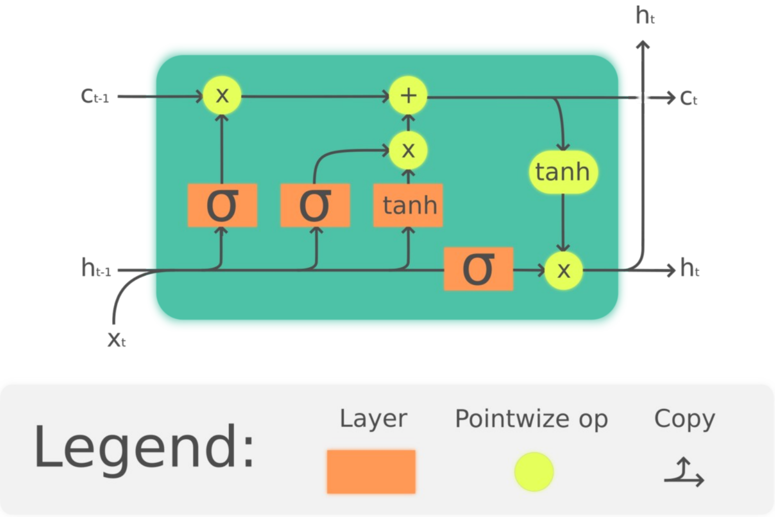
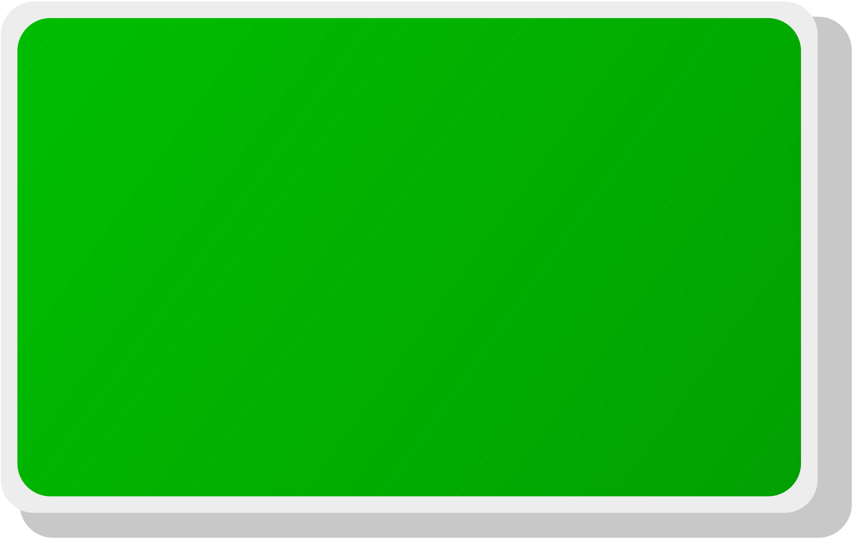


LSTM

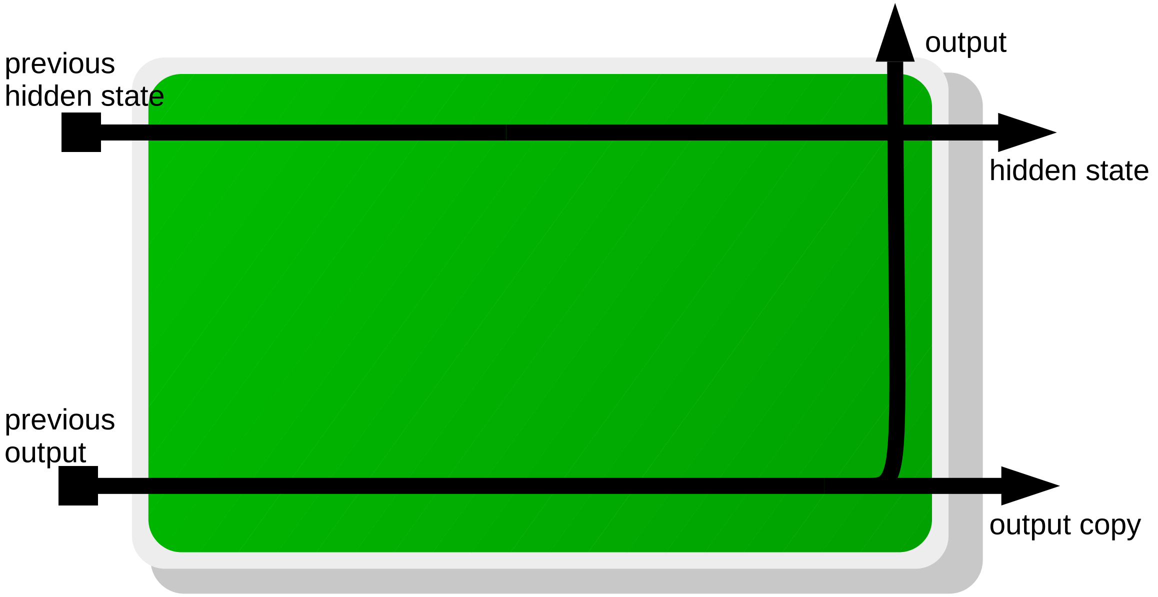
and applications

Understanding LSTM cell



```python
def LSTM_cell(weights, hidden_state, output, input):
```

Understanding LSTM cell



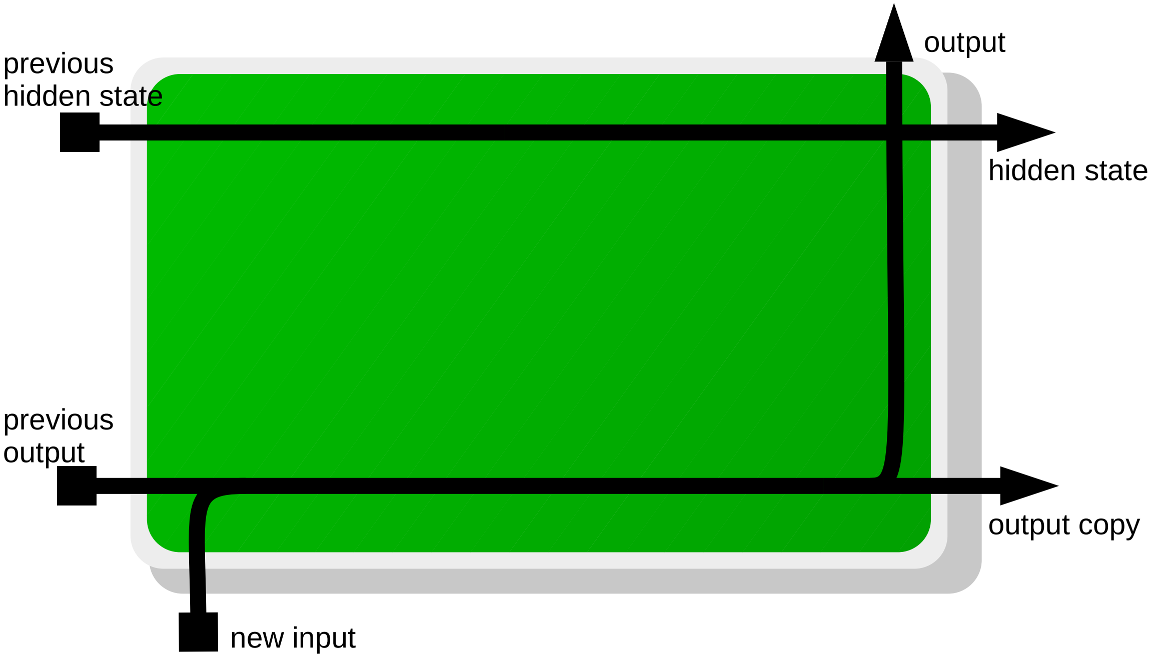
```
def LSTM_cell(weights, hidden_state, output, input):
    forget, update, candidate, output_gate = weights.split()






    return output, hidden_state
```

Understanding LSTM cell



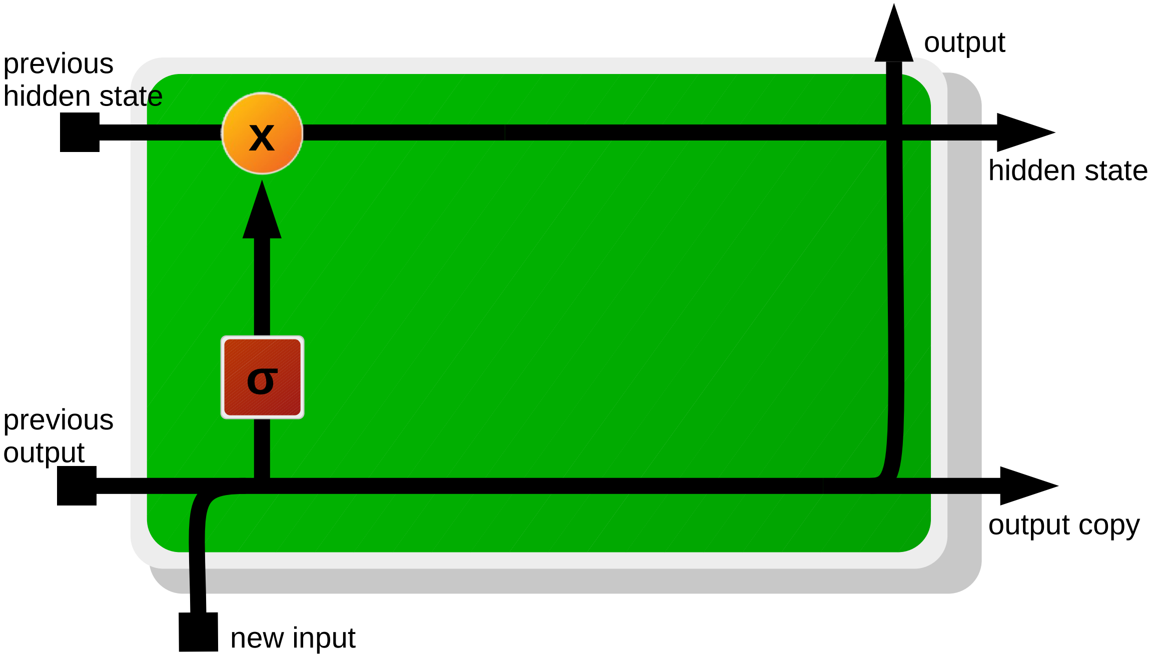
```
def LSTM_cell(weights, hidden_state, output, input):
    forget, update, candidate, output_gate = weights.split()
    in_out = concatenate(input, output)





    return output, hidden_state
```

Understanding LSTM cell



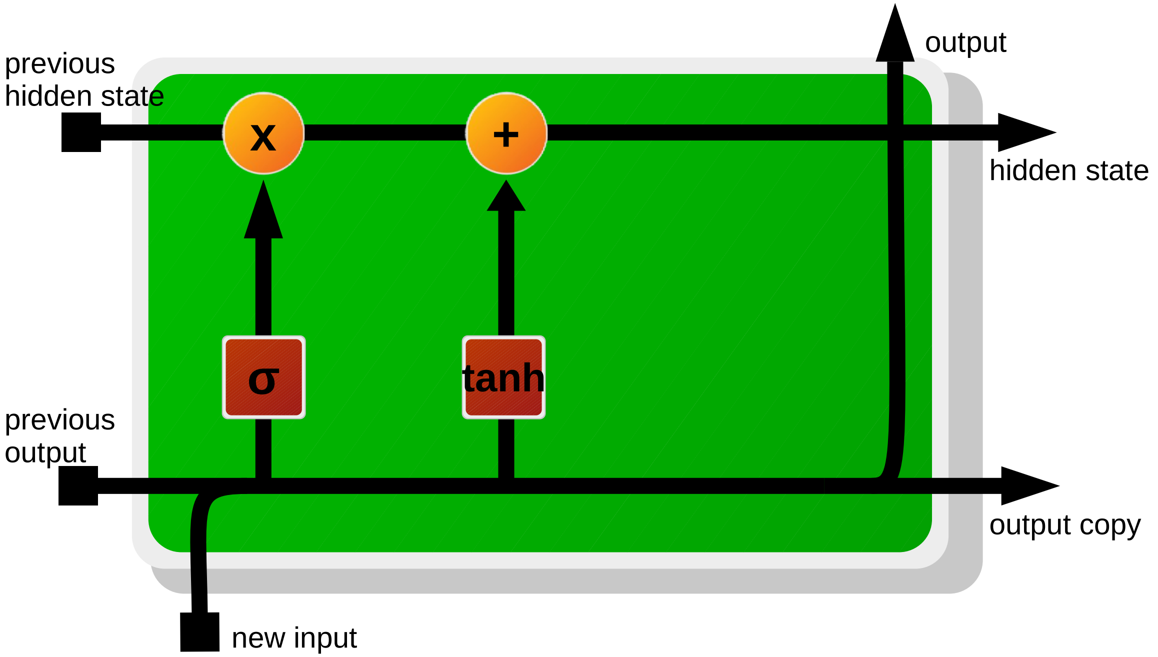
```python
def LSTM_cell(weights, hidden_state, output, input):
    forget, update, candidate, output_gate = weights.split()
    in_out = concatenate(input, output)

    hidden_state *= sigmoid(in_out @ forget)



    return output, hidden_state
```

Understanding LSTM cell

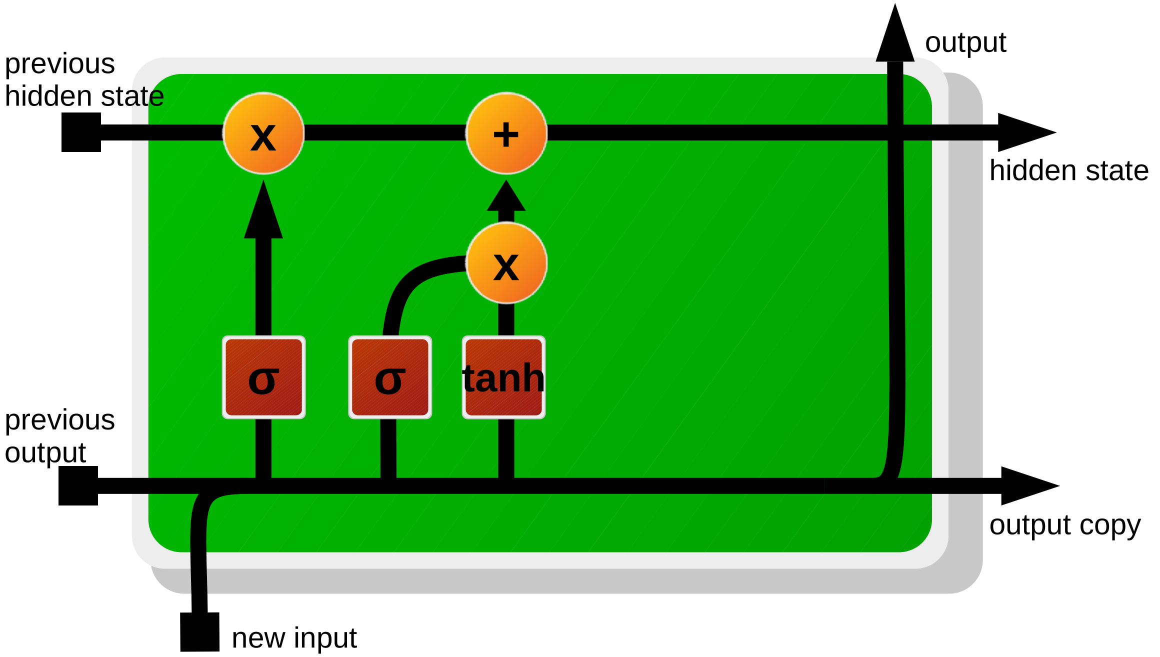


```
def LSTM_cell(weights, hidden_state, output, input):
    forget, update, candidate, output_gate = weights.split()
    in_out = concatenate(input, output)

    hidden_state *= sigmoid(in_out @ forget)
    hidden_state += tanh(in_out @ candidate)


    return output, hidden_state
```

Understanding LSTM cell

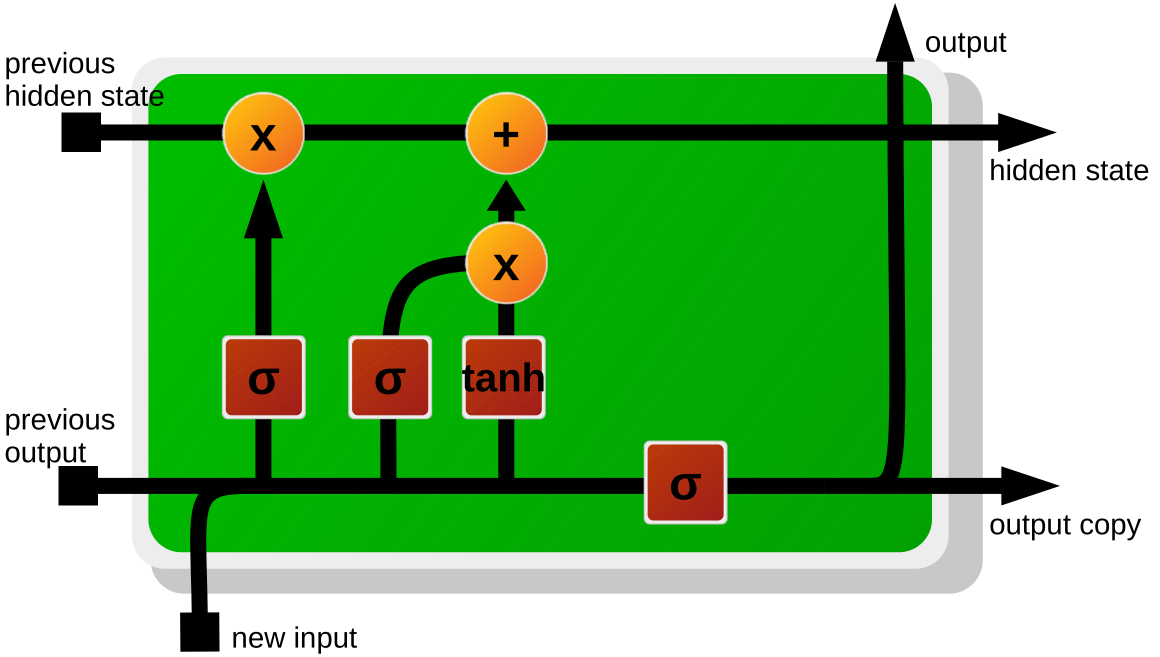


```
def LSTM_cell(weights, hidden_state, output, input):
    forget, update, candidate, output_gate = weights.split()
    in_out = concatenate(input, output)

    hidden_state *= sigmoid(in_out @ forget)
    hidden_state += tanh(in_out @ candidate) * sigmoid(in_out @ update)


    return output, hidden_state
```

Understanding LSTM cell

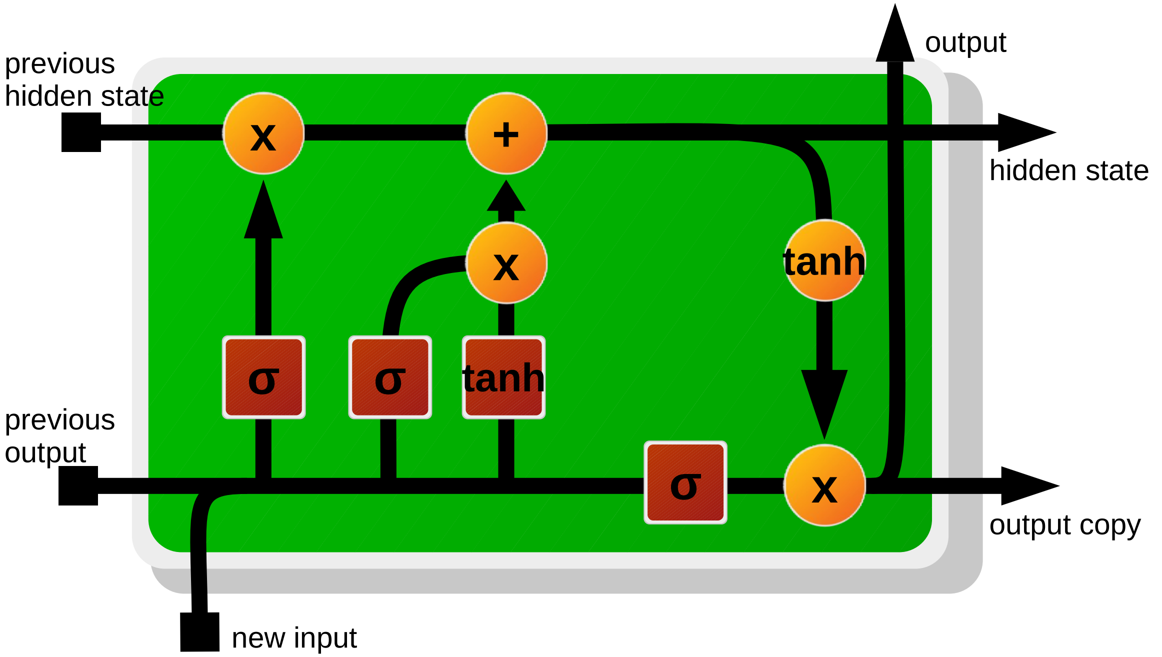


```
def LSTM_cell(weights, hidden_state, output, input):
    forget, update, candidate, output_gate = weights.split()
    in_out = concatenate(input, output)

    hidden_state *= sigmoid(in_out @ forget)
    hidden_state += tanh(in_out @ candidate) * sigmoid(in_out @ update)

    output = sigmoid(in_out @ output_gate)
    return output, hidden_state
```

Understanding LSTM cell

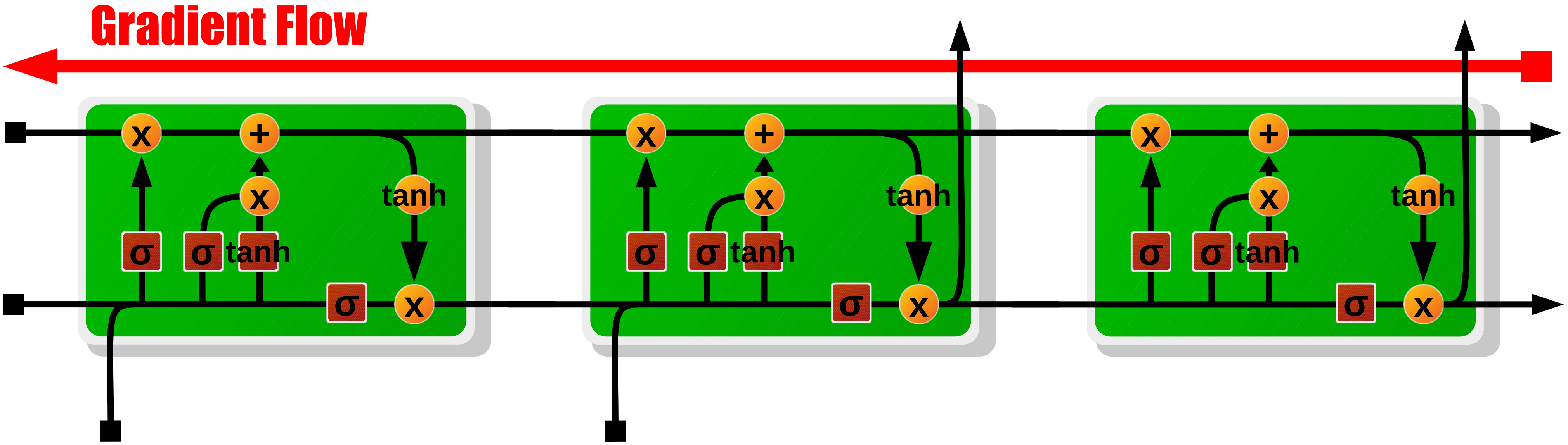


```
def LSTM_cell(weights, hidden_state, output, input):
     forget, update, candidate, output_gate = weights.split()
     in_out = concatenate(input, output)

     hidden_state *= sigmoid(in_out @ forget)
     hidden_state += tanh(in_out @ candidate) * sigmoid(in_out @ update)

     output = sigmoid(in_out @ output_gate) * tanh(hidden_state)
     return output, hidden_state
```

Backpropagation through time

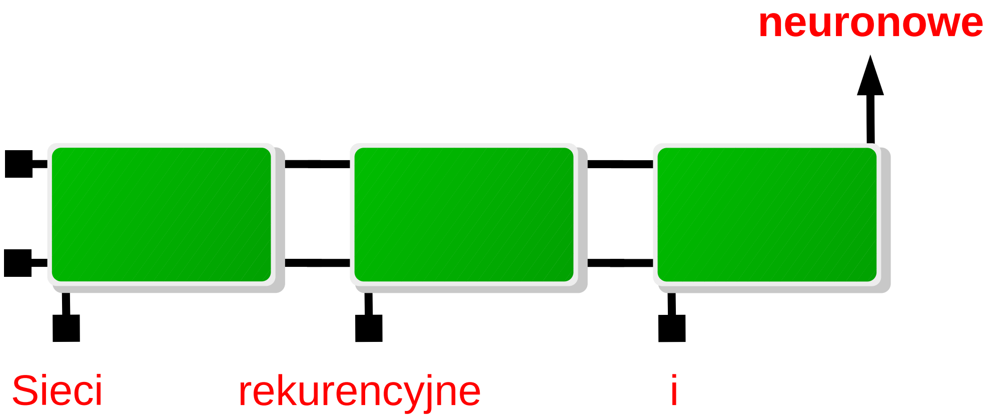


Predicting next word

A typical task for language understading

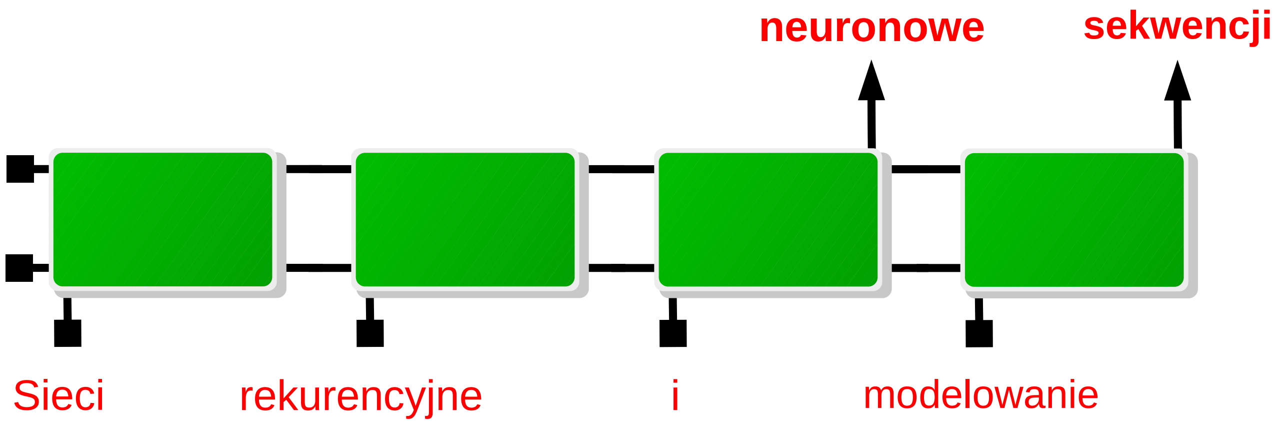
Predicting next word

A typical task for language understading



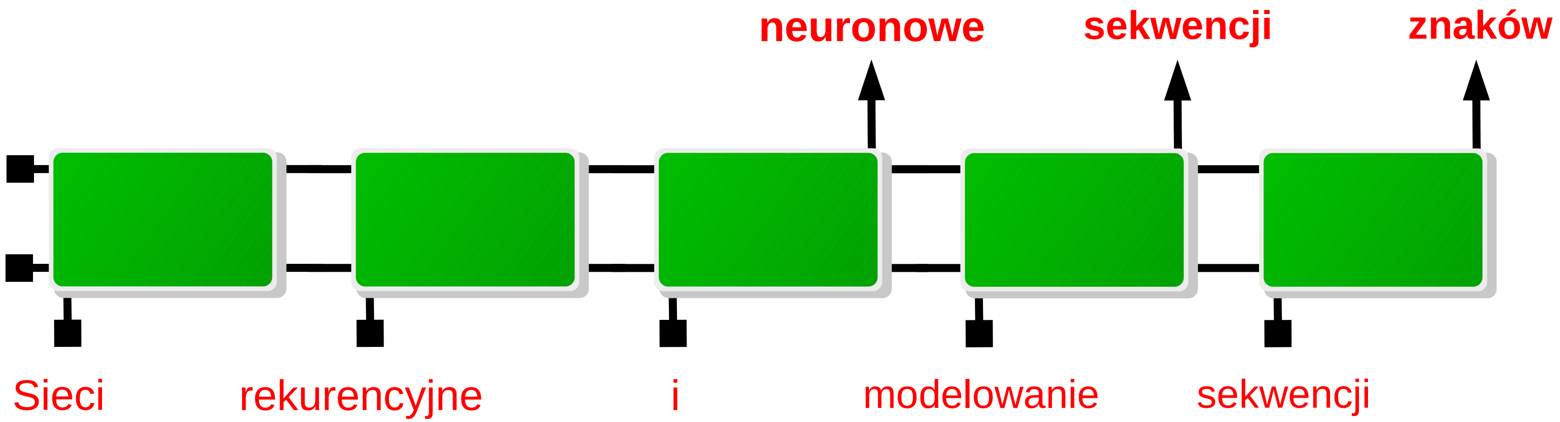
Predicting next word

A typical task for language understading



Predicting next word

A typical task for language understading

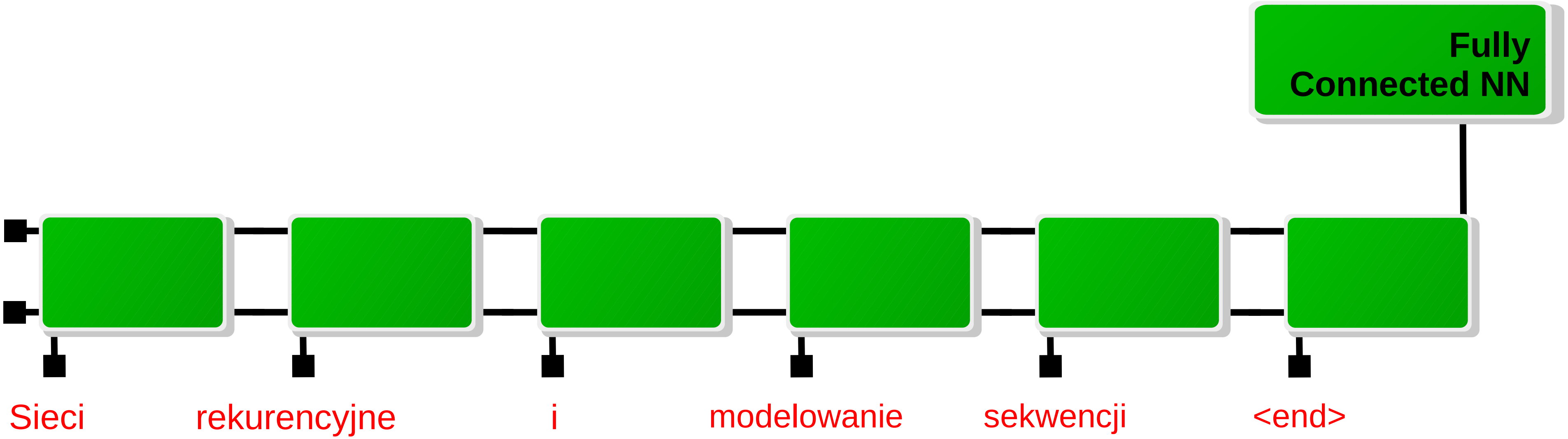


Classification

Another typical task

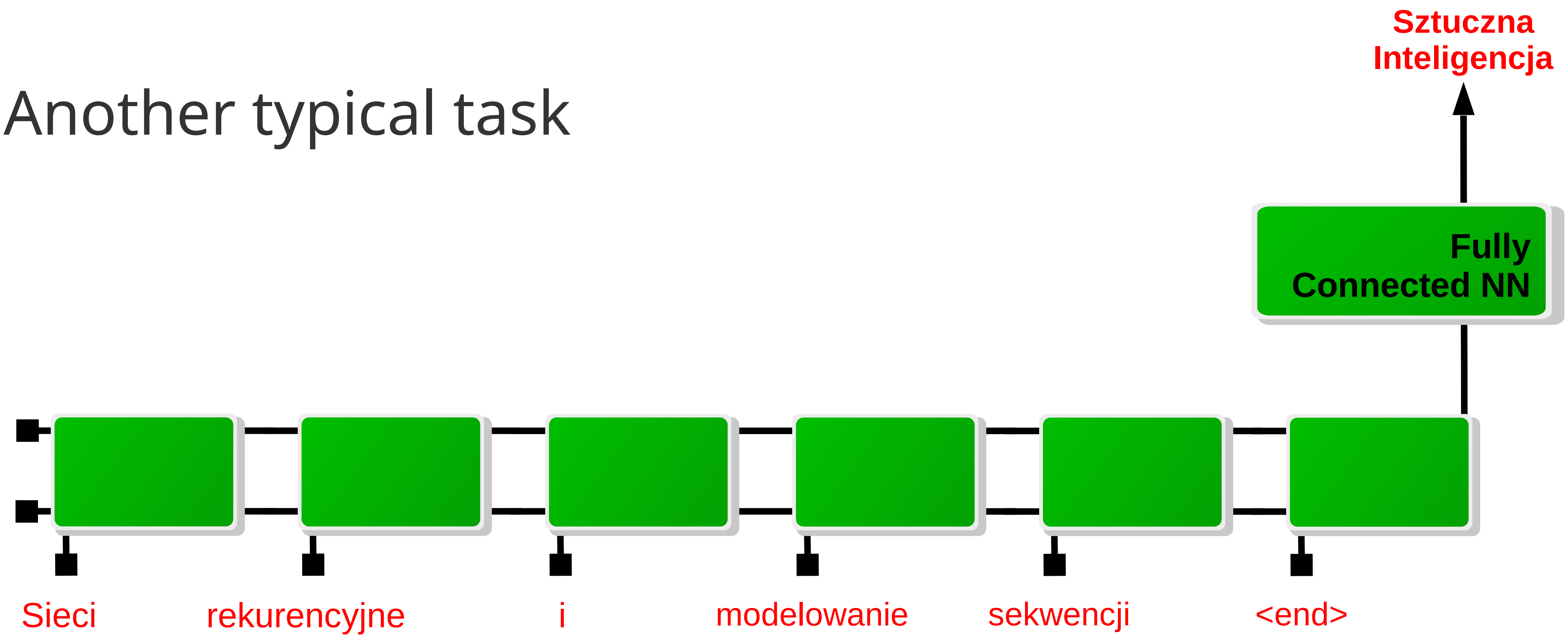
Classification

Another typical task



Classification

Another typical task



Machine Translation

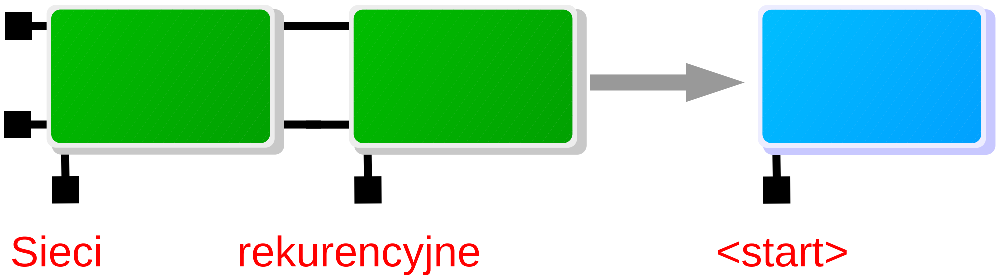
Yet another typical task

Two networks: encoder and decoder

Machine Translation

Yet another typical task

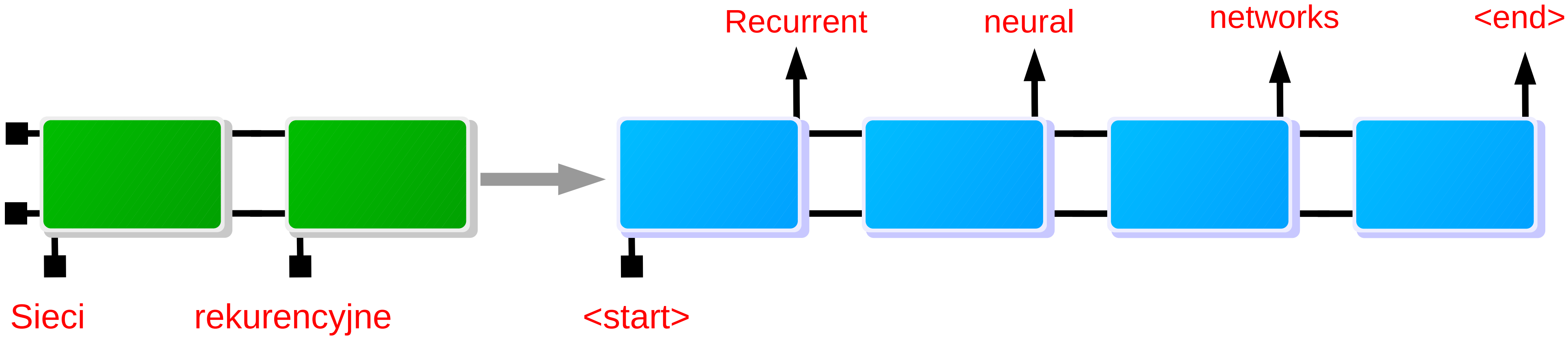
Two networks: encoder and decoder



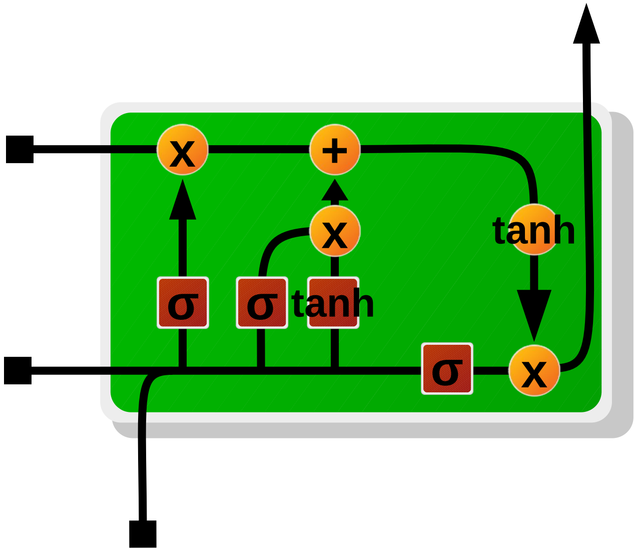
Machine Translation

Yet another typical task

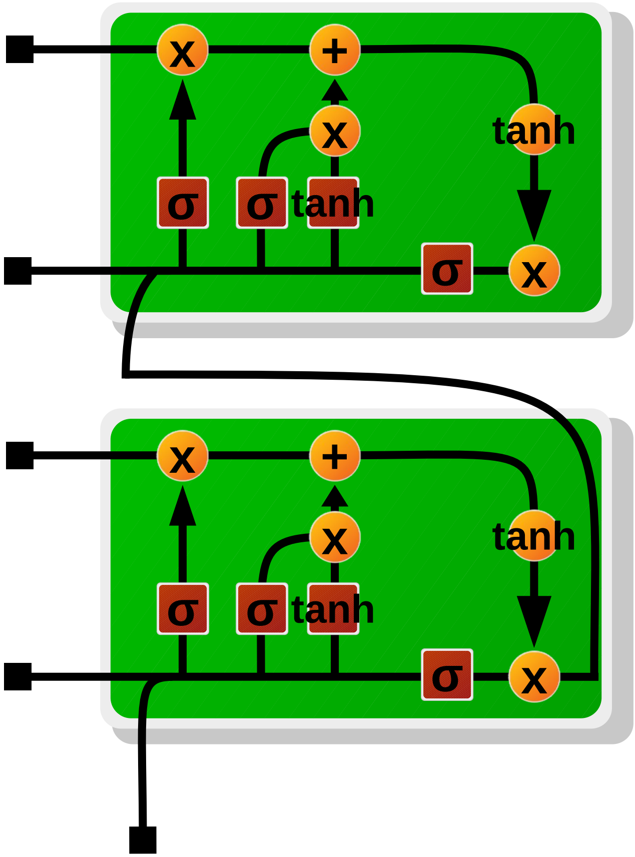
Two networks: encoder and decoder



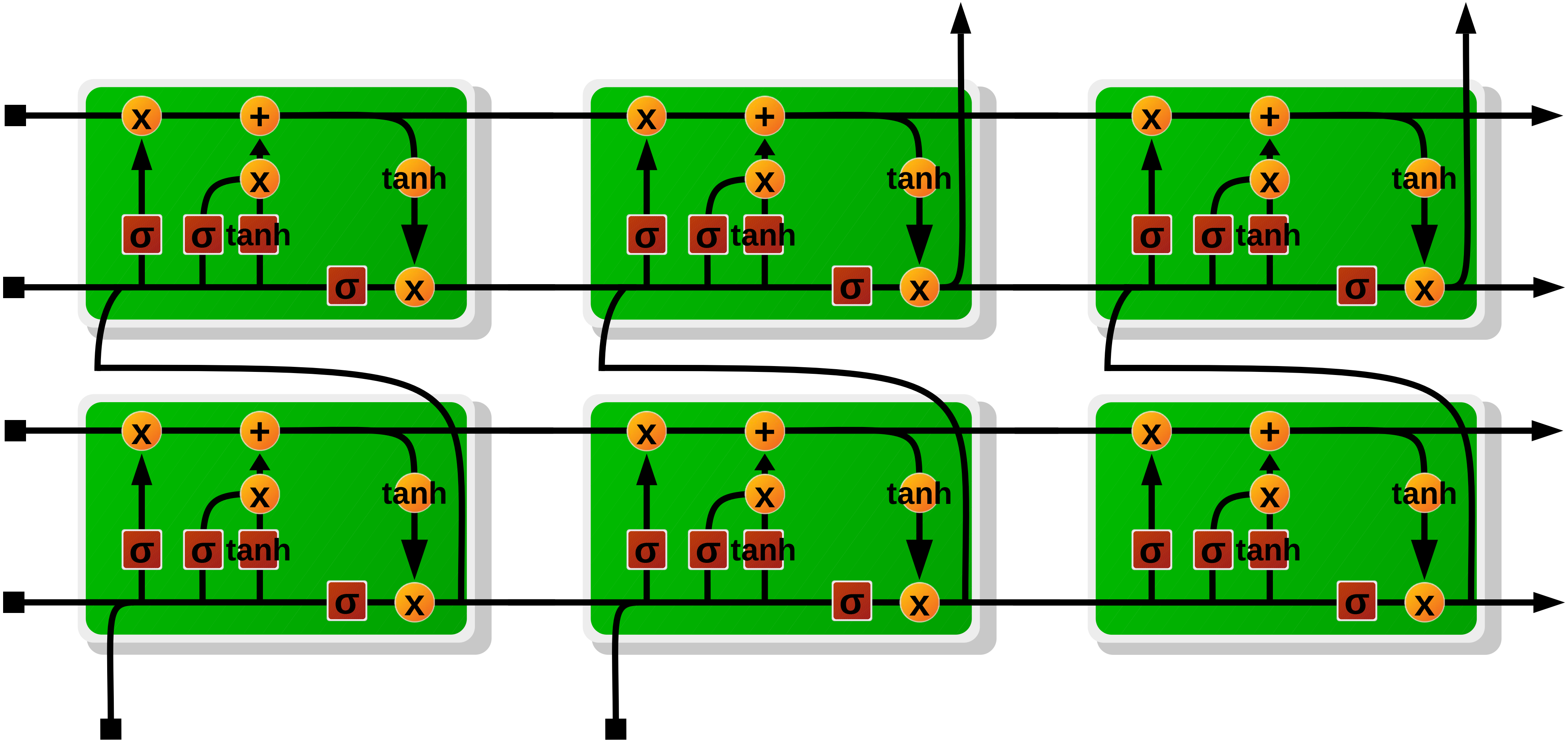
Multi-layer LSTM



Multi-layer LSTM

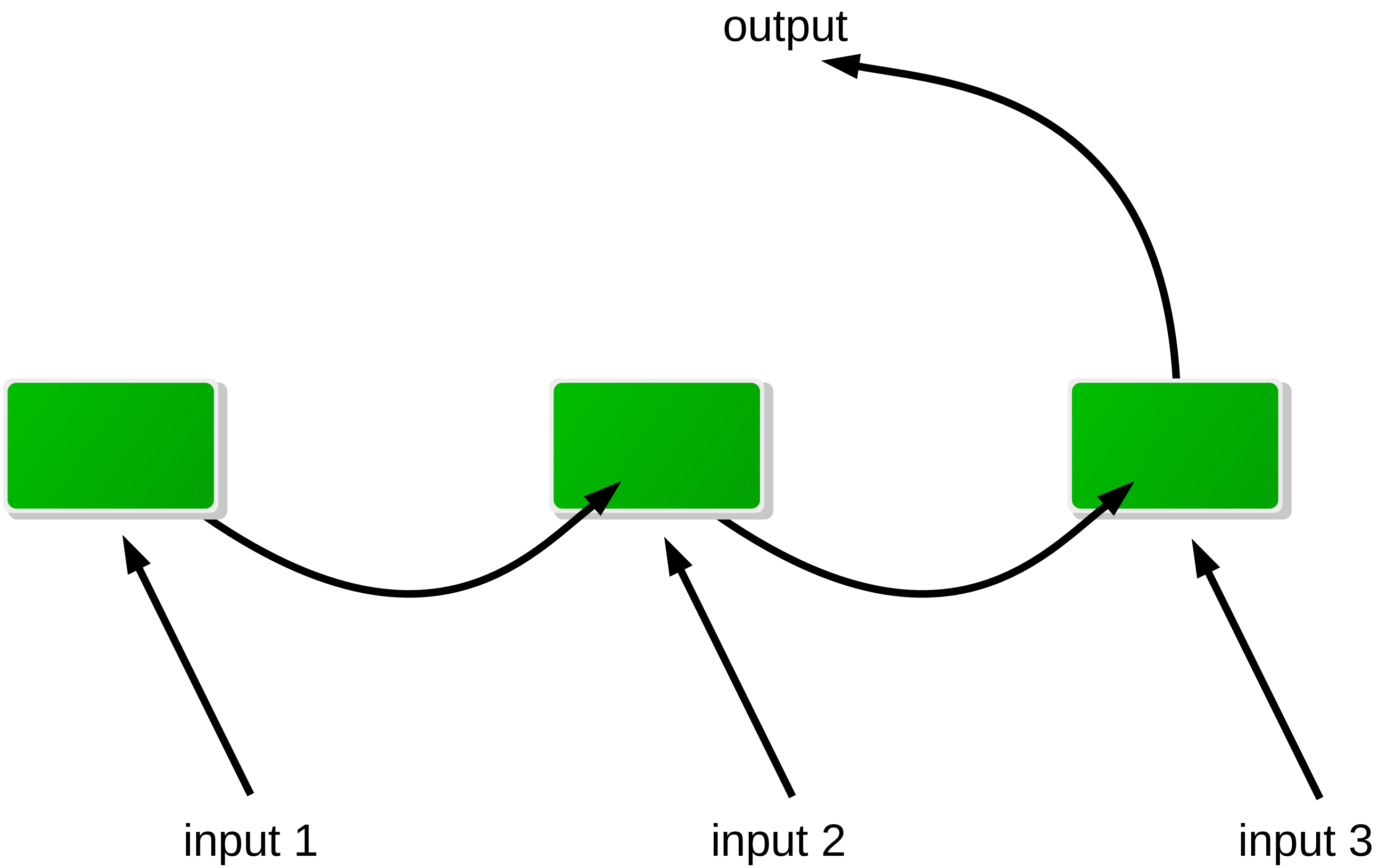


Multi-layer LSTM

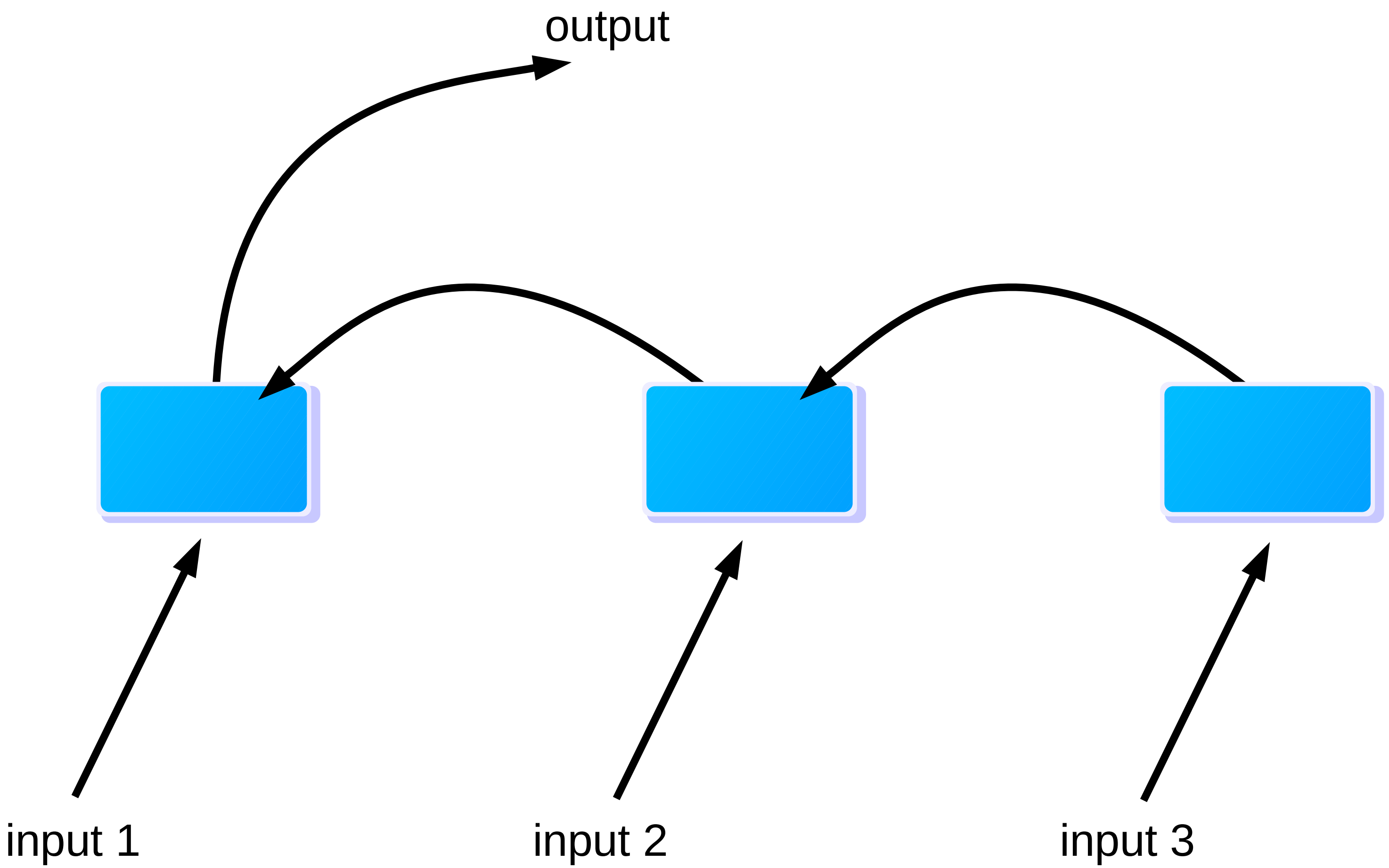


Bi-directional LSTM

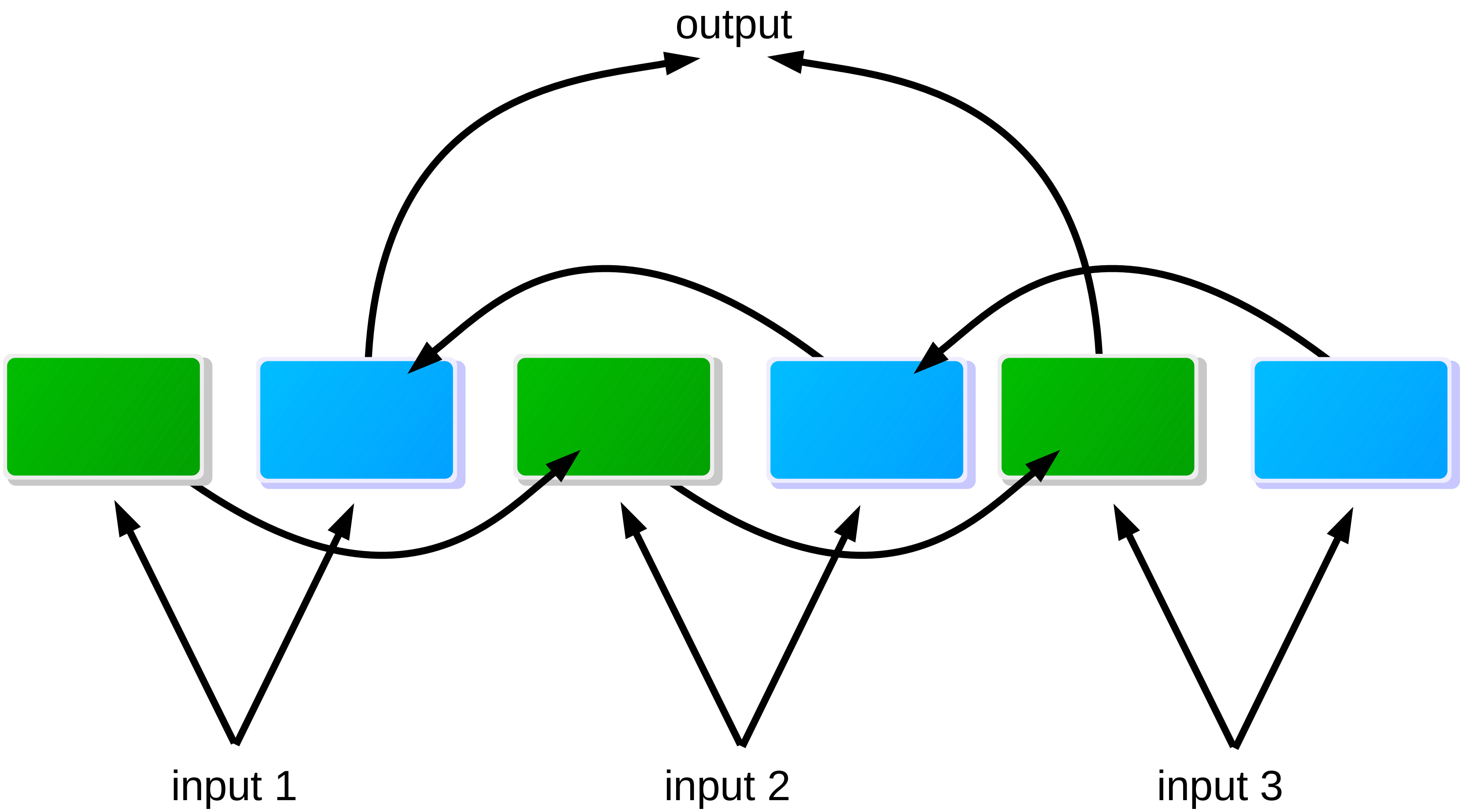
Bi-directional LSTM



Bi-directional LSTM



Bi-directional LSTM



Applications

Tasks involving sequence processing

Examples:

- Natural Language Tasks
- Reinforcement Learning
- Image Generation (Pixel RNN)

Language: Text Generation

Pomyliłaś niebo z gwiazdami odbitymi nocą ...

- 2 layer LSTM after reading polish fantasy

Language: Text Generation

Pomyliłaś niebo z gwiazdami odbitymi nocą i grupą przeszłością. A właśnie miało bladą chmur człowieka i siła woli i dodadko

- Jeszcze raz, że miała brudny i mocnyej i wszystko słyszy i zmężnie, to w porządku, było poruszać kiełbaś kretyna padała wokół sali. Pamiętaj: jak zaczyna marnować. Coś się stało

Byłem w przyczegowiez za wyśmieżym wszystkiego biedaka, jaki urządził pody tym nowe miasto pod człowiekiem, obiecał za rękę i rozbił go do środka. Magicznie ostrożnie. Doprowadził do tej różdżki wokół umysłu. Pozostali czarna tajemnicza rodziga jest wcale nie wpadłem, bo pierwszy z trzech dobrych dzieciaków odnał by Jacek, tak jak pan się przy maga, miała szarpać. Droga była zakładaca sytuacja

- 2 layer LSTM after reading polish fantasy

Language: Text Generation

PixelRNN is a neural network that allows to …

- GPT-2, OpenAI, explaining PixelRNN

Language: Text Generation

PixelRNN is a neural network that allows to training and learning (real data) efficiently. But before we go to the basic idea of the network, let us see some pictures to understand some things:
The hardest part is the layer with very different units and different weight. The idea is to apply gradient descent and that is one way to train and learn an deep neural network. But we have more than 2.5k+ of a representation for each bit to use. So we are faced with several problems.

- GPT-2, OpenAI, explaining PixelRNN

Language: Translation

We've got no choice.	Nie mamy wyboru.

- 30 mln parameters RNN with attention, validating after ~1h of training

Language: Translation

We've got no choice.	Nie mamy wyboru.
Why would you say that?	Dlaczego tak mówisz?
It’s too expensive.	Jest zbyt drogie.
I saw him run away.	Widziałem go jak uciekał.
They’ll come.	Przyszedł już.
Tom hasn’t talked to the manager yet.	Tomek jeszcze nie do końca Tom uciekł.

- 30 mln parameters RNN with attention, validating after ~1h of training

Language: Translation

We've got no choice.	Nie mamy wyboru.
Why would you say that?	Dlaczego tak mówisz?
It’s too expensive.	Jest zbyt drogie.
I saw him run away.	Widziałem go jak uciekał.
They’ll come.	Przyszedł już.
Tom hasn’t talked to the manager yet.	Tomek jeszcze nie do końca Tom uciekł.
He will return from Europe soon.	On niedługo wrócił do Bostonu.
It’s unlikely she’s as stupid as you think.	Chyba nie jest tak bezdradna jak sądzisz.
It hurts so much.	To bardzo źle.
You caused this.	Ty to spowodowałeś.
Tom is very old.	Tom jest bardzo stary.
He hasn’t read the boot yet.	Jeszcze nie napisał książkę.
Tell Tom that I want to go, too.	Powiedz Tomowi, że chcę iść, a to.

- 30 mln parameters RNN with attention, validating after ~1h of training

Reinforcement Learning

Partially Observable Markov Decision Process

> Standard **MDP** provides us with States, Actions and Rewards and those are observed by the agent. In **POMDP** everything is determined as well, but the agent cannot directly observe the state. Agent makes a belief about the environment based on the observations.

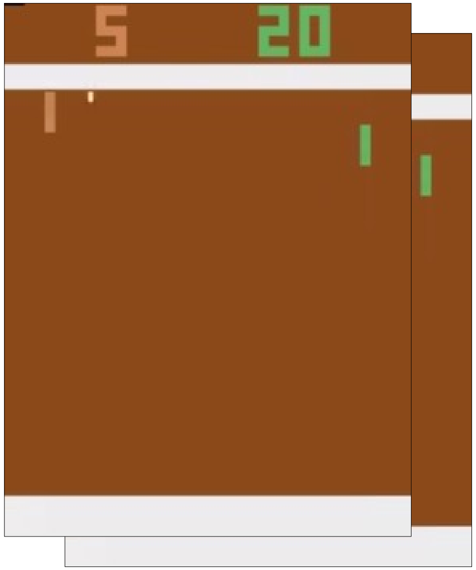
We have:

States, Actions, Transitions, Rewards – from standard MDP
Observations, Observation probabilities – from POMDP

Reinforcement Learning

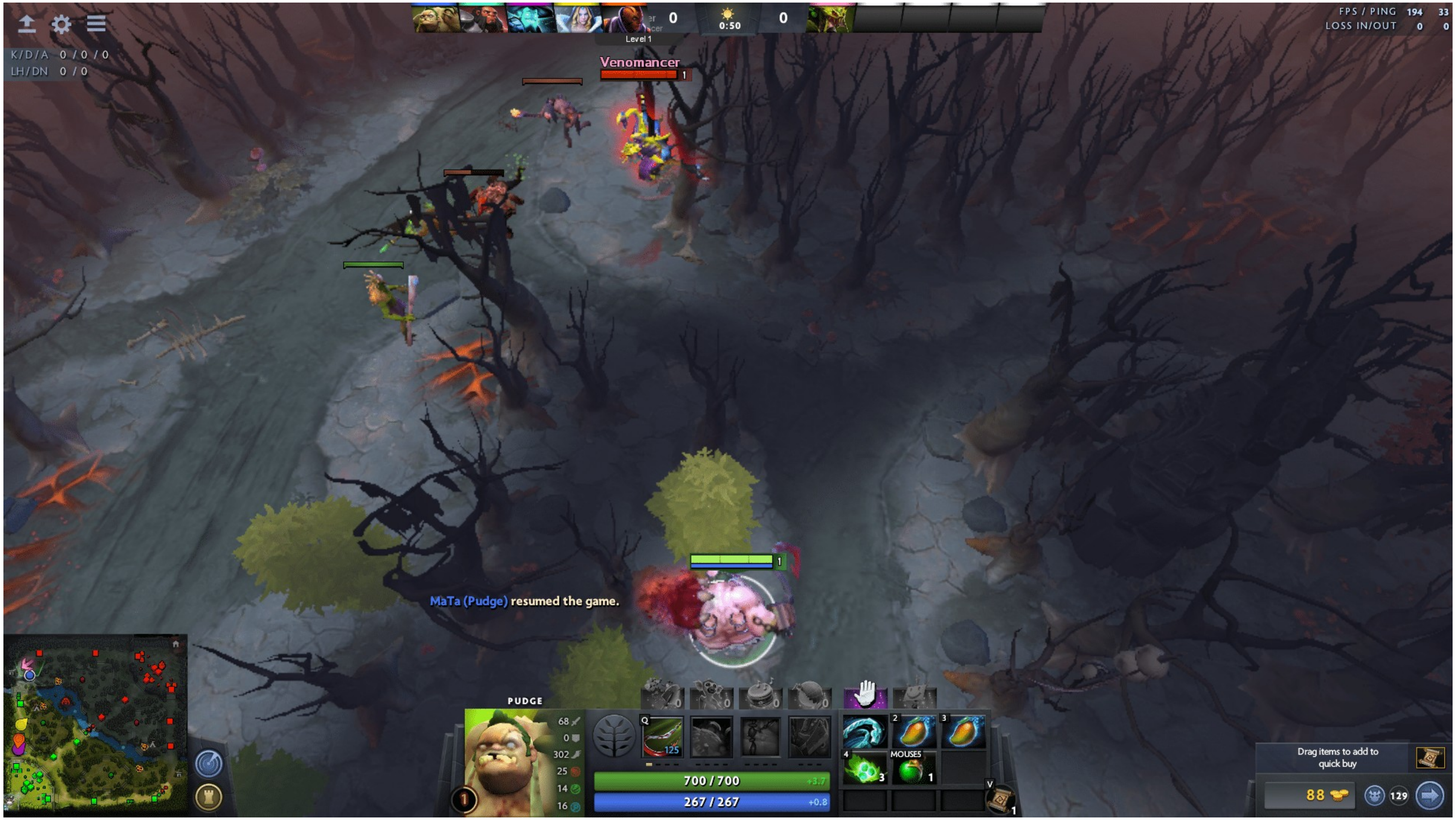
In Pong two consecutive frames provides us with all the informations

We don’t neet to remember anything.



Reinforcement Learning

In real life scenario, we hardly ever know everything from the observations.



MaTa (Pudge) resumed the game.

Reinforcement Learning

How to Dota 2 (three simple steps):

- Proximal Policy Optimization
- **One-layer LSTM** as policy estimator
- 45,000 years of self-play