Guidebook to Exercise 3: *Computational Graphs*

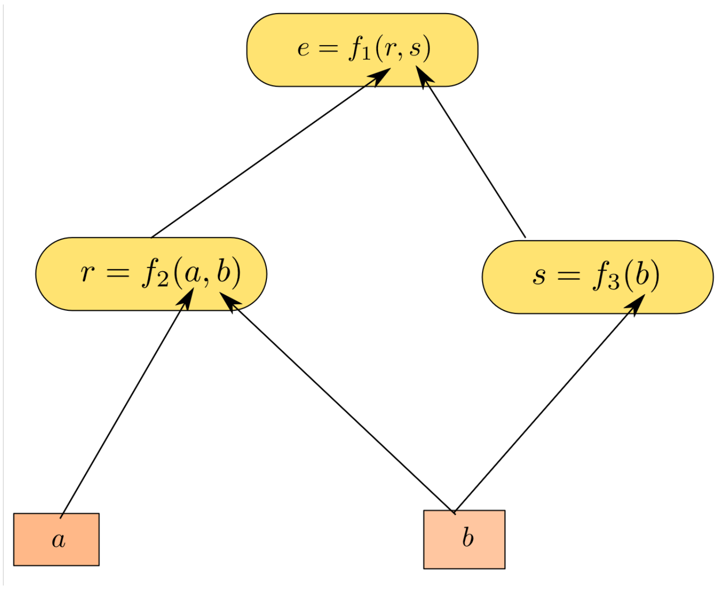
Prof. Dr. Ivan Habernal
Yassine Thlija

2025-10-30

1 Introduction

A *computational graph* is a directed acyclic graph (DAG) where nodes represent elementary operations or values (constants, parameters) and edges represent the flow of scalar (or vector) values between operations. Computational graphs are the foundation of automatic differentiation and backpropagation in neural networks.

Example from Lecture 3:



2 Some definitions

2.1 Nodes and values

We will work with scalar computational graphs. A *node* n computes a scalar value v_n. There are two elementary node types:

- **Leaf node (constant node)**: represents a numeric constant or input variable, e.g. a, b, x. Its value is fixed for a particular forward pass.
- **Operation node**: computes a function of its argument nodes. Examples include sums, products, differences and compositions (In the exercise we will only be using Sum and Product Nodes).

Edges are directed from argument nodes toward the operation node. The root node is the final scalar we are interested in.

2.2 Forward evaluation

Given values for leaf nodes, we compute values for internal nodes by evaluating their defining expressions in topological order.

Example: Let $s = a + b$, with $a = 1$, $b = 2$. Then the forward evaluation gives $s = 3$.

3 Local partial derivatives

For an operation node p that takes arguments $x_1, \ldots, x_k$ and computes $p = f(x_1, \ldots, x_k)$, we define the *local partial derivatives*

$$\frac{\partial p}{\partial x_i} = \frac{\partial f(x_1, \ldots, x_k)}{\partial x_i}.$$

These are *local* because they describe how p changes when only a single argument x_i changes, with the other arguments held fixed.

Examples

1. **Sum node:** $s = x_1 + \cdots + x_k$ gives

$$\frac{\partial s}{\partial x_i} = 1 \quad \forall i.$$

2. **Product node:** $p = \prod_{i=1}^{k} x_i$ gives

$$\frac{\partial p}{\partial x_i} = \prod_{j \neq i} x_j.$$

Example: For $p = a \cdot b \cdot c$, we have:

$$\frac{\partial p}{\partial a} = b \cdot c, \quad \frac{\partial p}{\partial b} = a \cdot c, \quad \frac{\partial p}{\partial c} = a \cdot b.$$

4 Global derivatives and the generalized chain rule

Suppose the final scalar of interest (the root) is y and a node in the graph is u with value v_u. We want the derivative of y with respect to the value at u, denoted

$$\frac{dy}{du}.$$

Because the graph is a DAG, y may depend on u through many distinct paths. The **generalized chain rule** (also used in backpropagation) states that the global derivative of y with respect to u is the sum over all immediate parents p of u of:

$$\frac{dy}{du} = \sum_{p \in \mathrm{Parents}(u)} \frac{\partial p}{\partial u} \cdot \frac{dy}{dp}.$$

If u is the root node, then $\frac{dy}{du} = 1$.

5 Some Examples

5.1 Example 1: A Simple Graph

Let

$$a = 2, \quad b = 3, \quad c = a + b, \quad y = c^2.$$

Forward pass:

$$c = 5, \quad y = 25.$$

Backward pass (chain rule demonstration):

$$\frac{dy}{dc} = 2c = 10, \quad \frac{dc}{da} = 1, \quad \frac{dc}{db} = 1.$$

Thus:

$$\frac{dy}{da} = \frac{dy}{dc} \cdot \frac{dc}{da} = 10, \quad \frac{dy}{db} = \frac{dy}{dc} \cdot \frac{dc}{db} = 10.$$

5.2 Example 2: Multiplicative Graph

Let

$$x = 1, \quad y = 2, \quad z = x \cdot y, \quad r = z + y.$$

Forward:

$$z = 2, \quad r = 4.$$

Backward:

$$\frac{dr}{dz} = 1, \quad \frac{dr}{dy} = 1.$$

Also,

$$\frac{dz}{dx} = y = 2, \quad \frac{dz}{dy} = x = 1.$$

Then:

$$\frac{dr}{dx} = \frac{dr}{dz}\frac{dz}{dx} = 2, \quad \frac{dr}{dy} = \frac{dr}{dz}\frac{dz}{dy} + 1 = 1 + 1 = 2.$$

6 If you want to dive deeper

Try building computational graphs for the following:

- $y = (a + b)(b + c)$
- $y = (x_1 + x_2 + x_3)^2$
- $y = x \cdot (x + 2)$

For each, perform both forward and backward passes and verify your gradients symbolically.