

自動微分

柏木 雅英

例題

関数

$$f = (x - y)(x^2 + y) + xy$$

計算順序

$$t_1 = x$$

$$t_2 = y$$

$$t_3 = t_1 - t_2 \quad (x - y)$$

$$t_4 = t_1^2 \quad (x^2)$$

$$t_5 = t_4 + t_2 \quad (x^2 + y)$$

$$t_6 = t_3 * t_5 \quad ((x - y)(x^2 + y))$$

$$t_7 = t_1 * t_2 \quad (xy)$$

$$t_8 = t_6 + t_7 \quad (f)$$

目的

与えられた x, y に対して $\frac{\partial f}{\partial x}$ と $\frac{\partial f}{\partial y}$ を計算すること。(計算例では $x = 2, y = 1$ 。)

計算ルール (bottom up)

$i = 3$ から 8 まで、

二項演算

$t_i = g_i(t_{p_i}, t_{q_i})$ という行に対して、

$$v_i = g_i(v_{p_i}, v_{q_i})$$

$$d_i = \frac{\partial g_i}{\partial t_{p_i}}(v_{p_i}, v_{q_i})d_{p_i} + \frac{\partial g_i}{\partial t_{q_i}}(v_{p_i}, v_{q_i})d_{q_i}$$

単項演算

$t_i = h_i(t_{p_i})$ という行に対して、

$$v_i = h_i(v_{p_i})$$

$$d_i = \frac{dh_i}{dt_{p_i}}(v_{p_i})d_{p_i}$$

計算ルール (top down)

まず、 $i = 3$ から 8 まで v_i と微分値を bottom up 法と同様に計算する。
次に、逆順に $i = 8$ から 3 まで、

二項演算

$t_i = g_i(t_{p_i}, t_{q_i})$ という行に対して、

$$d_{p_i} + = \frac{\partial g_i}{\partial t_{p_i}}(v_{p_i}, v_{q_i}) d_i$$

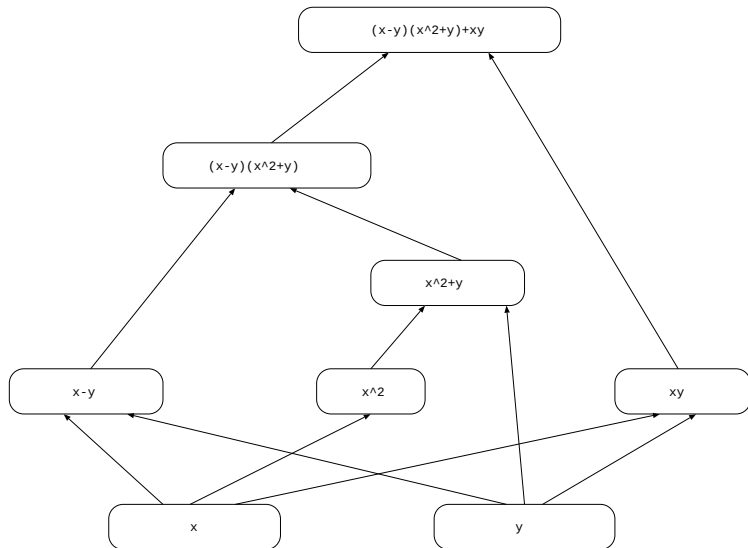
$$d_{q_i} + = \frac{\partial g_i}{\partial t_{q_i}}(v_{p_i}, v_{q_i}) d_i$$

単項演算

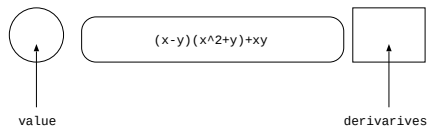
$t_i = h_i(t_{p_i})$ という行に対して、

$$d_{p_i} + = \frac{dh_i}{dt_{p_i}}(v_{p_i}) d_i$$

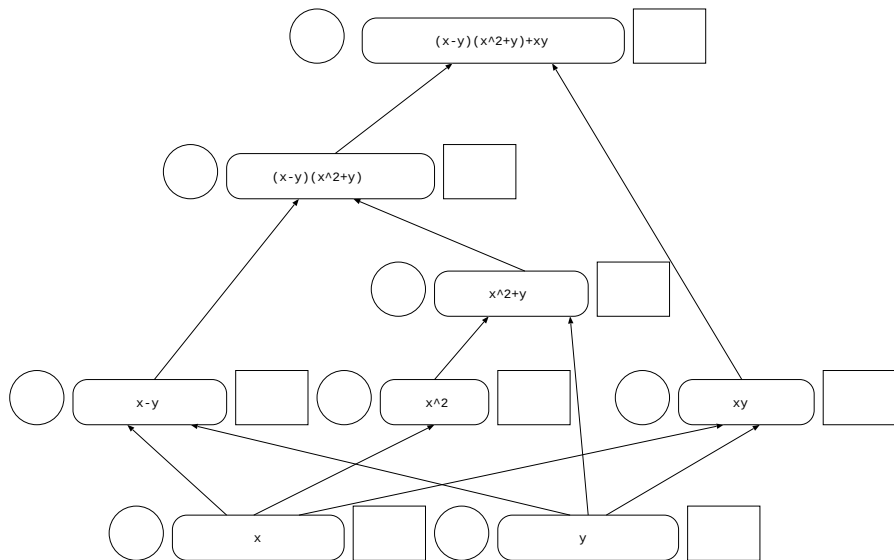
計算グラフ



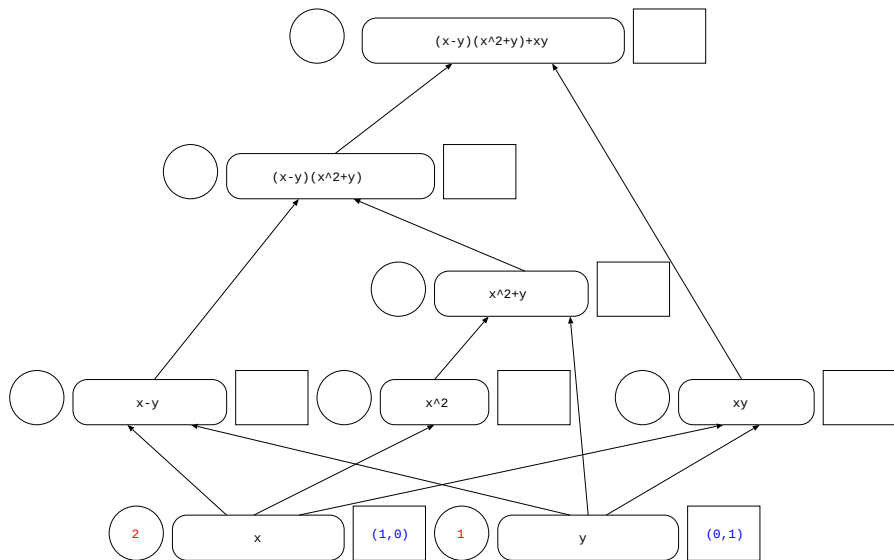
作業変数



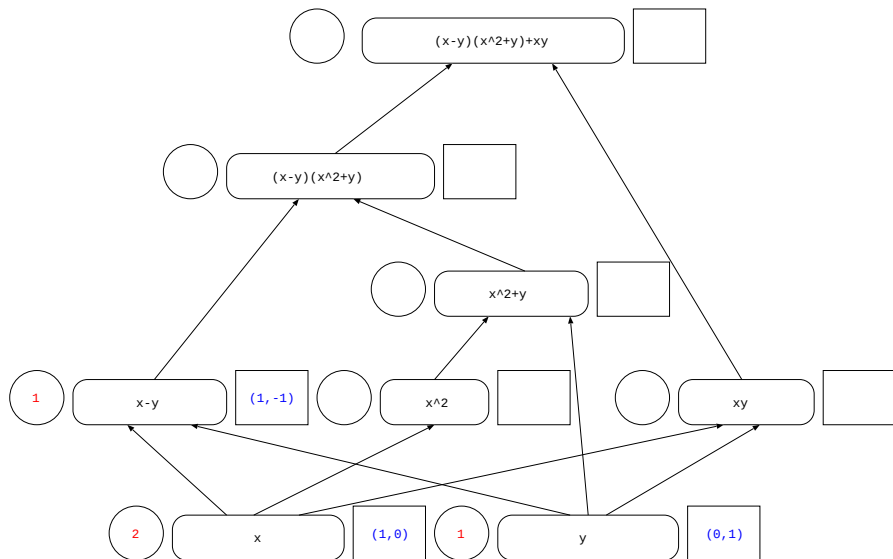
bottom up アルゴリズム (0)



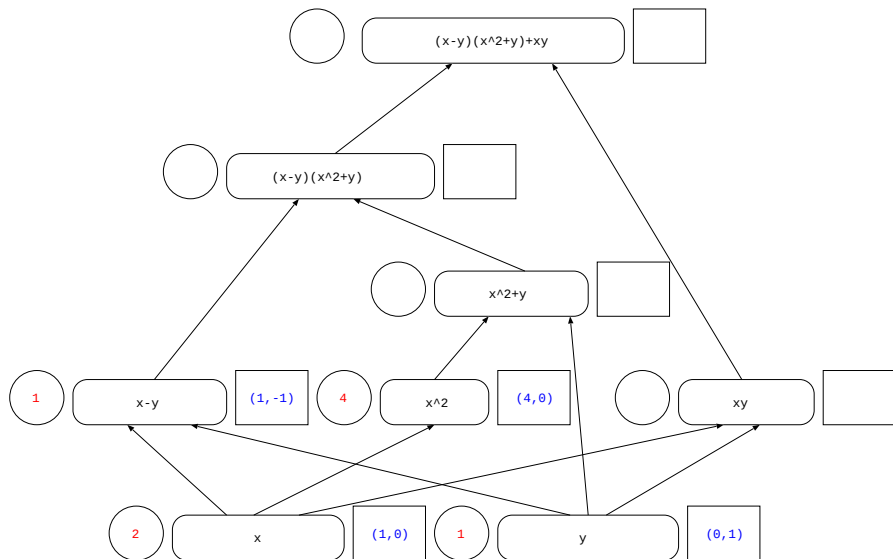
bottom up アルゴリズム (1) (初期化)



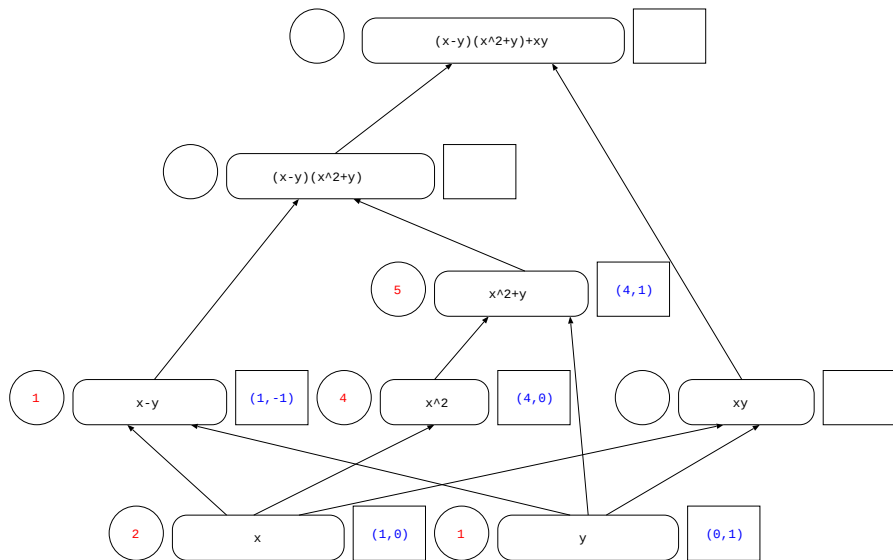
bottom up アルゴリズム (2)



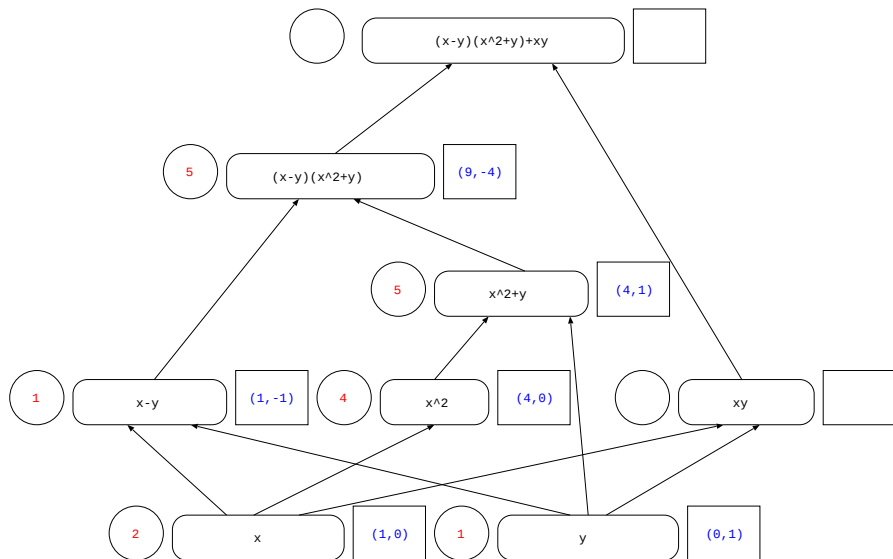
bottom up アルゴリズム (3)



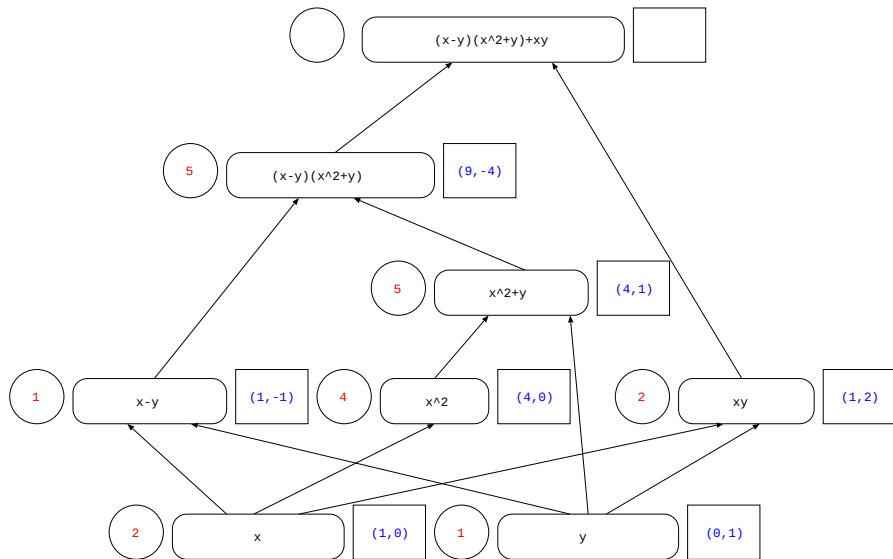
bottom up アルゴリズム (4)



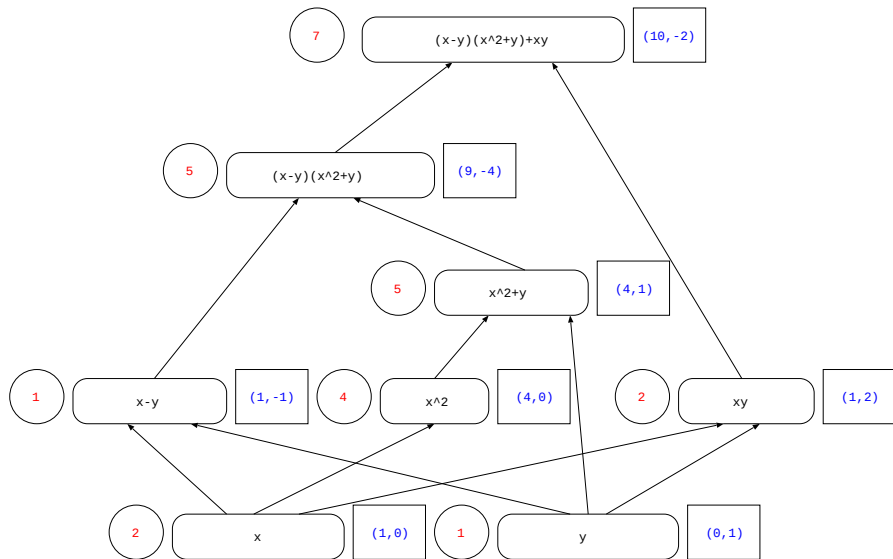
bottom up アルゴリズム (5)



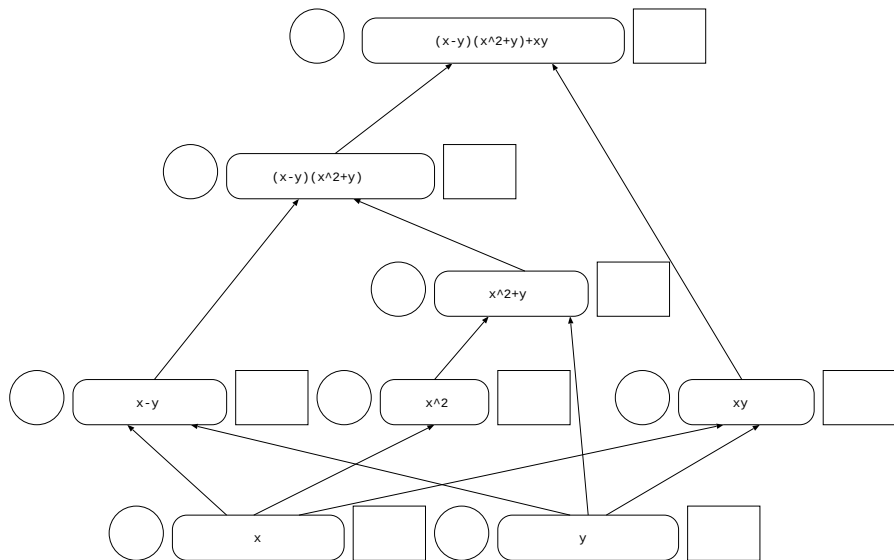
bottom up アルゴリズム (6)



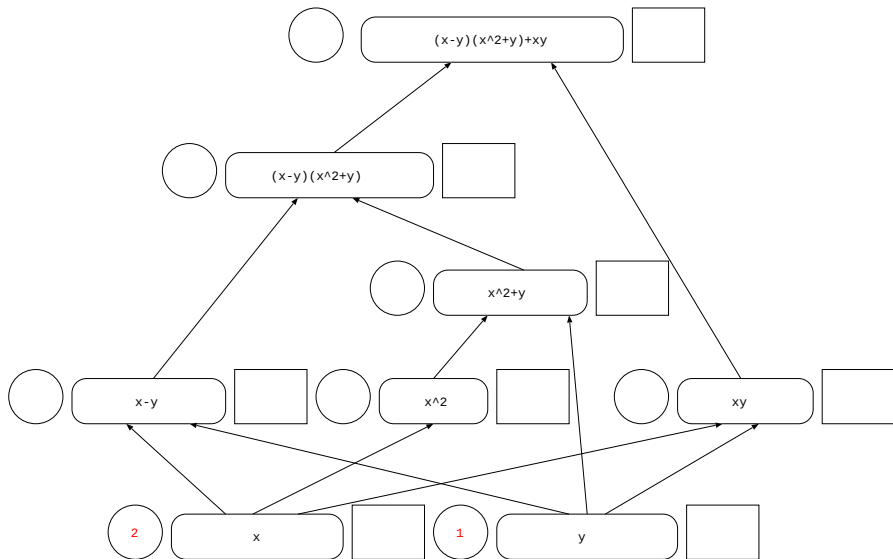
bottom up アルゴリズム (7)



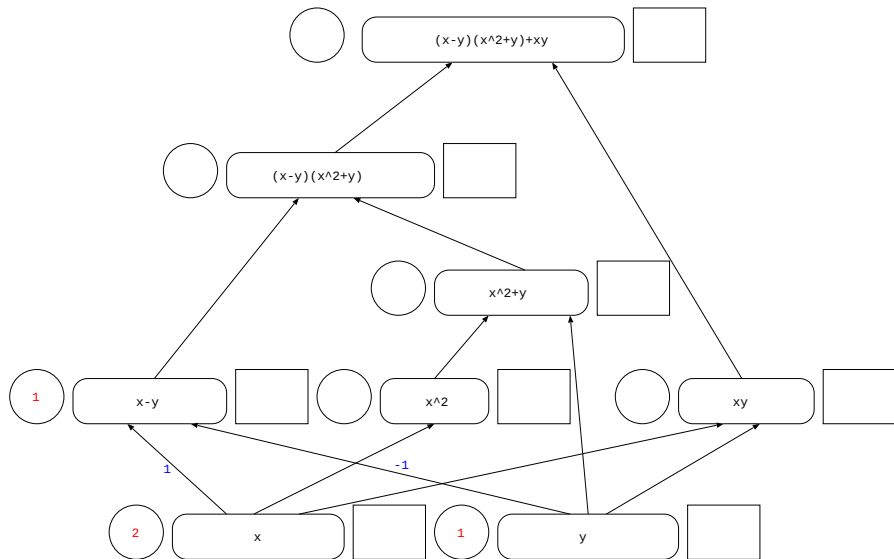
top down アルゴリズム (0)



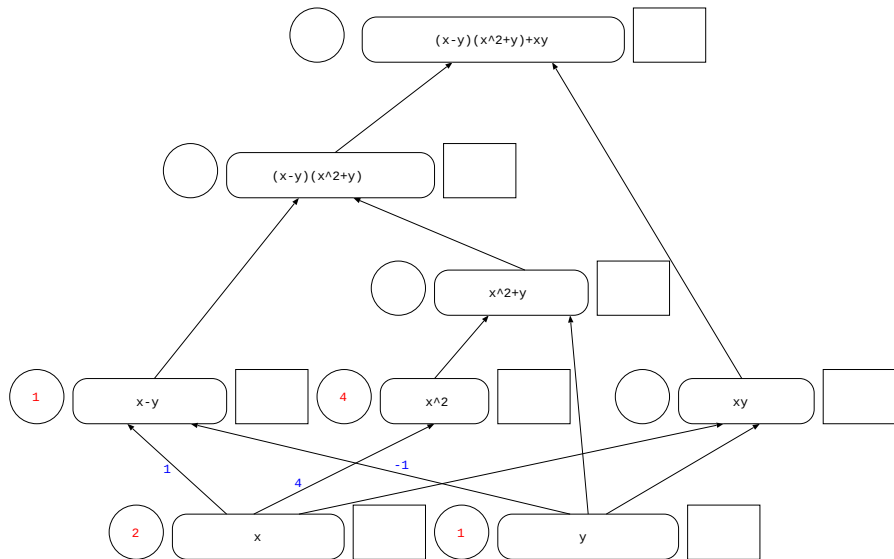
top down アルゴリズム (1) (前進開始)



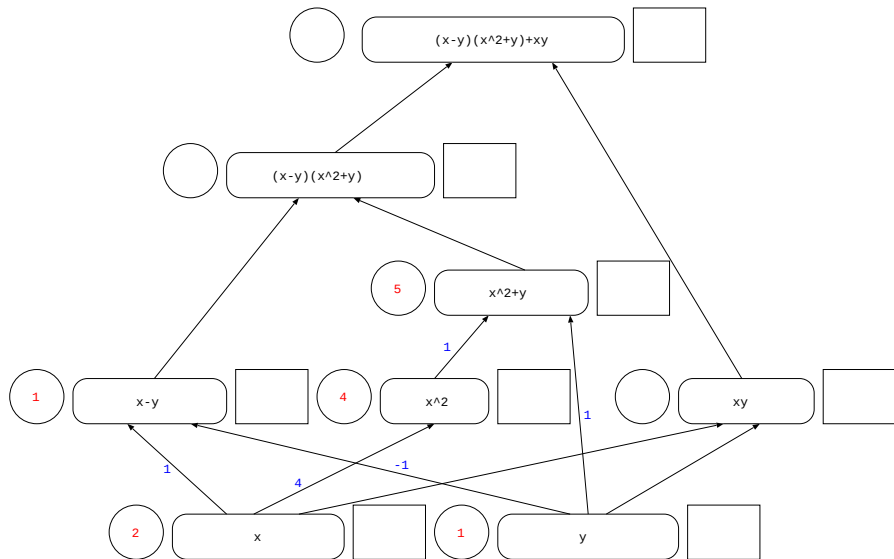
top down アルゴリズム (2)



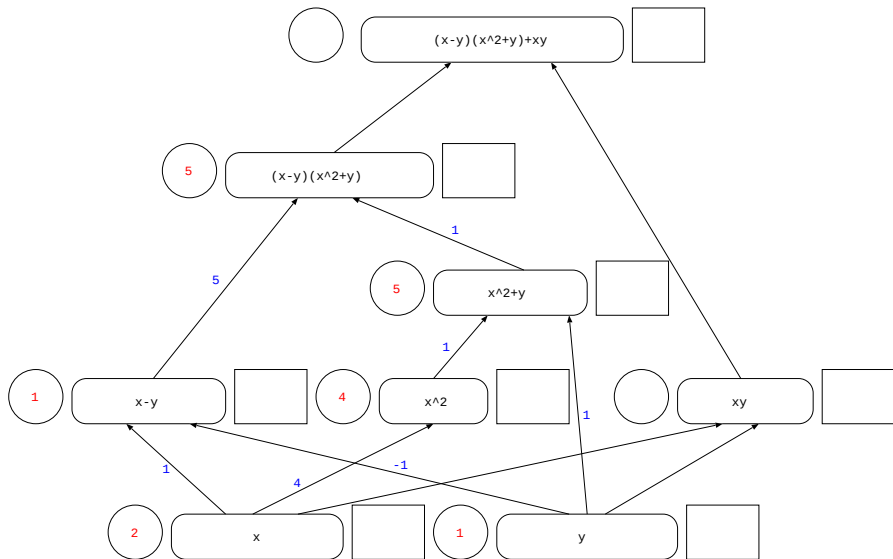
top down アルゴリズム (3)



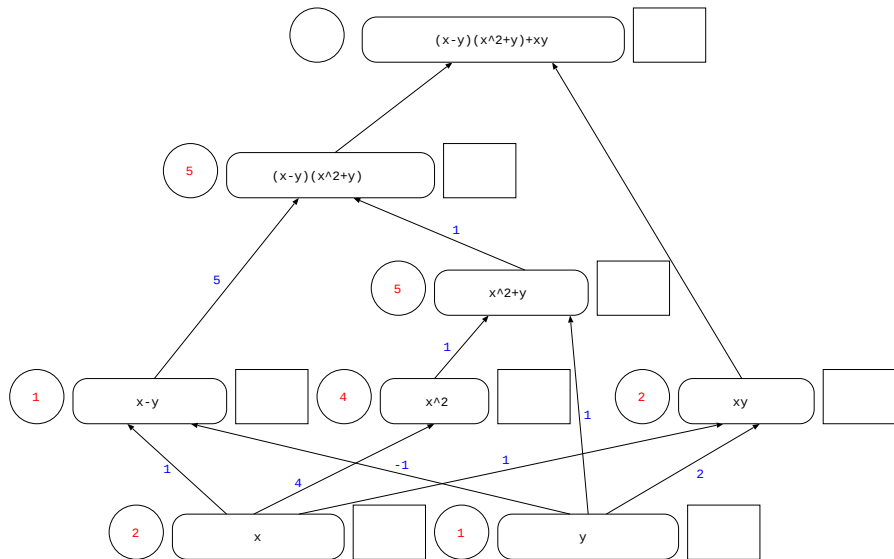
top down アルゴリズム (4)



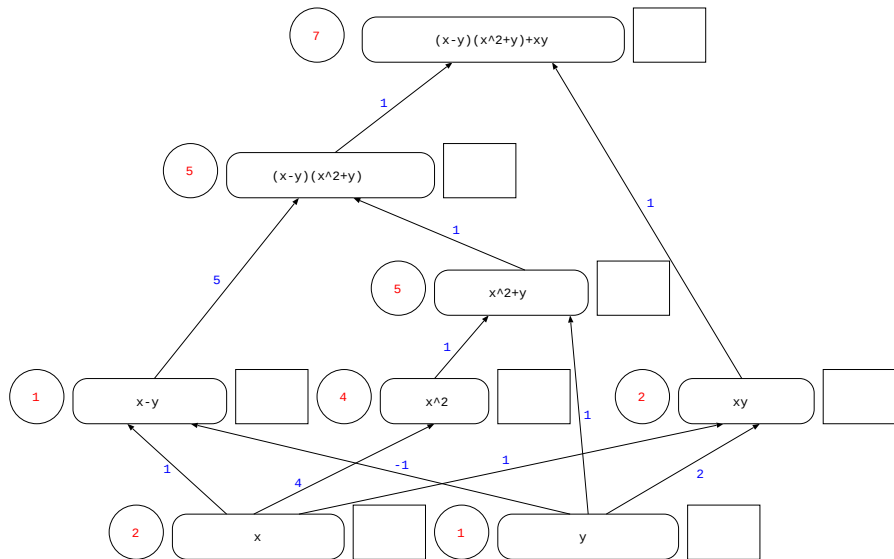
top down アルゴリズム (5)



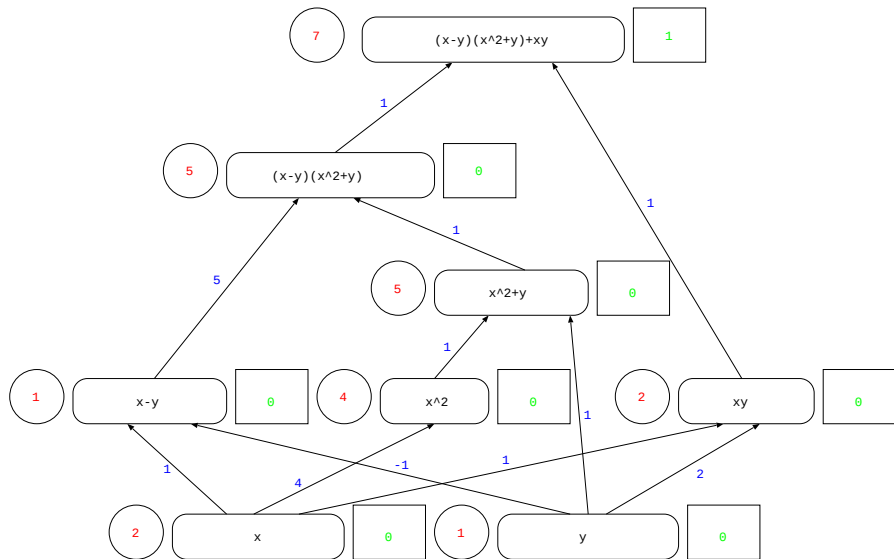
top down アルゴリズム (6)



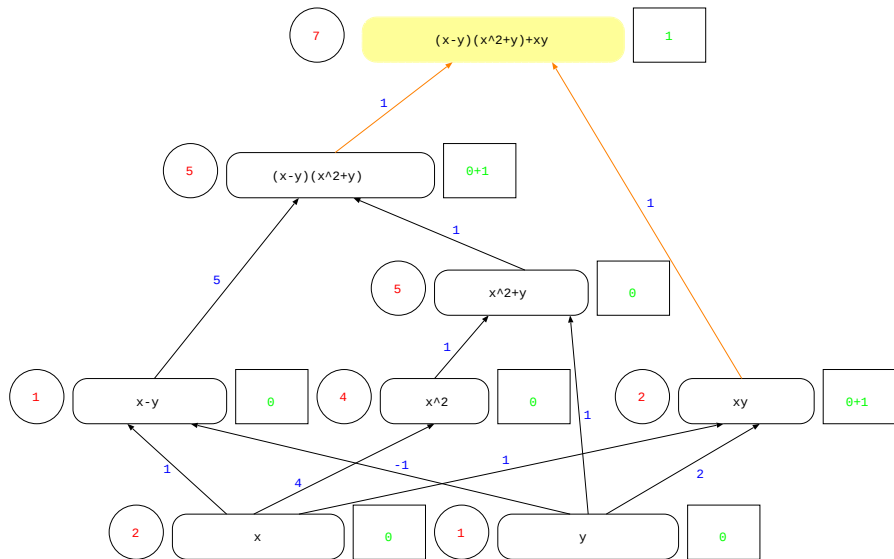
top down アルゴリズム (7)



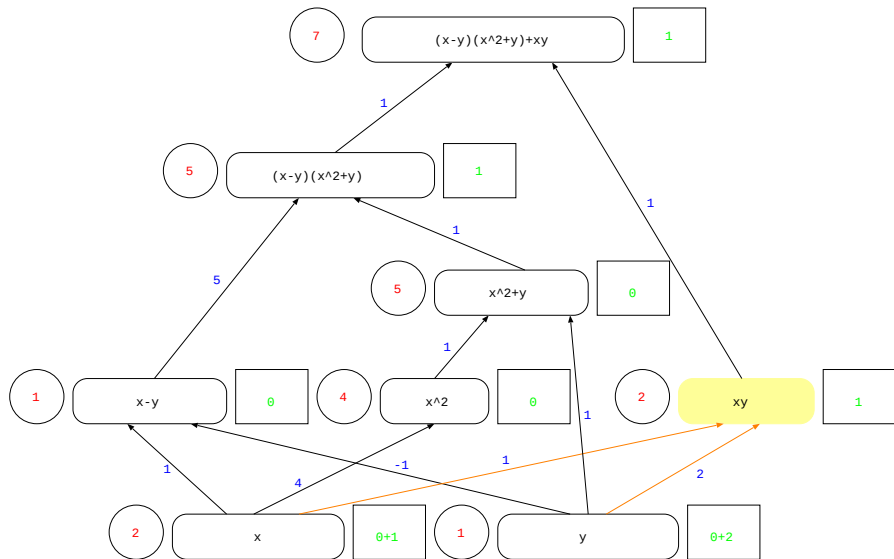
top down アルゴリズム (8) (後退開始)



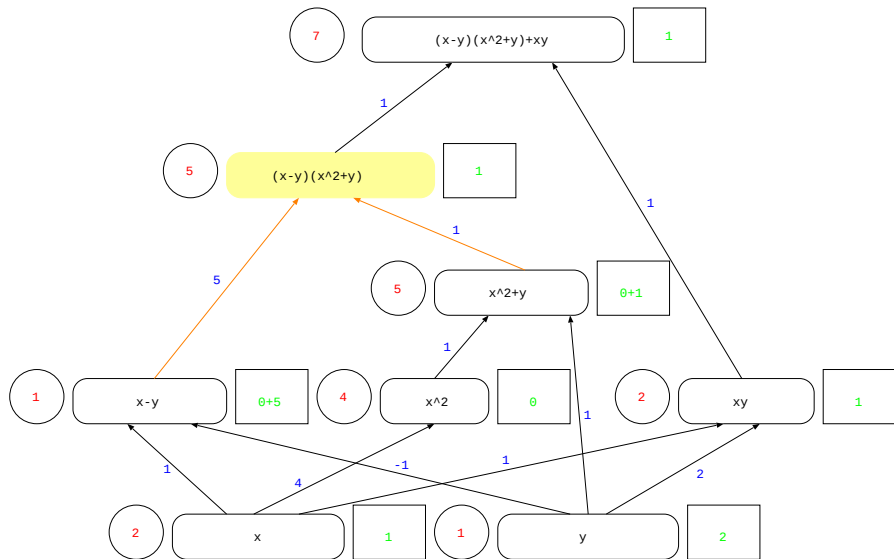
top down アルゴリズム (9)



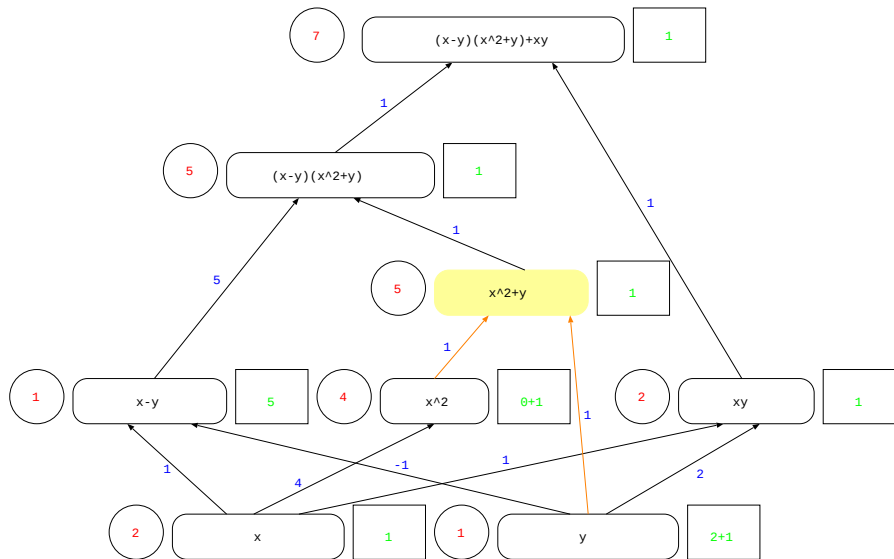
top down アルゴリズム (10)



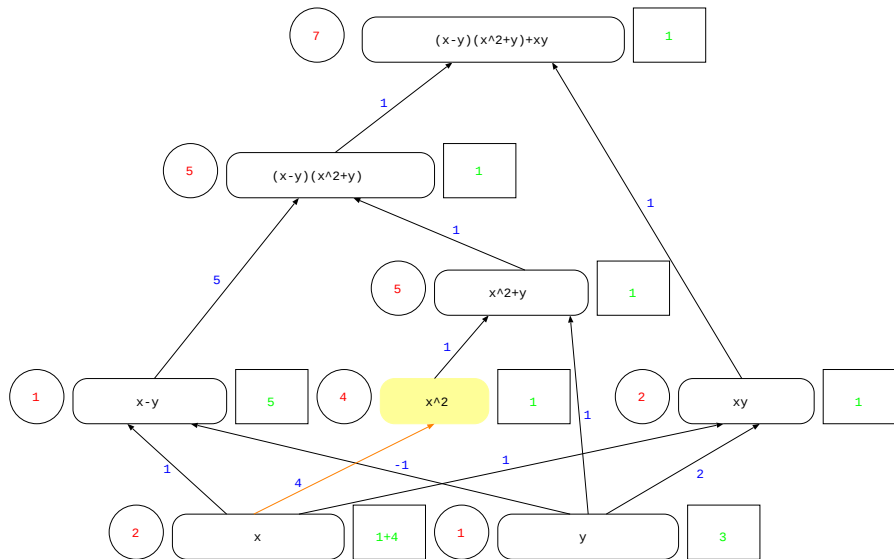
top down アルゴリズム (11)



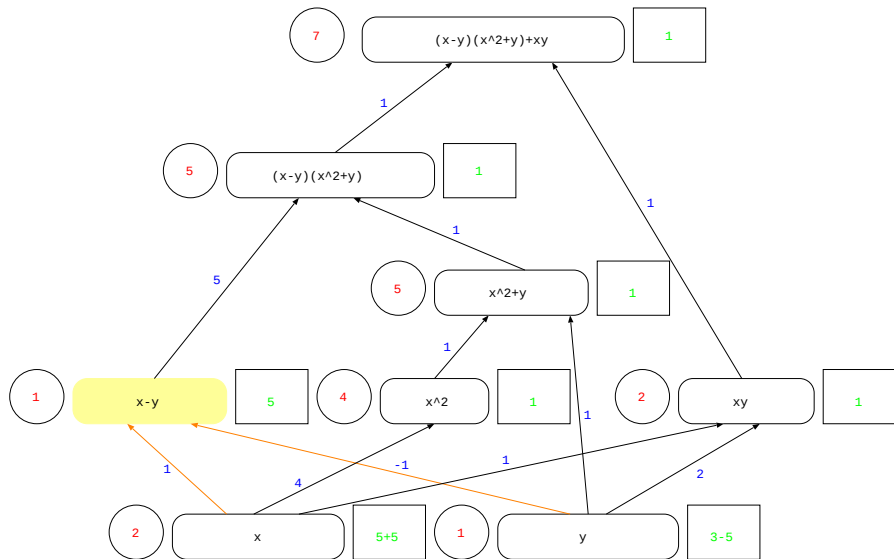
top down アルゴリズム (12)



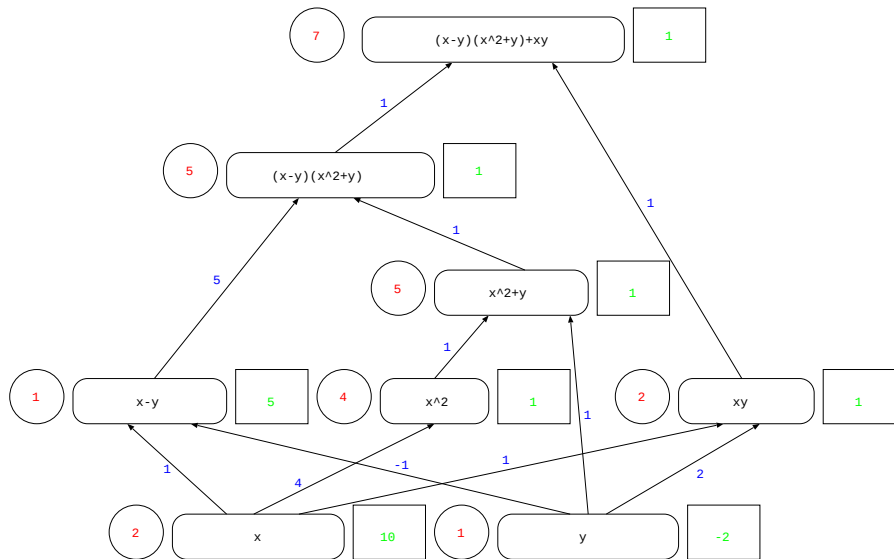
top down アルゴリズム (13)



top down アルゴリズム (14)



top down アルゴリズム (15)



それぞれの手法の特徴

bottom up

- forward mode と呼ばれる。
- 計算グラフを作成する必要がない = 実装が簡単
- 微分を格納する作業変数の数 = 入力変数の数

top down

- reverse mode と呼ばれる。
- 計算グラフを作成する必要がある = 実装が面倒
- 微分を格納する作業変数の数 = 出力変数の数

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

- $n \gg m$ なら top down が適している。
- $n \ll m$ なら bottom up が適している。