实作-2
Lab 2

明玉瑞 Yurui Ming

yrming@gmail.com

声明
Disclaimer

- 本讲义在准备过程中由于时间所限，所用材料来源并未规范标示引用来源。所引材料仅用于教学所用，作者无意侵犯原著者之知识版权，所引材料之知识版权均归原著者所有；若原著者介意之，请联系作者更正及删除。

  The time limit during the preparation of these slides incurs the situation that not all the sources of the used materials (texts or images) are properly referenced or clearly manifested. However, all materials in these slides are solely for teaching and the author is with no intention to infringe the copyright bestowed on the original authors or manufacturers. All credits go to corresponding IP holders. Please address the author for any concern for remedy including deletion.

背景
Background

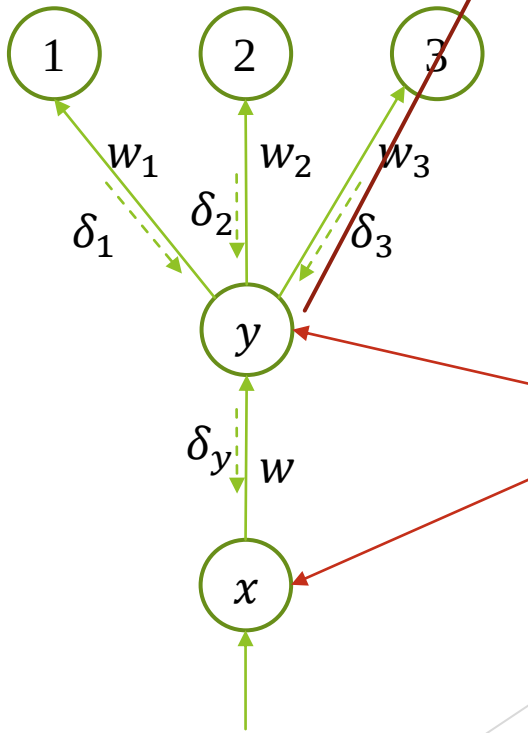
- 反向传播算法是逐层传播的，如果有中间值，计算起来没有技术难度，但会很繁琐：

  Back-propagation takes place in a layer-wise manner. If all intermediate values are maintained, there is no technical difficulties, however, it is time- and resource-consuming:

- 一般性规则（General rule）

  $\delta_y \triangleq (\delta_1 w_1 + \delta_2 w_2 + \delta_3 w_3) y'$

  $w(t+1) = w(t) - \eta \delta_y x$



$$y = \sigma\left(\sum wx\right)$$

σ的导数的解析表达均存在，即容易计算y'。
The analytical form of the derivative of σ always exists, which makes commutation of y' easier.

中间值
（Intermediate Values）

背景
Background

- 回顾机器学习四要素：数据，模型，损失函数，学习方法，上面的反向传播属于学习方法，同时，我们需要一个库来帮助我们管理其他三个要素。

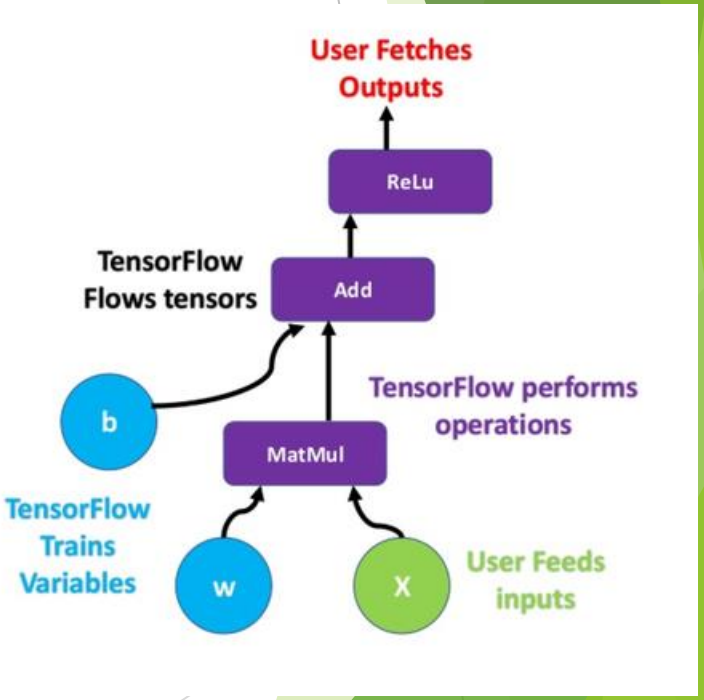
  The four elements constituting machine learning include data, model, loss function and learning method. Back-propagation belongs to learning method category; however, a library is still needed to help manage other three elements.

- 由谷歌开发与开源的TensorFlow便是这样一个开源库，其为研究者及工业界快速实现深度学习模型提供了有力支持。

  The library TensorFlow which is developed and open-sourced by Google offers scholars and industries the most powerful tool for conducting deep learning research.

TensorFlow

- TensorFlow从字面意思上讲，可以认为是张量（Tensor）与流（Flow）的加总。这里的张量与数学上的张量是不同的，可以认为是向量和矩阵的推广，即更高维度数组的一般化。TensorFlow在内部将张量表示为基本数据类型的n维数组。因为TensorFlow是将对Tensor的操作以单向图的形式表示的，流可以理解为数据流或操作流。
- Superficially speaking, TensorFlow can be treated as the combination of Tensor and Flow. The definition of Tensor here is quite different from the counterpart in mathematics. The tensor in DL context can be regarded as the extension of high dimensional arrays, which is the internal representation of Tensors in TensorFlow. Since TensorFlow arranges tensor in a manner of directed graph, flow is interpreted as data flow or operation flow.

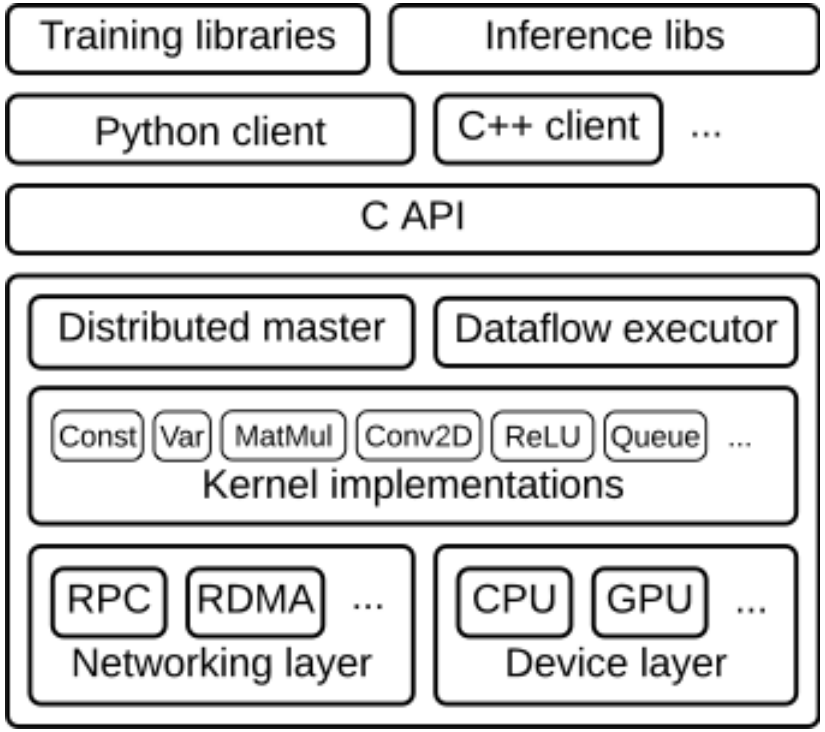


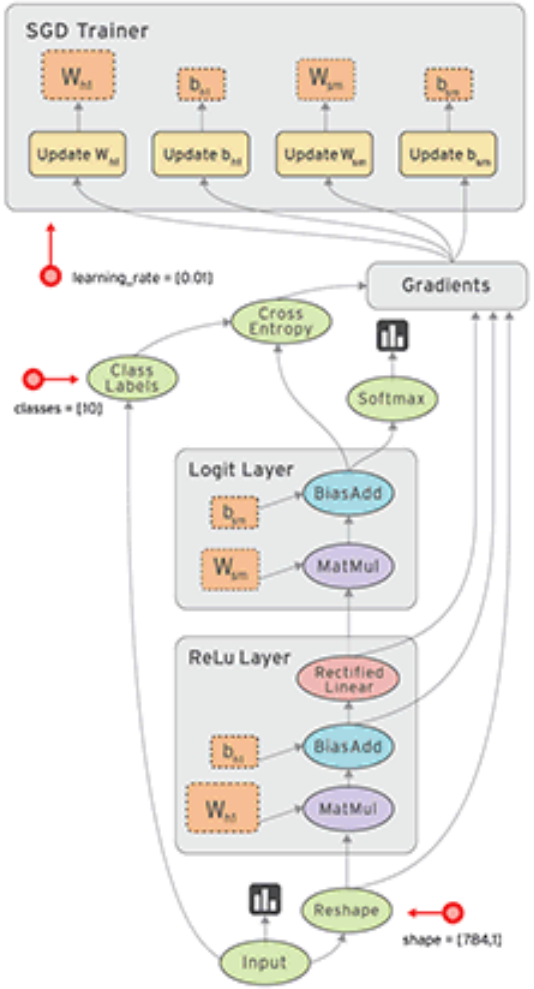
TensorFlow

- TensorFlow 1.x版本基于静态图，即先由Python语言构造好静态计算图，然后交由底层计算引擎执行，返回最后结果；2.x版本基于动态图，即计算图可以交互构建，部分地交由计算引擎执行， 返回中间结果，这样有利于开发者调试程序。1.x更偏重于库的形式，即开发者要利用TensorFlow提供的API构建深度学习的整个流程涉及的方方面面；2.x更偏重于框架，即开发者在API中提供适当的回调，即可以完成深度学习的基本流程。

  The 1.x version of TensorFlow is based on static computation graph, which is constructed by scripts language such as python, then delivered to the level engine for execution to have the final results return. Conversely, version 2.x is based on the dynamic computation graph, which means the graph can be partially constructed before send to the engine for computation. The intermediate result is also visible to the developer, which can ease the frustration when error occurs. Notably, 1.x is library-oriented, which means the developer has to utilize the provided API to construct each aspect of deep learning, while 2.x is more framework-oriented, which means the process can be automated on condition that proper callbacks are provided.

TensorFlow


Training libraries
Inference libs
Python client
C++ client
...
C API
Distributed master
Dataflow executor
Const
Var
MatMul
Conv2D
ReLU
Queue
...
Kernel implementations
RPC
RDMA
...
Networking layer
CPU
GPU
...
Device layer


SGD Trainer
Update W
Update b
Update W
Update b
learning_rate = [0.01]
Gradients
Cross Entropy
Class Labels
classes = [10]
Softmax
Logit Layer
BiasAdd
MatMul
ReLu Layer
Rectified Linear
BiasAdd
MatMul
Reshape
shape = [784,1]
Input

TensorFlow

- 由于TensorFlow构建计算图由前端完成，而依据计算图进行计算由后端完成，因此需要有相关机制来请求相关资源进行相关计算，TensorFlow中定义了Session这一概念来完成对应功能。我们可以将计算机科学中进程的概念与TensorFlow中Session的概念进行类比，两者均有在特定的上下文中，进行资源管理与计算执行等动能。

  To cater to the case that the graph construction is performed by the front-end whilst computation based on the graph is performed by the backend, it desires a mechanism to allocate the resources and manage the computations. To fulfill these demands, TensorFlow defines session, which can be compared with process in computer science. Both of them manages the resources and execute the computations in specific context.

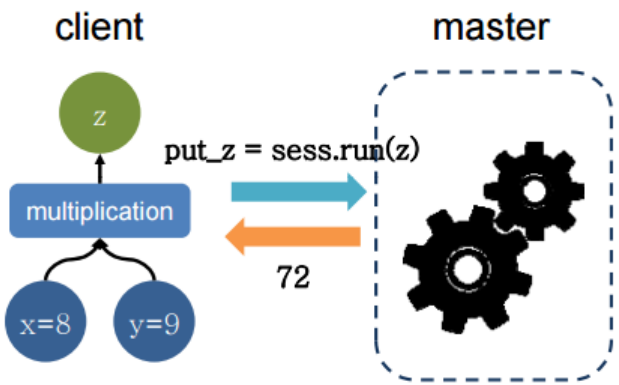
TensorFlow

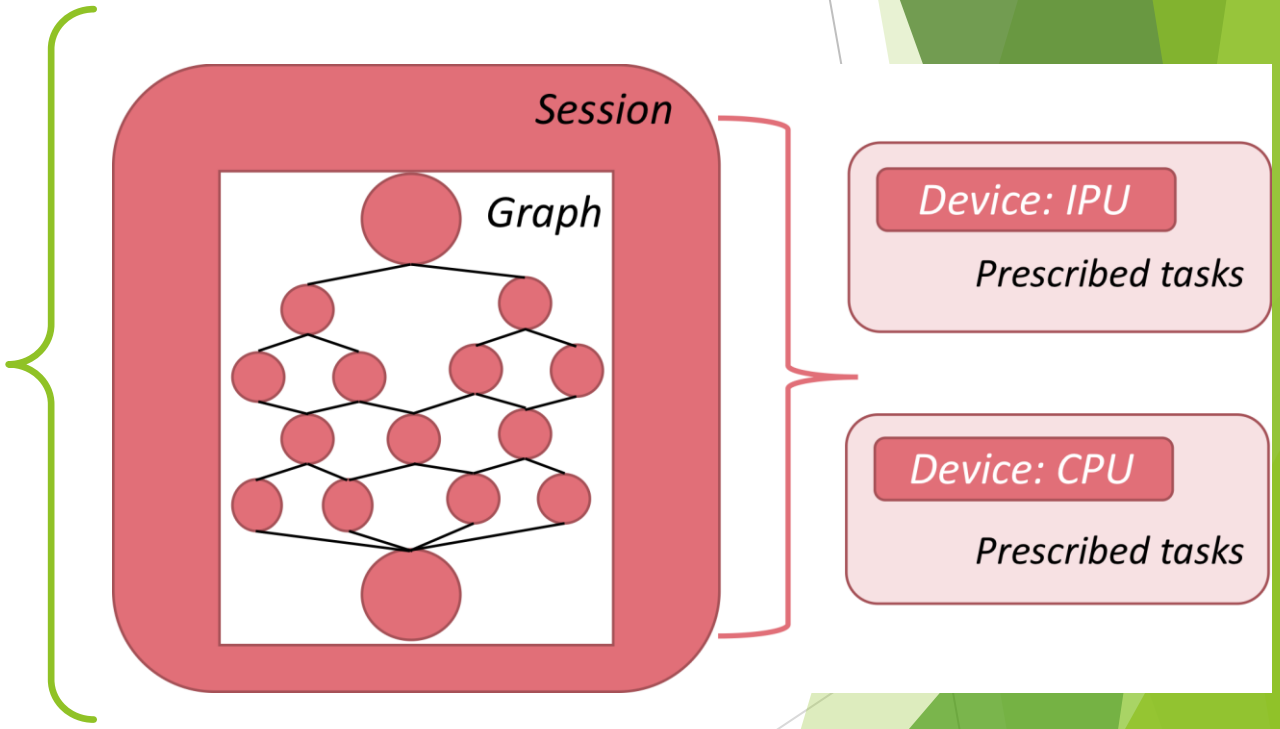
source code

```
import tensorflow as tf

x = tf.constant(8)
y = tf.constant(9)
z = tf.mul(x,y)

sess = tf.Session()

out_z = sess.run(z)

print('out_z: %d' % out_z)
```

output

```
out_z: 72
```





线性回归
Linear Regression

- 回归最简单的定义是，给出一个点集D，用一个函数去拟合这个点集，并且使得点集与拟合函数间的误差最小，如果这个函数曲线是一条直线，那就被称为线性回归。什么是线性，线性在一维时指的是一条直线，在二维时是一个平面的，也可以是三维或者更高的维度，只是较难想象而已。当样本特征列为1或者说点集中的点的坐标只有一个维度或者只有一个变元时，称为一元回归，多于一个维度时，称为多元回归。本实作仅考虑一个变元的情况，即$y = wx + b$。

  The simplest implication for regression is to find a function to match the points in a given point set D with the criterion that the error between the points and corresponding locations in the equation is minimal among candidate solutions. If the function is expressed as a linear function, it is called linear regression. Linearity means it is a line in one dimension, while it occurs in a plane in two dimension. It can be in higher dimensions which incurs the difficulty in imagination. When the number of feature columns, or coordinates of the points, or variables, is just one, it is called one-variable regression, otherwise it is called multivariable regression. Here we just consider the one-variable case, that's $y = wx + b$.

习题
Problems

- 请按照实作-2实验文档，完成各步骤操作并观察结果

  Follow the steps in the document of lab-2, finish each procedures and check the corresponding outputs.