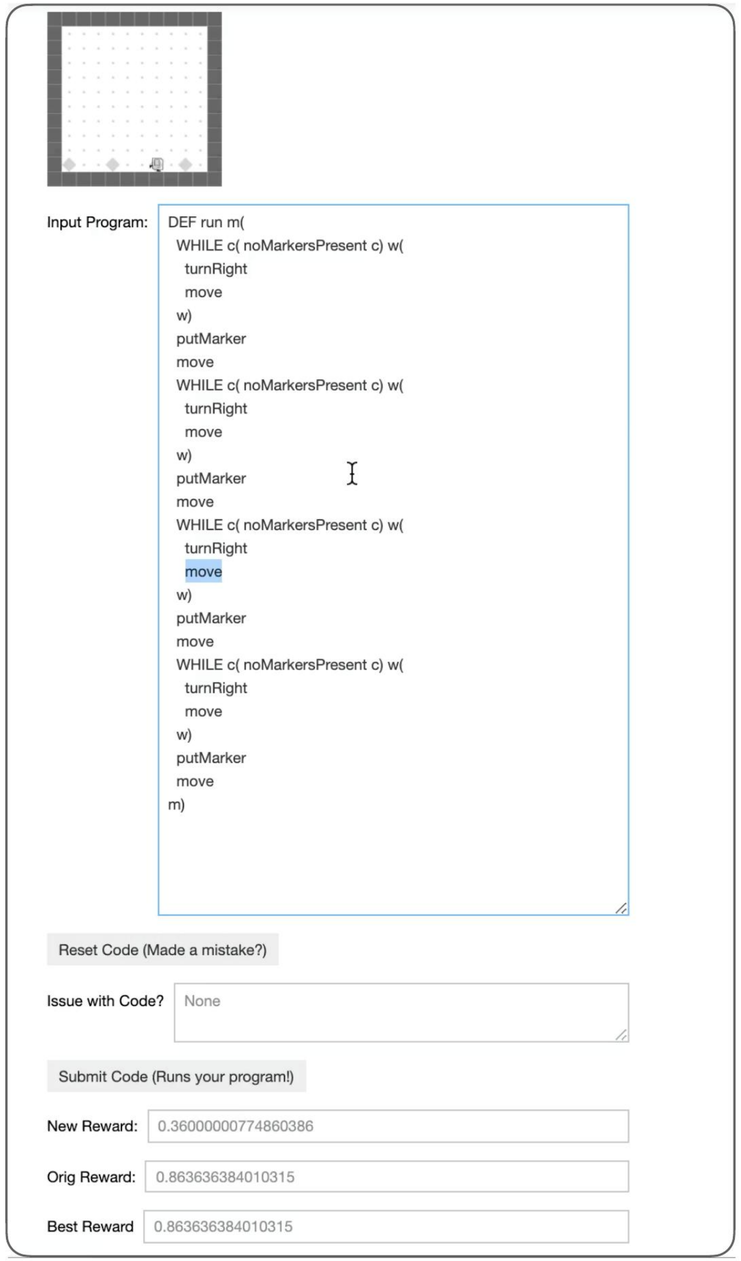


Experimental Results - Zero-shot Generalization

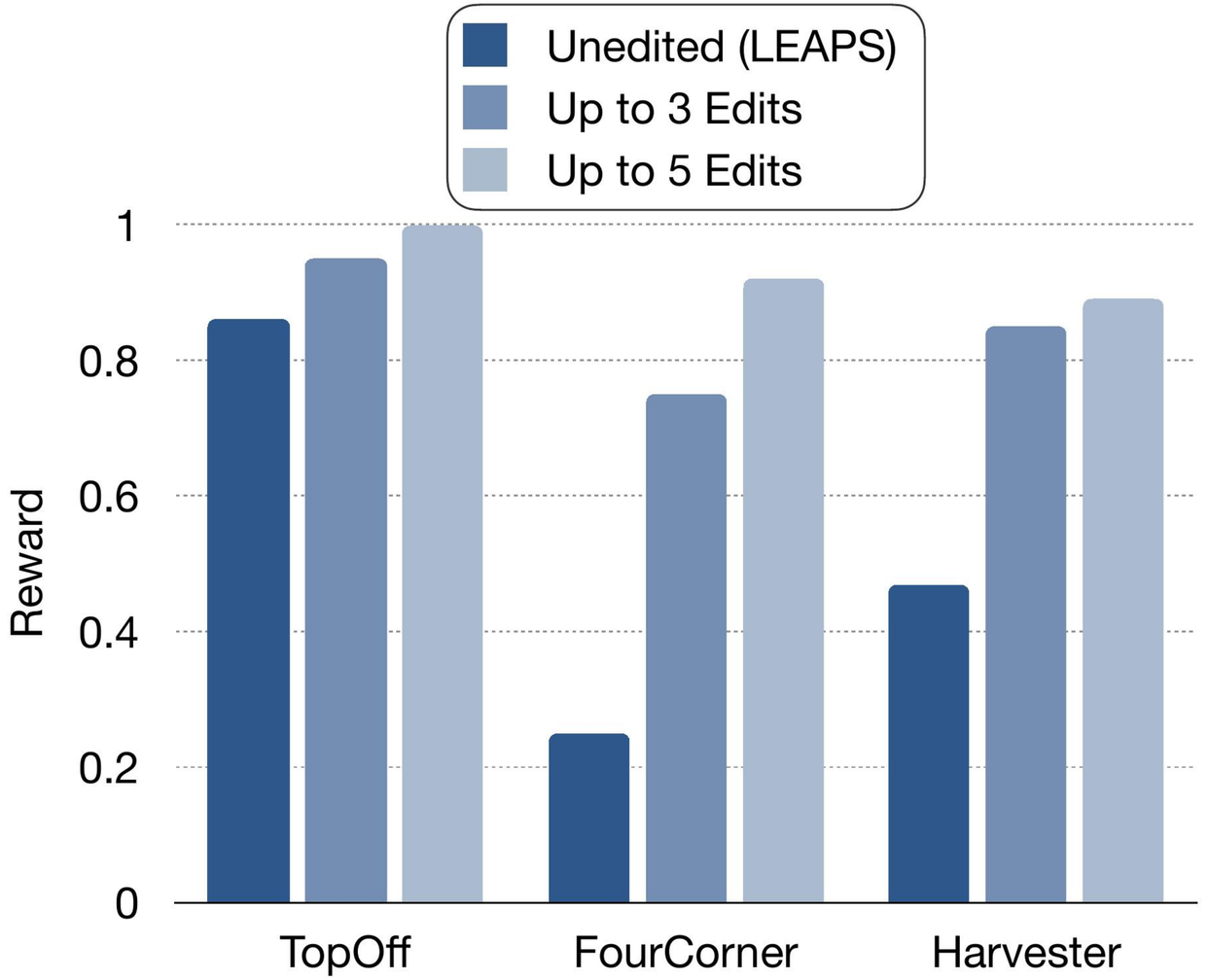


Interpretability & Interactability

Interactive Debugging Interface

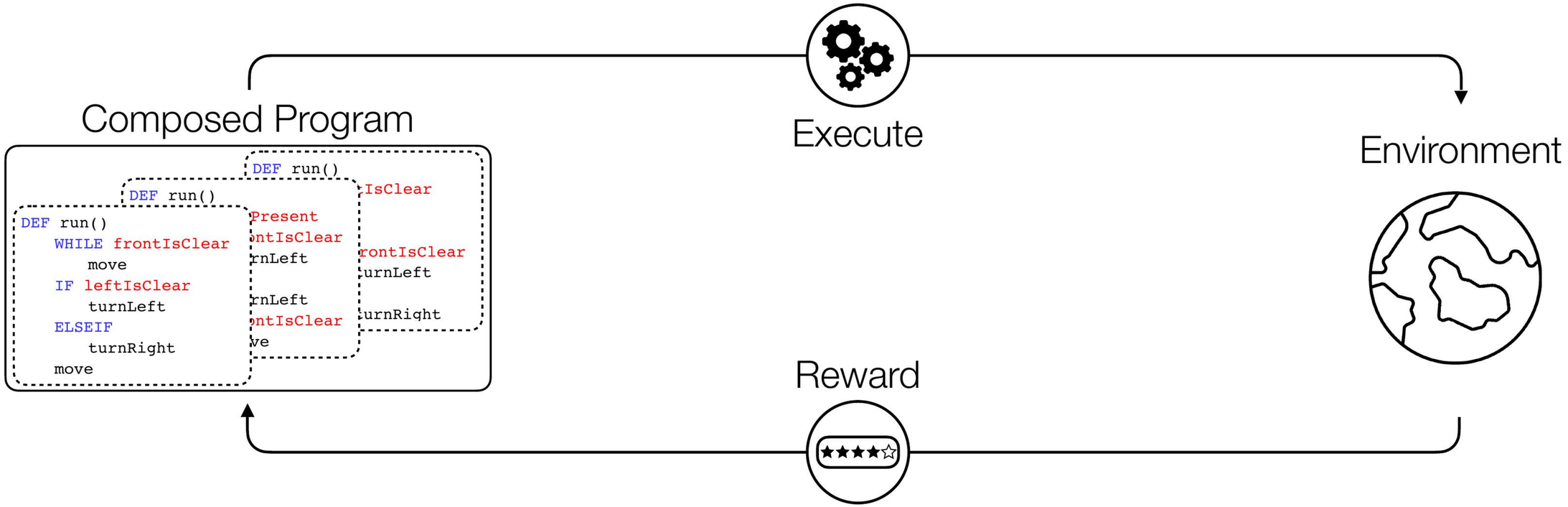


Performance Improvement



Hierarchical Programmatic Reinforcement Learning via Learning to Compose Programs

ICML 2023




Guan-Ting Liu


En-Pei Hu

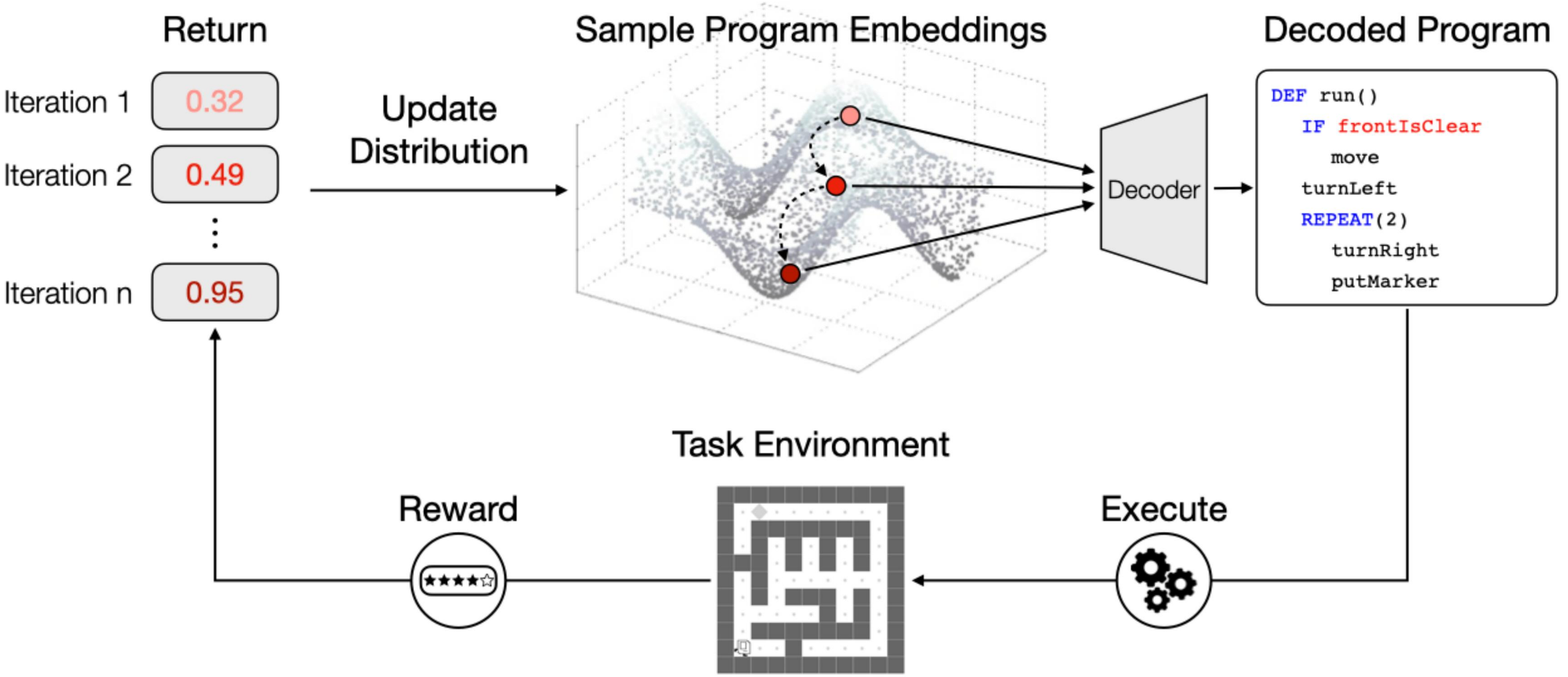

Pu-Jen Cheng


Hung-Yi Lee


Shao-Hua Sun

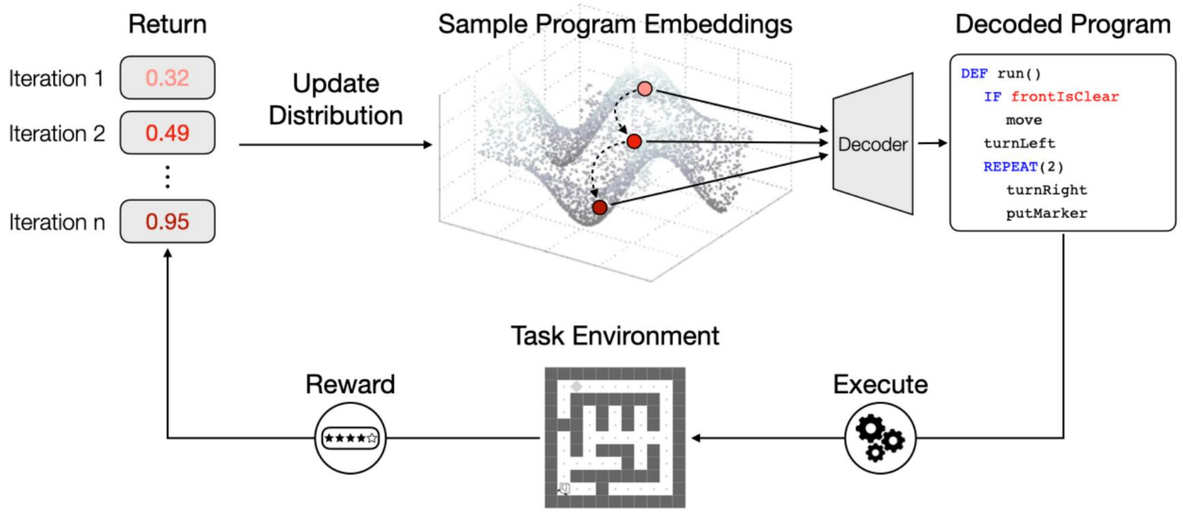
LEAPS: Learning Embeddings for IAtent Program Synthesis

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



LEAPS: Learning Embeddings for IAtent 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

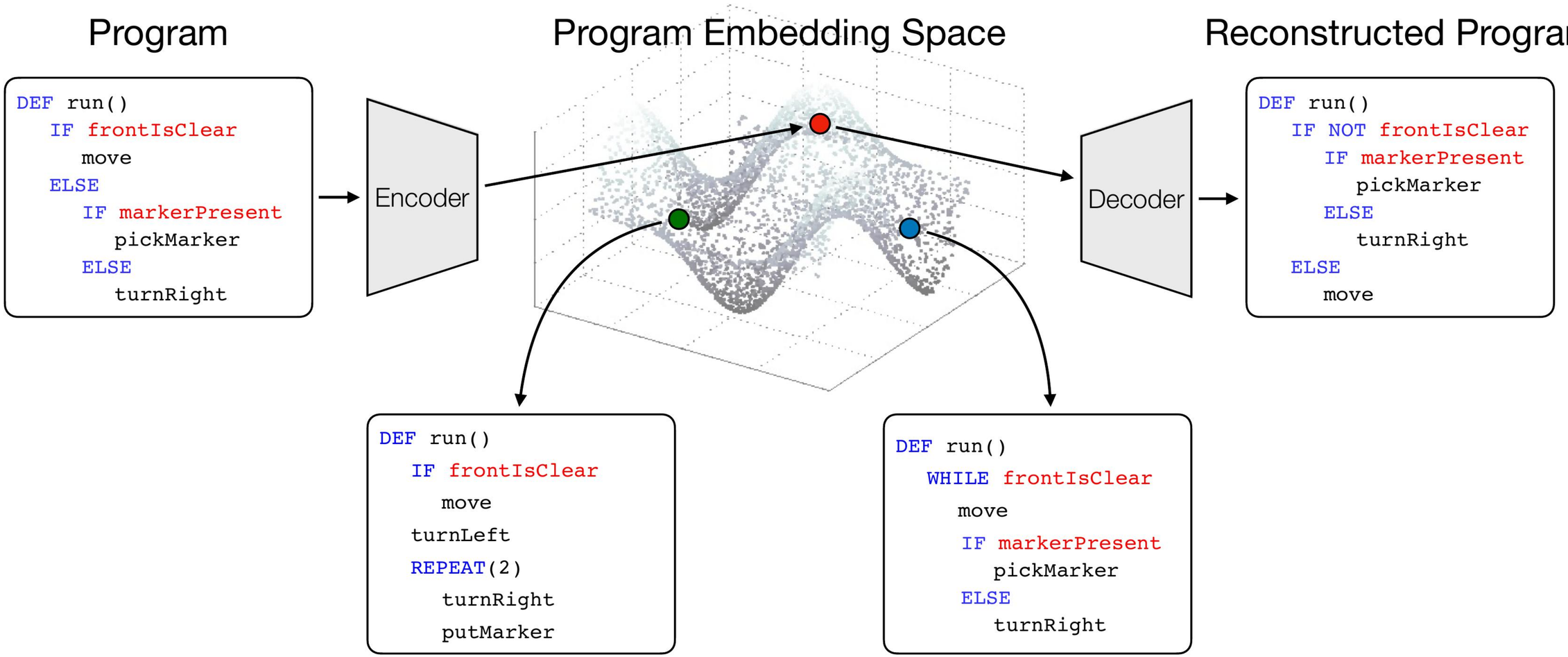
Poor credit assignment

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

Cannot accurately **attribute rewards** to corresponding program parts

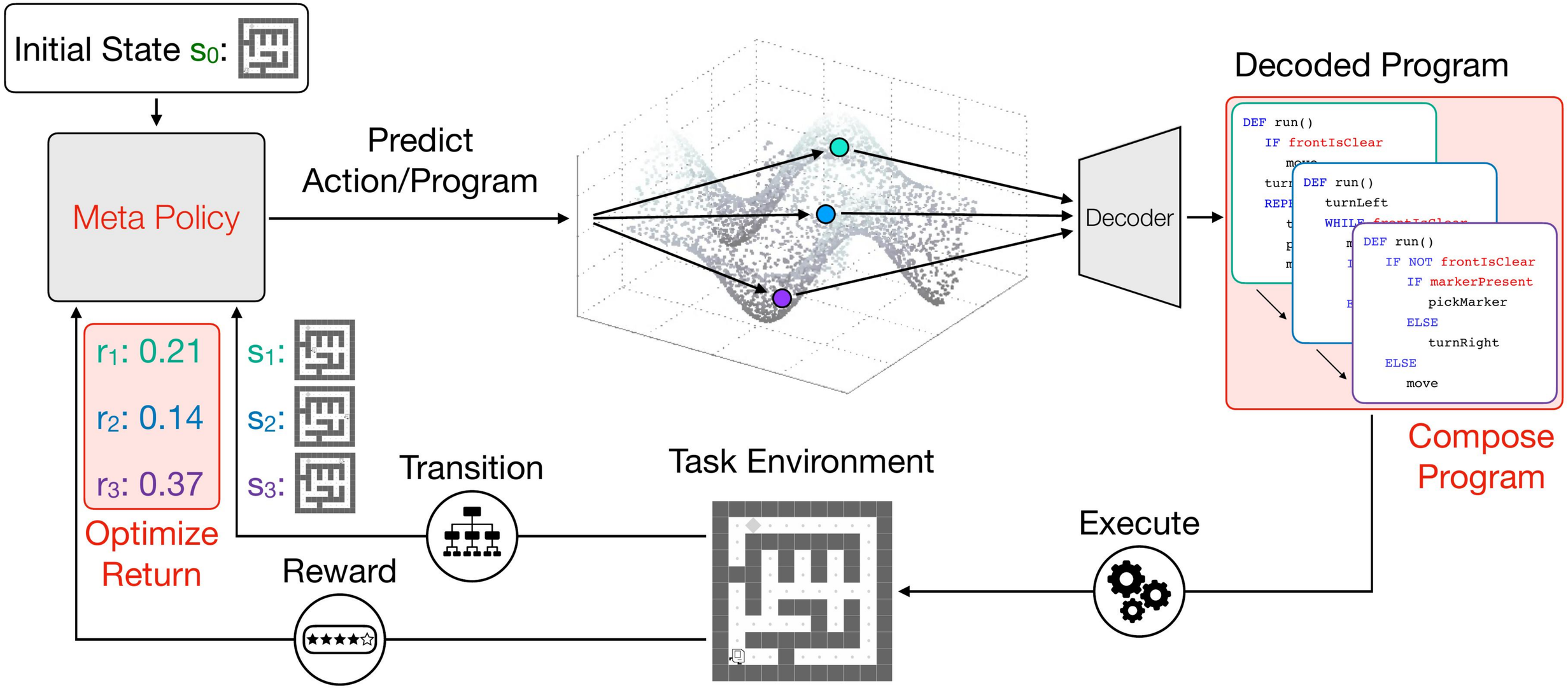
HPRL: Hierarchical Programmatic Reinforcement Learning

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

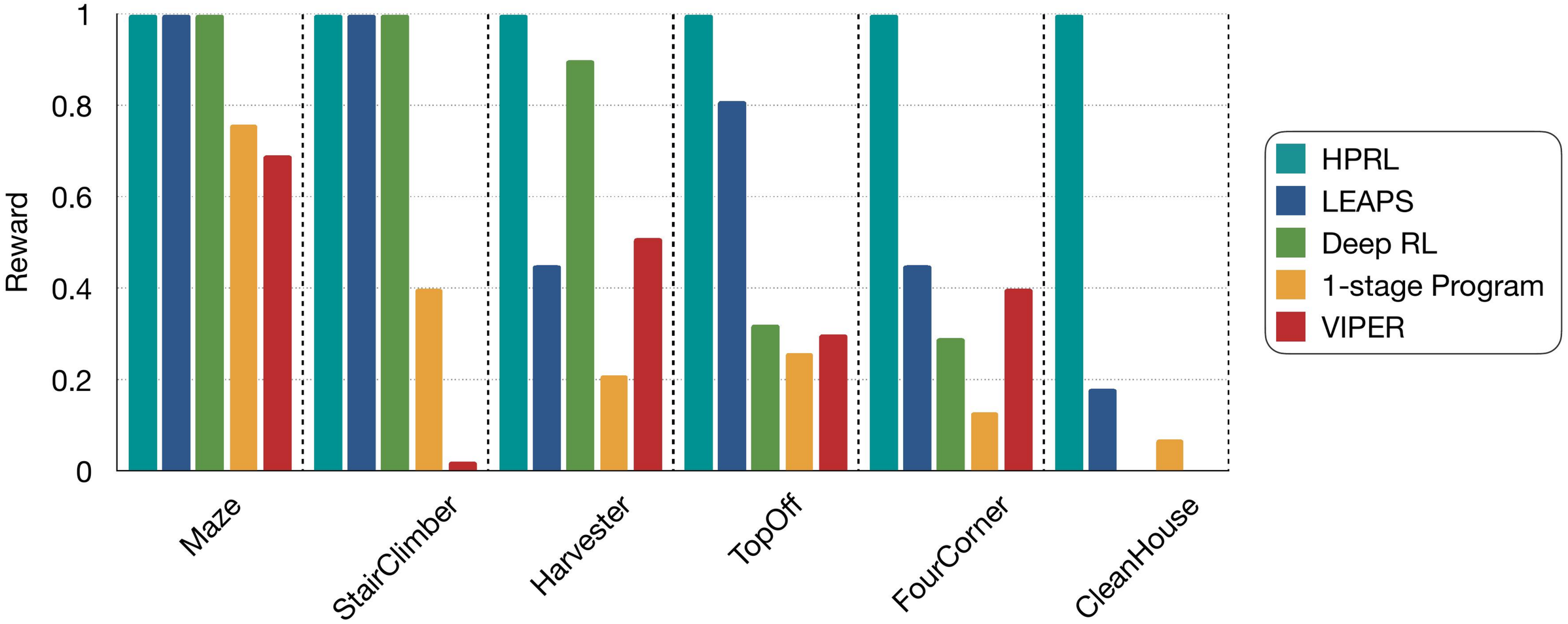


HPRL: Hierarchical Programmatic Reinforcement Learning

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**

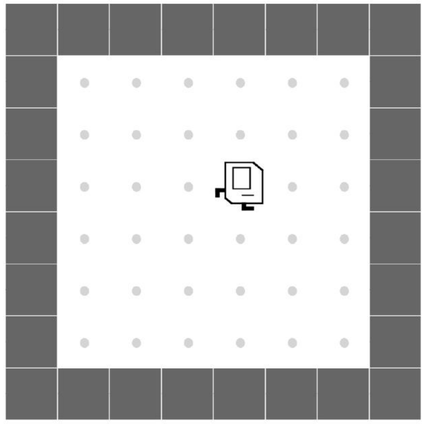


Quantitative Results - Karel Tasks

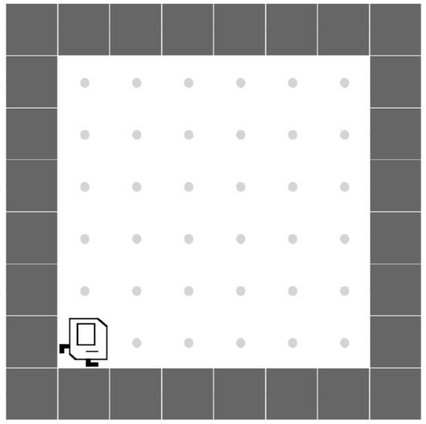


Karel-Hard Tasks

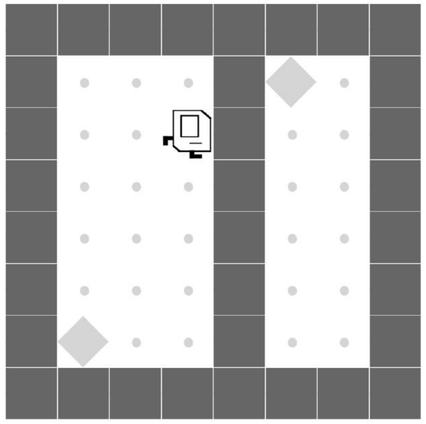
OneStroke



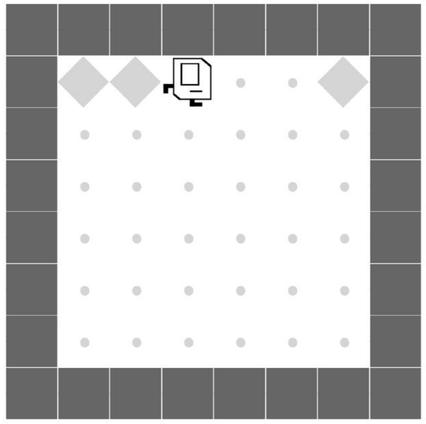
Seeder

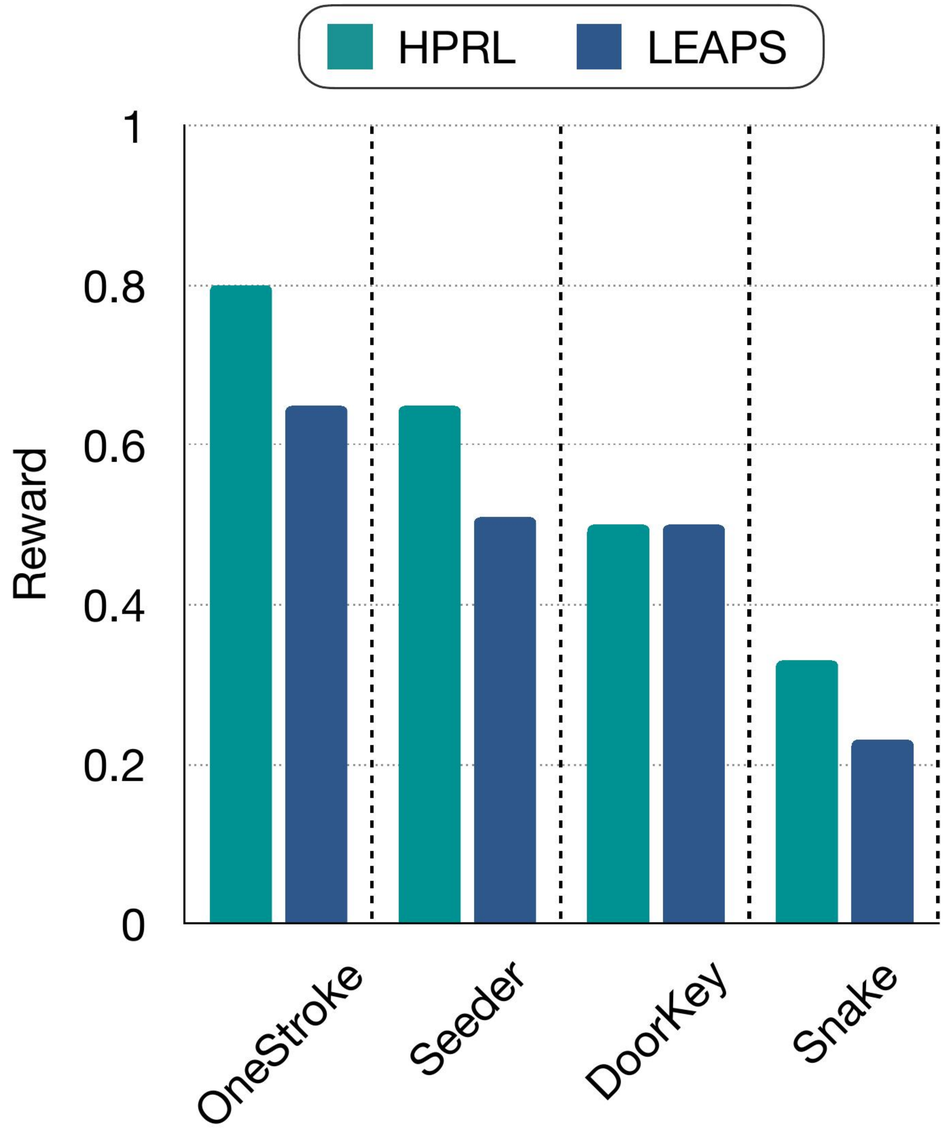


DoorKey



Snake

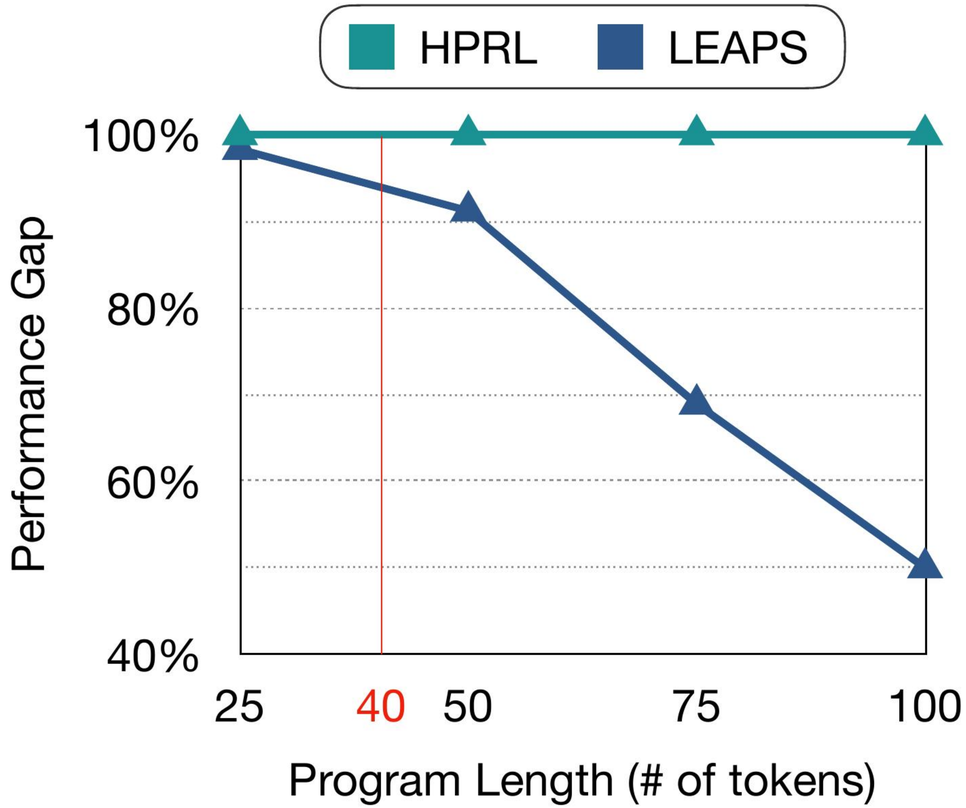




Additional Experiments

Limited program distribution

Synthesize out-of-distributionally long programs

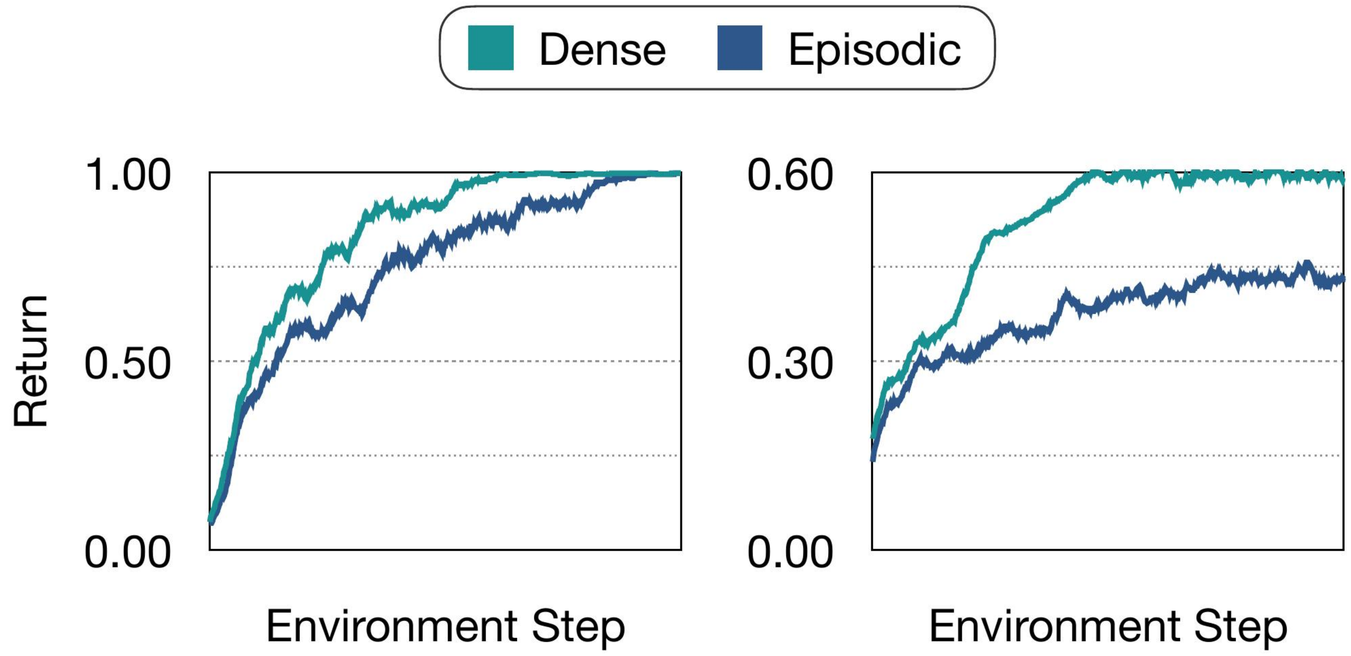


- HPRL can synthesize programs **longer** than the dataset programs (< 40 tokens) better than LEAPS

Poor credit assignment

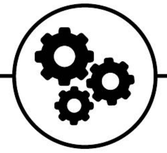
Learning from episodic reward

- **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)
```

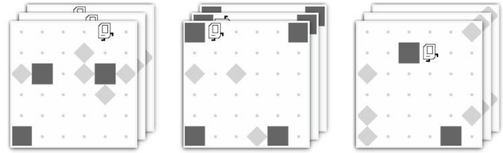
Takeaways

Program Synthesis × Reinforcement Learning

= Interpretable and Generalizable Policies

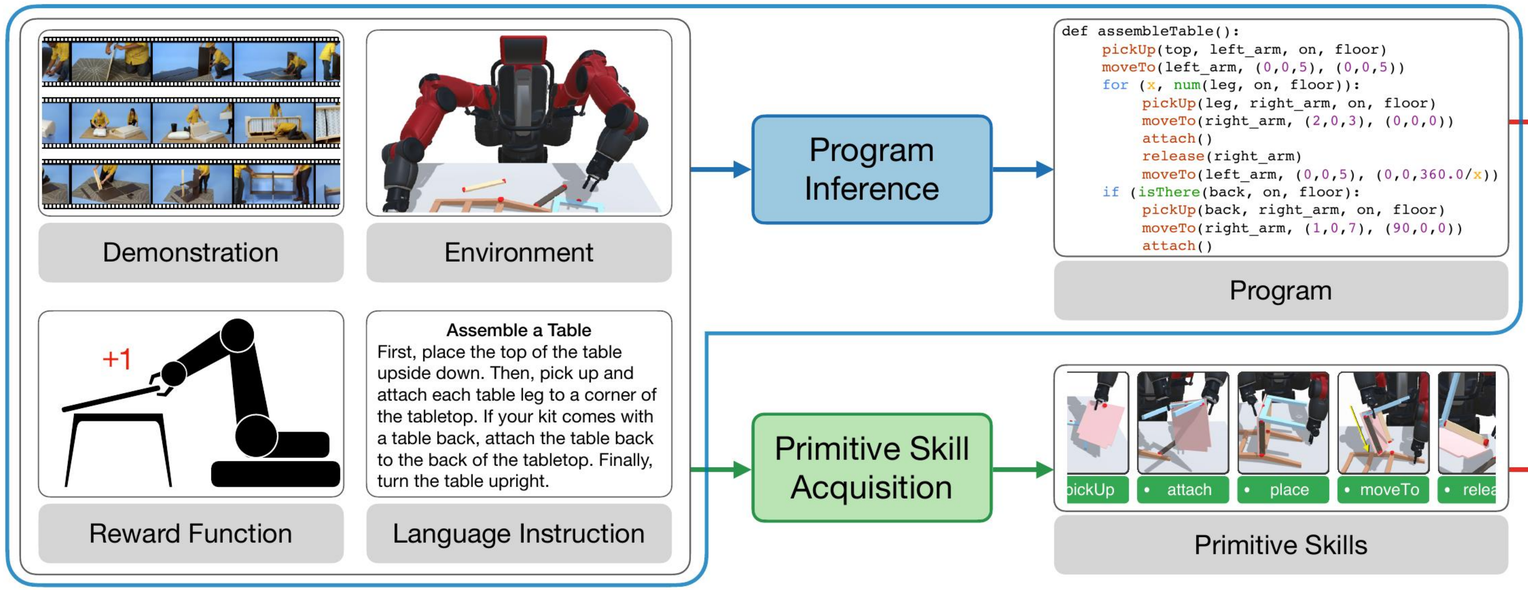
Environment

Demonstrations



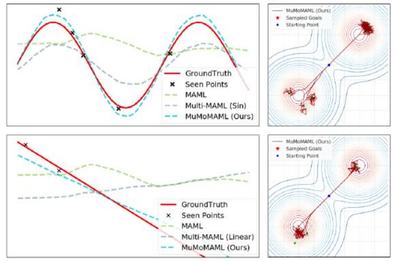
Reward

This Talk



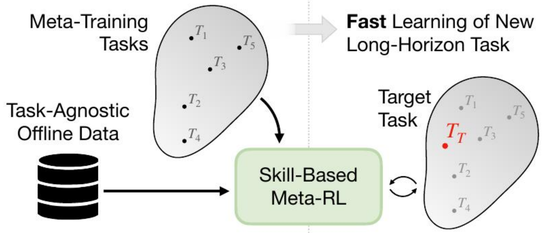
Primitive Skill Acquisition

Toward Multimodal Model-Agnostic Meta-Learning



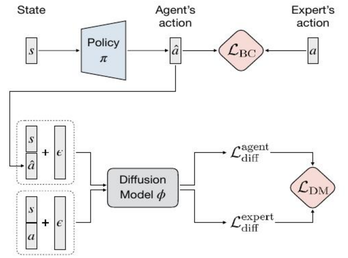
NeurIPS-W 2018

Skill-based Meta-Reinforcement Learning



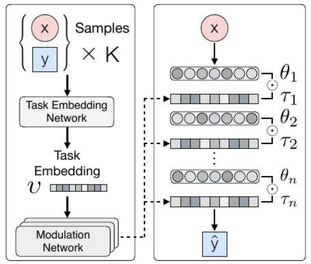
ICLR 2022

Diffusion Model-Augmented Behavioral Cloning



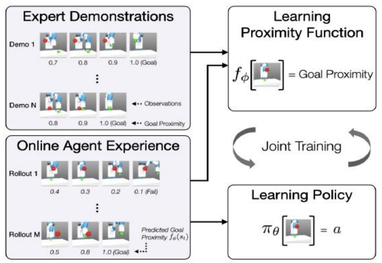
Technical Report 2023

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



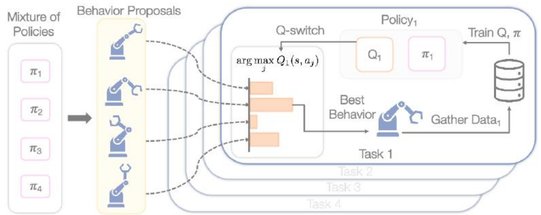
NeurIPS 2019

Generalizable Imitation Learning from Observation via Inferring Goal Proximity



NeurIPS 2021

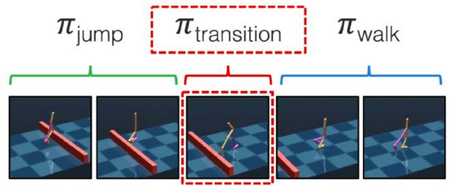
Efficient Multi-Task Reinforcement Learning via Selective Behavior Sharing



NeurIPS-W 2022

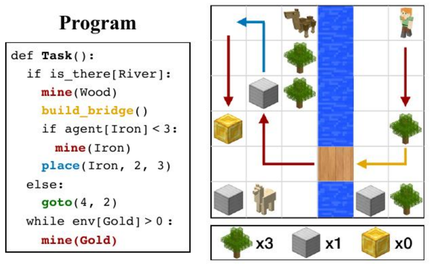
Task Execution

Composing Complex Skills by Learning Transition Policies



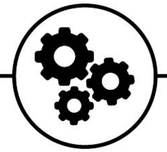
ICLR 2019

Program Guided Agent



ICLR 2020

Execute



Program

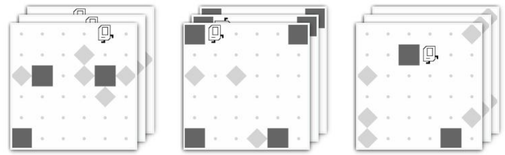
```
DEF run() m(
    WHILE c( markerPresent c) w(
        WHILE c( markerPresent c) w(
            pickMarker
            move w)
        turnRight
        move
        turnLeft
        WHILE c( markerPresent c) w(
            pickMarker
            move w)
        turnLeft
        move
        turnRight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

Thank You

Questions?

Environment

Demonstrations



Reward