APS360 Fundamentals of AI

Lisa Zhang

Lecture 11; Feb 14, 2019

Announcement

- Additional PyTorch tutorials
- Assignment 4 Corrections
- Project Teams
- Midterm (next Thursday, room RW117)

Agenda

- Recurrent Neural Network
- Sentiment analysis

Recurrent Neural Networks

Variable Length Input

- When working with text (and other inputs), we are interested in working with variable-sized inputs.

When have we seen variable-sized inputs before?

Variable Length Input

- When working with text (and other inputs), we are interested in working with variable-sized inputs.

When have we seen variable-sized inputs before?

Convolutional Neural Networks!

Long Dependency

Today has been the best day of my life . . . not.

Want:

An architecture that

- Can take in variable-sized **sequential** input
- Can remember things over time: has some sort of **memory** or **state**

Want:

An architecture that

- Can take in variable-sized **sequential** input
- Can remember things over time: has some sort of **memory** or **state**

Recurrent Neural Networks!

Recurrent Neural Networks (RNN)

- Make predictions based on a sequence (**input** is a sequence)
- Generate a sequence (**output** is a sequence)
- Or both! The setting where both input and output are sequences is called a **sequence-to-sequence prediction**
  - Example: Machine translation

RNN as a dynamical system

We can think of RNN as a dynamical system, that updates the hidden units based on the next input:

```
hidden = f(last_hidden, input)
output = g(hidden)
```

Where `f` and `g` are a multi-layer perceptron (neural networks).

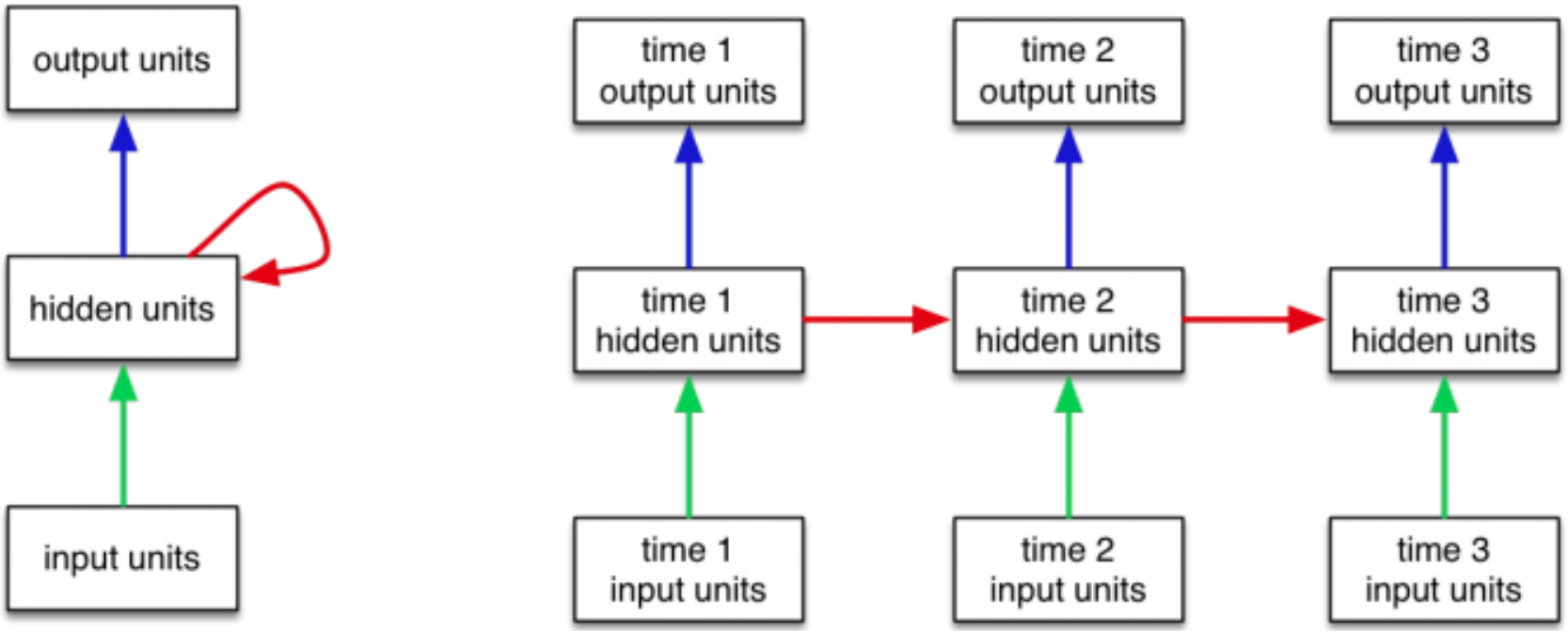
Reuse of `f` and `g`

The functions `f` and `g` are reused across all "time steps"

```
next_hidden = f(hidden, next_input)
next_output = g(next_hidden)
```

Where `f` and `g` are a multi-layer perceptron (neural networks).

Picture

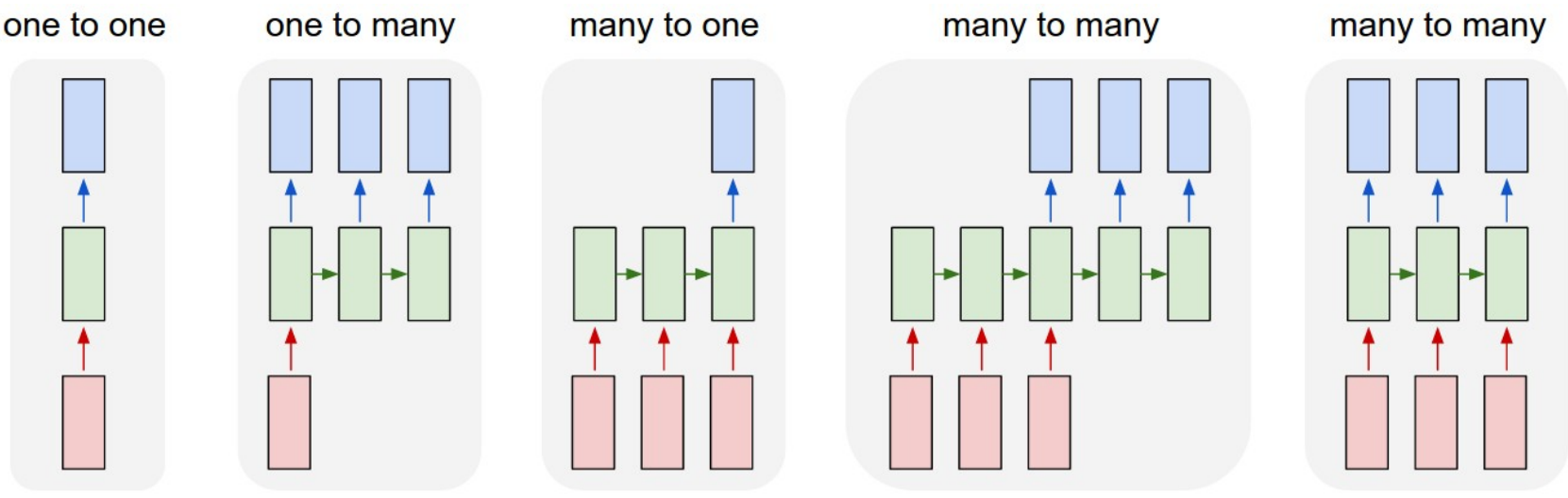

output units
hidden units
input units
time 1
output units
time 2
output units
time 3
output units
time 1
hidden units
time 2
hidden units
time 3
hidden units
time 1
input units
time 2
input units
time 3
input units

Variable Sequences

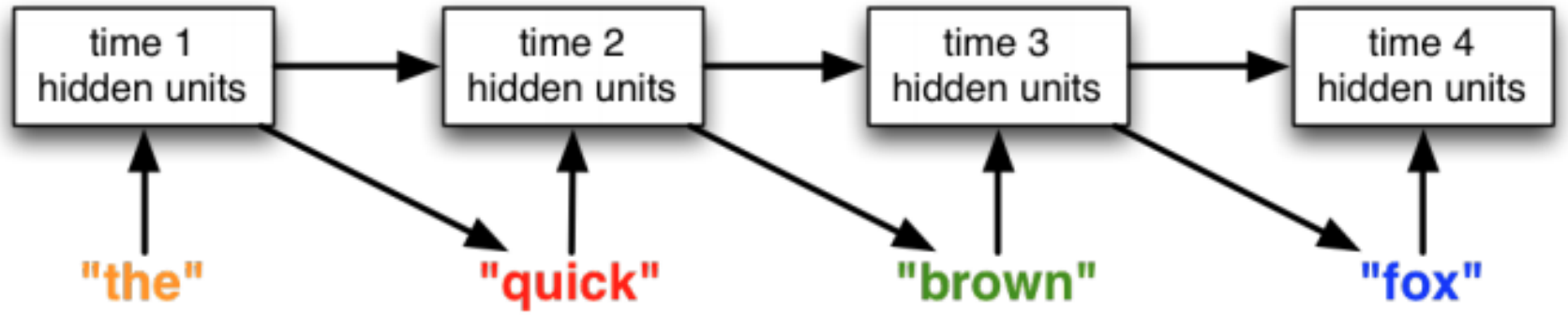
For inputs of various lengths. . .

- Apply `f` and `g` multiple times
- Apply `f` and `g` **different** number of times for **different** inputs

RNN Applications



RNN Generating Sequences



When we generate from the model, the **output** feed back into the network as **inputs**.

Computation Considerations

- Vocabularies very large, how do we encode the input?
- Harder to implement than MLP and CNN
  - e.g. harder to do things like batching

Let's move to PyTorch!