OpenAI: our goal

Build safe AGI

Make it beneficial, and make its benefits be widely distributed

Components of AGI

Do hard things in simulation

Transfer skills from the simulation to the real world

Learn world models

Safety and deployment

Components of AGI

Do hard things in simulation

Transfer skills from the simulation to the real world

Learn world models

Safety and deployment

Our recent results

- OpenAI Five
- Dactyl
- Unsupervised language understanding

OpenAI Five

Dota

Dota is popular

Largest professional scene

Annual prize pool of $40M+

5 seconds

Dota is popular

Dota is hard

Strategy, tactics

Partial observability

Games are long

120 heroes, surprising interactions

20,000 actions per game, massive action space

Pros dedicate their lives to the game, 10K+ hrs of deliberate practice

Our approach

Very large scale reinforcement learning: millennia of practice

LSTM policy = honeybee brain

Self play

Reward shaping

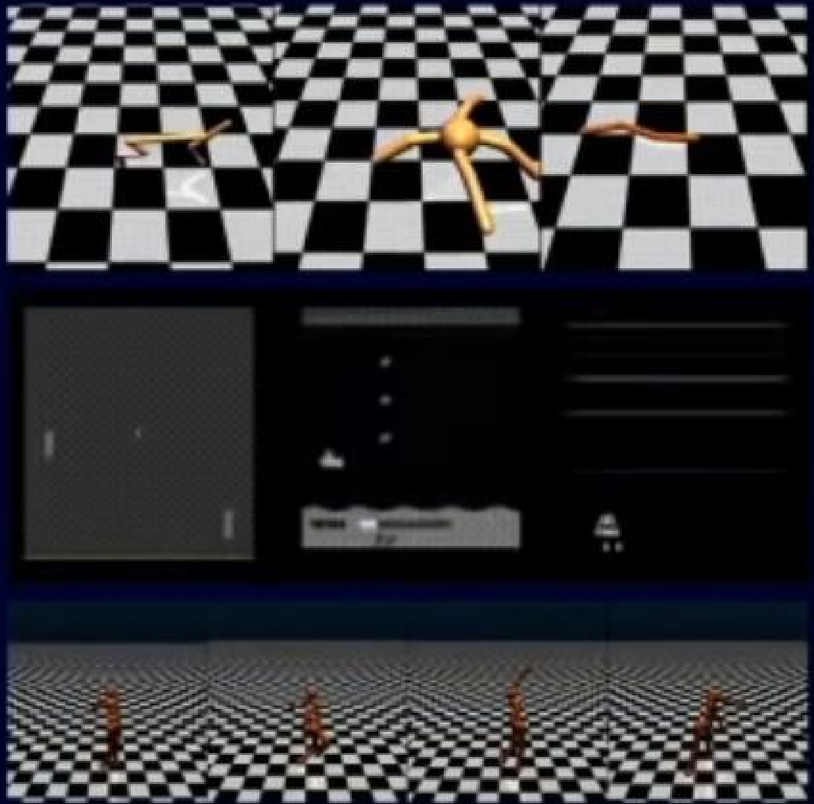
Reinforcement learning (RL) actually works!

Nearly all RL experts believed that RL can't solve tasks as hard as Dota

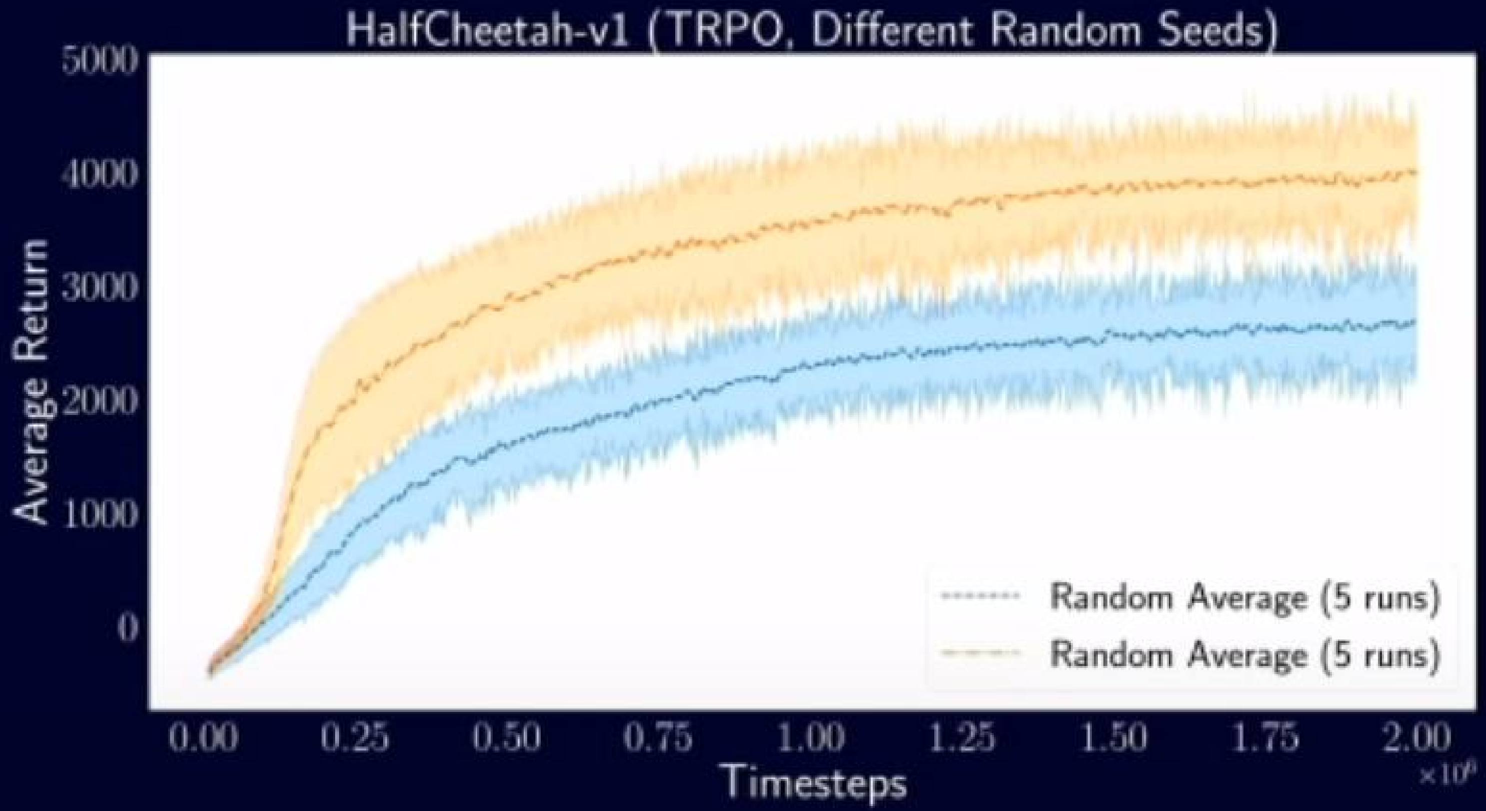
Horizon too long

Reinforcement learning (RL) actually works

Pure RL had been applied only to simple games and simple simulated robotics



Skepticism about RL



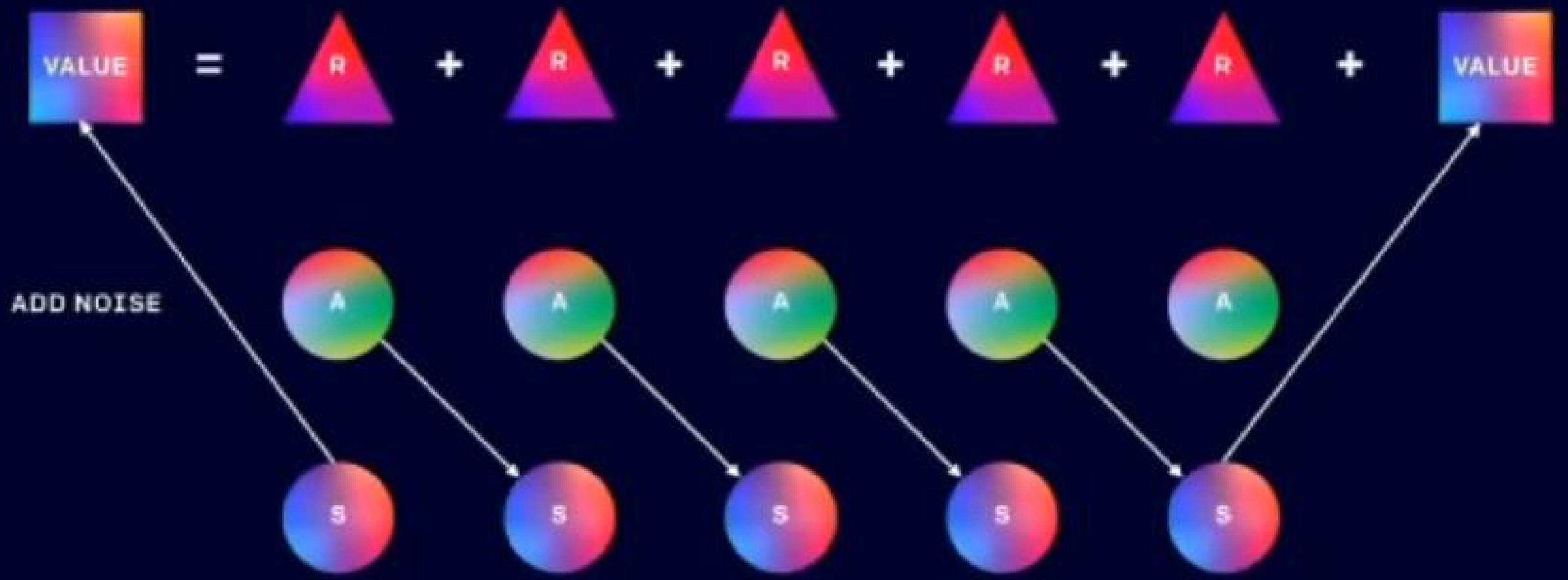
From Henderson et al., 2017

RL: the basics

Add noise to actions

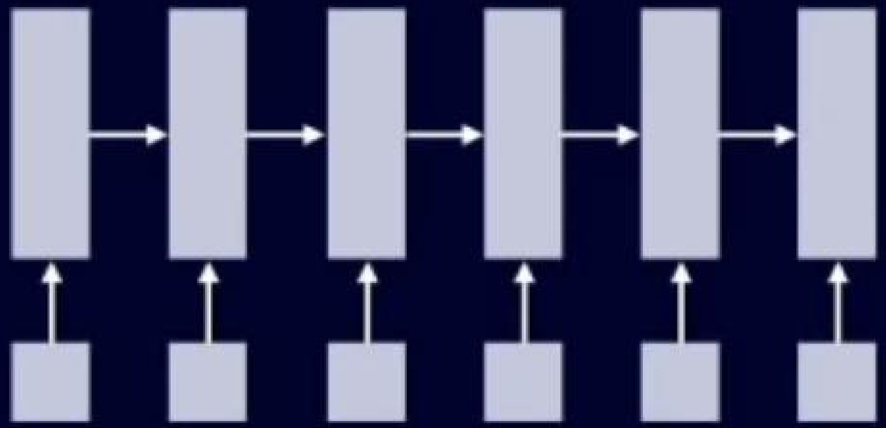
Use reward to tell if action was good

The key: Actor critic

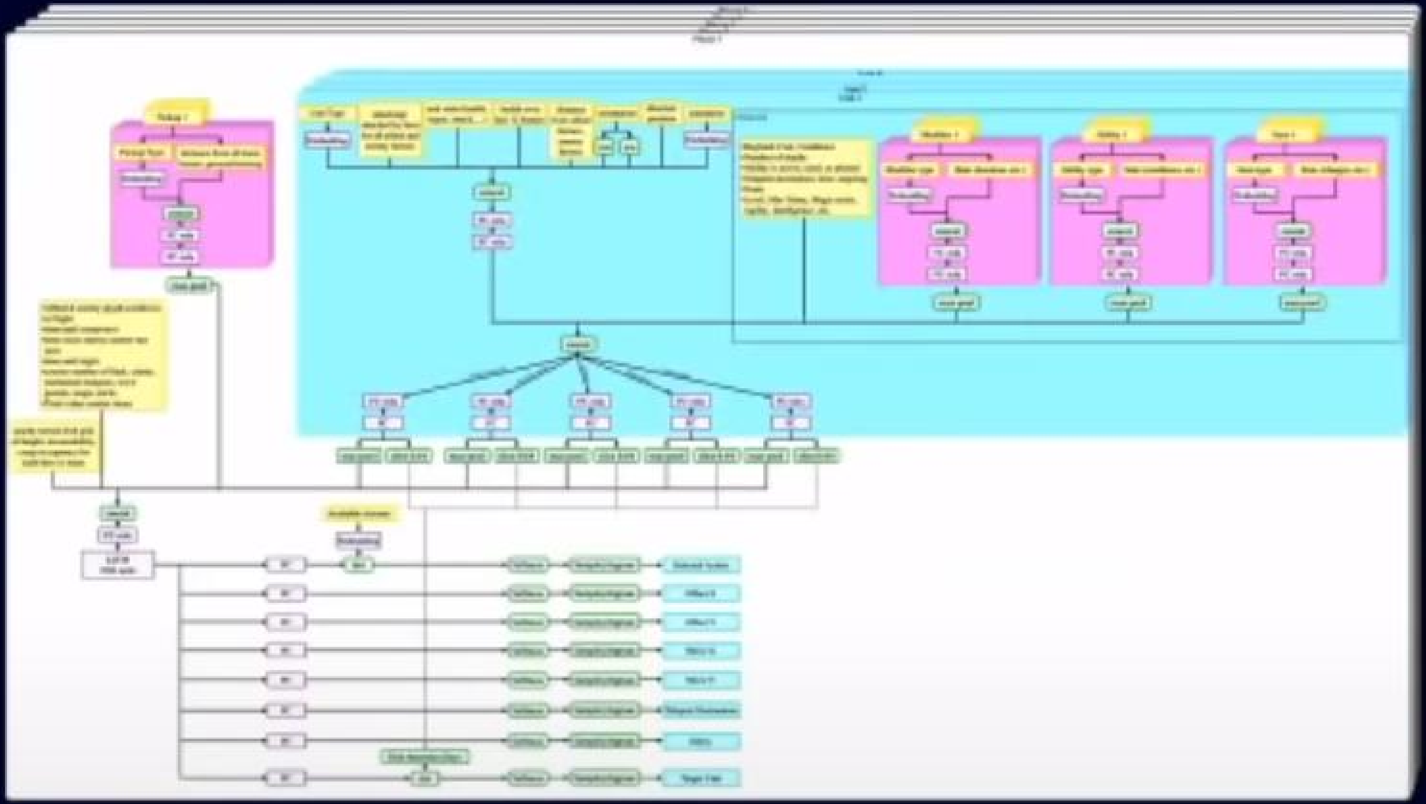


LSTM policy

A recurrent neural network that's easy to train



LSTM policy



Self play

80% of the time: play against self

20% of the time: against past versions

Cool facts

100k+ CPU cores

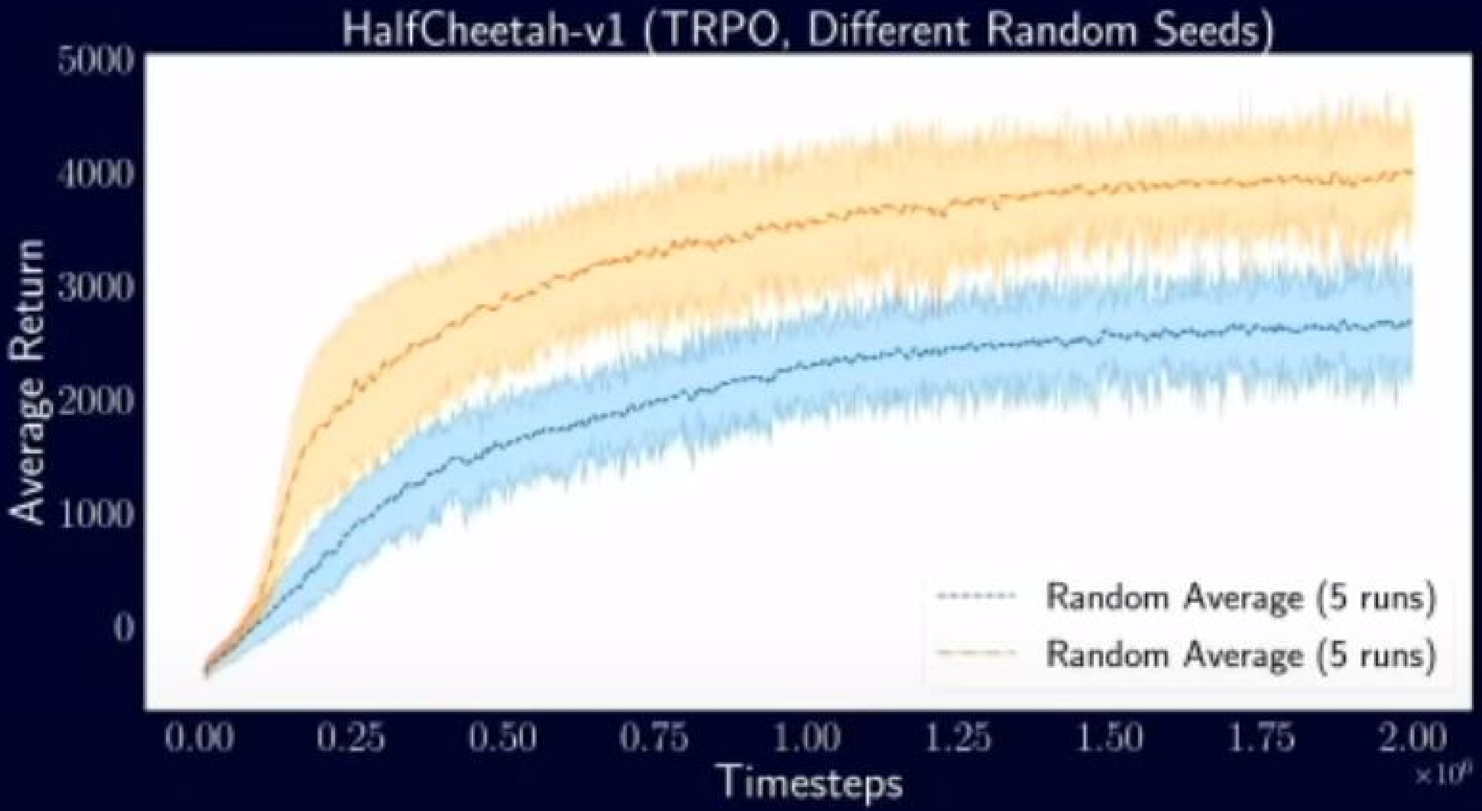
1k+ GPUs

RL time horizon of 5 minutes

- gamma = .9997

Games last for 20,000 moves

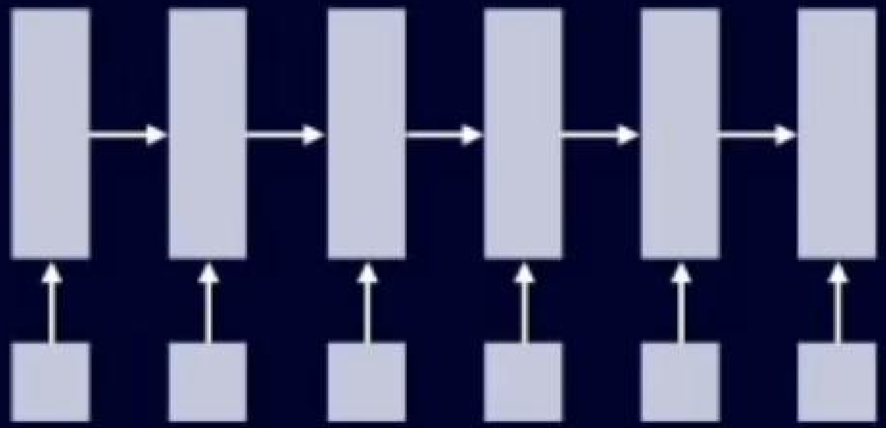
Skepticism about RL



From Henderson et al., 2017

LSTM policy

A recurrent neural network that's easy to train

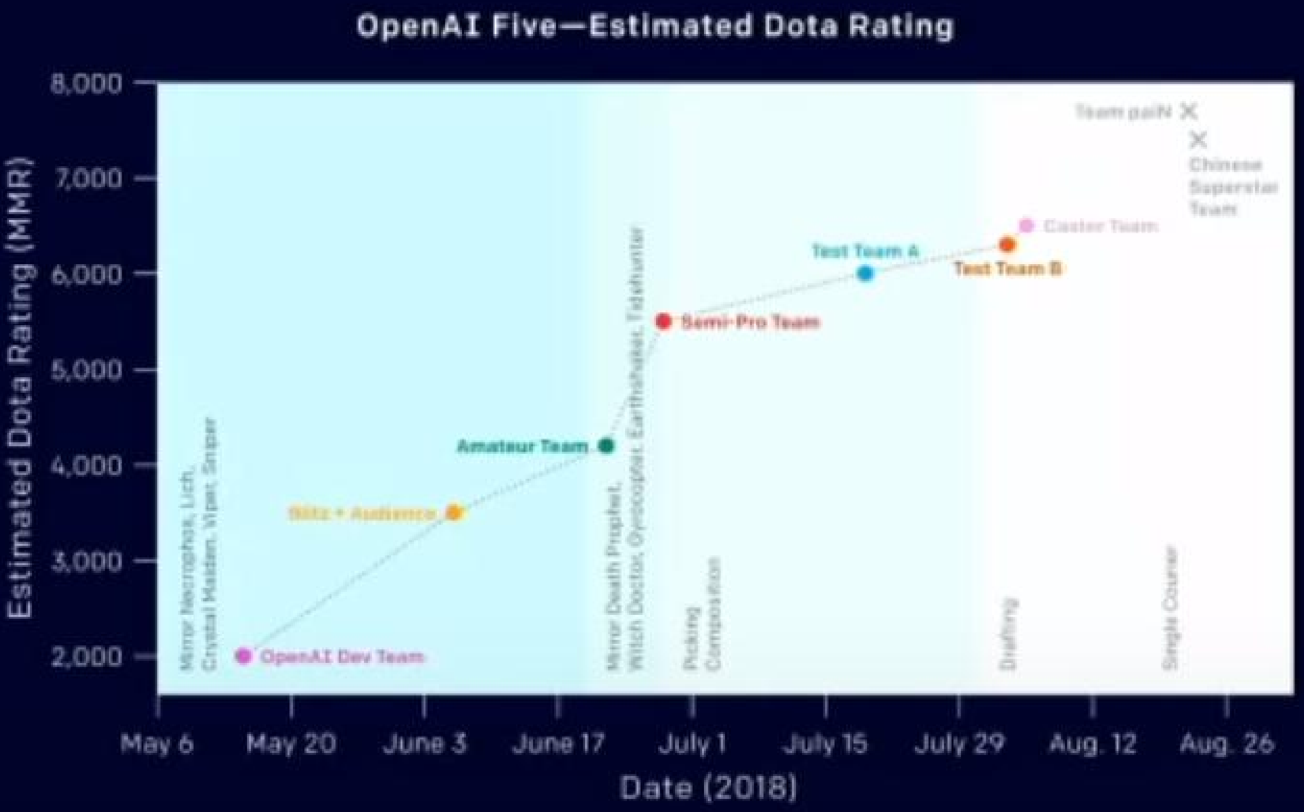


Team spirit

At first, each LSTM greedily maximized its own reward

Over time, the reward of each LSTM was made equal to the reward of the team

Results


OpenAI Five—Estimated Dota Rating
Estimated Dota Rating (MMR)
8,000
7,000
6,000
5,000
4,000
3,000
2,000
Team paiN
Chinese Superstar Team
Caster Team
Test Team A
Test Team B
Semi-Pro Team
Amateur Team
Blitz + Audience
OpenAI Dev Team
Mirror Necrophos, Lich, Crystal Maiden, Viper, Sniper
Mirror Death Prophet, Witch Doctor, Gyrocopter, Earthshaker, Tidehunter
Picking Composition
Drafting
Single Courier
May 6
May 20
June 3
June 17
July 1
July 15
July 29
Aug. 12
Aug. 26
Date (2018)

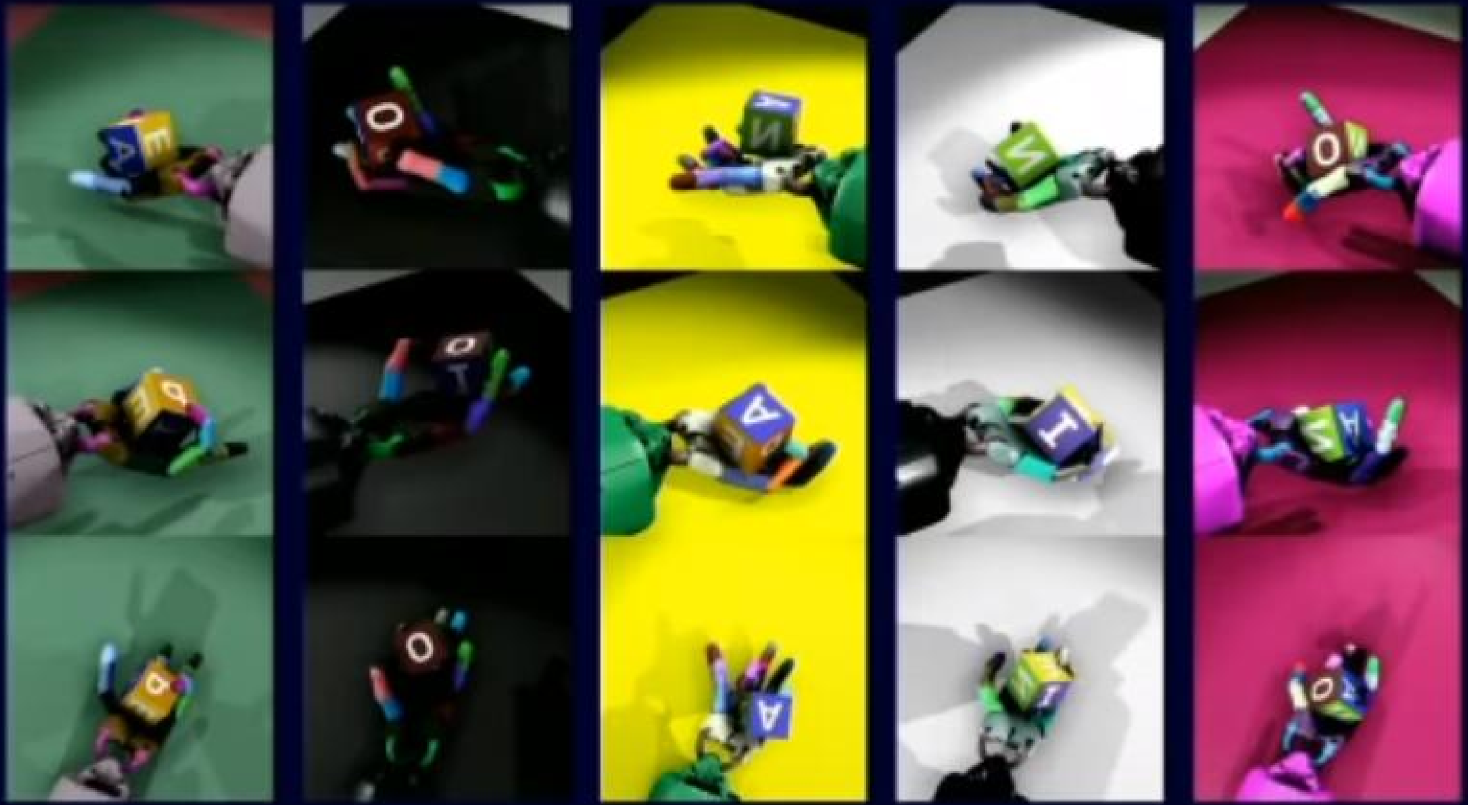
Remaining tasks

Beat the strongest teams

Dactyl: learning dexterous manipulation

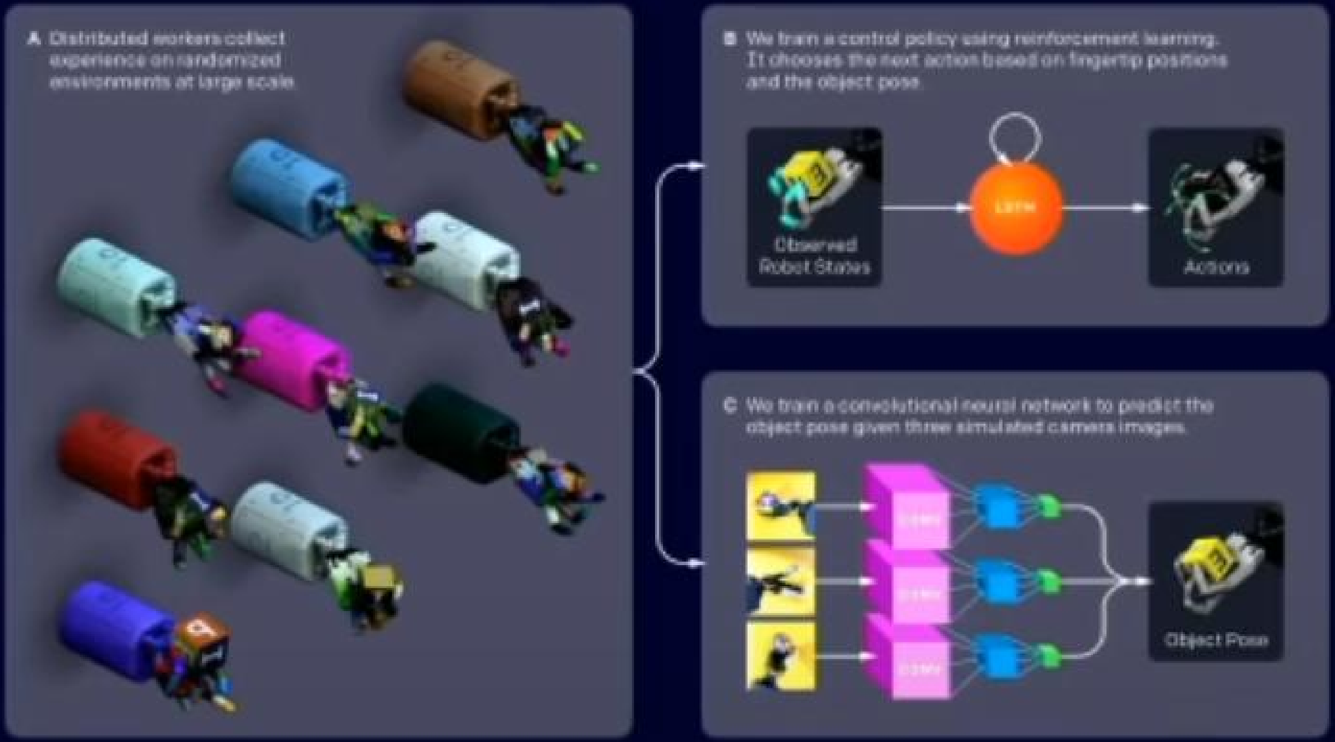
Diverse objects

Domain randomization



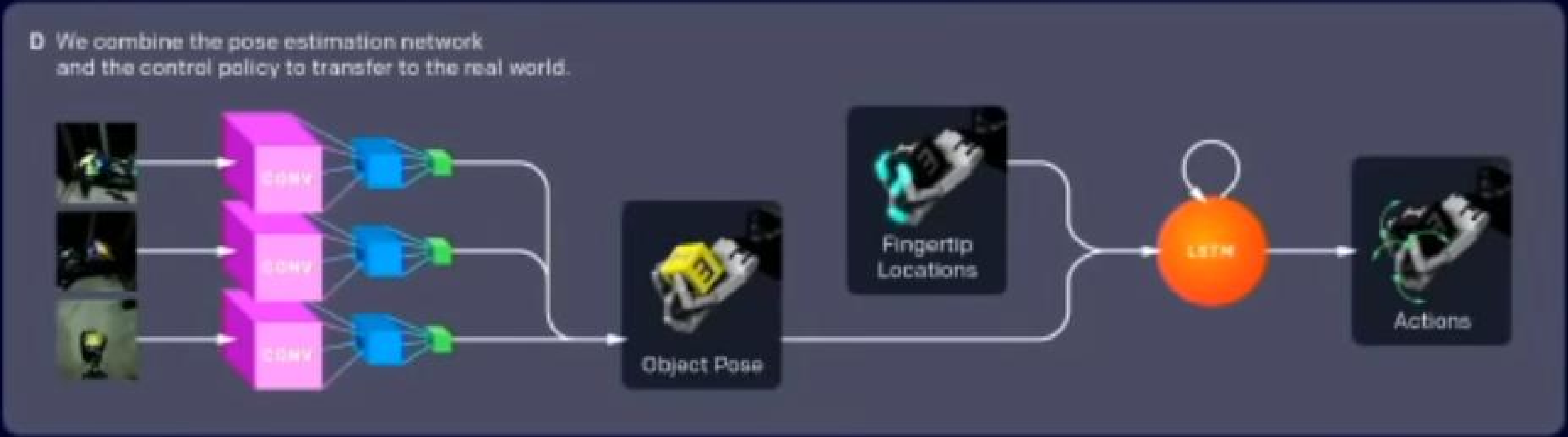
Domain randomization

Train in Simulation

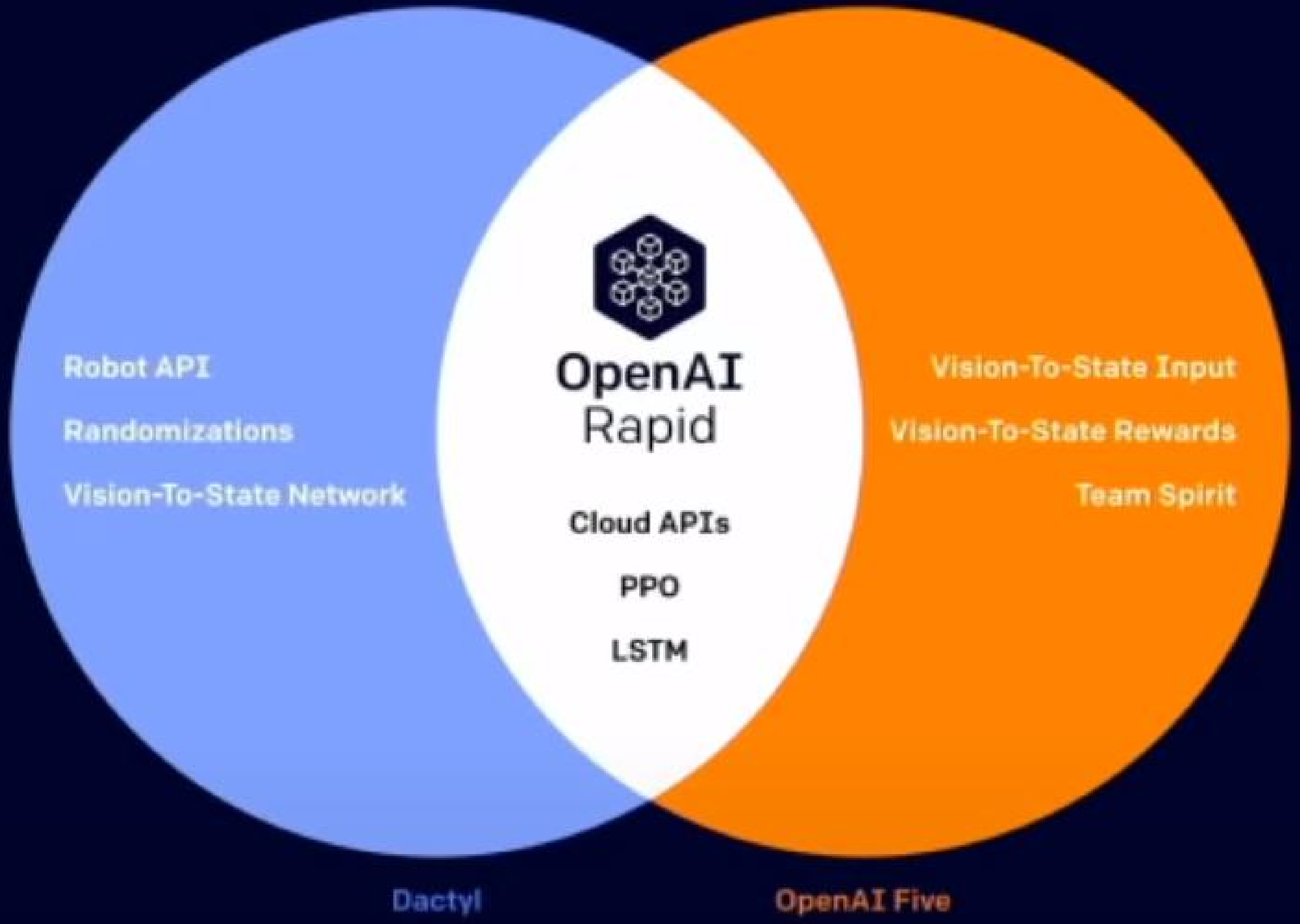


Domain randomization

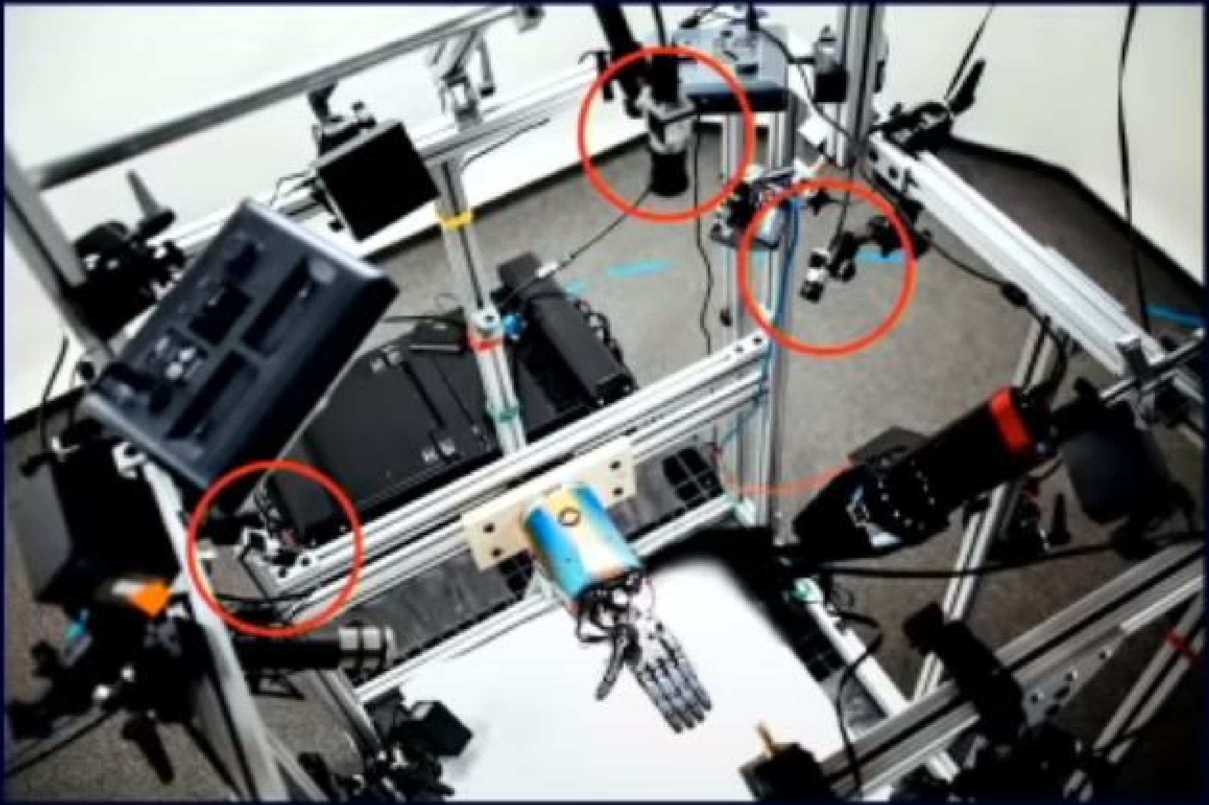
Transfer to the Real World



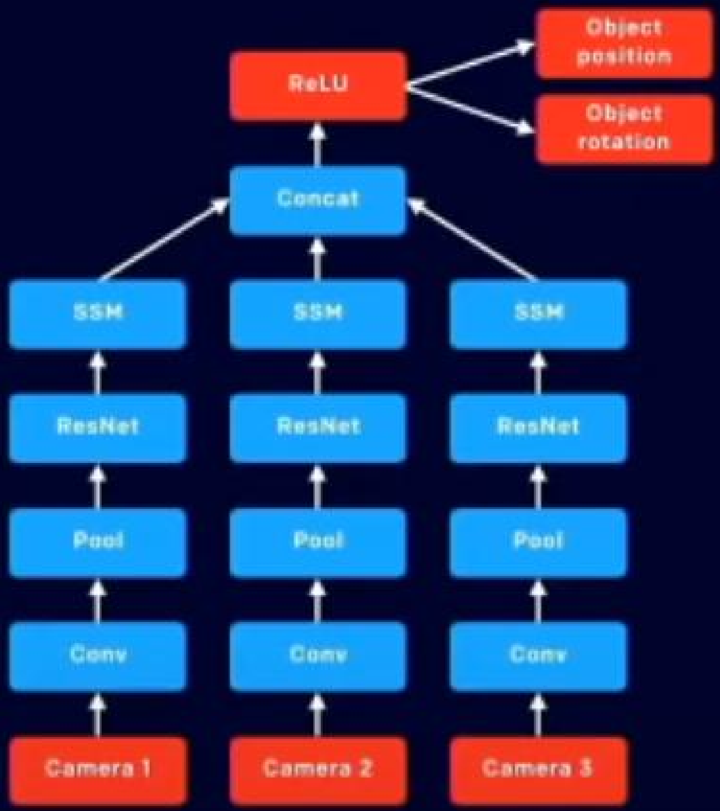
OpenAI Rapid


Robot API
Randomizations
Vision-To-State Network
OpenAI
Rapid
Cloud APIs
PPO
LSTM
Vision-To-State Input
Vision-To-State Rewards
Team Spirit
Dactyl
OpenAI Five

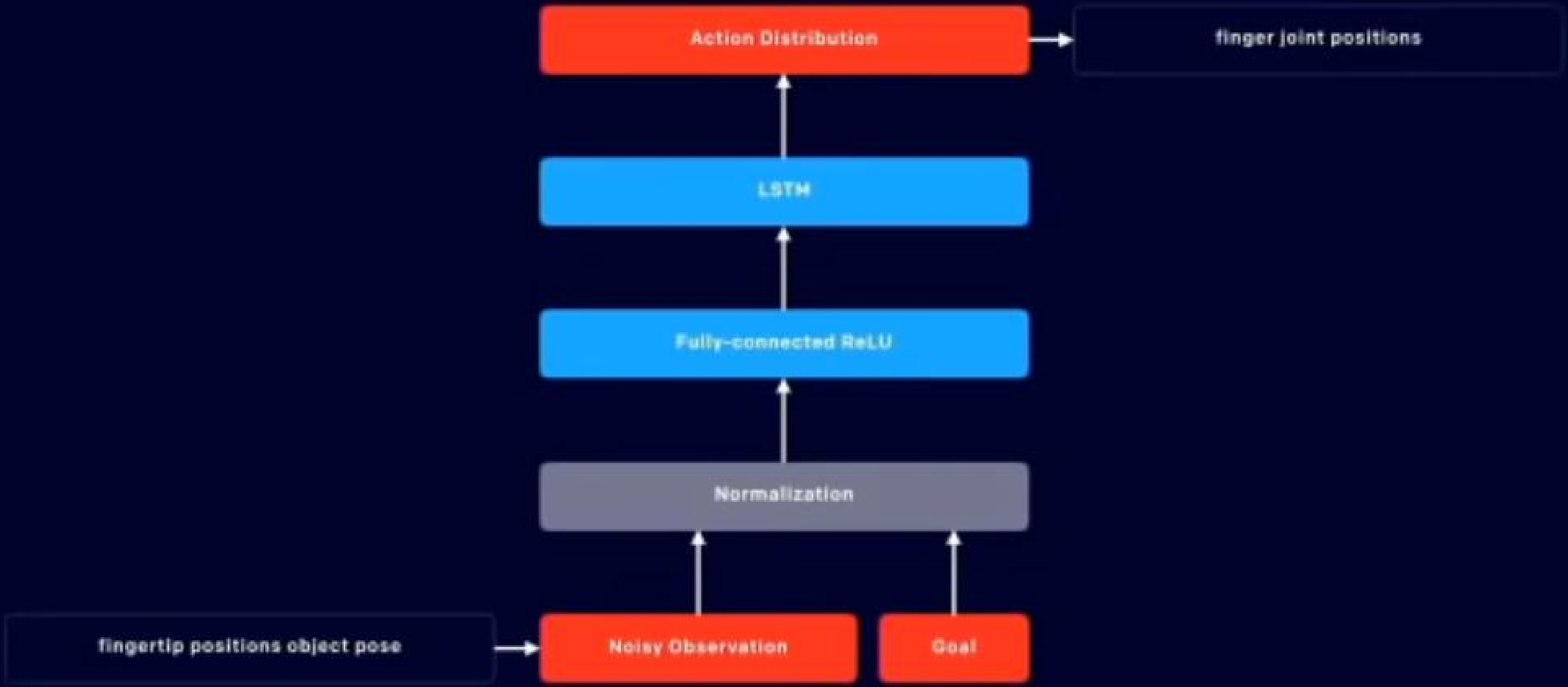
Vision-based system



Vision Architecture



Policy Architecture



Improving Language Understanding with Unsupervised Learning

Dataset	Task	SOTA	Ours
SNLI	Textual Entailment	89.3	**89.9**
MNLI Matched	Textual Entailment	80.6	**82.1**
MNLI Mismatched	Textual Entailment	80.1	**81.4**
SciTail	Textual Entailment	83.3	**88.3**
QNLI	Textual Entailment	82.3	**88.1**
RTE	Textual Entailment	**61.7**	56.0
STS-B	Semantic Similarity	81.0	**82.0**
QQP	Semantic Similarity	66.1	**70.3**
MRPC	Semantic Similarity	**86.0**	82.3
RACE	Reading Comprehension	53.3	**59.0**
ROCStories	Commonsense Reasoning	77.6	**86.5**
COPA	Commonsense Reasoning	71.2	**78.6**
SST-2	Sentiment Analysis	**93.2**	91.3
CoLA	Linguistic Acceptability	35.0	**45.4**
GLUE	Multi Task Benchmark	68.9	**72.8**

Biggest improvement:

Details

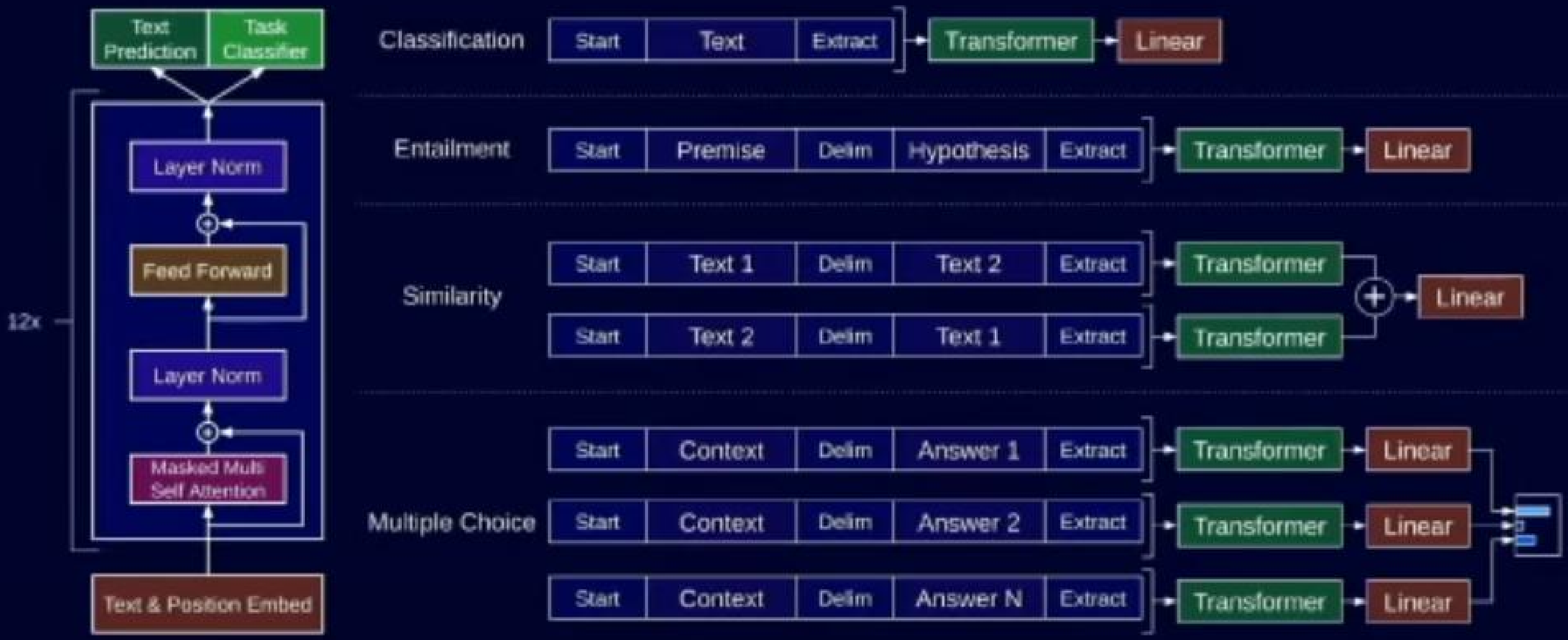
The model: a transformer

Dataset: a corpus of books

Context size: 512

Training time: 8 P100s for 1 month

Details



Text Prediction
Task Classifier
Layer Norm
Feed Forward
Layer Norm
Masked Multi Self Attention
12x
Text & Position Embed
Classification
Start
Text
Extract
Transformer
Linear
Entailment
Start
Premise
Delim
Hypothesis
Extract
Transformer
Linear
Similarity
Start
Text 1
Delim
Text 2
Extract
Transformer
Start
Text 2
Delim
Text 1
Extract
Transformer
Linear
Multiple Choice
Start
Context
Delim
Answer 1
Extract
Transformer
Linear
Start
Context
Delim
Answer 2
Extract
Transformer
Linear
Start
Context
Delim
Answer N
Extract
Transformer
Linear

Can the current AI boom scale to AGI?

GOAL IS TO PRESENT EVIDENCE THAT:

While highly uncertain, near-term AGI should be taken as a serious possibility.

Lessons from history of science

Lessons from history of AI

Fundamental limits of deep learning

Practical limits on compute

"With few exceptions, scientists seem to make rather poor prophets; this is rather surprising, for imagination is one of the first requirements of a good scientist. Yet, time and again, distinguished astronomers and physicists have made utter fools of themselves by declaring publicly that such-and-such a project was impossible."

— *Profiles of the Future* (1962)

Moving goalposts

Simon Newcomb

1901: heavier-than-air flight is impossible

1908: it's possible, but won't be important since flying machines will never scale to both pilot and a passenger

First airplane flight, at Kitty Hawk (December 17, 1903)

Aversion to scale

On extending V-2 (**14-ton** rocket) technology to 5-ton payload (**200-ton** rocket):

America: never going to happen

Russia: let's do this

Result: Russia first to space

Replica German V-2 rocket (invented 1942)

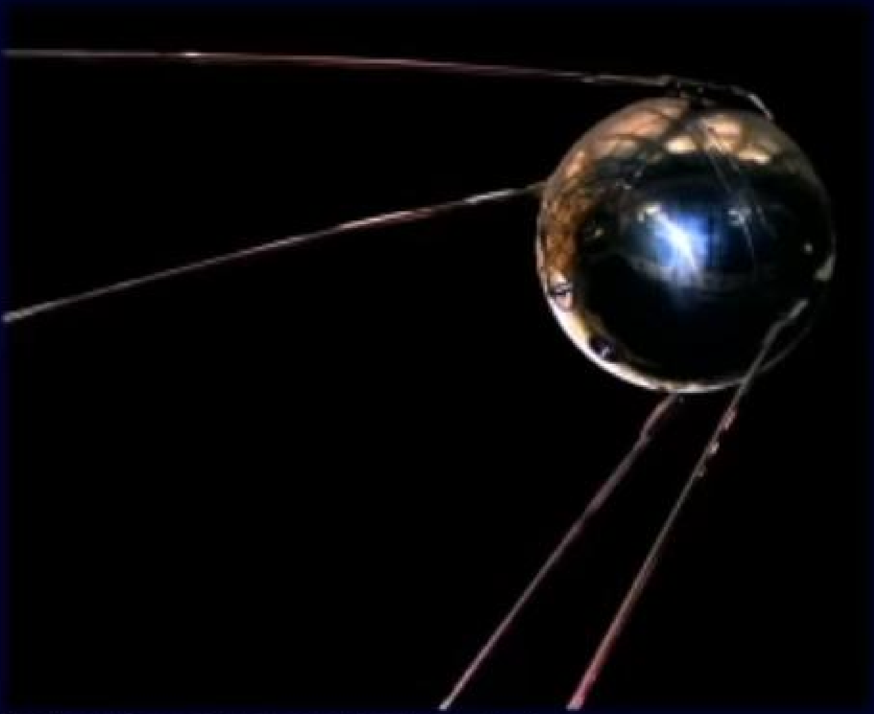
Replica Russian R-7 rocket, first reliable means to transport objects into Earth orbit (launched 1957)

Habitual detractors

1936: "It must be said at once that the whole procedure sketched in the present volume presents difficulties so fundamental a nature that we are forced to dismiss the notion as essentially impracticable, in spite of the author's insistent appeal to put aside prejudice and to recollect the supposed impossibility of heavier-than-air flight before it was actually accomplished."

1956: "Space travel is utter bilge"

—Riet Woolley, Astronomer Royal of the UK



Replica of Sputnik 1 (Launched October 4, 1957)

Lessons from history of science

Lessons from history of AI

Fundamental limits of deep learning

Practical limits on compute

"We have the impression that many people in the connectionist community do not understand that [back-propagation] is merely a particular way to compute a gradient and have assumed that back-propagation is a new learning scheme that somehow gets around the basic limitations of hill-climbing."

—Minsky & Papert (1988)

History of AI—narrative I'd heard

1960s: Perceptron

1970s: expert systems

1980s: backprop

1990s: SVM + kernel trick

2012-: ImageNet

2020-: ??

Guyon

Vapnik

Cortes

Sutskever, Krizhevsky, and Hinton

Community reacted to hype around perceptrons

Rosenblatt & perceptron (1950s)

"The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence. Later perceptrons will be able to recognize people and call out their names and instantly translate speech in one language to speech and writing in another language, it was predicted."

—New York Times (1959)

Perceptrons (1969)

Papert

“There was *some* hostility in the energy behind the research reported in *Perceptrons*...Part of our drive came, as we quite plainly acknowledged in our book, from the fact that funding and research energy were being dissipated on...misleading attempts to use connectionist methods in practical applications.”

—Papert (1988)

Where was the money going?

Minsky

> "In the late 1950s and early 1960s, after Rosenblatt's work, there was a great wave of neural network research activity. There were maybe thousands of projects. For example Stanford Research Institute had a good project. But nothing happened. The machines were very limited. So I would say by 1965 people were getting worried. They were trying to get money to **build bigger machines**, but they didn't seem to be going anywhere."
>
> —Minsky (1989)

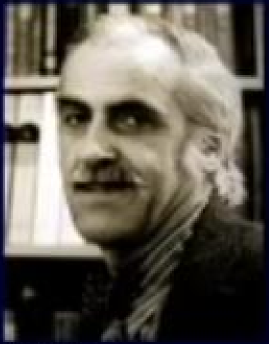
Neural networks revival

"In the early 1980s, dramatic decreases in computing costs brought about a 'democratization' in the access to computing resources."

—A Sociological Study of the Official History of the Perceptrons Controversy (1996)

Rummelhart

Hinton



McClelland

Anderson

History of AI—alternate narrative

Tesauro with TD-Gammon (1992)

Deep learning has consistently scaled with compute for **60 years**

New levels of compute steers researchers to develop new algorithms such as backpropagation, not the other way around

Fads have been more political than technical

"It took great chutzpah for Gerald Tesauro to start **wasting computer cycles** on temporal difference learning in the game of Backgammon" [32k parameters; total training: **5s on modern GPU**]

—Pollack & Blair (1997)

Lessons from history of science

Lessons from history of AI

Fundamental limits of deep learning

Practical limits on compute

"The speed with which those who once declaimed, 'It's impossible' can switch to, 'I said it could be done all the time' is really astounding."

—Profiles of the Future (1962)

AlexNet (2012)

Before

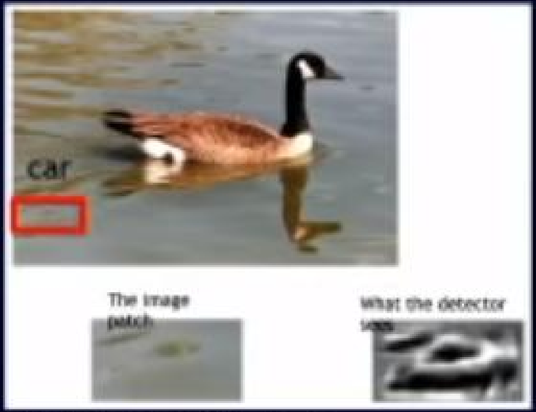
Fantasy: one algorithm to solve speech recognition, machine translation, object recognition better than decades of domain-specific ingenuity

The result

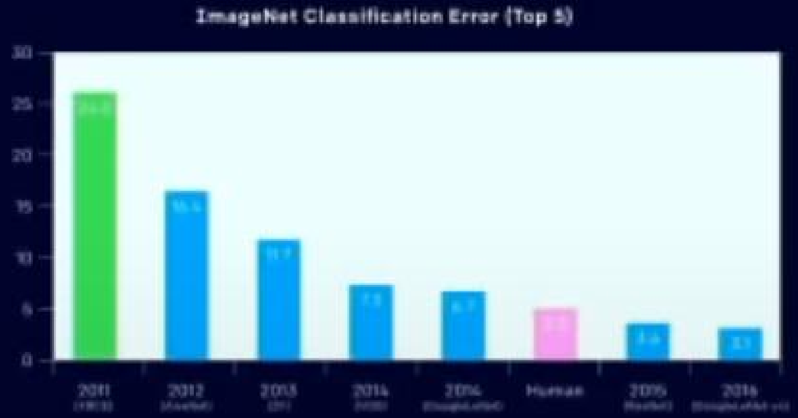
60M-parameter neural network which got way better performance on the ImageNet dataset than any other approach

After

Neural networks dominant approach in these fields



HOG features (2005)— figure by Torrabla



DQN (2013)

Before

Deep learning is about static datasets

The result

150,000-parameter feed-forward neural network which learned to play a number of Atari games purely from pixels and score

After

RL + deep networks might actually be able to observe and act in the world

Neural Machine Translation (2015)

Before

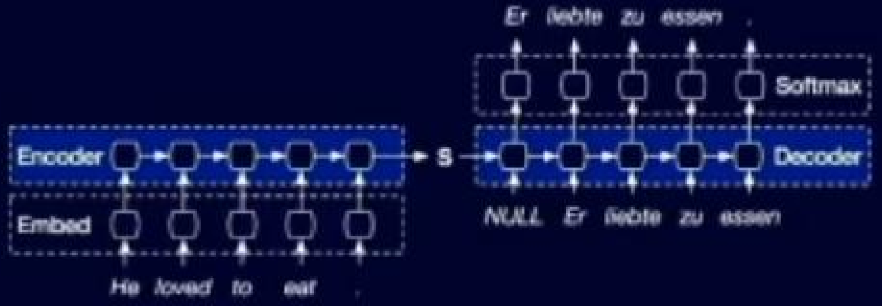
Deep learning is just perception

The result

380M-parameter neural network mapping input sequence in one language to output sequence in another

After

Deep learning can solve even the hardest supervised learning problems, regardless of type signature



AlphaGo (2016)

Before

RL can't actually solve hard tasks

The result

72-million parameter feedforward neural network, coupled with MCTS, to defeat top humans at Go

After

RL + MCTS can solve hard problems, given: discrete actions, modest action space, simulator at test time

OpenAI Five (2018)

Before

RL can't solve hard tasks on its own—long-term planning is fundamental barrier

The result

100M-parameter LSTM competitive with (not yet exceeding) top humans in the esports game Dota 2

After

RL can solve extremely hard problems with long-term planning, given only training-time simulator

Dactyl (2018)

Before

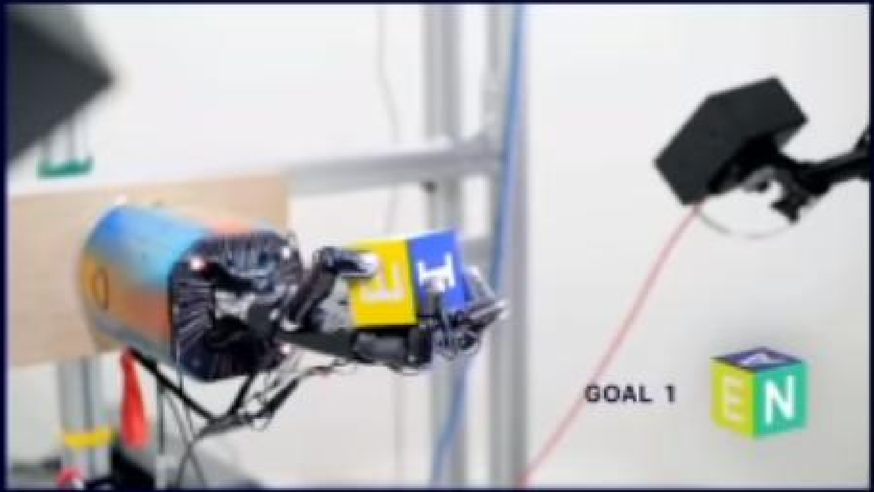
Deep RL is limited to games and other perfectly-simulatable problems

The result

1.5M-parameter LSTM, trained in simulation by the OpenAI Five training system & learning algorithm, deployed on a physical robot hand to manipulate a cube

After

Deep RL can cross reality gap given only a "good enough" simulator at training time



Unsupervised NLP (2018)

Before

AI progress is driven by labeled datasets.

The result

117M-parameter transformer trained by reading 7,000 self-published books which, with a small amount of supervised fine-tuning, sets state-of-the-art on a huge variety of NLP datasets

After

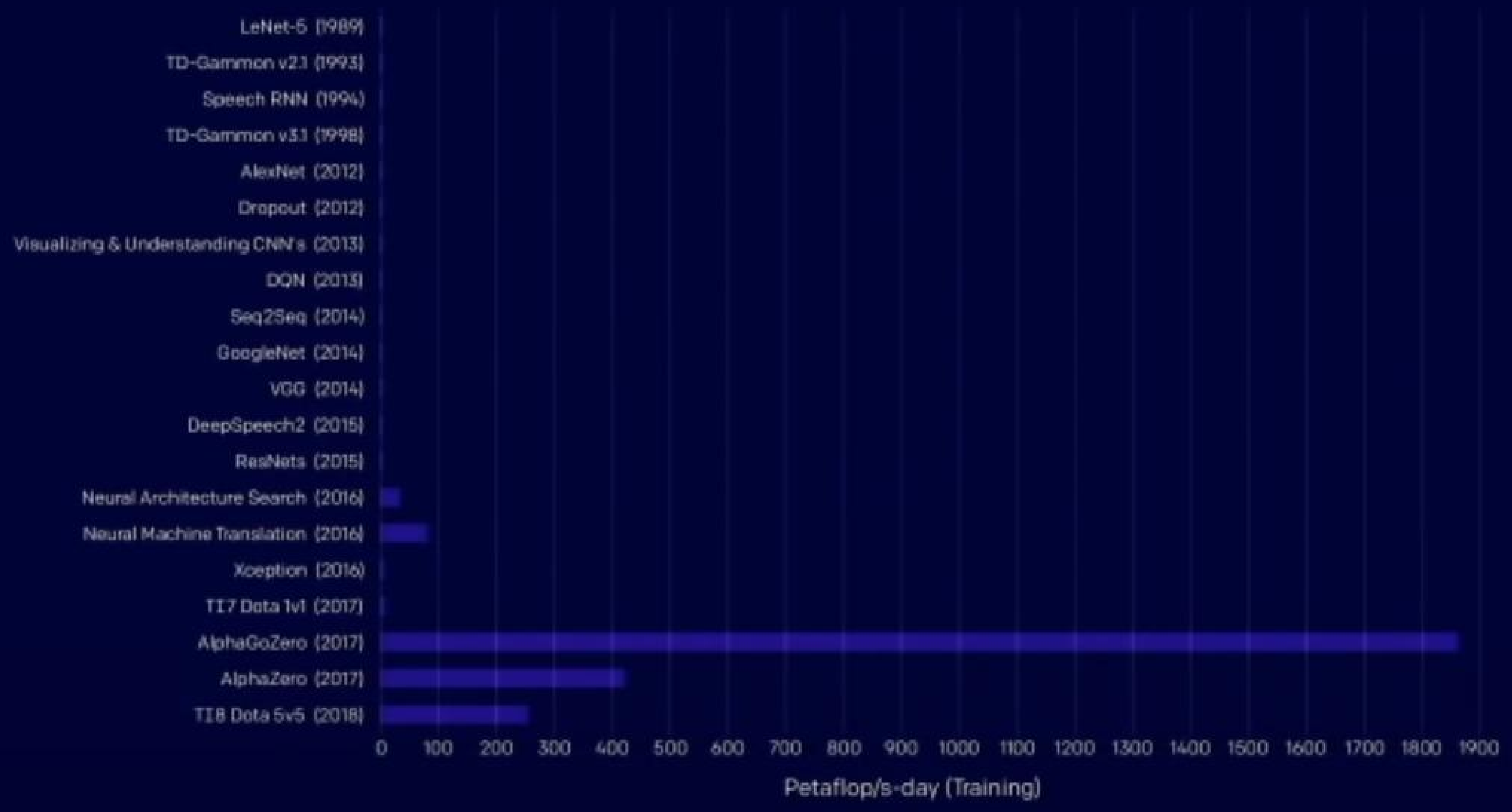
Unlabeled data can be even more important than labeled.

Dataset	Task	SOTA	Ours
SNLI	Textual Entailment	89.3	**89.9**
MNLI Matched	Textual Entailment	80.6	**82.1**
MNLI Mismatched	Textual Entailment	80.1	**81.4**
SciTail	Textual Entailment	83.3	**88.3**
QNLI	Textual Entailment	82.3	**88.1**
RTE	Textual Entailment	**61.7**	56.0
STS-B	Semantic Similarity	81.0	**82.0**
QQP	Semantic Similarity	66.1	**70.3**
MRPC	Semantic Similarity	**86.0**	82.3
RACE	Reading Comprehension	53.3	**59.0**
ROCStories	Commonsense Reasoning	77.6	**86.5**
COPA	Commonsense Reasoning	71.2	**78.6**
SST-2	Sentiment Analysis	**93.2**	91.3
CoLA	Linguistic Acceptability	35.0	**45.4**
GLUE	Multi Task Benchmark	68.9	**72.8**

Can we put a confident bound on near-term compute progress?

Doubling period: 3.5 months

AlexNet to AlphaGo Zero: A 300,000x Increase in Compute



LeNet-5 (1989)
TD-Gammon v2.1 (1993)
Speech RNN (1994)
TD-Gammon v3.1 (1998)
AlexNet (2012)
Dropout (2012)
Visualizing & Understanding CNN's (2013)
DQN (2013)
Seq2Seq (2014)
GoogleNet (2014)
VGG (2014)
DeepSpeech2 (2015)
ResNets (2015)
Neural Architecture Search (2016)
Neural Machine Translation (2016)
Xception (2016)
TI7 Dota 1v1 (2017)
AlphaGoZero (2017)
AlphaZero (2017)
TI8 Dota 5v5 (2018)
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
Petaflop/s-day (Training)

THIS TALK'S GOAL IS TO PRESENT EVIDENCE THAT:

While highly uncertain,
near-term AGI should be taken
as a serious possibility.

Means proactively thinking about risks:

- Machines pursuing goals misspecified by their operator
- Malicious humans subverting deployed systems
- Out-of-control economy that grows without resulting in improvements to human lives

Thank You