数式微分と自動微分

Kohei Asano

whoami

所属: 北大数学科 B4

- 卒業ポスターは多項式プログラムから不変量を自動で計算するなにか

最近興味のあること: 型, 論理など

数式の微分をしていたら面白かったので

今日は計算機の微分の話をします

TL;DR

- ☀️ 高校数学的な**数式微分**からはじめて**自動微分**をわかった気になる
    - 自動微分はPyTorch, TensorFlowでも使われている, 機械学習ブームを支える技術
- ☀️ Rustで数式微分の最適化をしてみました
- ☀️ 数式いじりの自動化はいいぞ

もくじ

- **計算機での微分**
- 数式微分から自動微分
- 微分グラフの最適化と実験
- その他参考
- まとめ

計算機での微分

- 数値微分
  - 前進差分法, 中心差分法

! 勾配を計算する時, 多変数関数は変数の数だけ関数の評価をしないといけないので計算量がネック

計算機での微分

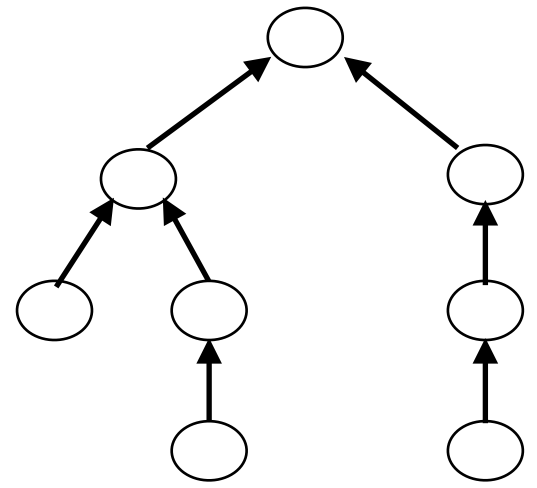
- 数値微分
  - 前進差分法, 中心差分法

! 勾配を計算する時, 多変数関数は変数の数だけ関数の評価をしないといけないので計算量がネック

- **数式微分(今日のヒロイン)**
  - 高校数学的な**式木**の書き換え
  - 微分以外の操作を表せたりする(制約とか)

! データ構造が大きくなるの(Expression Swell)が問題と言われる (本当？)

式木 := 数式を表す木



計算機での微分

数値微分

- 前進差分法, 中心差分法

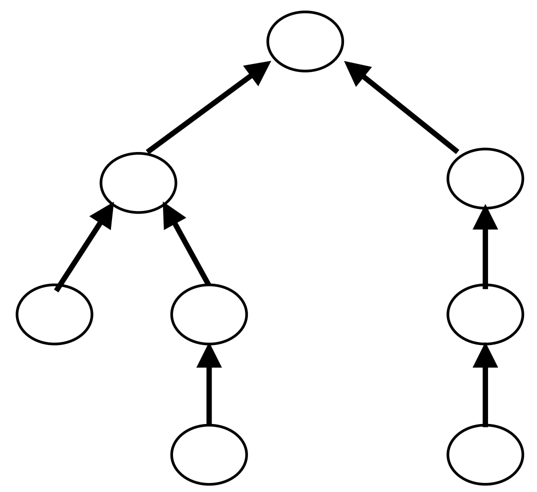
! 勾配を計算する時, 多変数関数は変数の数だけ関数の評価をしないといけないので計算量がネック

数式微分(今日のヒロイン)

- 高校数学的な**式木**の書き換え
- 微分以外の操作を表せたりする(制約とか)

! データ構造が大きくなるの(Expression Swell)が問題と言われる (本当？)

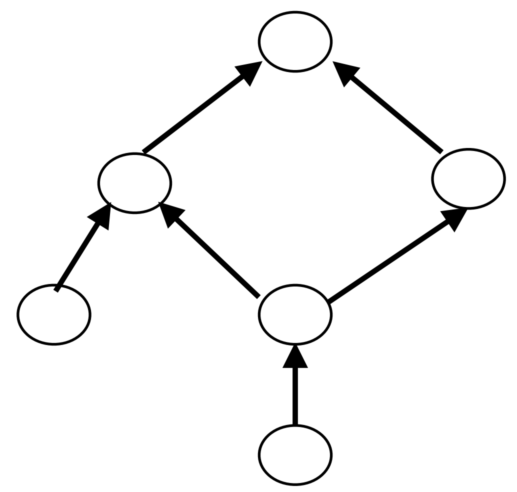
式木 := 数式を表す木



自動微分(今日の主役)

- **計算グラフ**上で高速に微分する. 物理でも機械学習でも人気.
- 機械学習の基礎の**誤差逆伝播法はReverseモードの自動微分**

計算グラフ := 計算構造を表すDAG



もくじ

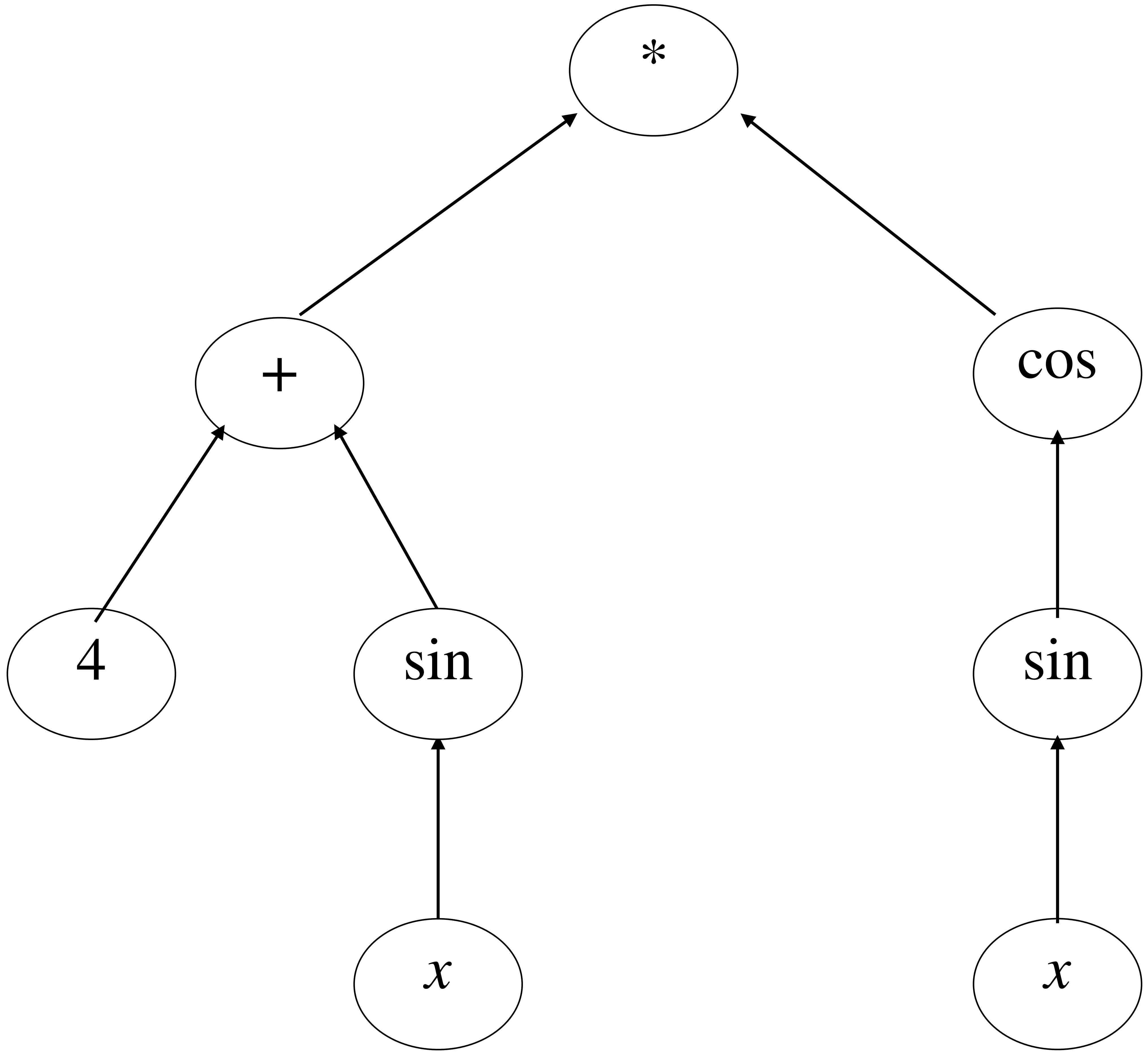
数式微分から自動微分

☀ 愚直な数式微分から, 少しずつ工夫していって自動微分を理解する

- 式木 := 数式を表す木構造
- 式DAG := 数式を表すグラフ
- 微分グラフ := 式DAGの辺に微分の**式**を載せたもの
- 計算グラフ := 式DAGの辺に微分の**値**を載せたもの

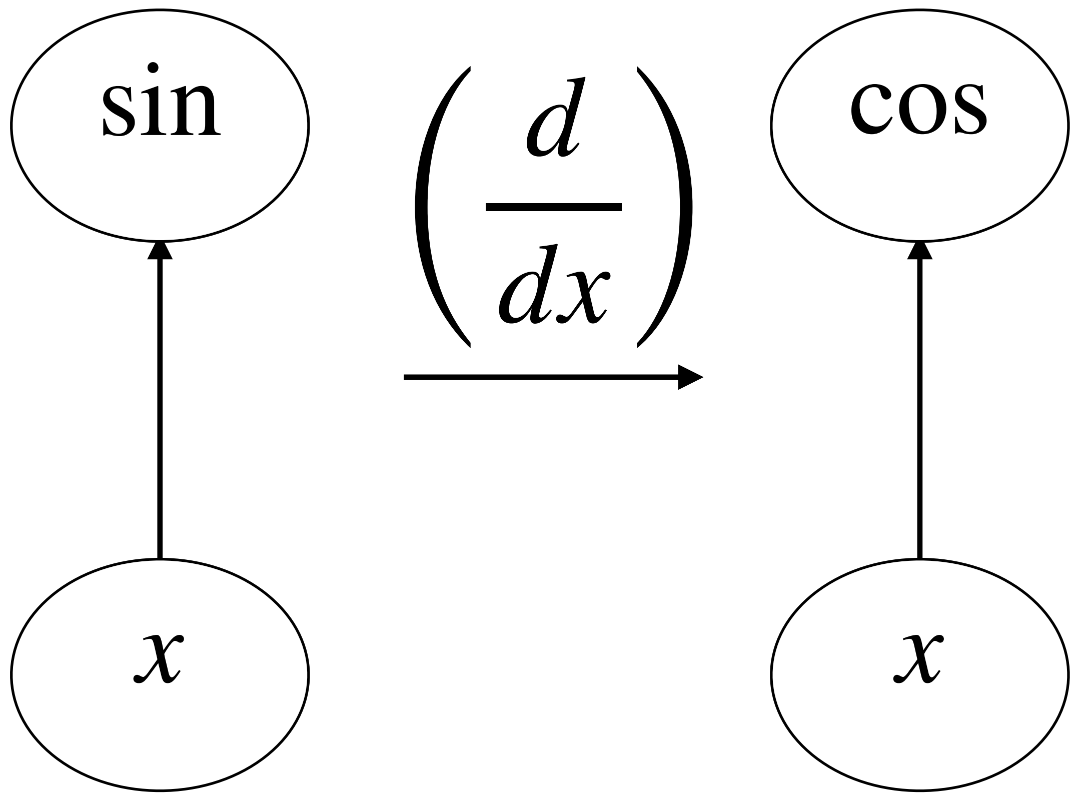
式木

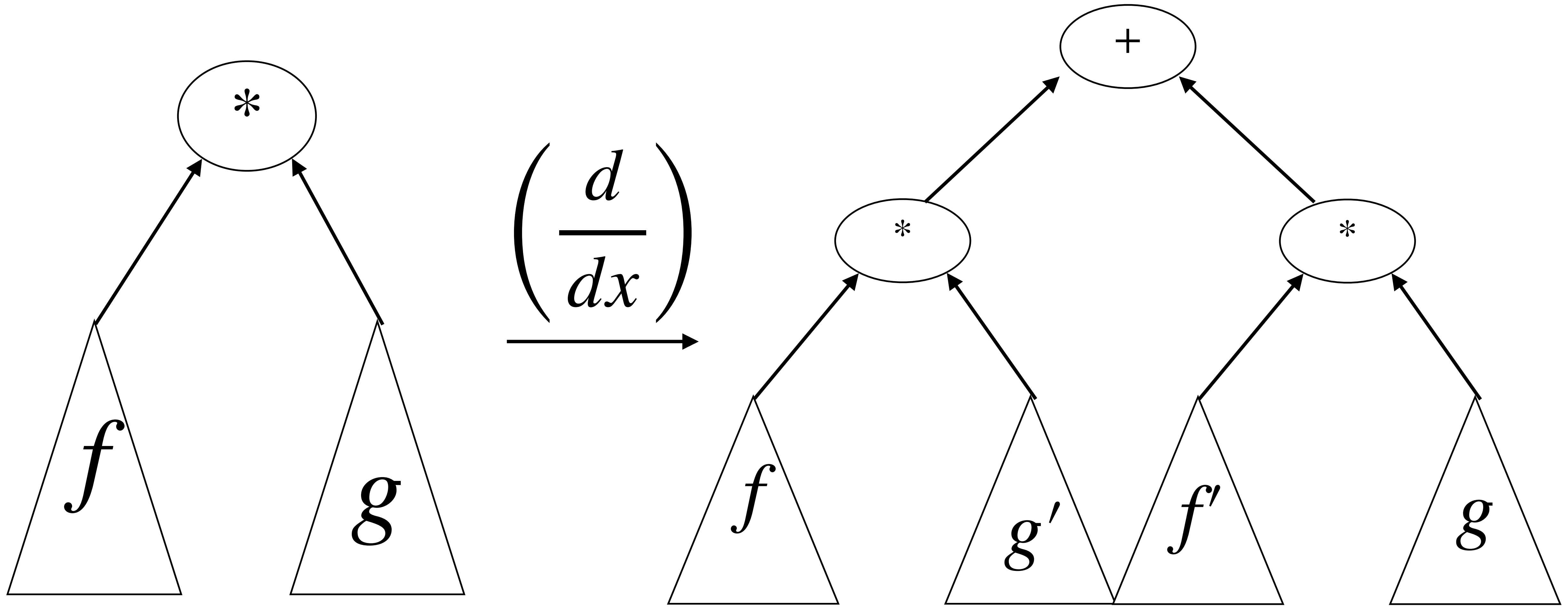
$f(x) = (4 + \sin(x)) \cdot \cos(\sin(x)) =$

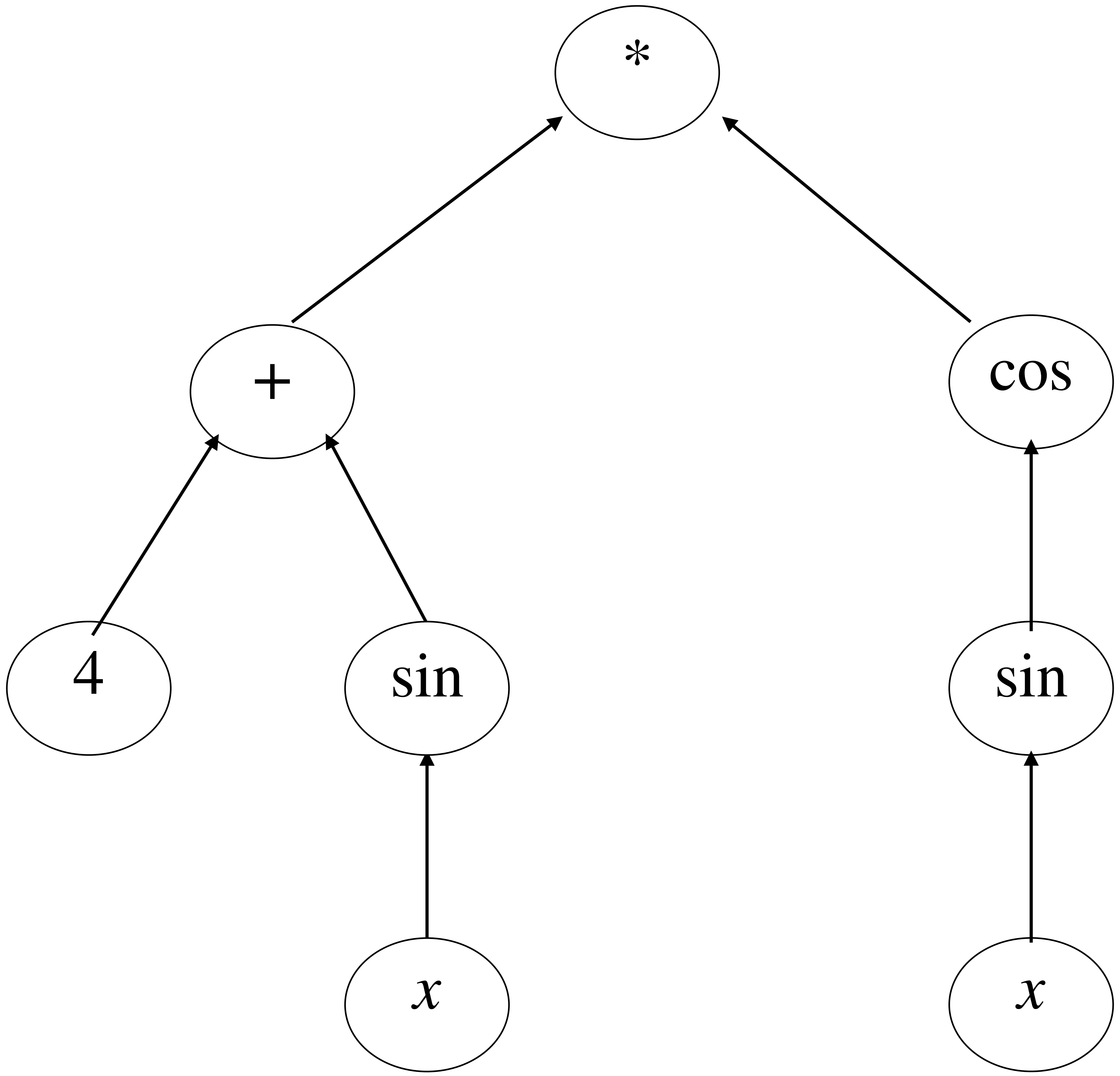


式木の微分

$$\left(\frac{d}{dx}\right)(4+\sin(x))\cdot\cos(\sin(x))$$



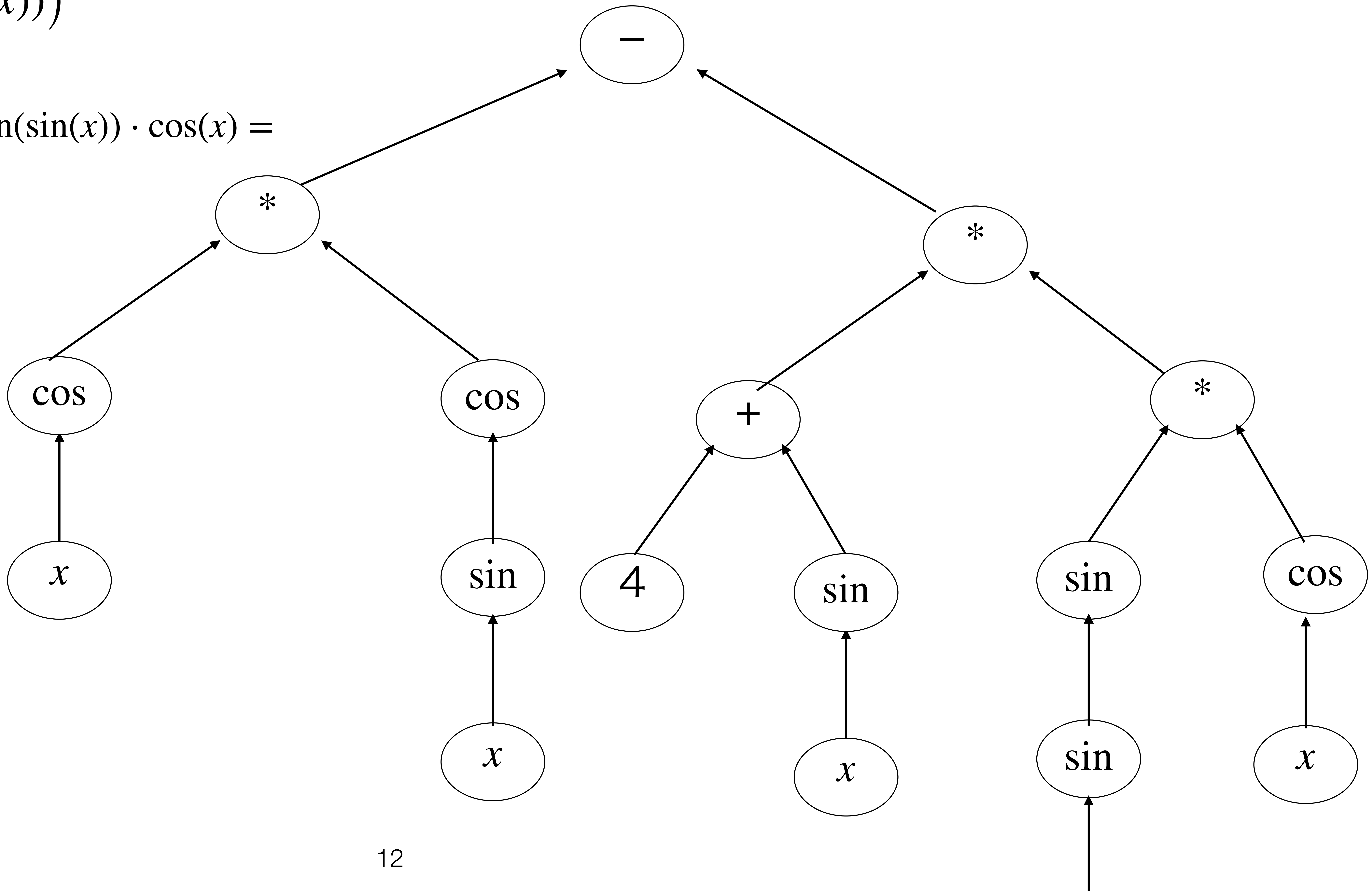




式木の微分

$$\left(\frac{d}{dx}\right)\big((4+\sin(x))\cdot\cos(\sin(x))\big)$$

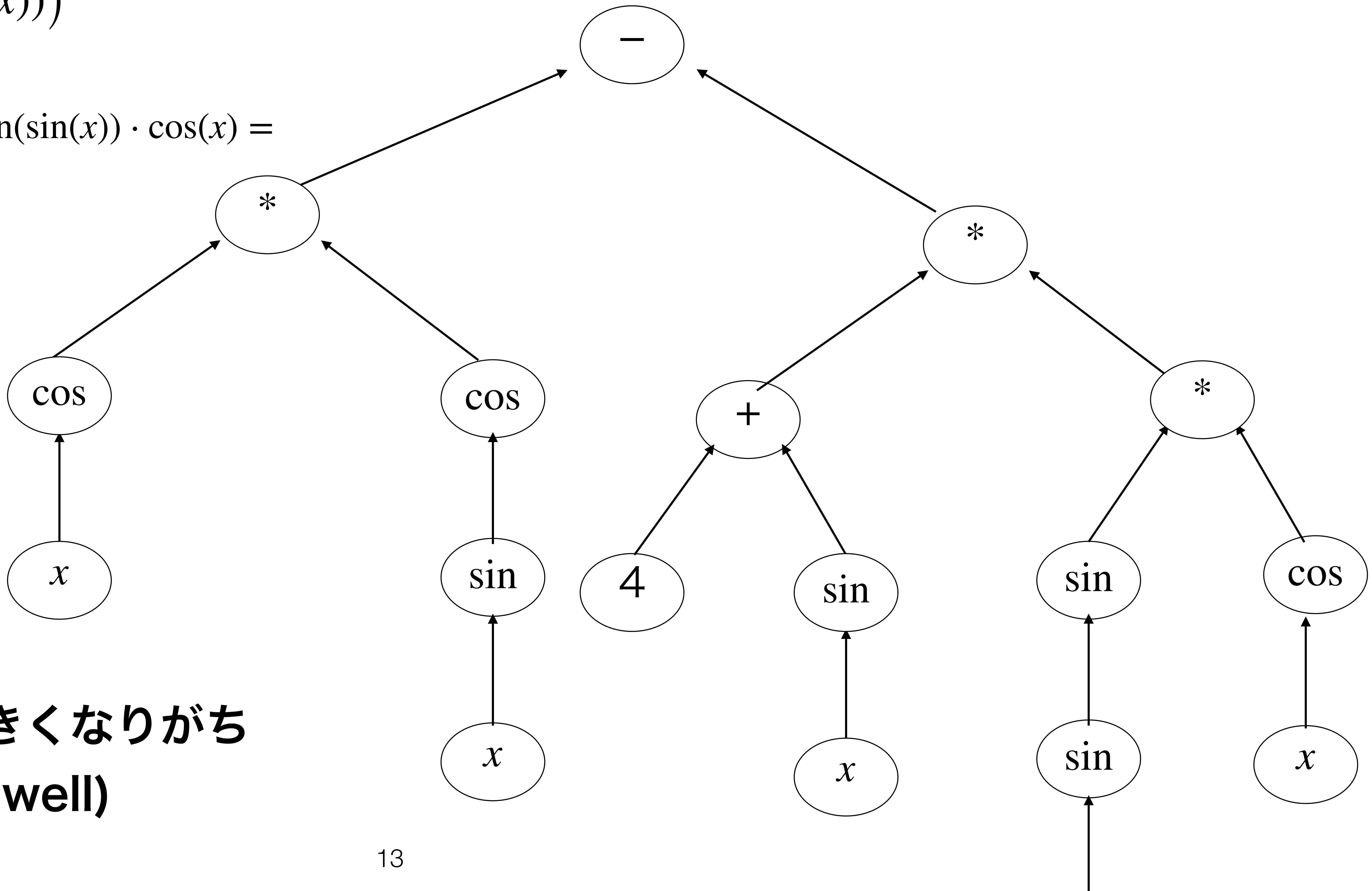
$$=\cos(x)\cdot\cos(\sin(x)) - (4+\sin(x))\cdot\sin(\sin(x))\cdot\cos(x) =$$

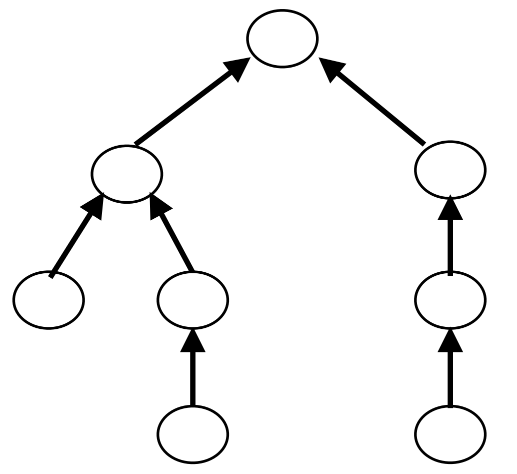


式木の微分

$$\left(\frac{d}{dx}\right)\big((4+\sin(x))\cdot\cos(\sin(x))\big)$$

$$= \cos(x)\cdot\cos(\sin(x)) - (4+\sin(x))\cdot\sin(\sin(x))\cdot\cos(x) =$$





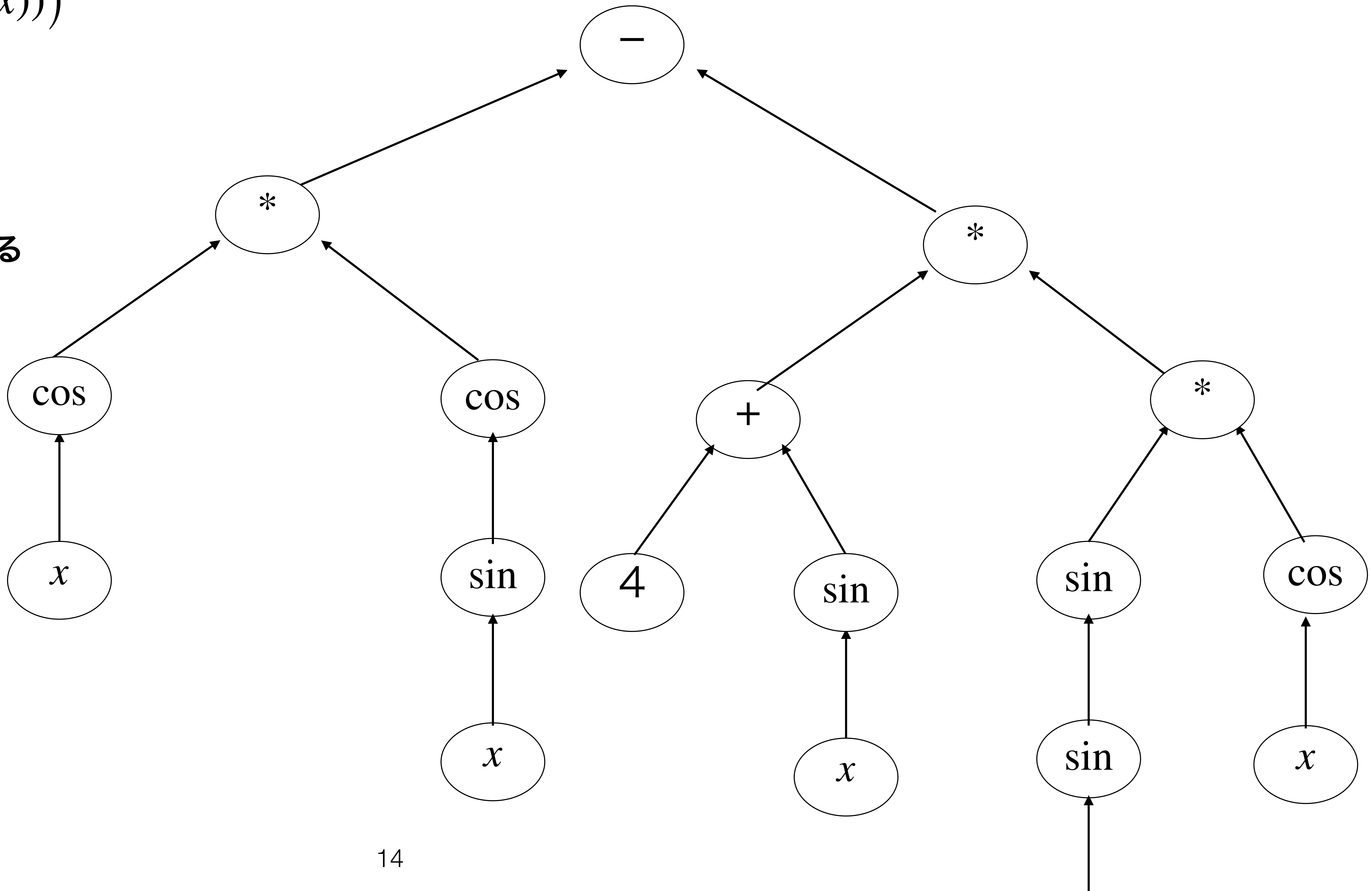
元のサイズに対して大きくなりがち
(Expression Swell)

式木の微分値を求める$(x = \pi)$

$$\left(\frac{d}{dx}\right)\big((4+\sin(x))\cdot\cos(\sin(x))\big)$$

葉から値を伝搬して求めることができる

$\mathcal{O}(V')$ $\quad V'$:微分後の頂点の数

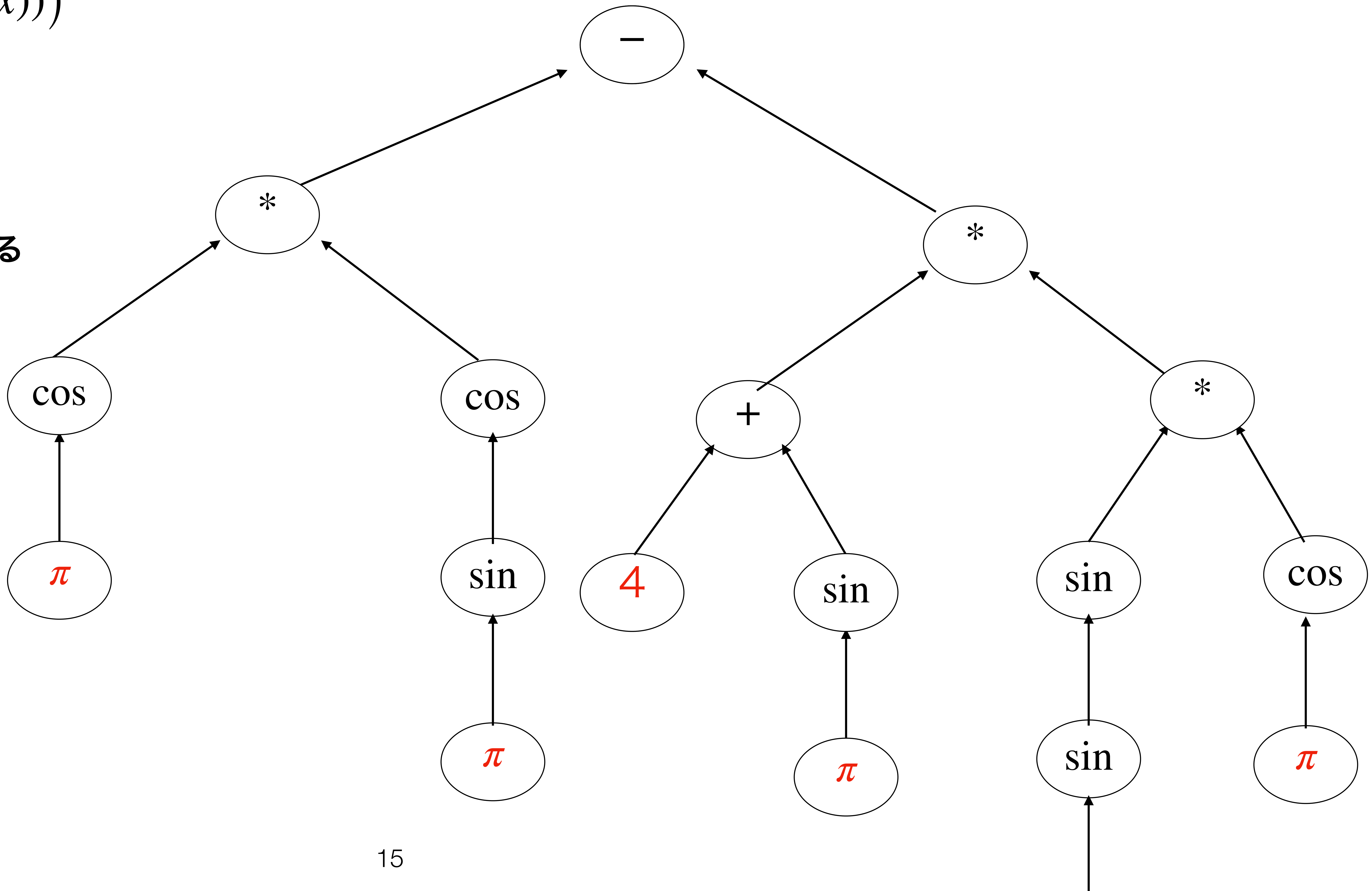


式木の微分値を求める$(x = \pi)$

$$\left(\frac{d}{dx}\right)\left((4 + \sin(x)) \cdot \cos(\sin(x))\right)$$

葉から値を伝搬して求めることができる

$\mathcal{O}(V')$ V' :微分後の頂点の数

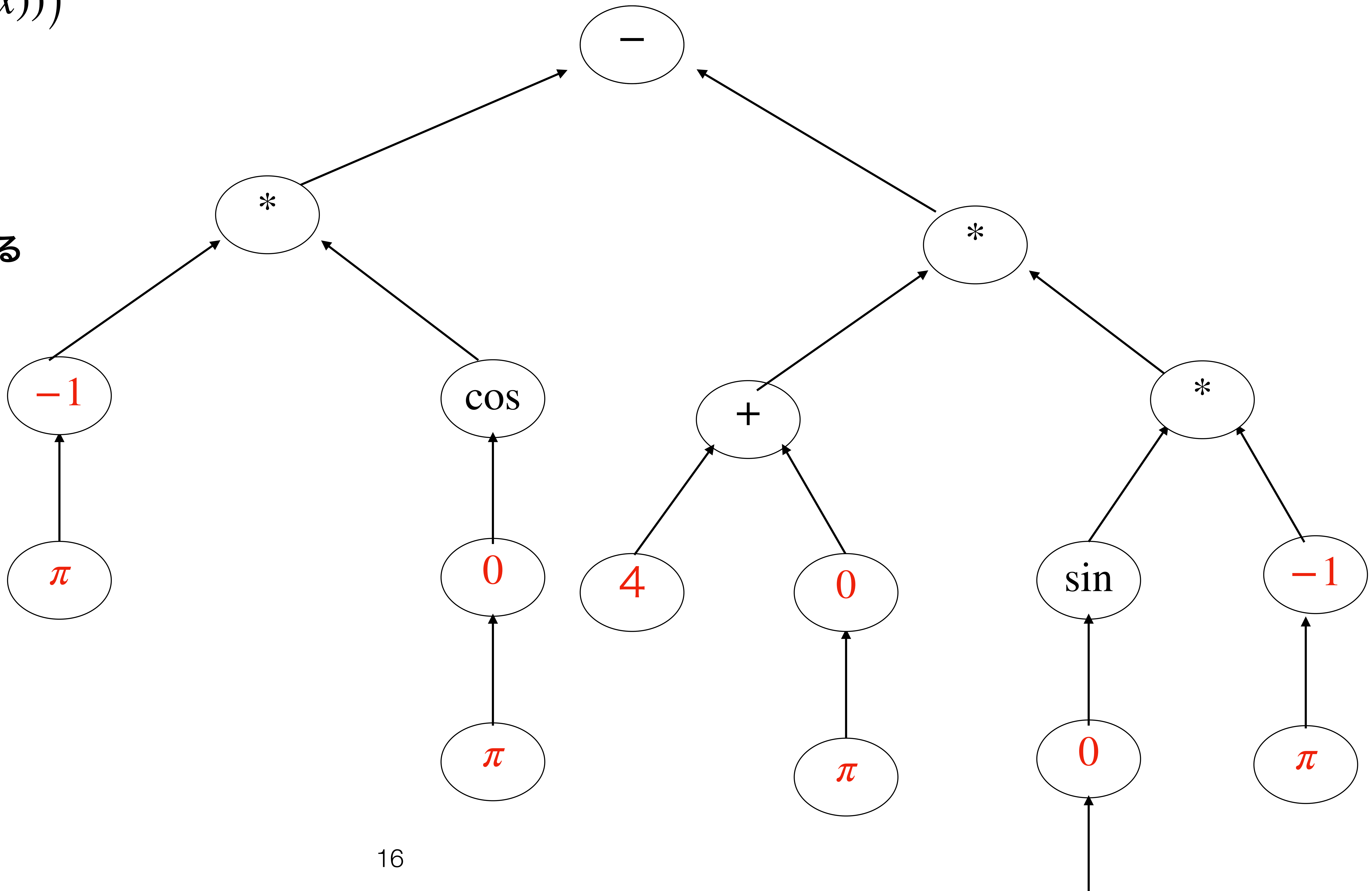


式木の微分値を求める$(x = \pi)$

$$\left(\frac{d}{dx}\right)\left((4 + \sin(x)) \cdot \cos(\sin(x))\right)$$

葉から値を伝搬して求めることができる

$\mathcal{O}(V')$　V' :微分後の頂点の数

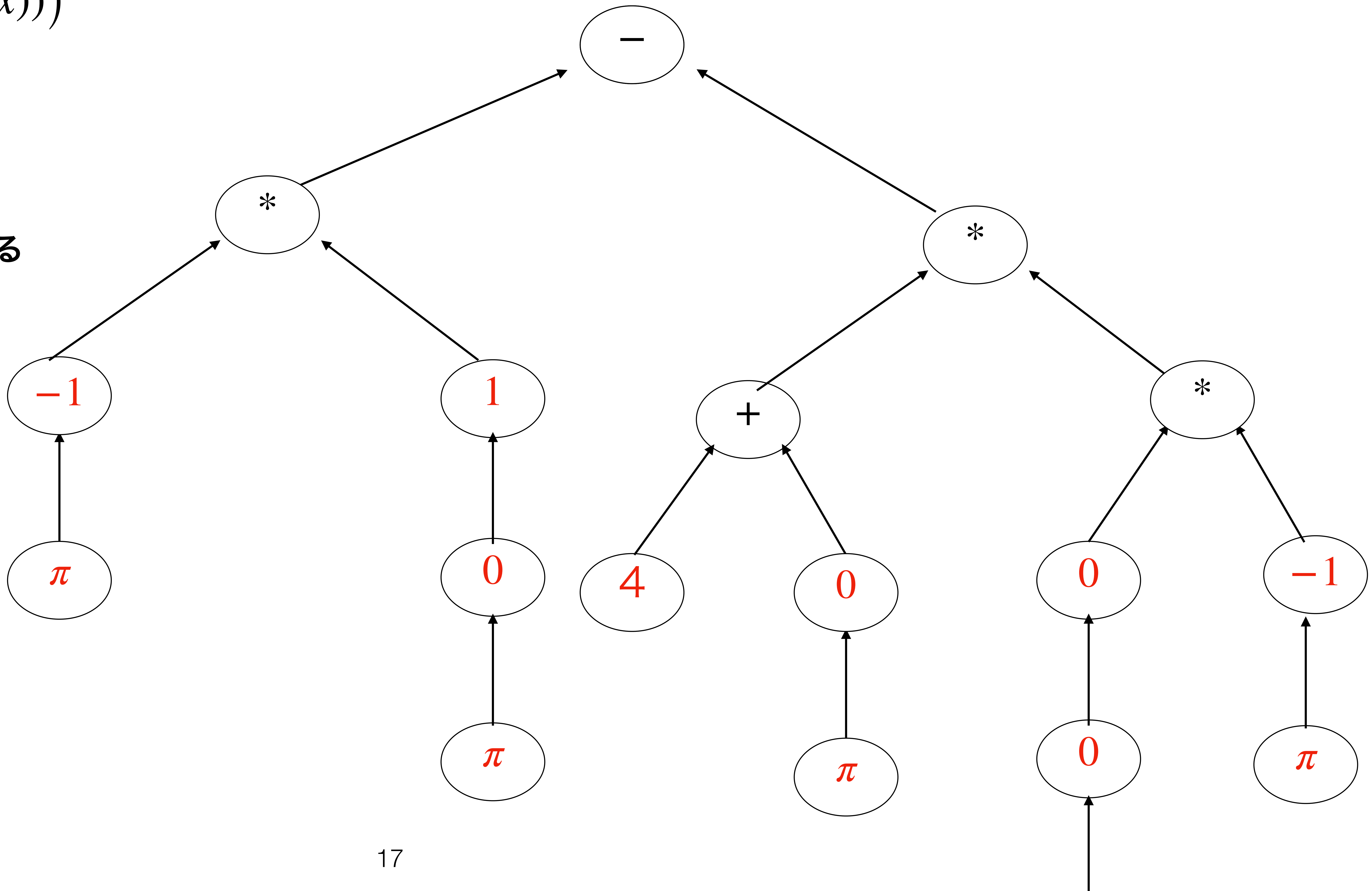


式木の微分値を求める($x = \pi$)

$$\left(\frac{d}{dx}\right)\left((4 + \sin(x)) \cdot \cos(\sin(x))\right)$$

葉から値を伝搬して求めることができる

$\mathcal{O}(V')$　V' :微分後の頂点の数

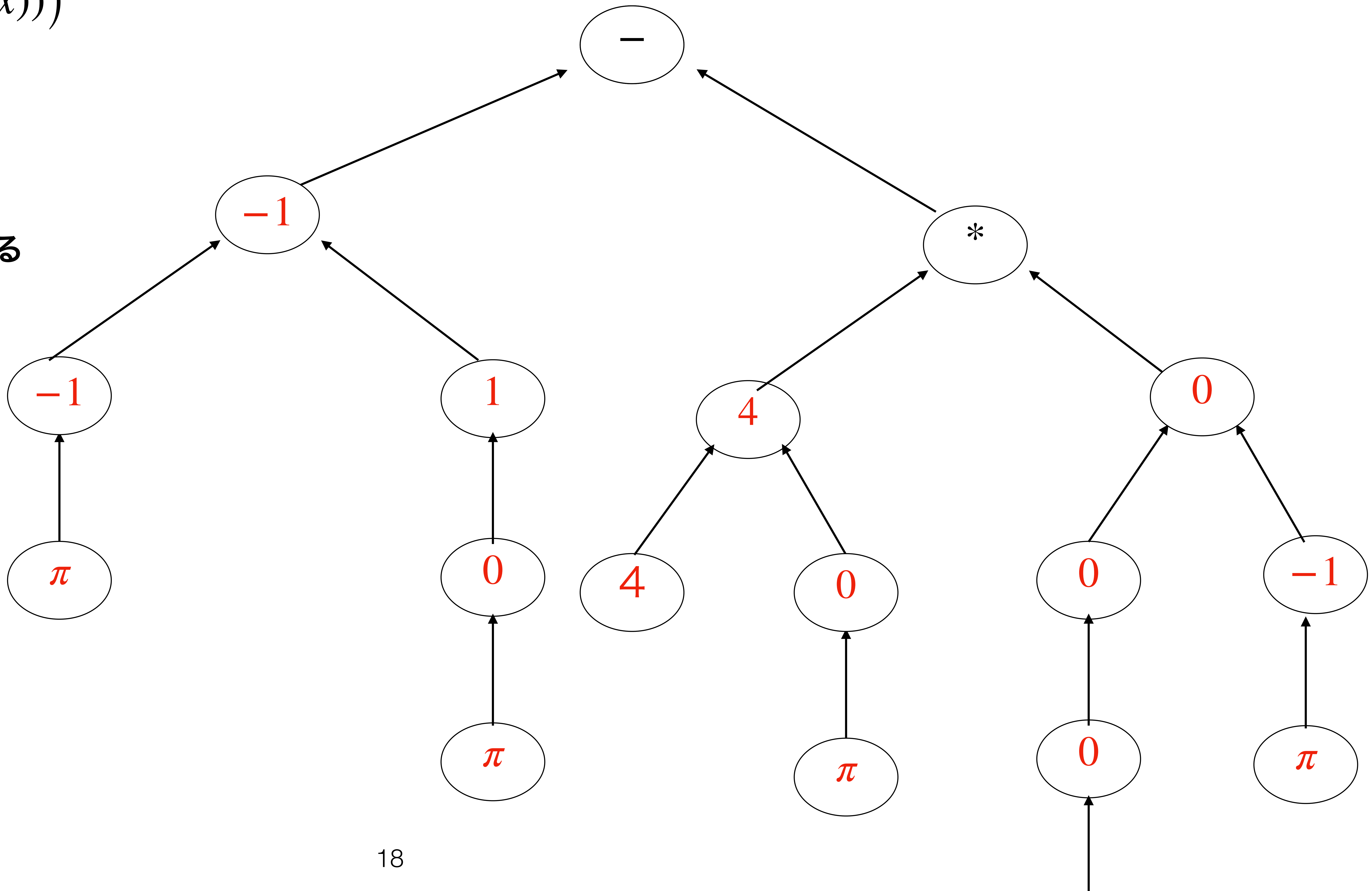


式木の微分値を求める$(x = \pi)$

$$\left(\frac{d}{dx}\right)\left((4 + \sin(x)) \cdot \cos(\sin(x))\right)$$

葉から値を伝搬して求めることができる

$\mathcal{O}(V')$　V' **:微分後の頂点の数**



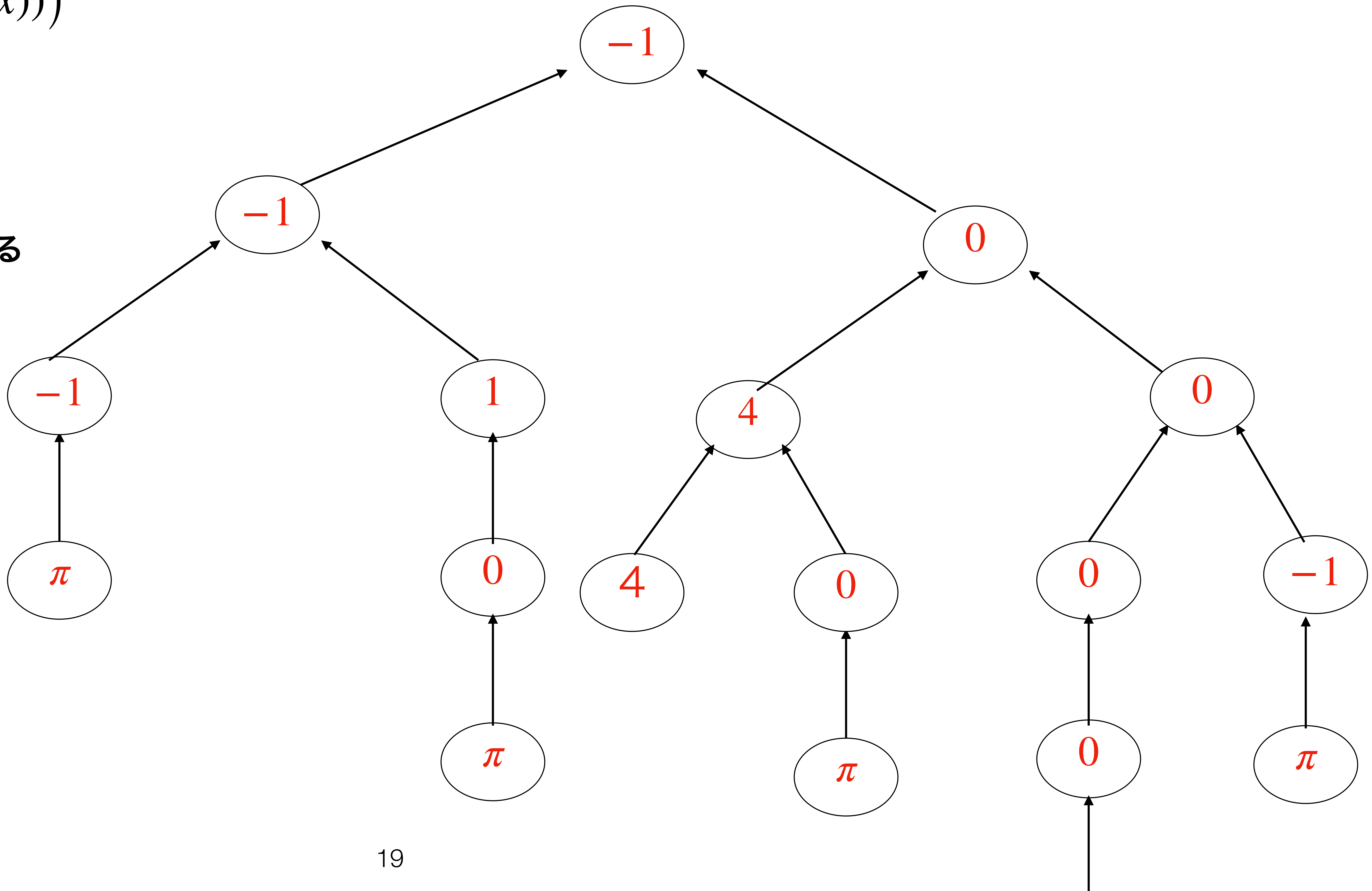
式木の微分値を求める$(x = \pi)$

$$\left(\frac{d}{dx}\right)\left((4+\sin(x))\cdot\cos(\sin(x))\right)$$

$$= -1$$

葉から値を伝搬して求めることができる

$\mathcal{O}(V')$　V' :**微分後の頂点の数**



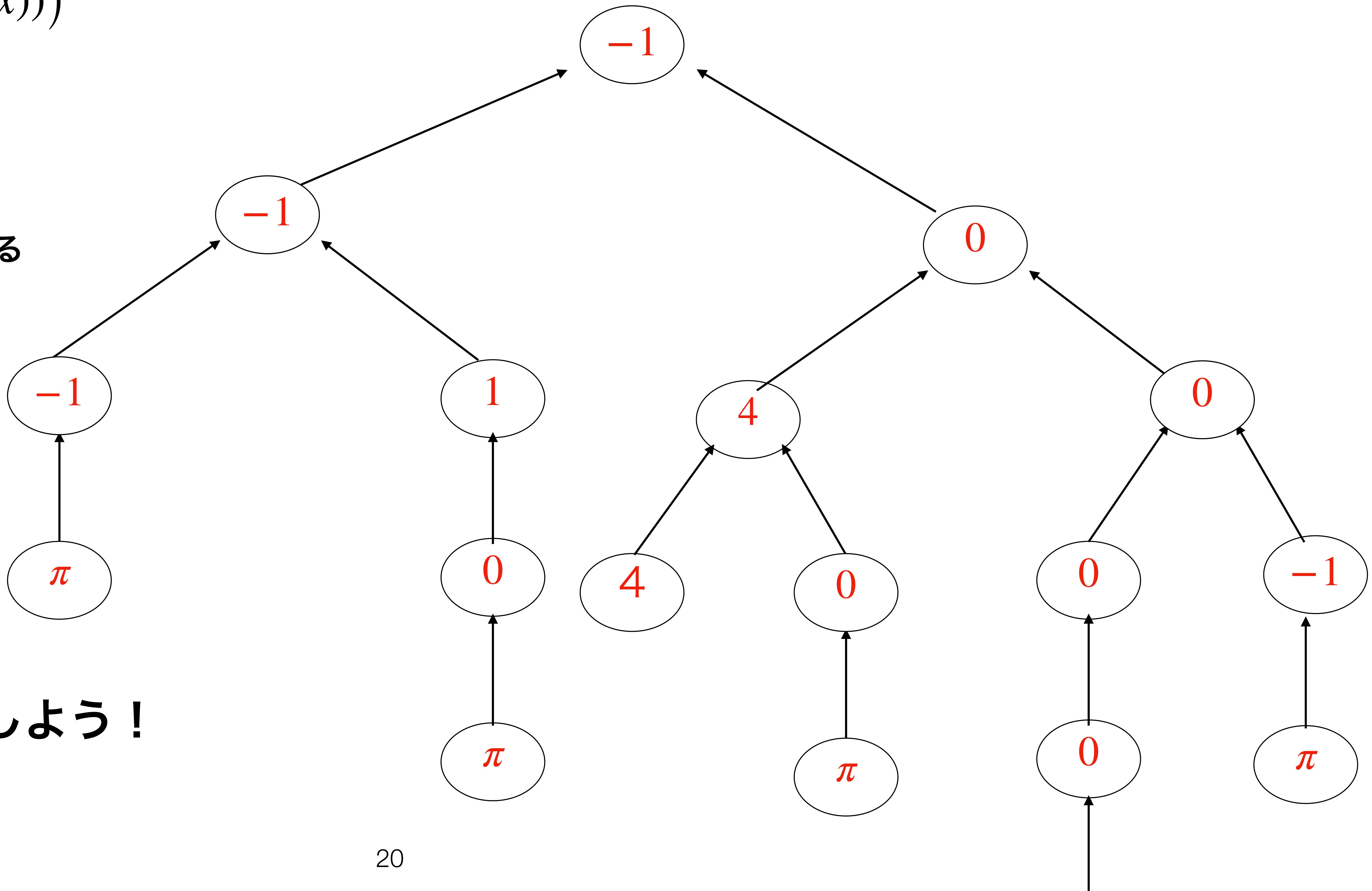
式木の微分値を求める$(x = \pi)$

$$\left(\frac{d}{dx}\right)\Big((4+\sin(x))\cdot\cos(\sin(x))\Big)$$

$$= -1$$

葉から値を伝搬して求めることができる

$\mathcal{O}(V')$　V' **:微分後の頂点の数**

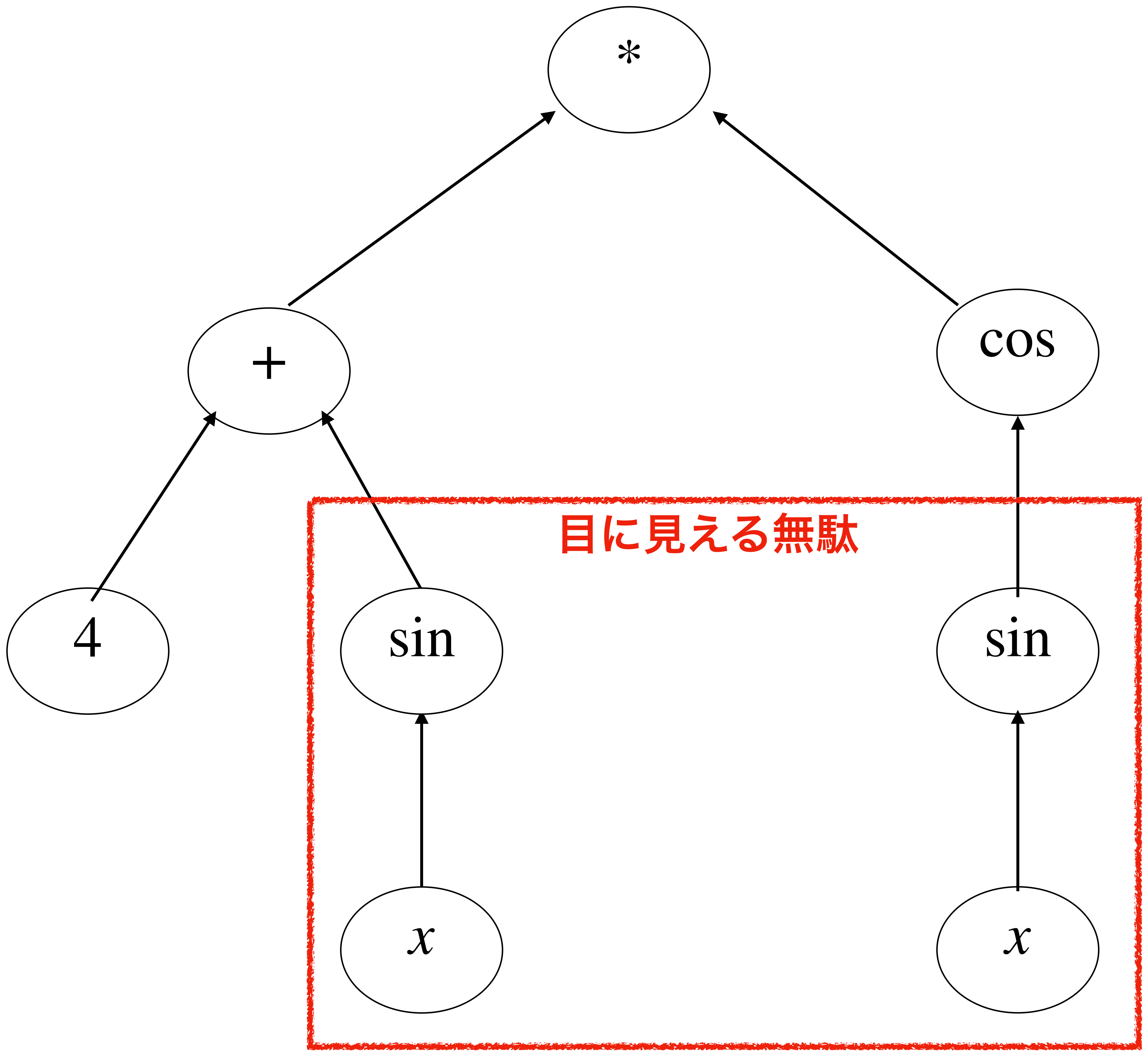


改善しよう！

式DAG(計算グラフ)

注) DAG := Directed Acyclic Graph
閉路がない有向グラフのこと

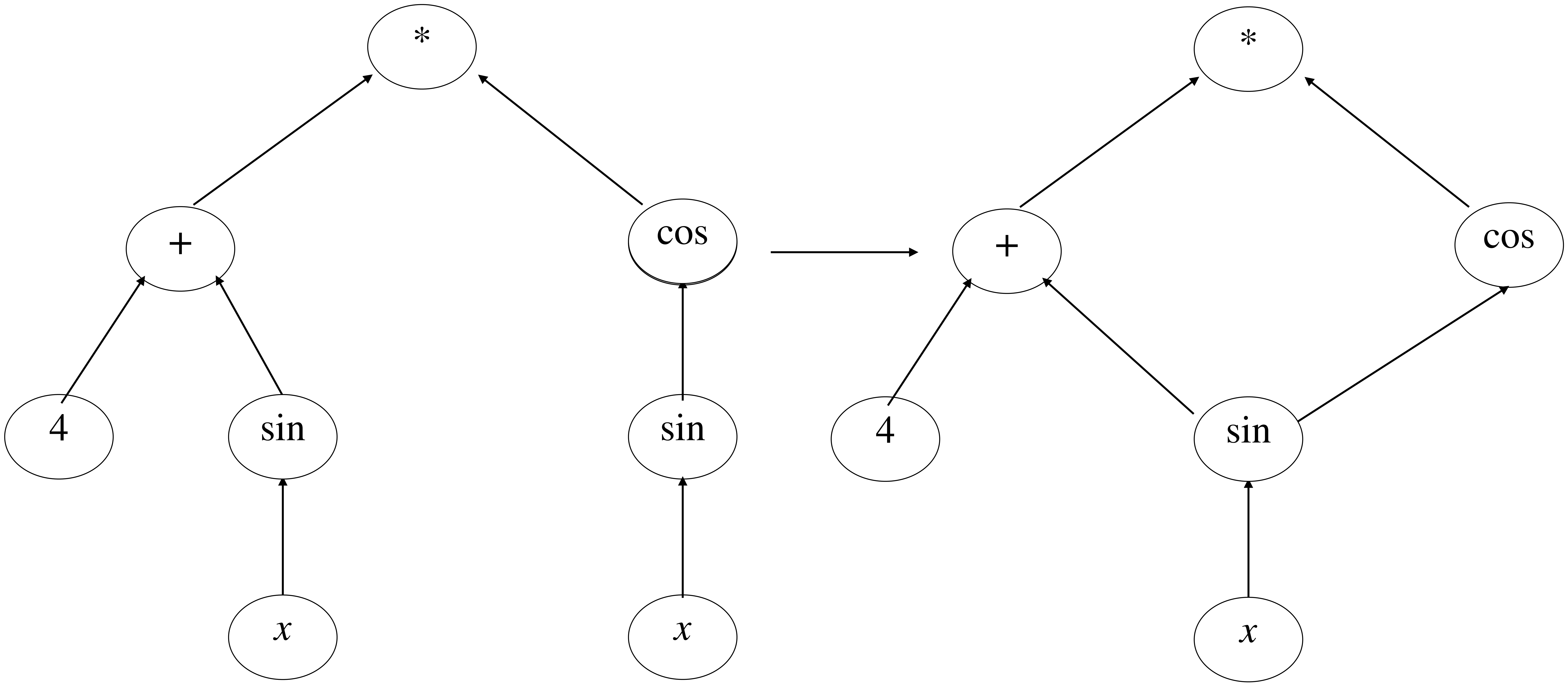
$$f(x) = (4 + \sin(x)) \cdot \cos(\sin(x))$$



式DAG(計算グラフ)

注) DAG := Directed Acyclic Graph
閉路がない有向グラフのこと

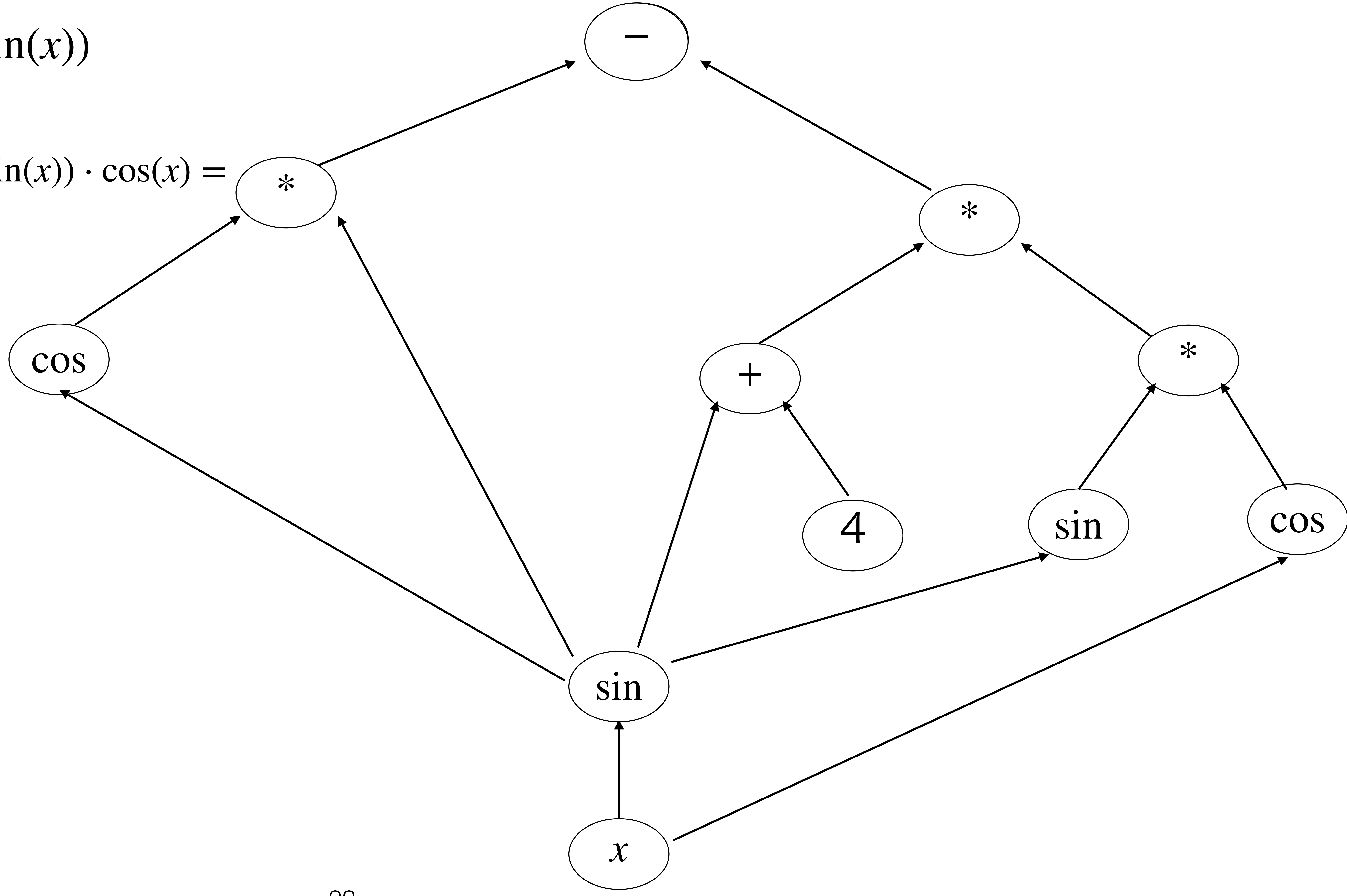
$$f(x) = (4 + \sin(x)) \cdot \cos(\sin(x))$$



式DAGで微分値を求める($x = \pi$)

$$\frac{df}{dx} = \left(\frac{d}{dx}\right)(4 + \sin(x)) \cdot \cos(\sin(x))$$

$$= \cos(x) \cdot \cos(\sin(x)) - (4 + \sin(x)) \cdot \sin(\sin(x)) \cdot \cos(x) =$$



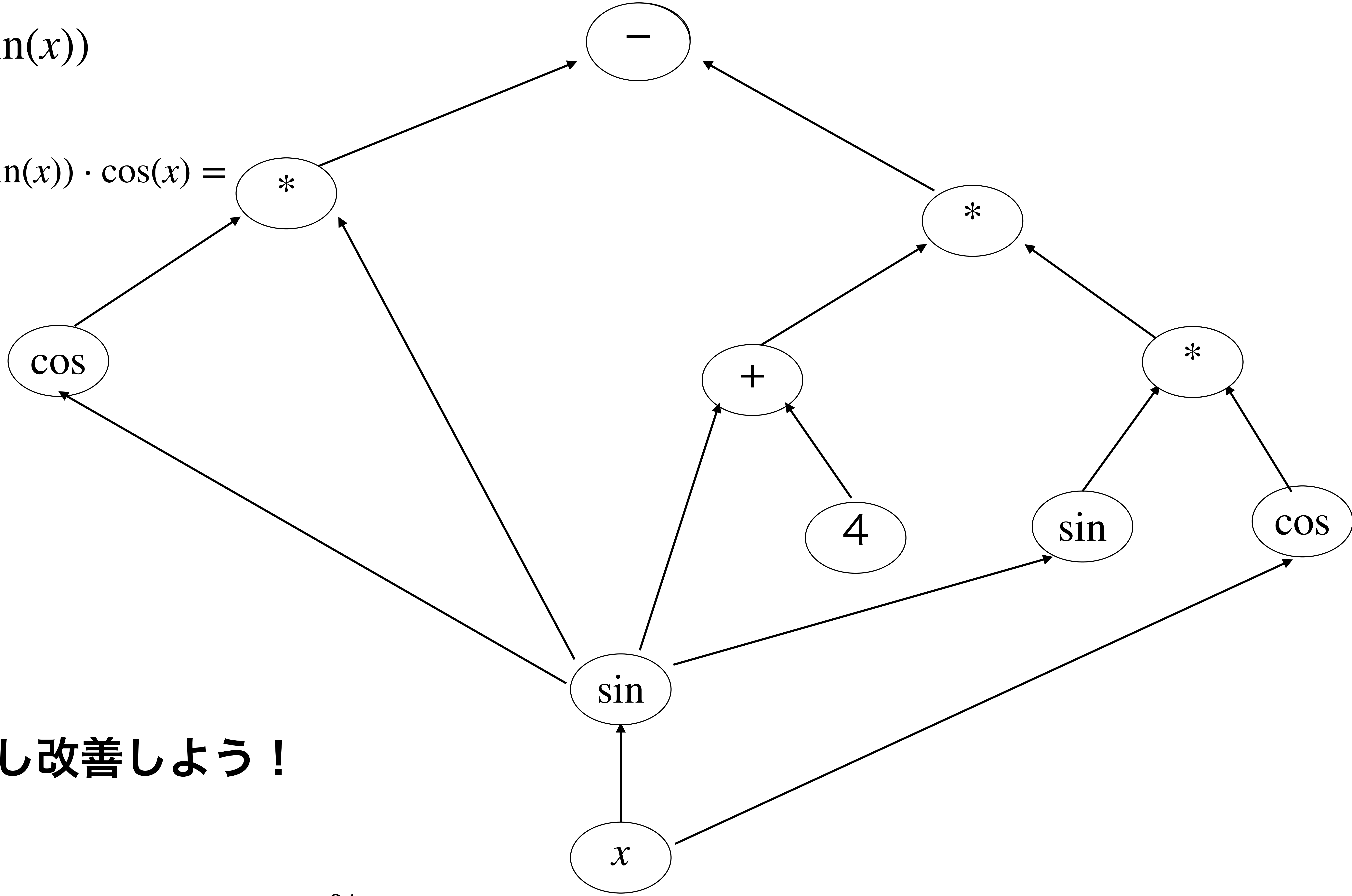
さっきと同じように微分規則で変形
メモ化して計算ができる

$\mathcal{O}(V)$ V :微分後のDAGの頂点数

式DAGで微分値を求める$(x = \pi)$

$$\frac{df}{dx} = \left(\frac{d}{dx}\right)(4 + \sin(x)) \cdot \cos(\sin(x))$$

$$= \cos(x) \cdot \cos(\sin(x)) - (4 + \sin(x)) \cdot \sin(\sin(x)) \cdot \cos(x) =$$



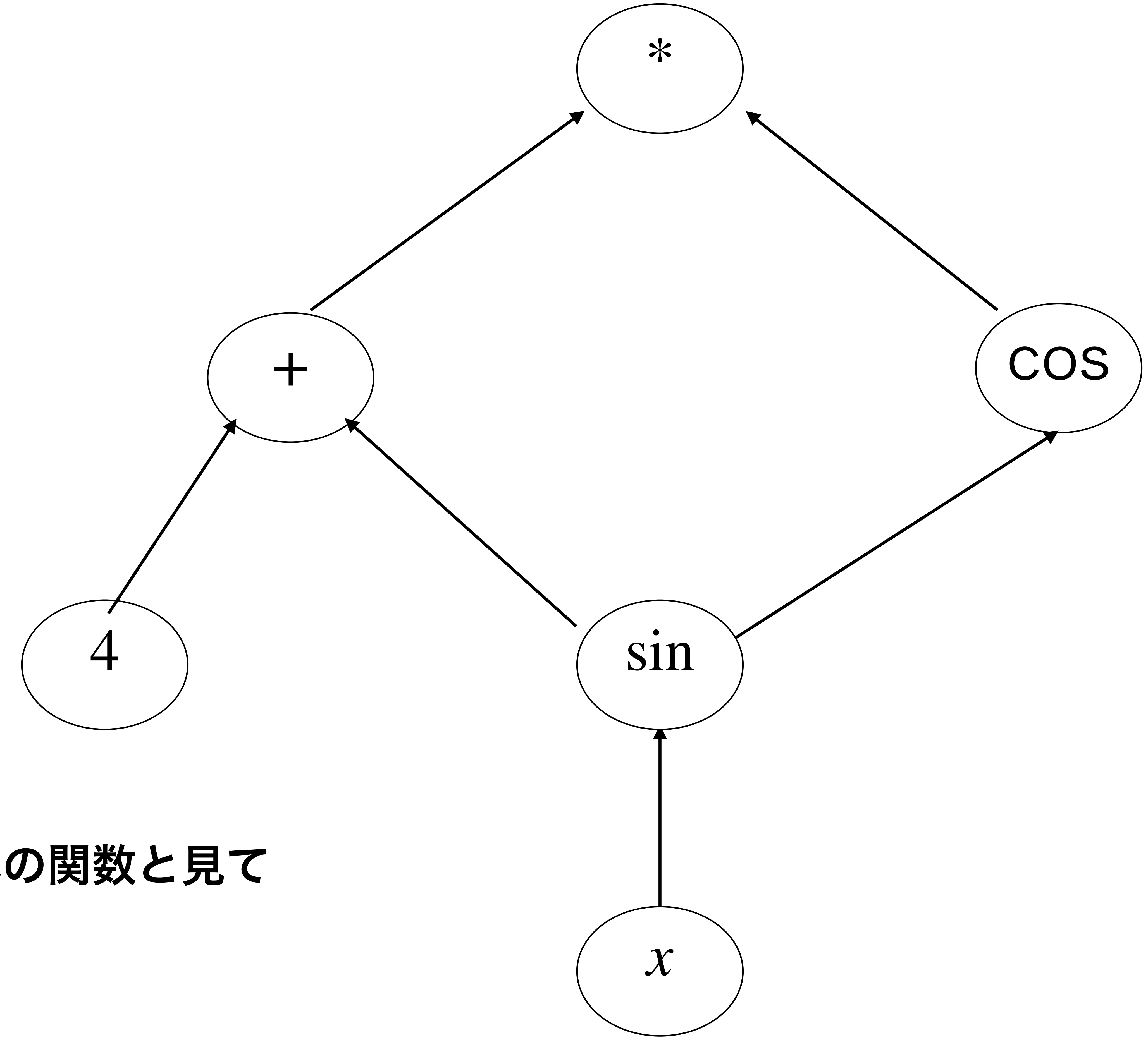
さっきと同じように微分規則で変形
メモ化して計算ができる

$\mathcal{O}(V)$　V :微分後のDAGの頂点数

もう少し改善しよう！

微分グラフ

$$f(x) = (4 + \sin(x)) \cdot \cos(\sin(x)) = a(x) \cdot b(x)$$



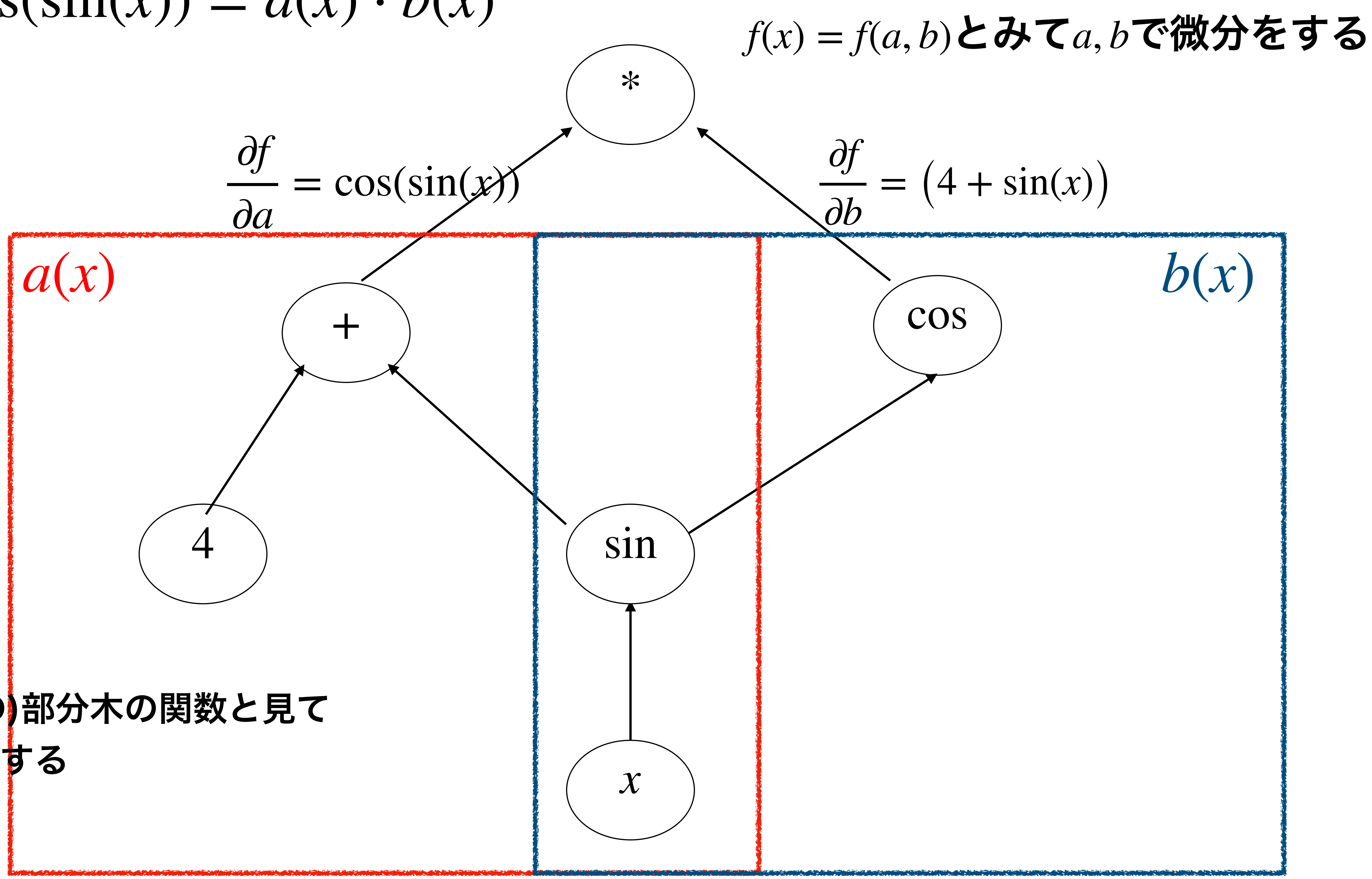
式DAGの各頂点を(元の式木の時の)部分木の関数と見て
各部分木で微分をする

微分グラフ

$$f(x) = (4 + \sin(x)) \cdot \cos(\sin(x)) = a(x) \cdot b(x)$$

$$a(x) = 4 + \sin(x)$$

$$b(x) = \cos(\sin(x))$$



式DAGの各頂点を(元の式木の時の)部分木の関数と見て
各部分木で微分をする

微分グラフ

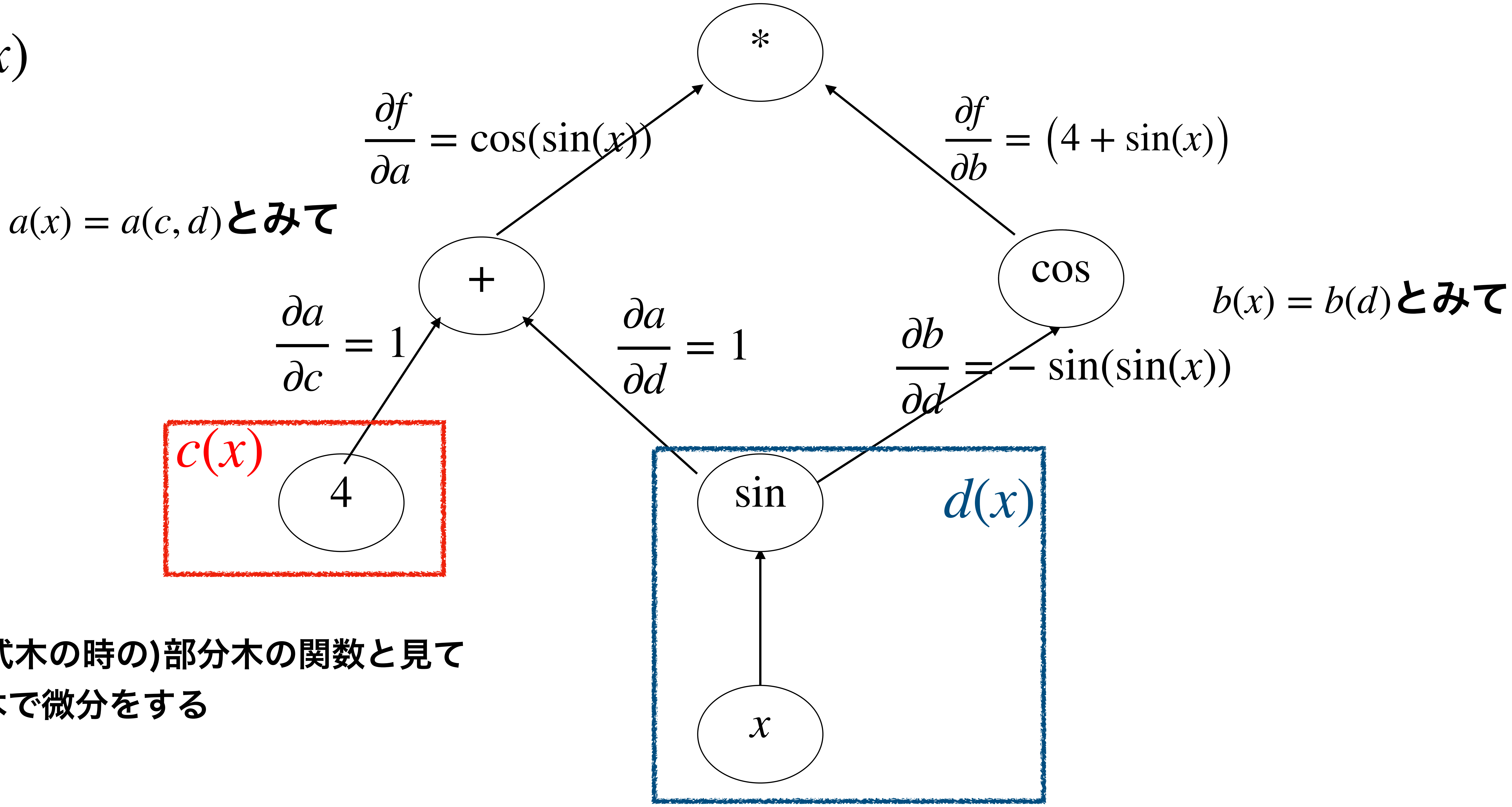
$$f(x) = (4 + \sin(x)) \cdot \cos(\sin(x)) = a(x) \cdot b(x)$$

$$a(x) = c(x) + d(x)$$

$$b(x) = \cos(d(x))$$

$$c(x) = 4$$

$$d(x) = \sin(x)$$



式DAGの各頂点を(元の式木の時の)部分木の関数と見て
各部分木で微分をする

微分グラフ

$$f(x) = (4 + \sin(x)) \cdot \cos(\sin(x)) = a(x) \cdot b(x)$$

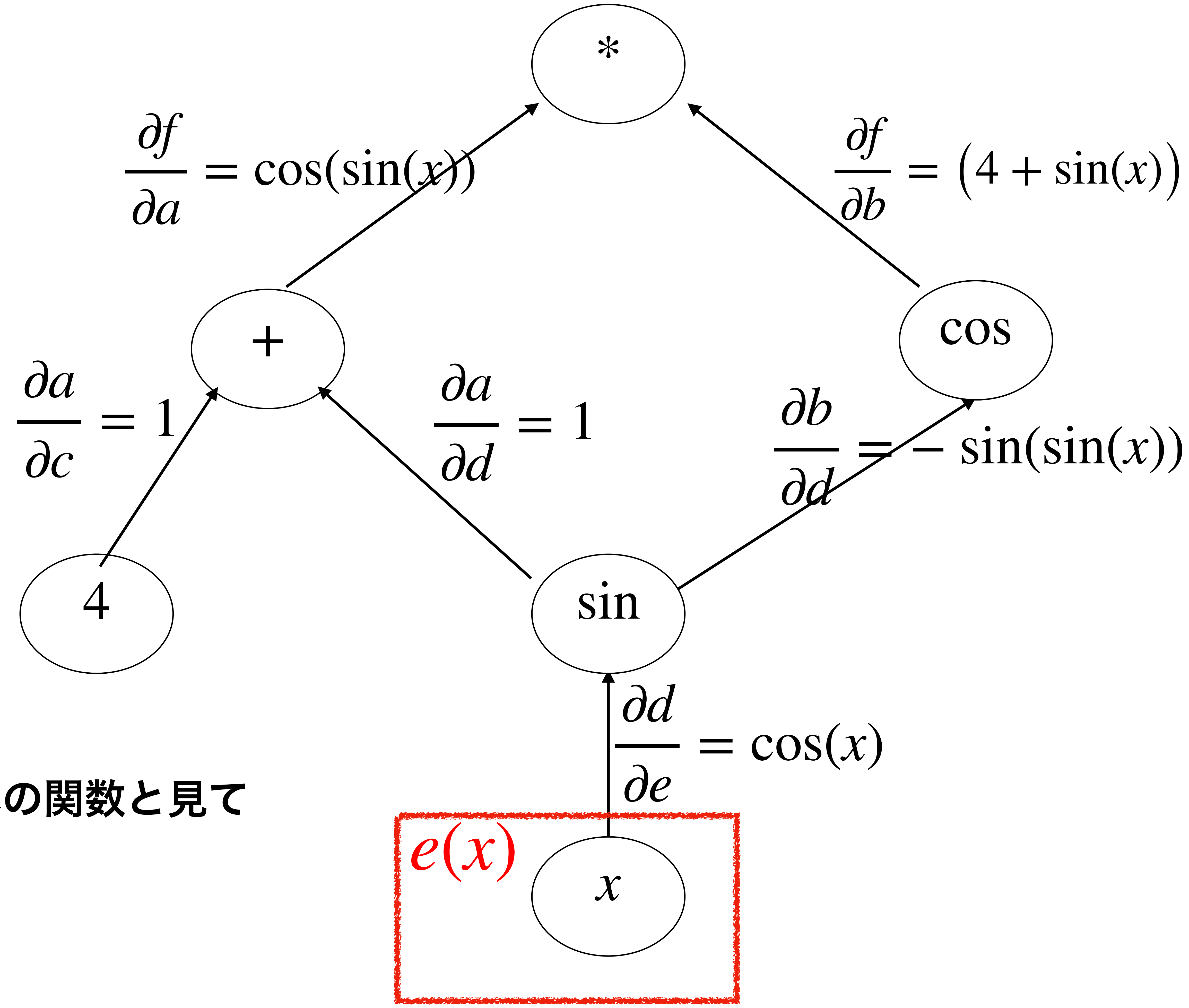
$$a(x) = c(x) + d(x)$$

$$b(x) = \cos(d(x))$$

$$c(x) = 4$$

$$d(x) = \sin(e(x))$$

$$e(x) = x$$

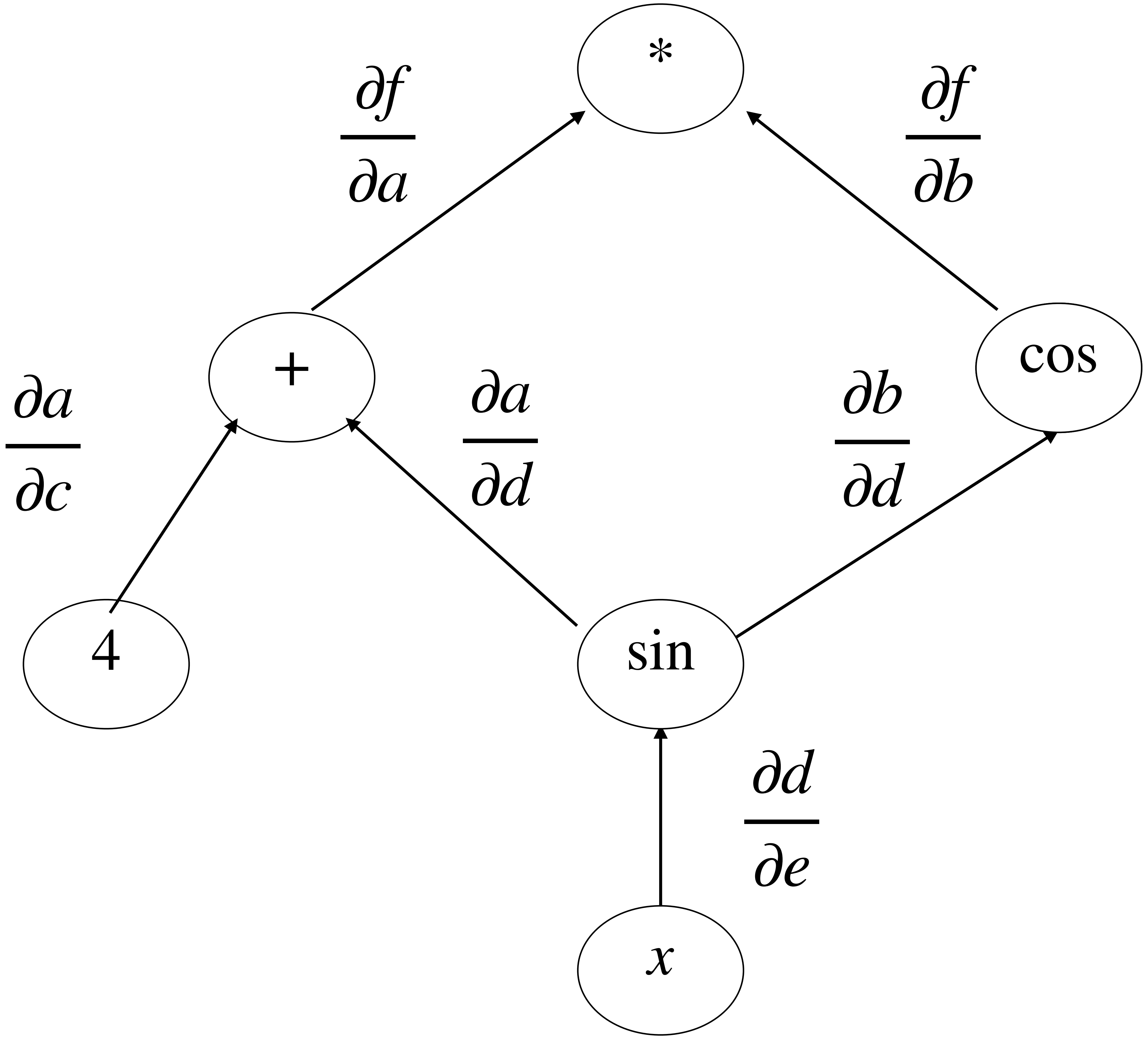


式DAGの各頂点を(元の式木の時の)部分木の関数と見て各部分木で微分をする

微分グラフ

連鎖律を使うと$\dfrac{\partial f}{\partial x}$を計算したい時

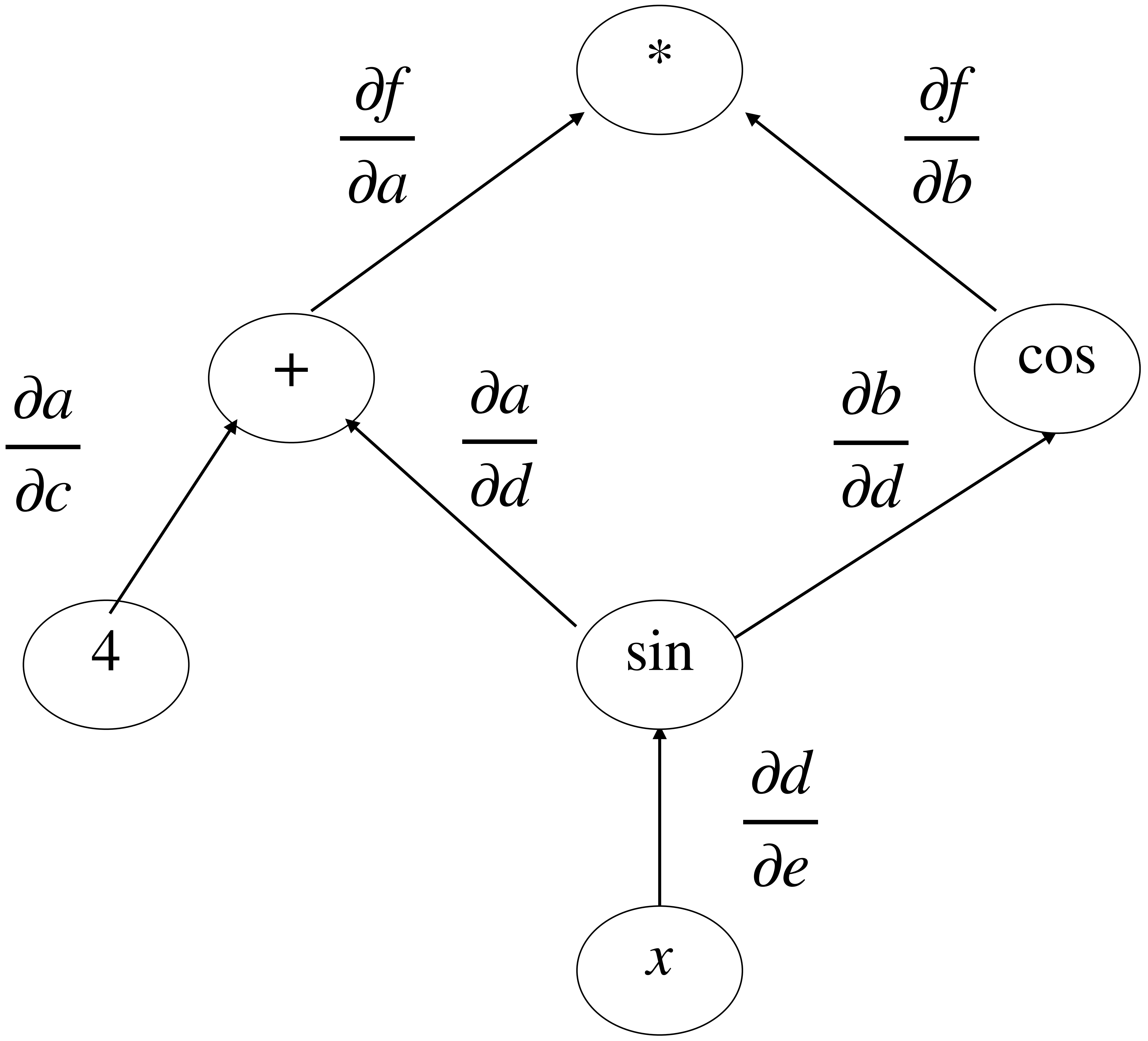
$$\frac{\partial f}{\partial x}$$



微分グラフで微分

連鎖律を使うと$\dfrac{\partial f}{\partial x}$を計算したい時

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial a}\frac{\partial a}{\partial x} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial x}$$

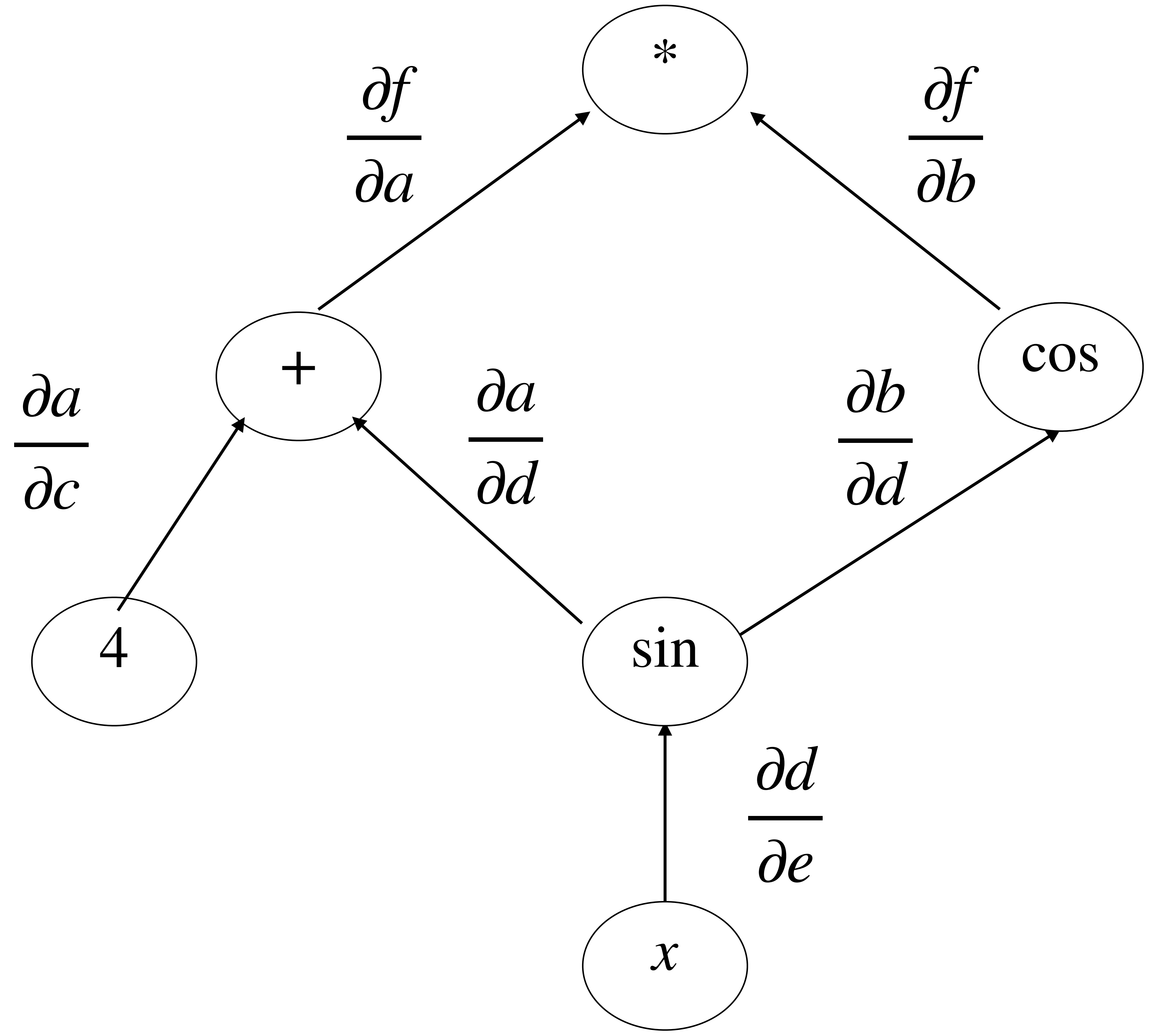


微分グラフで微分

連鎖律を使うと$\frac{\partial f}{\partial x}$を計算したい時

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial a}\frac{\partial a}{\partial x} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial x}$$

$$= \frac{\partial f}{\partial a}\left(\frac{\partial a}{\partial c}\frac{\partial c}{\partial x} + \frac{\partial a}{\partial d}\frac{\partial d}{\partial x}\right) + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\frac{\partial d}{\partial x}$$



微分グラフで微分

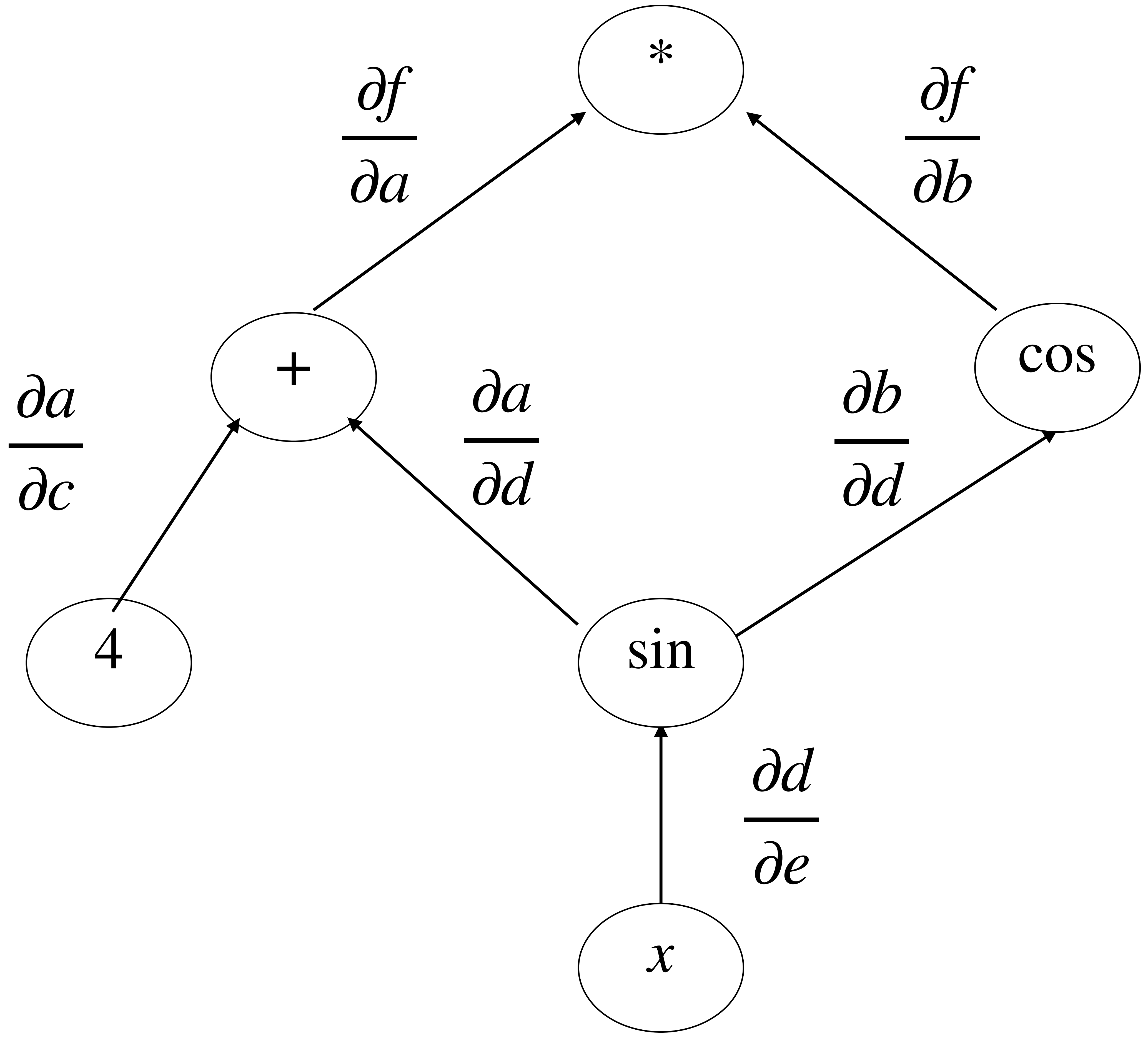
連鎖律を使うと$\frac{\partial f}{\partial x}$を計算したい時

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial a}\frac{\partial a}{\partial x} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial x}$$

$$= \frac{\partial f}{\partial a}\left(\frac{\partial a}{\partial c}\frac{\partial c}{\partial x} + \frac{\partial a}{\partial d}\frac{\partial d}{\partial x}\right) + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\frac{\partial d}{\partial x}$$

$$= \frac{\partial f}{\partial a}\left(\frac{\partial a}{\partial c}\frac{\partial c}{\partial x} + \frac{\partial a}{\partial d}\frac{\partial d}{\partial e}\frac{\partial e}{\partial x}\right) + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\frac{\partial d}{\partial e}\frac{\partial e}{\partial x}$$

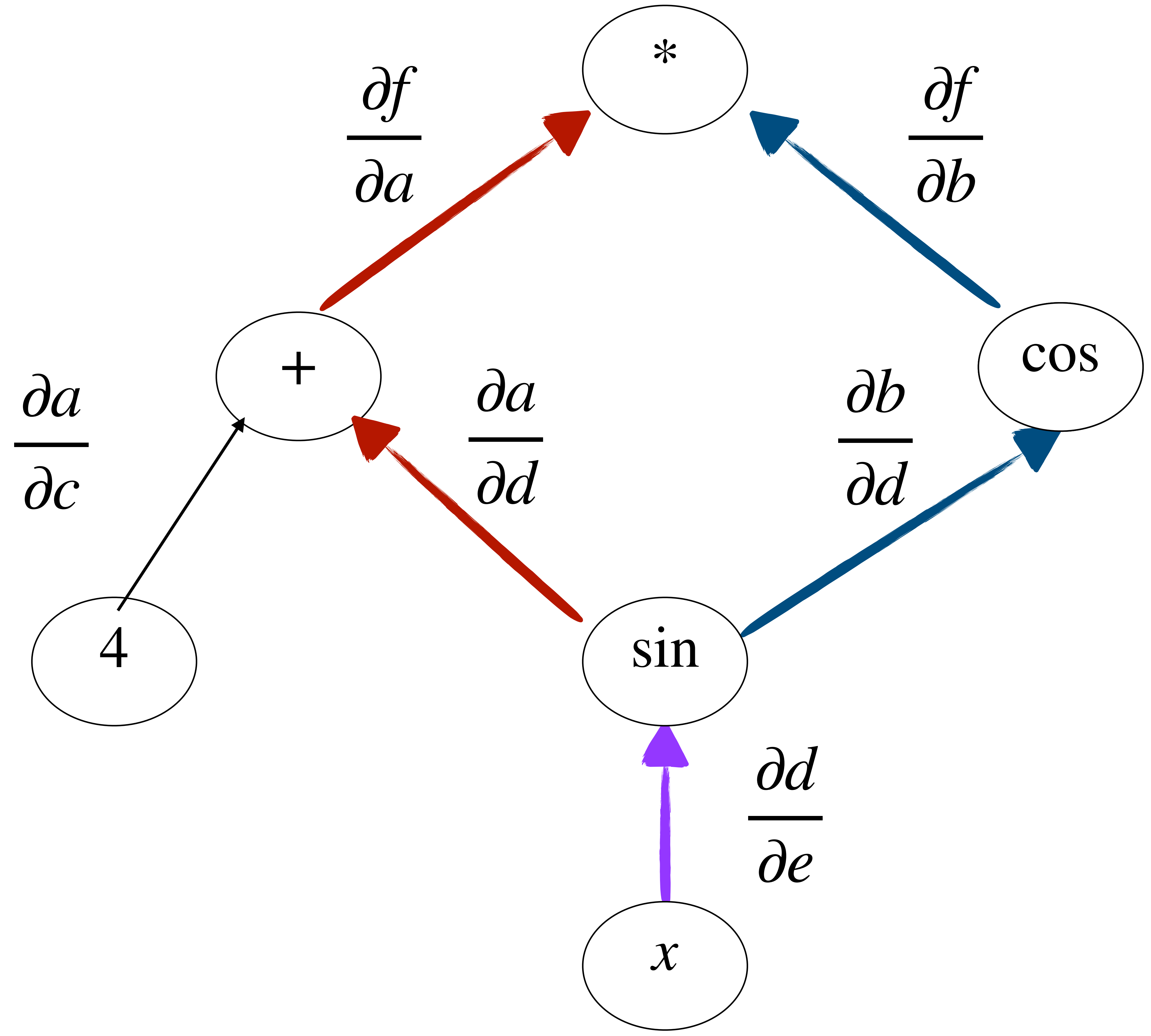
$$= \frac{\partial f}{\partial a}\frac{\partial a}{\partial d}\frac{\partial d}{\partial e} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\frac{\partial d}{\partial e} \qquad \left(\frac{\partial e}{\partial x} = 1, \frac{\partial c}{\partial x} = 0\right)$$



微分グラフで微分

連鎖律を使うと$\frac{\partial f}{\partial x}$を計算したい時

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial f}{\partial a}\frac{\partial a}{\partial x} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial x} \\
&= \frac{\partial f}{\partial a}\left(\frac{\partial a}{\partial c}\frac{\partial c}{\partial x} + \frac{\partial a}{\partial d}\frac{\partial d}{\partial x}\right) + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\frac{\partial d}{\partial x} \\
&= \frac{\partial f}{\partial a}\left(\frac{\partial a}{\partial c}\frac{\partial c}{\partial x} + \frac{\partial a}{\partial d}\frac{\partial d}{\partial e}\frac{\partial e}{\partial x}\right) + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\frac{\partial d}{\partial e}\frac{\partial e}{\partial x} \\
&= \underline{\frac{\partial f}{\partial a}\frac{\partial a}{\partial d}\frac{\partial d}{\partial e}} + \underline{\frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\frac{\partial d}{\partial e}} \qquad \left(\frac{\partial e}{\partial x} = 1, \frac{\partial c}{\partial x} = 0\right)
\end{aligned}
$$



xから, 根までの全てのパスの辺の積の和

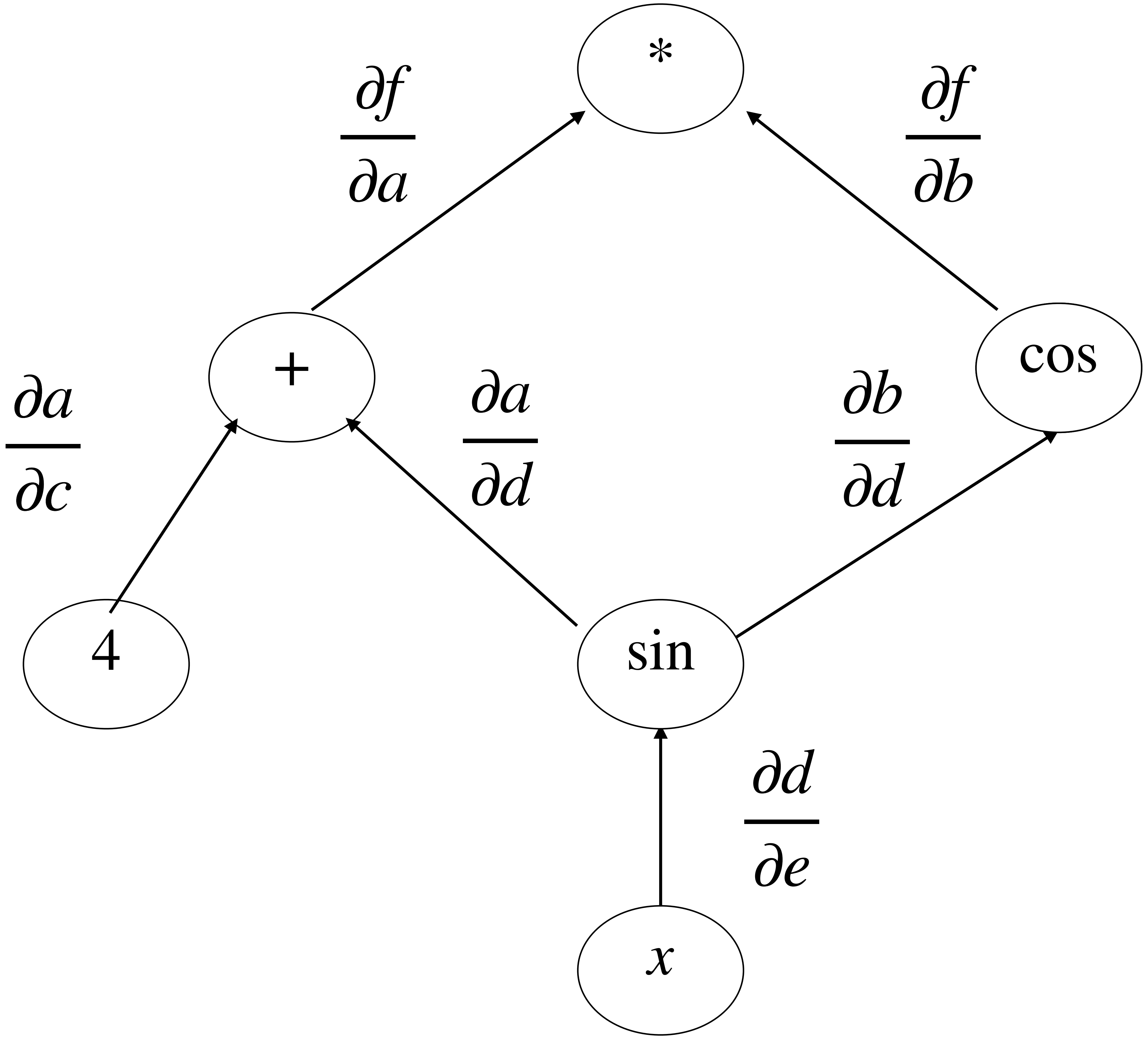
微分グラフで微分値を求める$(x = \pi)$

$$\frac{df}{dx} = \left(\frac{d}{dx}\right)(4 + \sin(x)) \cdot \cos(\sin(x))$$

葉から根まで,計算しながらパスを数えれば良い

$\mathcal{O}(V)$ $\quad V$ **:元のDAGの頂点数**

$$\frac{\partial d}{\partial e} = \cos(x)$$



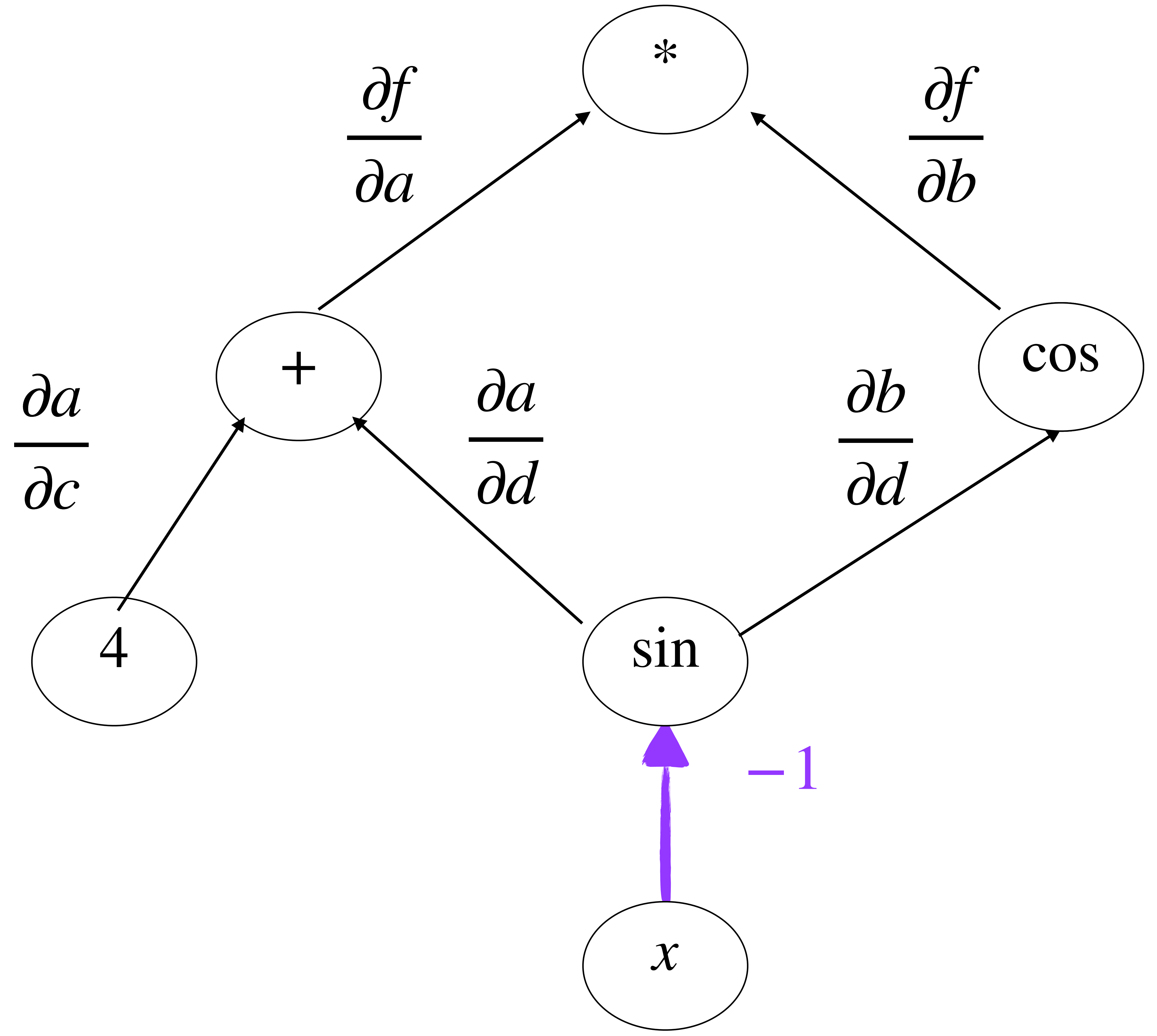
微分グラフで微分値を求める$(x = \pi)$

$$\frac{df}{dx} = \left(\frac{d}{dx}\right)(4 + \sin(x)) \cdot \cos(\sin(x))$$

葉から根まで,計算しながらパスを数えれば良い

$\mathcal{O}(V)$　V :元のDAGの頂点数

$$\frac{\partial d}{\partial e} = \cos(x) \qquad \frac{\partial d}{\partial e}(\pi) = -1$$



微分グラフで微分値を求める$(x = \pi)$

$$\frac{df}{dx} = \left(\frac{d}{dx}\right)(4 + \sin(x)) \cdot \cos(\sin(x))$$

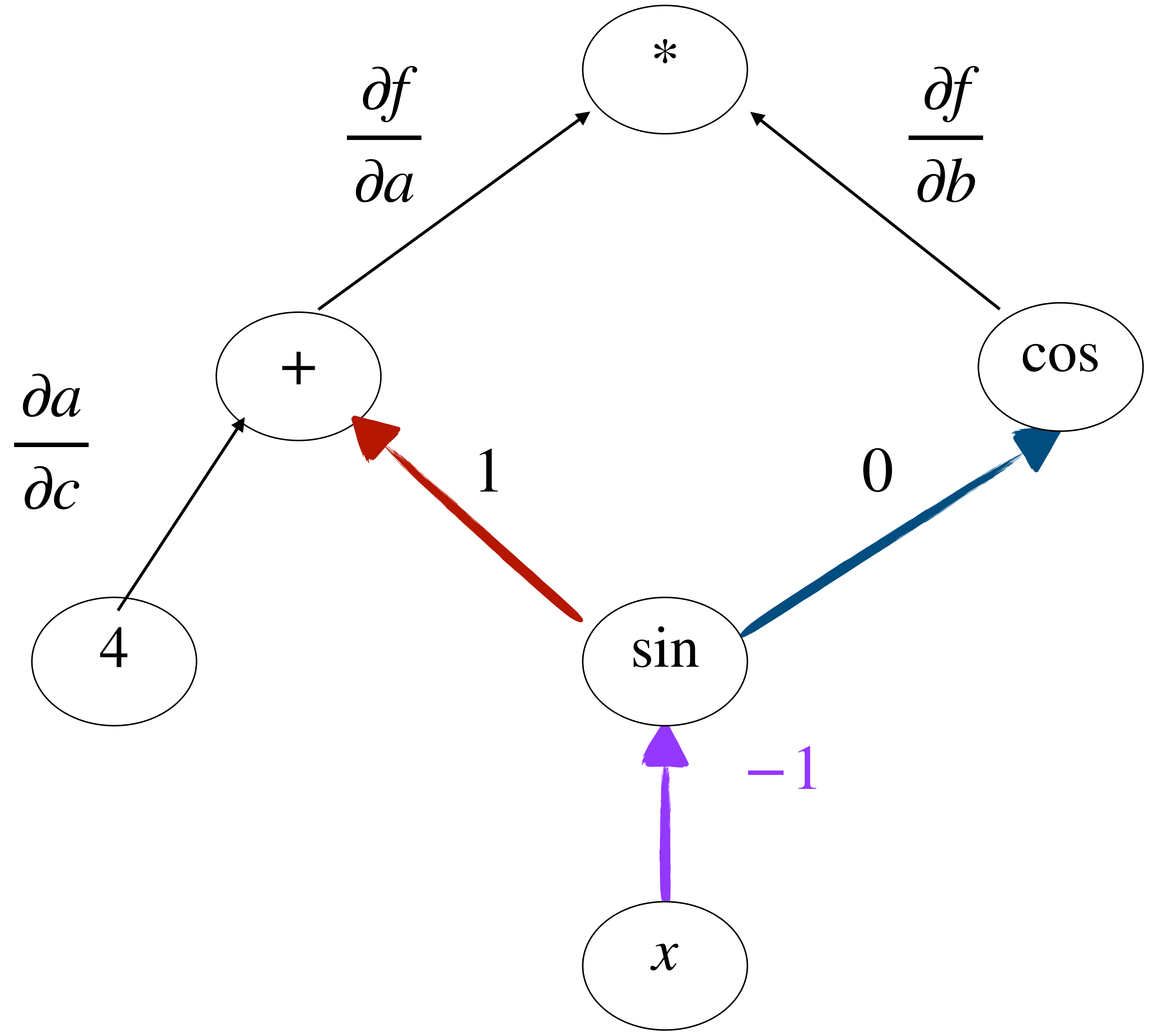
葉から根まで,計算しながらパスを数えれば良い

$\mathcal{O}(V)$　V :元のDAGの頂点数

$$\frac{\partial d}{\partial e} = \cos(x) \qquad \frac{\partial d}{\partial e}(\pi) = -1$$

$$\frac{\partial a}{\partial d} = 1 \qquad \frac{\partial a}{\partial d}(\pi) = 1$$

$$\frac{\partial b}{\partial d} = -\sin(\sin(x)) \longrightarrow \frac{\partial b}{\partial d}(\pi) = 0$$



微分グラフで微分値を求める$(x=\pi)$

$$\frac{df}{dx} = \left(\frac{d}{dx}\right)(4+\sin(x))\cdot\cos(\sin(x)) \;=\; -1 + 0 = -1$$

葉から根まで,計算しながらパスを数えれば良い

$\mathcal{O}(V)$　V **:元のDAGの頂点数**

$\frac{\partial d}{\partial e} = \cos(x)$　　$\frac{\partial d}{\partial e}(\pi) = -1$

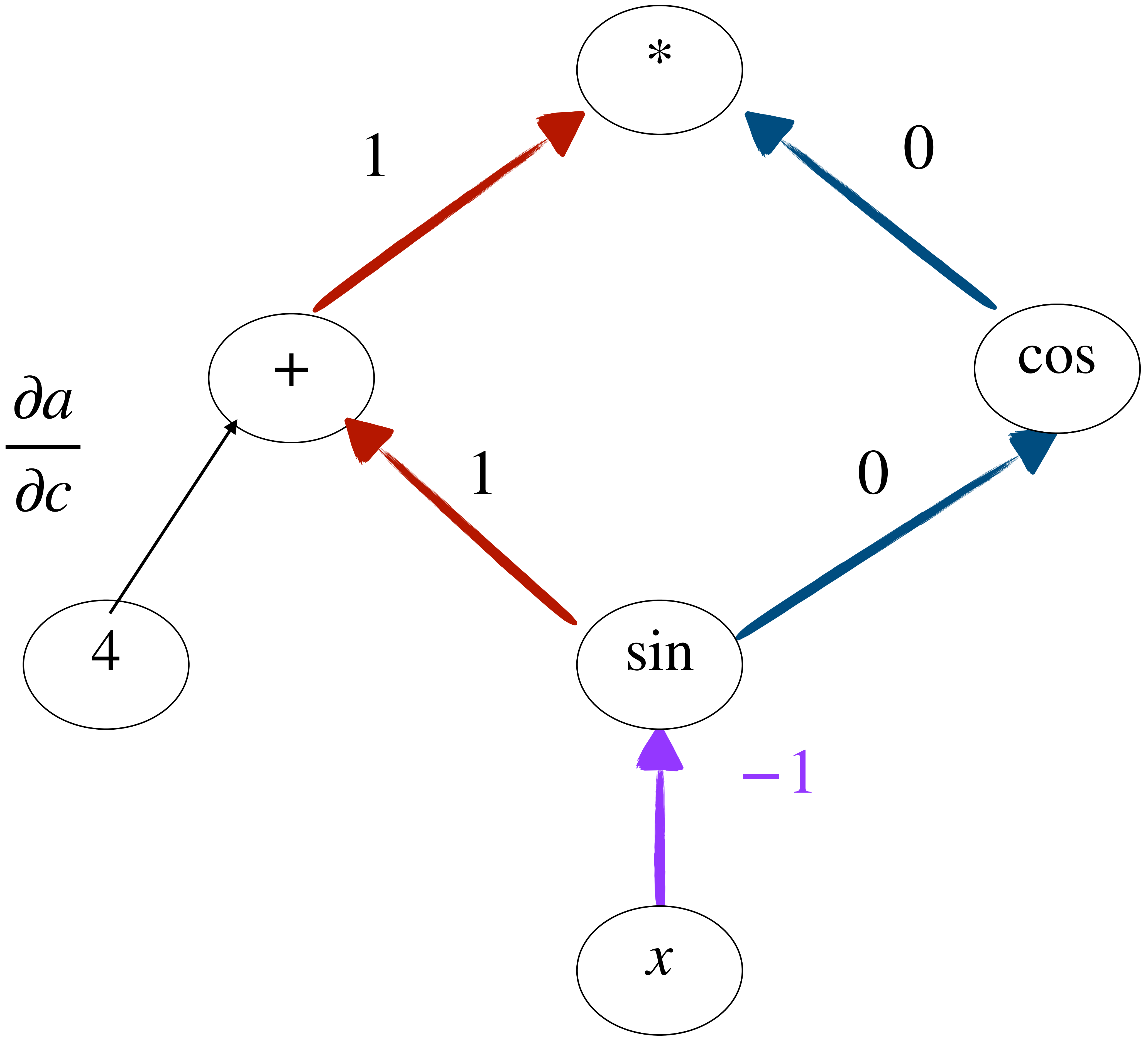
$\frac{\partial a}{\partial d} = 1$　　$\frac{\partial a}{\partial d}(\pi) = 1$

$\frac{\partial b}{\partial d} = -\sin(\sin(x))$ → $\frac{\partial b}{\partial d}(\pi) = 0$

$\frac{\partial f}{\partial a} = \cos(\sin(x))$　　$\frac{\partial f}{\partial a}(\pi) = 1$

$\frac{\partial f}{\partial b} = (4+\sin(x))$　　$\frac{\partial f}{\partial b}(\pi) = 0$

$\frac{\partial a}{\partial c}$



微分グラフで微分値を求める$(x = \pi)$

$$\frac{df}{dx} = \left(\frac{d}{dx}\right)(4 + \sin(x)) \cdot \cos(\sin(x)) \ = \ -1 \ + \ 0 \ = -1$$

葉から根まで,計算しながらパスを数えれば良い

$\mathcal{O}(V)$ 　 V :元のDAGの頂点数

$$\frac{\partial d}{\partial e} = \cos(x) \qquad \frac{\partial d}{\partial e}(\pi) = -1$$

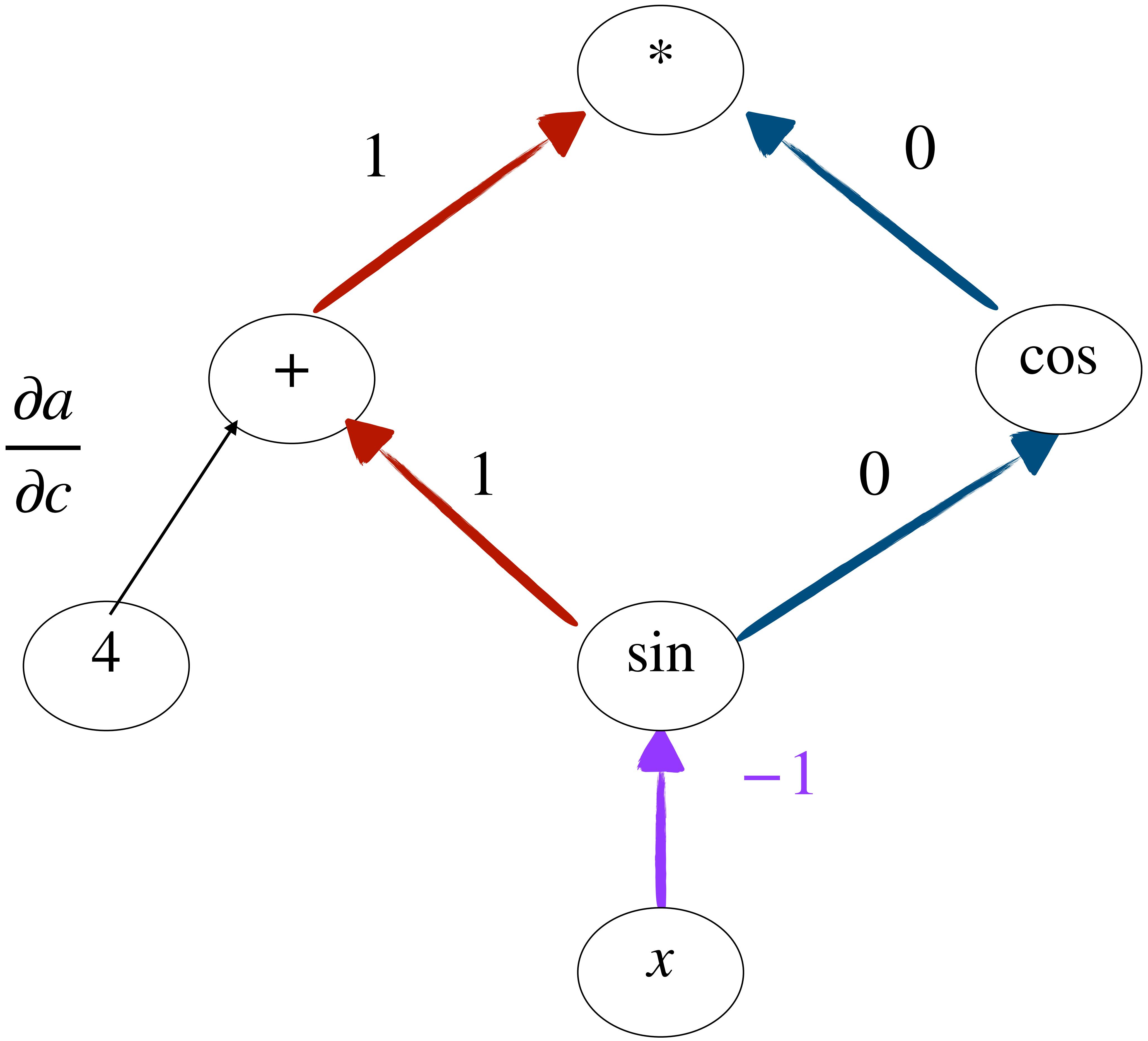
$$\frac{\partial a}{\partial d} = 1 \qquad \frac{\partial a}{\partial d}(\pi) = 1$$

$$\frac{\partial b}{\partial d} = -\sin(\sin(x)) \longrightarrow \frac{\partial b}{\partial d}(\pi) = 0$$

$$\frac{\partial f}{\partial a} = \cos(\sin(x)) \qquad \frac{\partial f}{\partial a}(\pi) = 1$$

$$\frac{\partial f}{\partial b} = \left(4 + \sin(x)\right) \qquad \frac{\partial f}{\partial b}(\pi) = 0$$

$$\frac{\partial a}{\partial c}$$

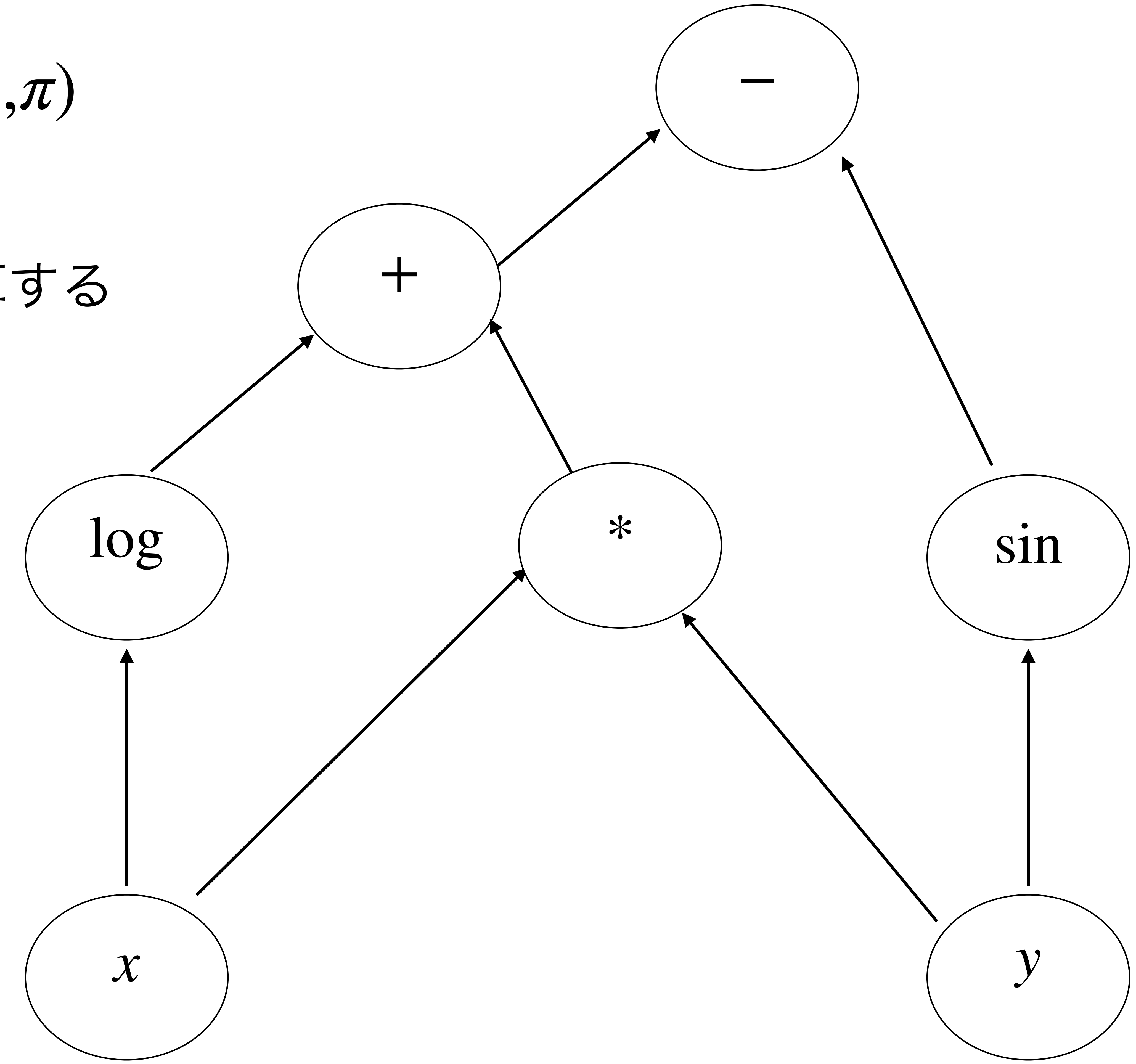


式の構造があれば式自体はいらなくて,
辺が値だけ持てればよさそう

自動微分(計算グラフ上の)

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$(x, y) = (1, \pi)$



1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 各辺の微分値を使って欲しい微分値を計算する

Forwardモード, Reverseモードがあり

$f: \mathbb{R}^n \to \mathbb{R}^m$ (n : 入力の次元, m : 出力の次元)として

$n << m$ならForwardモード

$n >> m$ならReverseモード

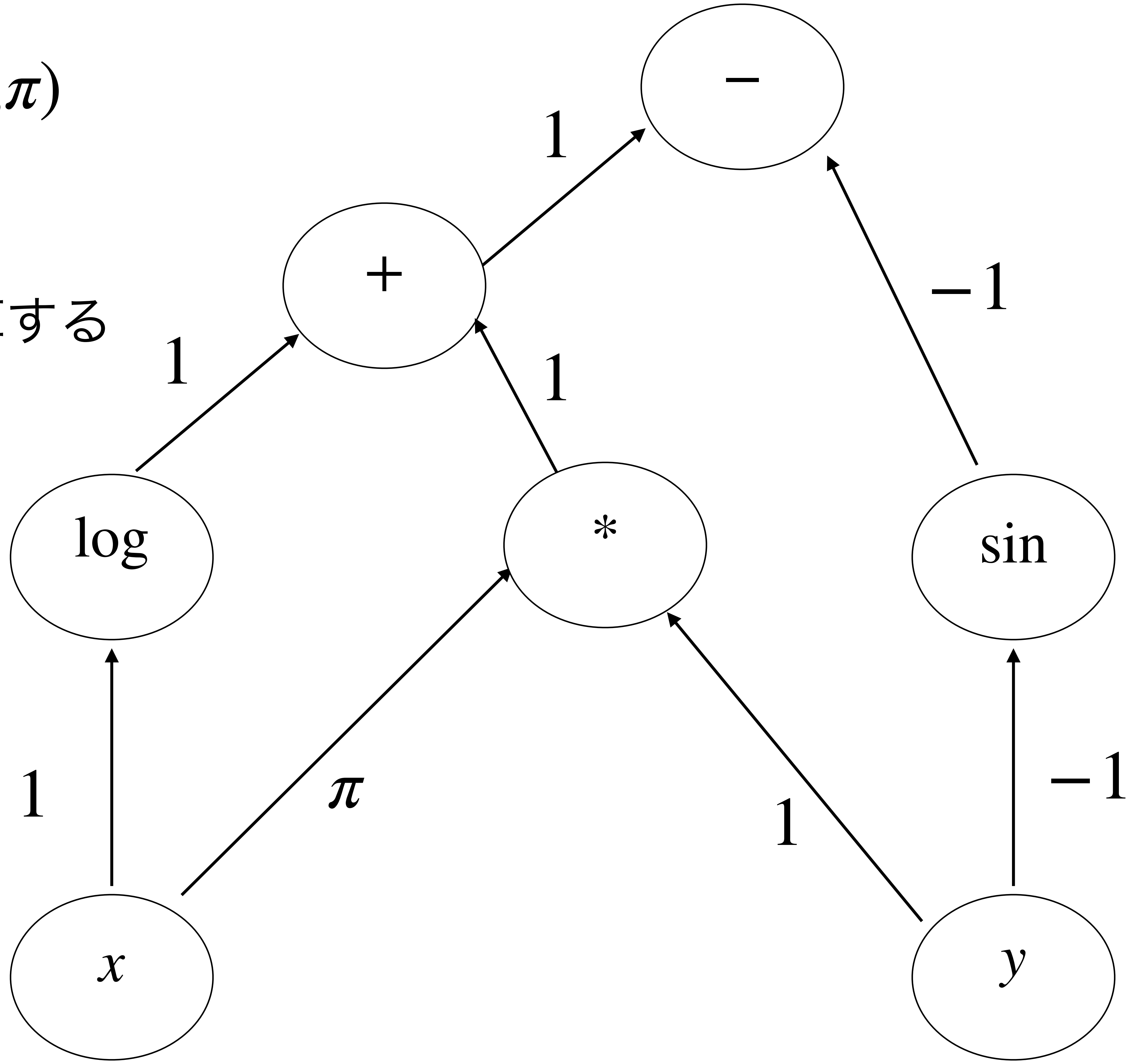
がよいと言われる

Forwardモードの自動微分

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

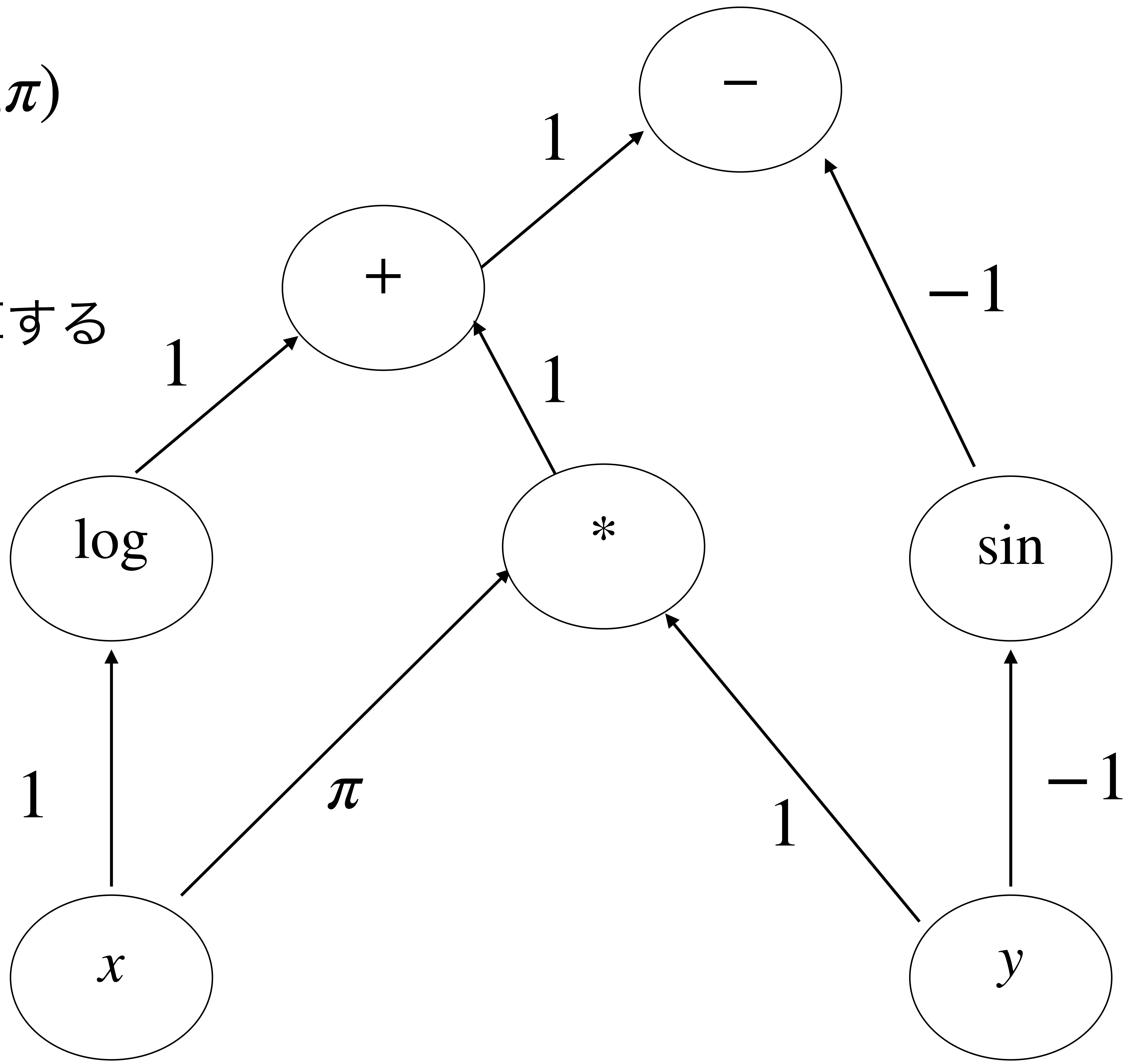
1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する



Forwardモードの自動微分

$$f(x,y) = \log(x) + x \cdot y - \sin(y)$$

$$(x,y) = (1,\pi)$$



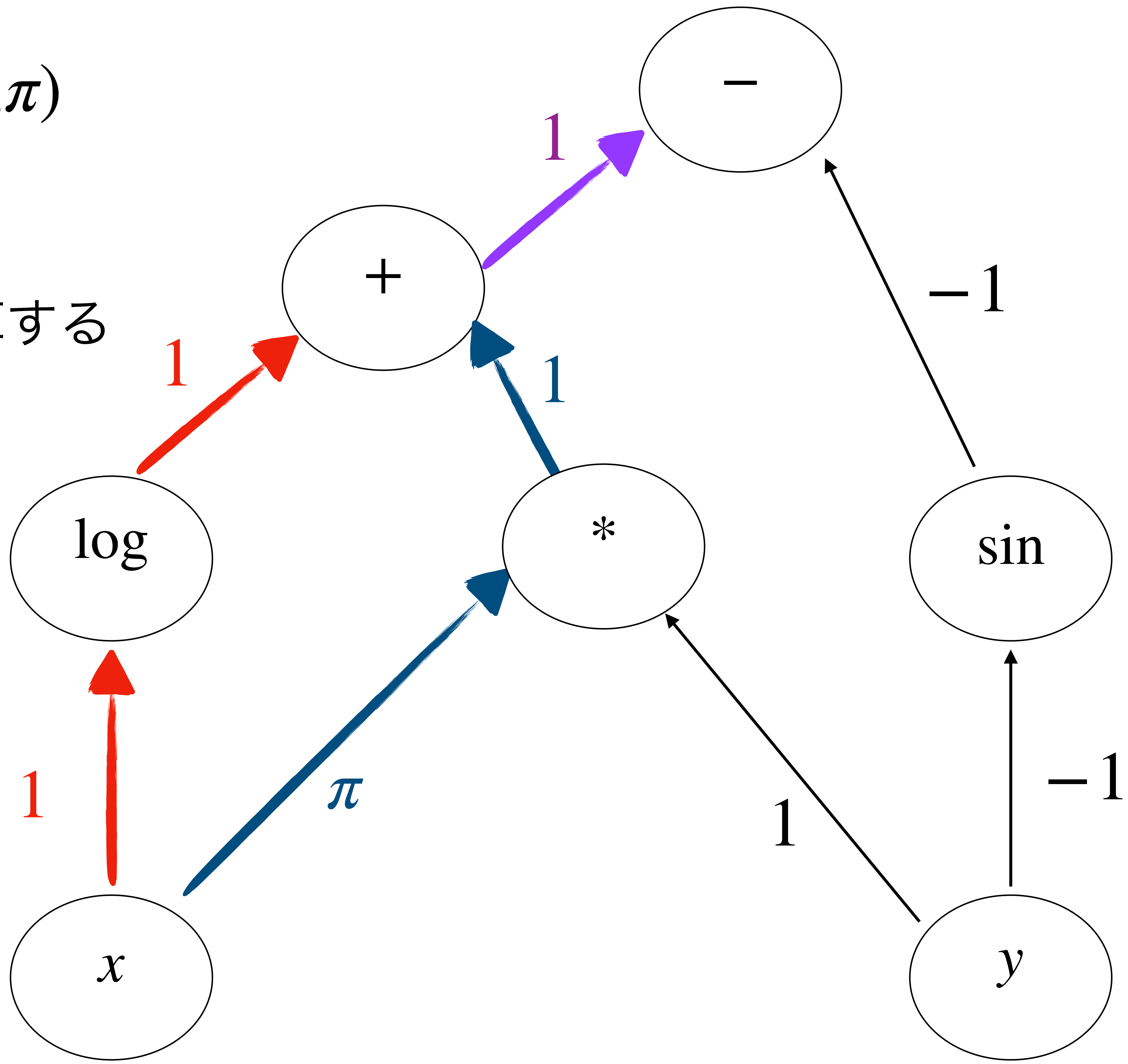
1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**

Forwardモードの自動微分

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$(x, y) = (1, \pi)$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**



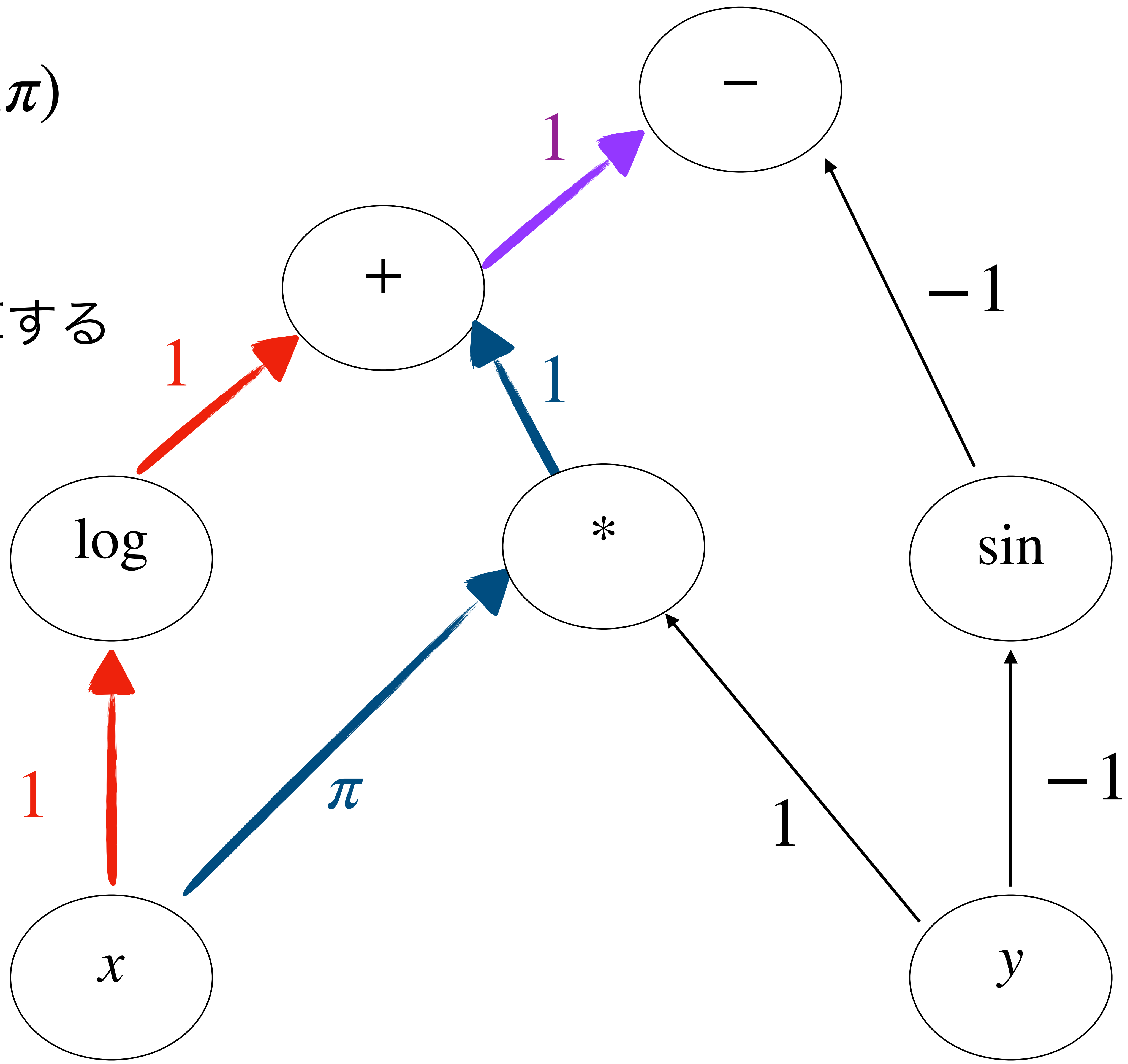
Forwardモードの自動微分

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**

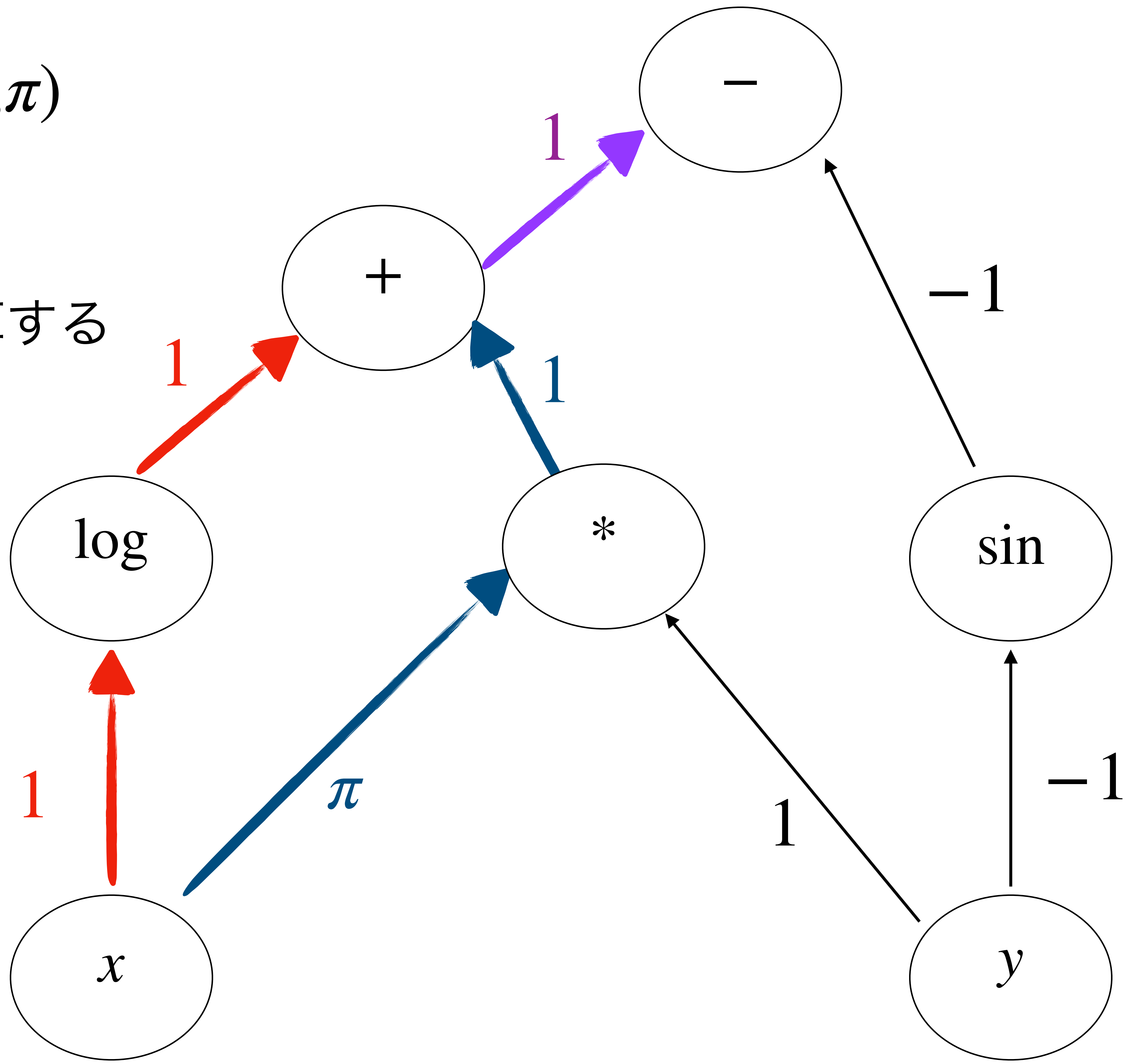
$$\frac{\partial f}{\partial x}(1, \pi) = 1 + \pi$$



Forwardモードの自動微分

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$



1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = \underline{1} + \underline{\pi} \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

Forwardモードの自動微分

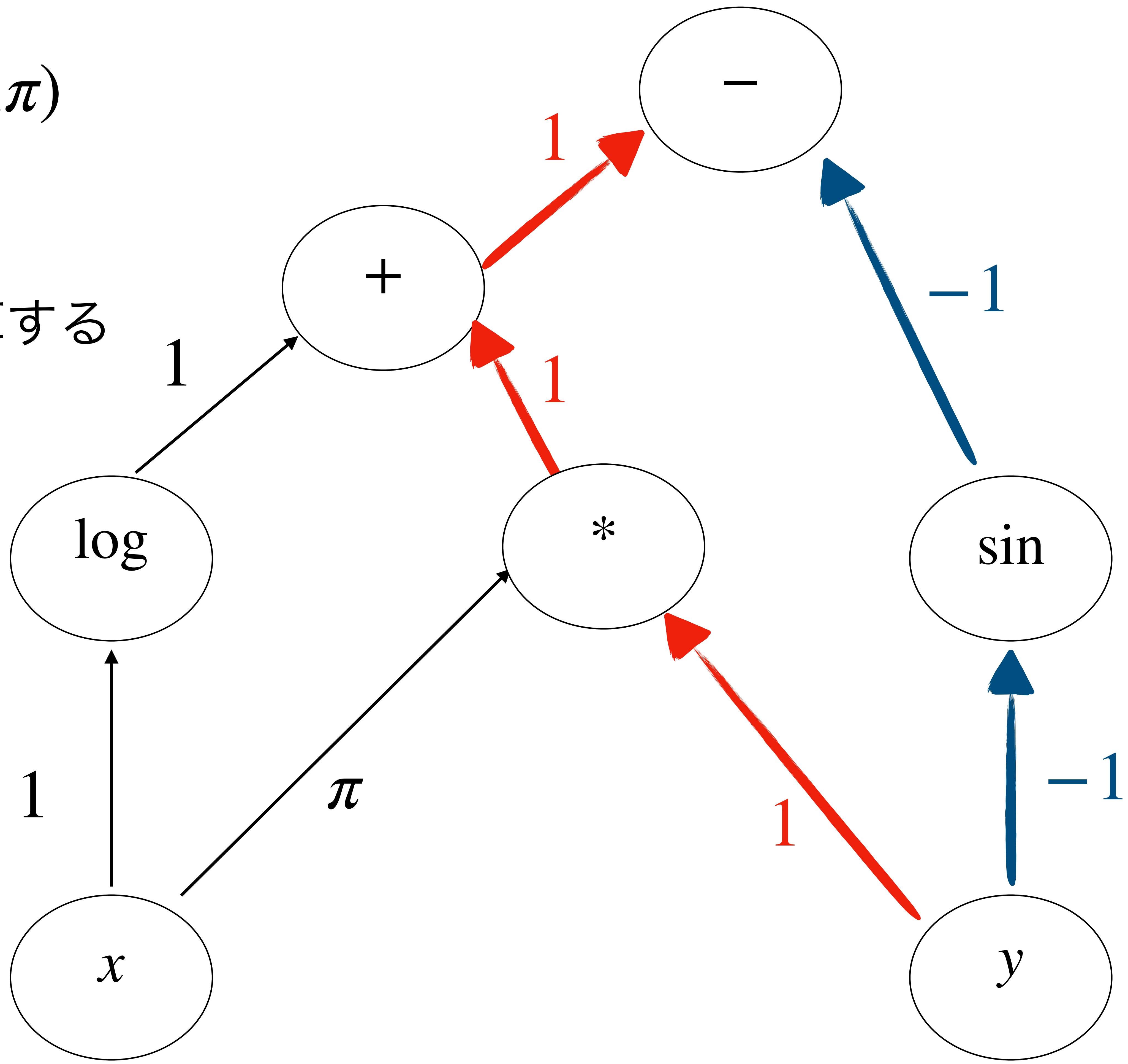
$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = 1 + \pi \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

$$\frac{\partial f}{\partial y}(1, \pi) = \underline{1} + \underline{1} = 2$$



Forwardモードの自動微分

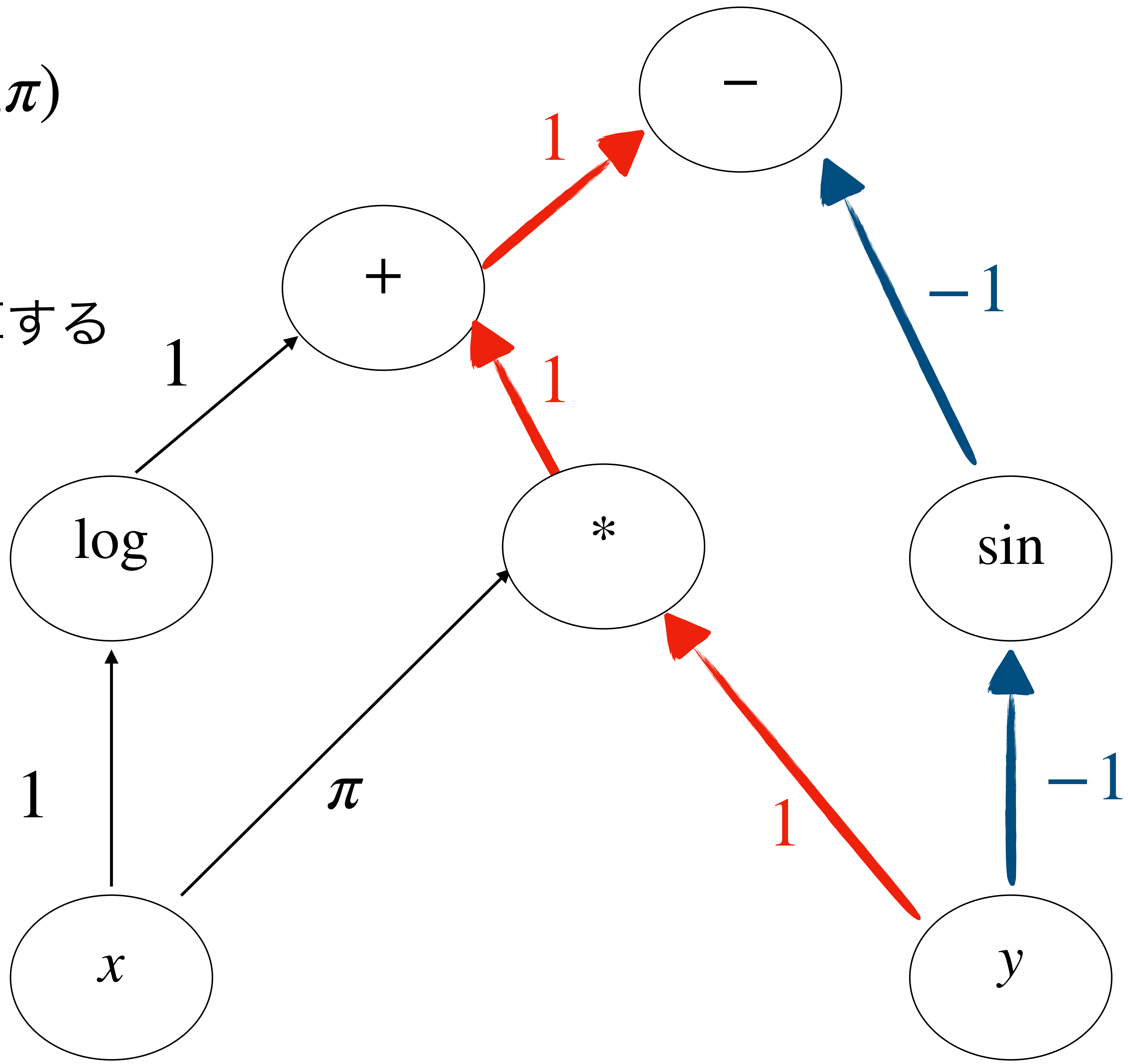
$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = 1 + \pi \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

$$\frac{\partial f}{\partial y}(1, \pi) = \underline{1} + \underline{1} = 2 \quad \Leftarrow \quad \frac{\partial f}{\partial y} = x - \cos(y)$$



Forwardモードの自動微分

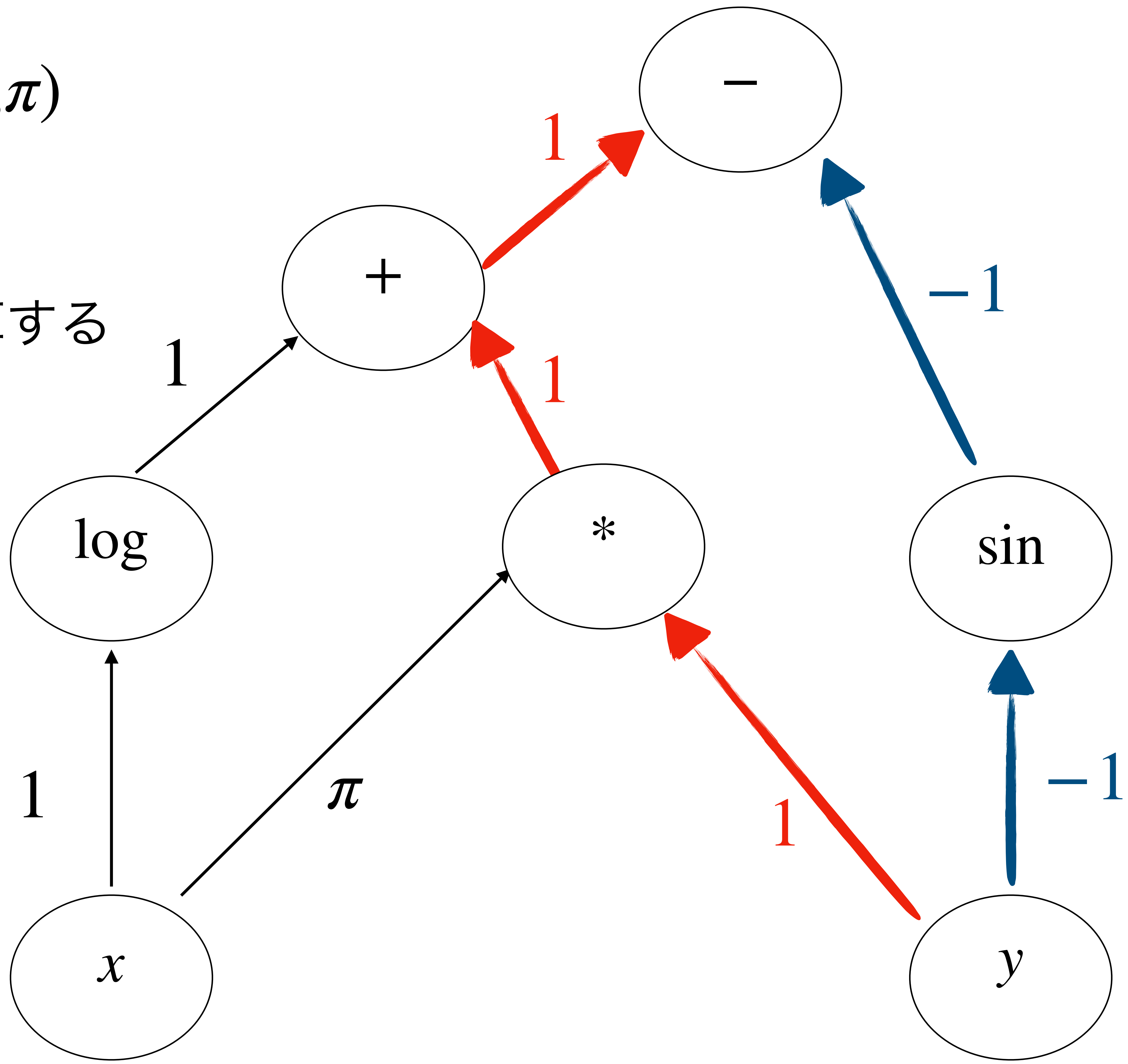
$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = 1 + \pi \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

$$\frac{\partial f}{\partial y}(1, \pi) = \underline{1} + \underline{1} = 2 \quad \Leftarrow \quad \frac{\partial f}{\partial y} = x - \cos(y)$$



勾配を求めるには, 変数の数だけ集める必要がある！

Forwardモードの自動微分

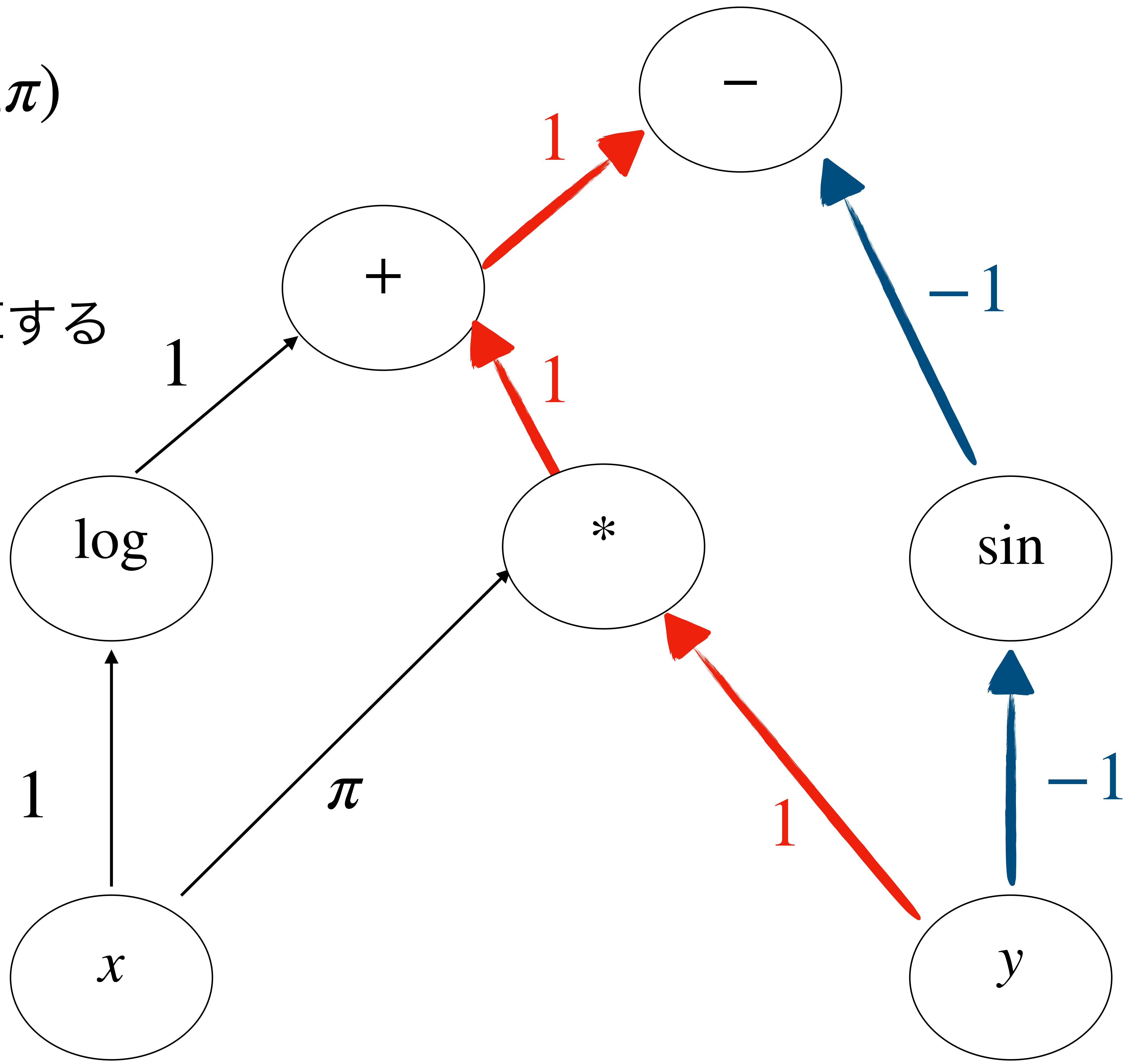
$$f(x,y) = \log(x) + x \cdot y - \sin(y)$$

$$(x,y) = (1,\pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**葉(変数頂点)から集める**

$$\frac{\partial f}{\partial x}(1,\pi) = 1 + \pi \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

$$\frac{\partial f}{\partial y}(1,\pi) = \underline{1} + \underline{1} = 2 \quad \Leftarrow \quad \frac{\partial f}{\partial y} = x - \cos(y)$$



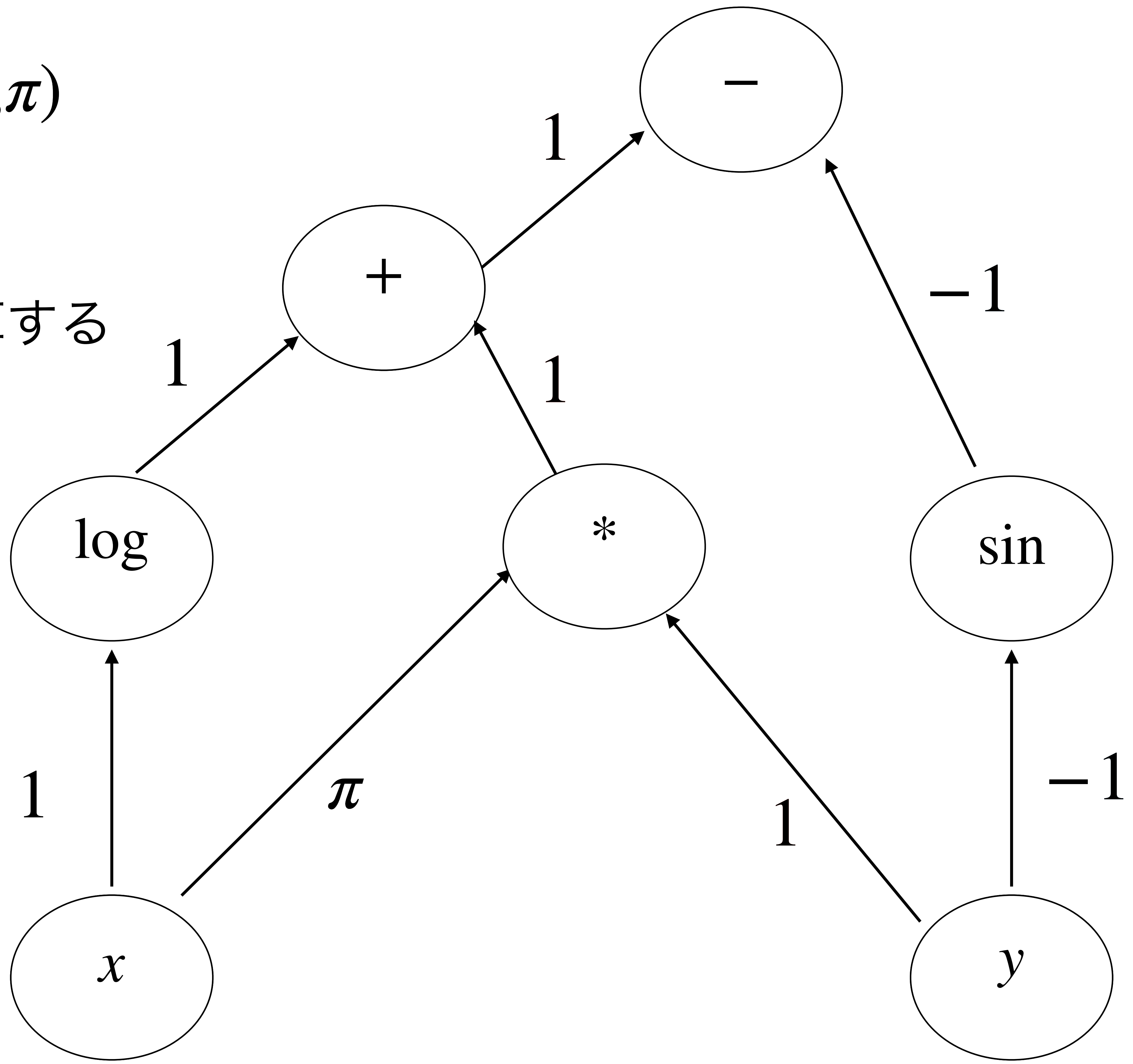
勾配を求めるには, 変数の数だけ集める必要がある！
→逆から辿った方が勾配を求めるにはよさそう

Reverseモードの自動微分

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$(x, y) = (1, \pi)$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**根から集める**

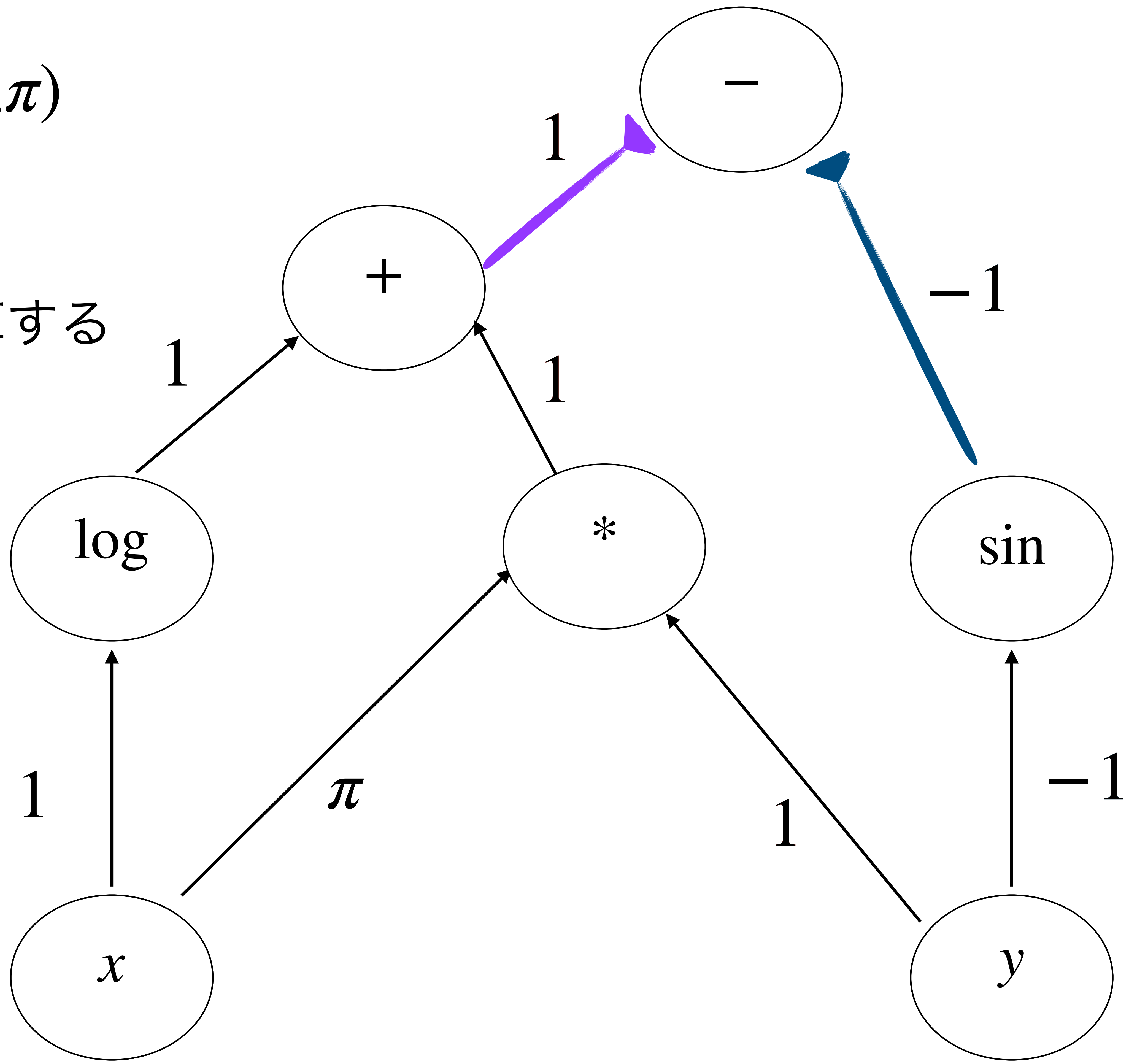


Reverseモードの自動微分

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**根から集める**

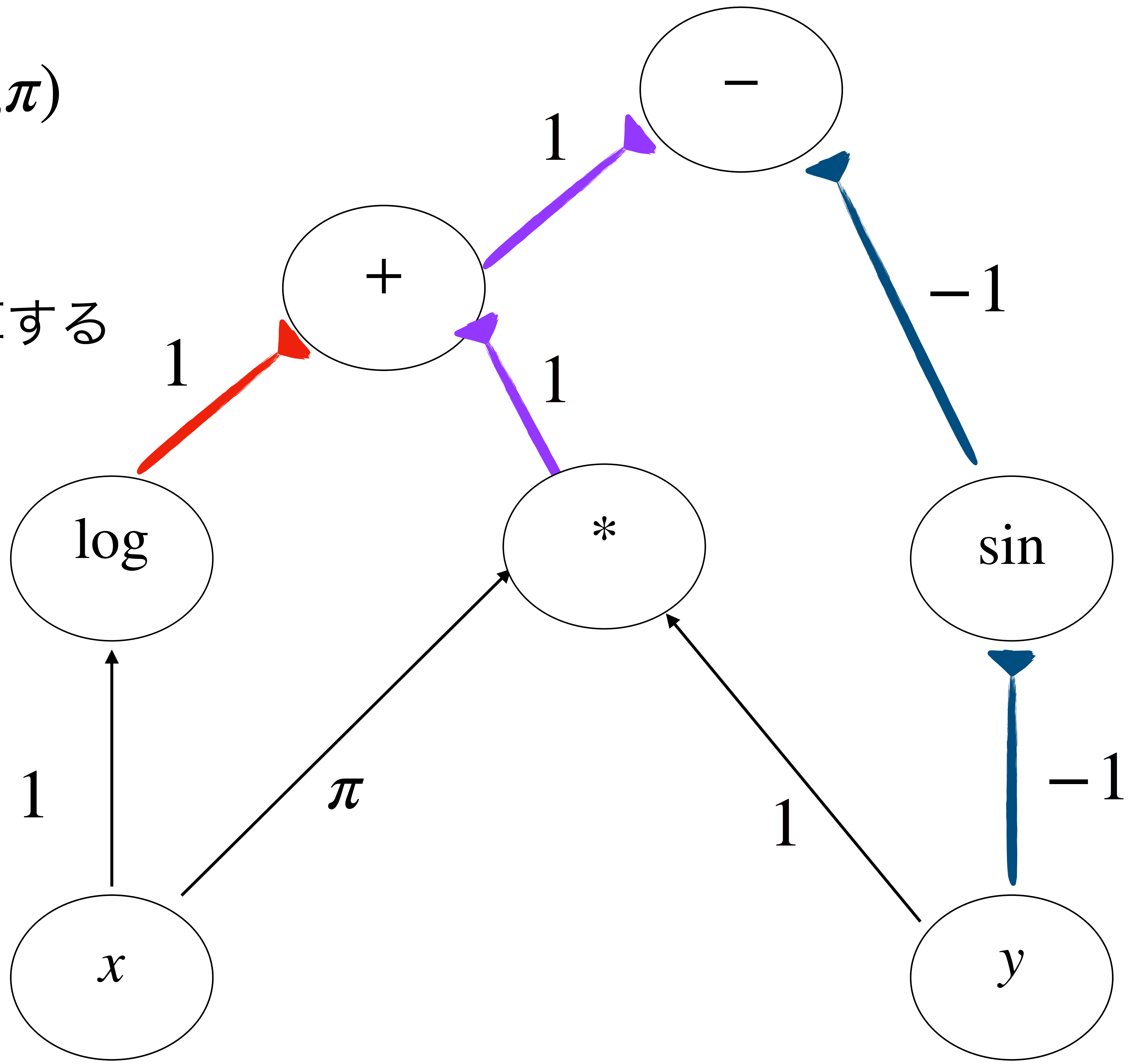


Reverseモードの自動微分

$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$(x, y) = (1, \pi)$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**根から集める**



Reverseモードの自動微分

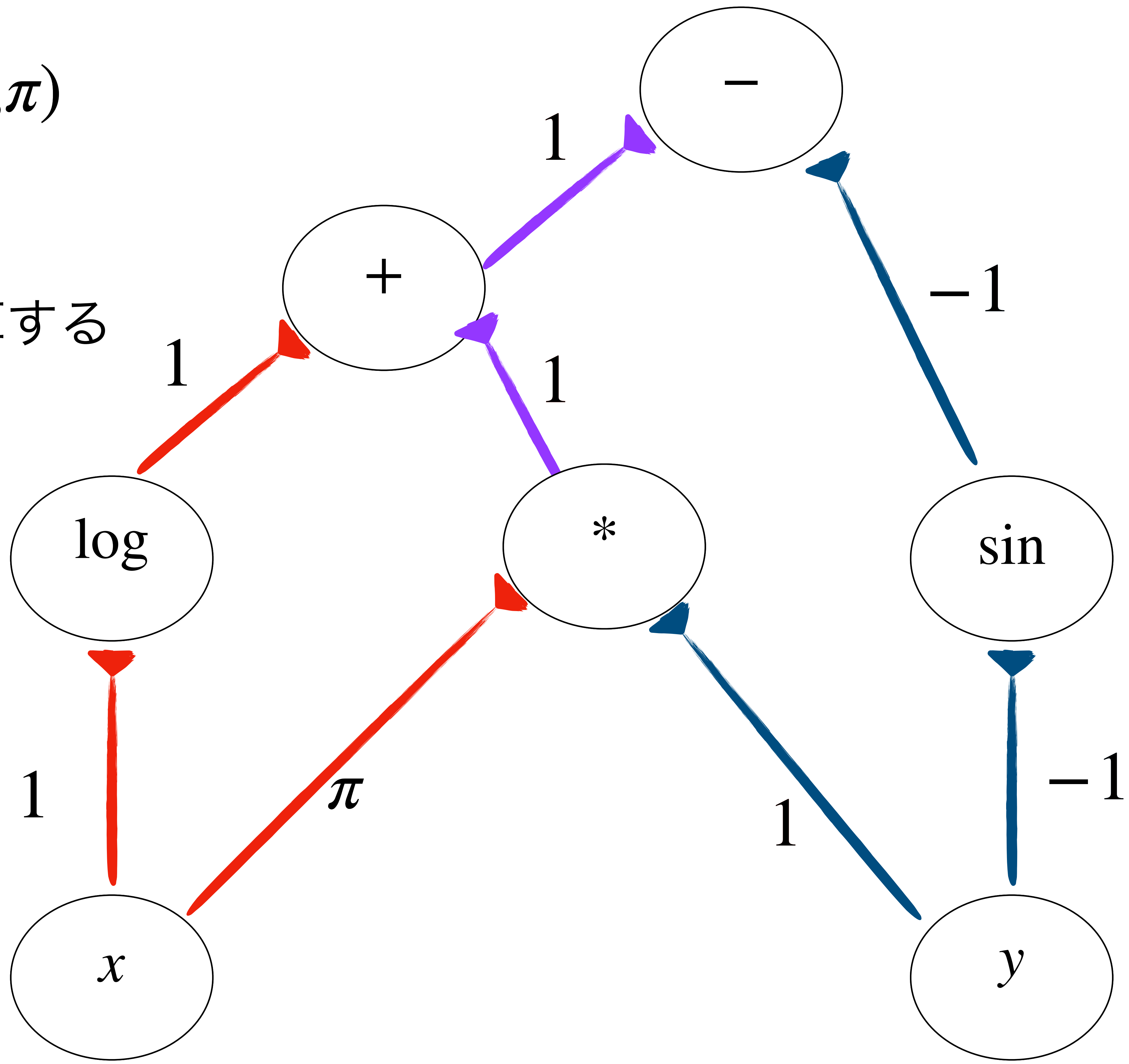
$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**根から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = \underline{1 + \pi}$$

$$\frac{\partial f}{\partial y}(1, \pi) = \underline{1 + 1} = 2$$



Reverseモードの自動微分

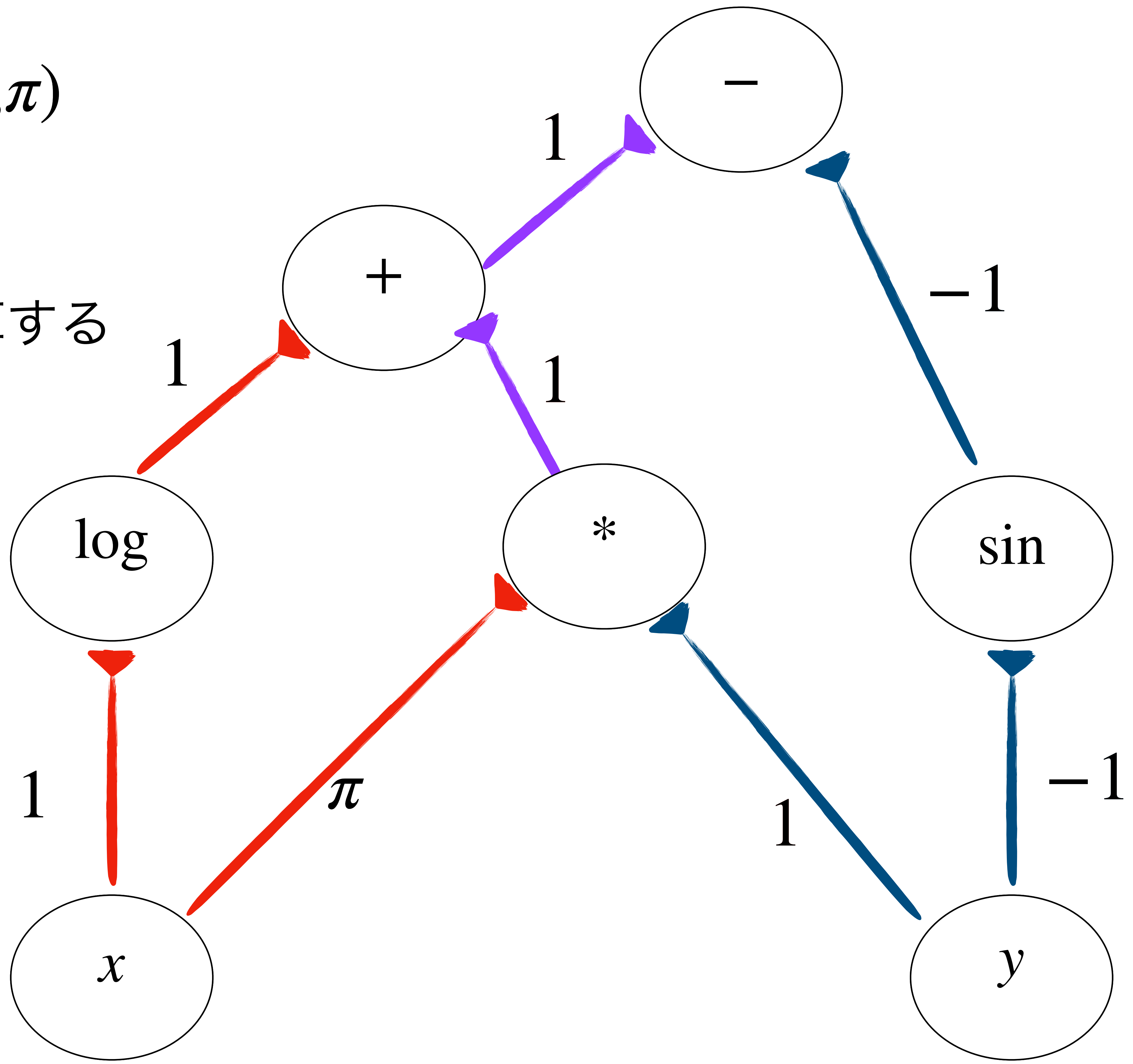
$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**根から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = \underline{1 + \pi} \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

$$\frac{\partial f}{\partial y}(1, \pi) = \underline{1 + 1} = 2 \quad \Leftarrow \quad \frac{\partial f}{\partial y} = x - \cos(y)$$



Reverseモードの自動微分

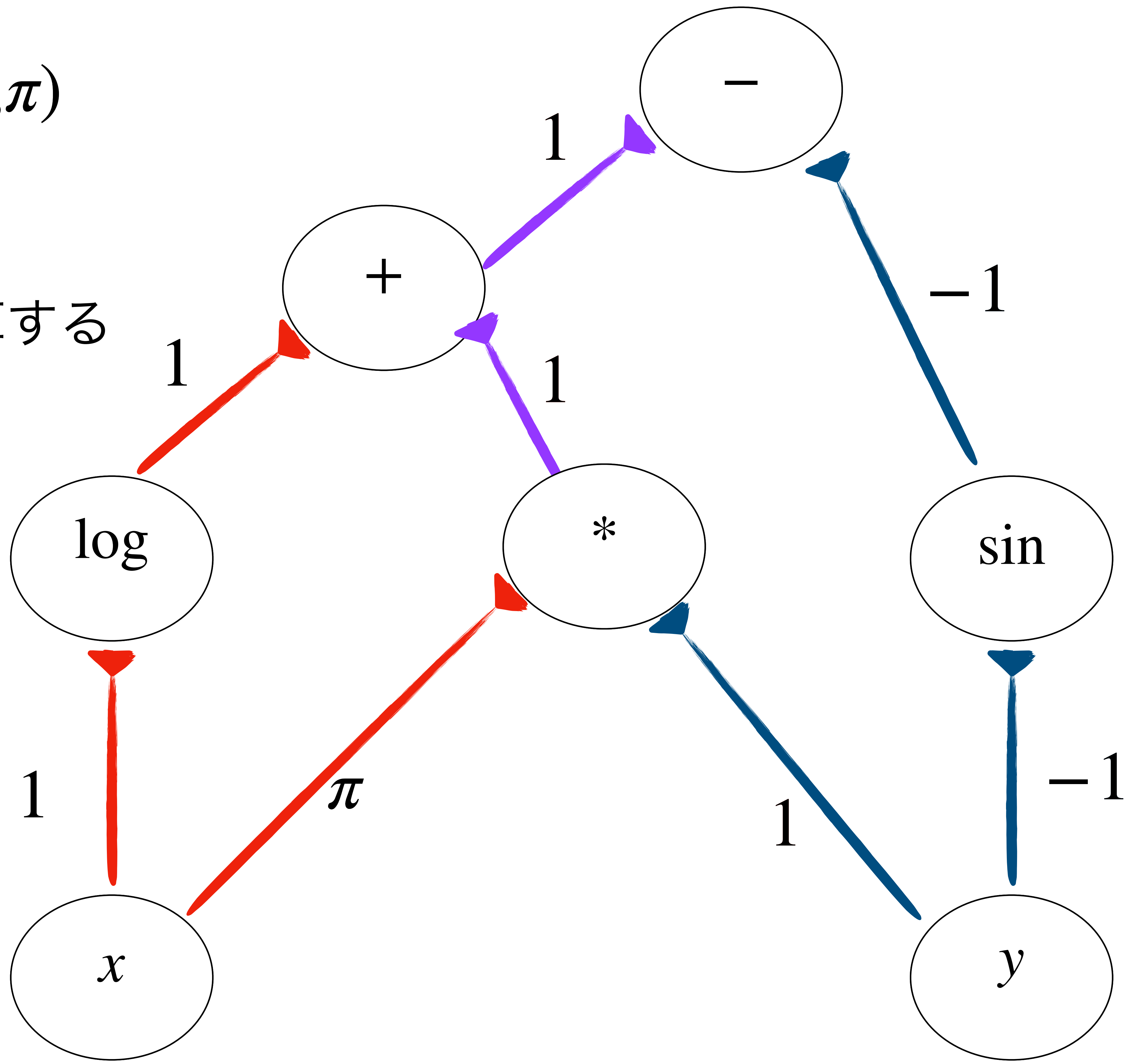
$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**根から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = \underline{1 + \pi} \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

$$\frac{\partial f}{\partial y}(1, \pi) = \underline{1 + 1} = 2 \quad \Leftarrow \quad \frac{\partial f}{\partial y} = x - \cos(y)$$



一回で, 一つの出力に関する勾配が求まる！

Reverseモードの自動微分

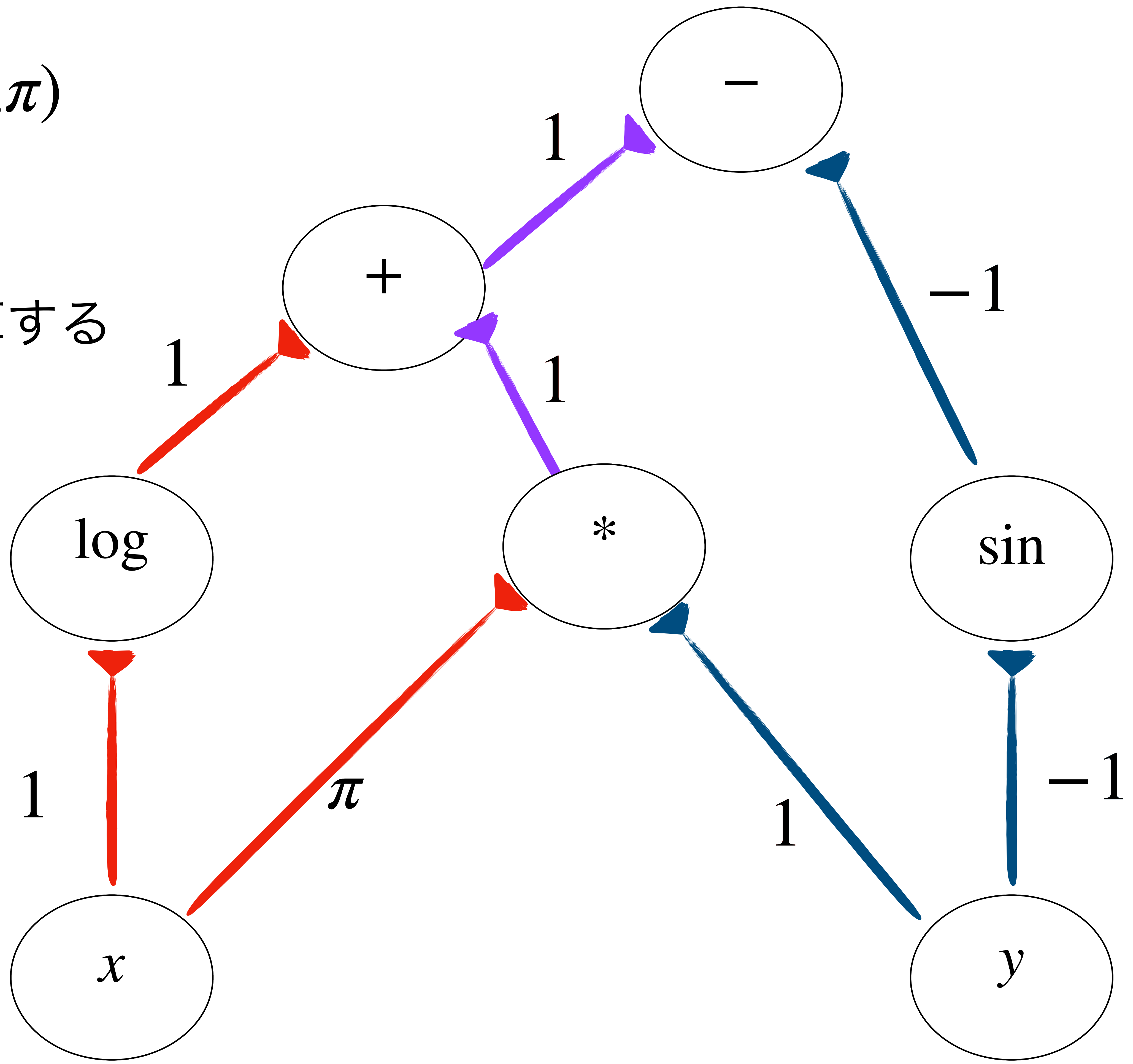
$$f(x, y) = \log(x) + x \cdot y - \sin(y)$$

$$(x, y) = (1, \pi)$$

1. 計算グラフに具体的な値を入れて計算しながら辺の微分値も計算する
2. 辺の微分値を**根から集める**

$$\frac{\partial f}{\partial x}(1, \pi) = \underline{1 + \pi} \quad \Leftarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x} + y$$

$$\frac{\partial f}{\partial y}(1, \pi) = \underline{1 + 1} = 2 \quad \Leftarrow \quad \frac{\partial f}{\partial y} = x - \cos(y)$$



一回で, 一つの出力に関する勾配が求まる！

機械学習の誤差逆伝播法と同じ！

もくじ

- 計算機での微分
- 数式微分から自動微分
- **微分グラフの最適化と実験**
  - 実験
  - 実装よもやま
- その他参考
- まとめ

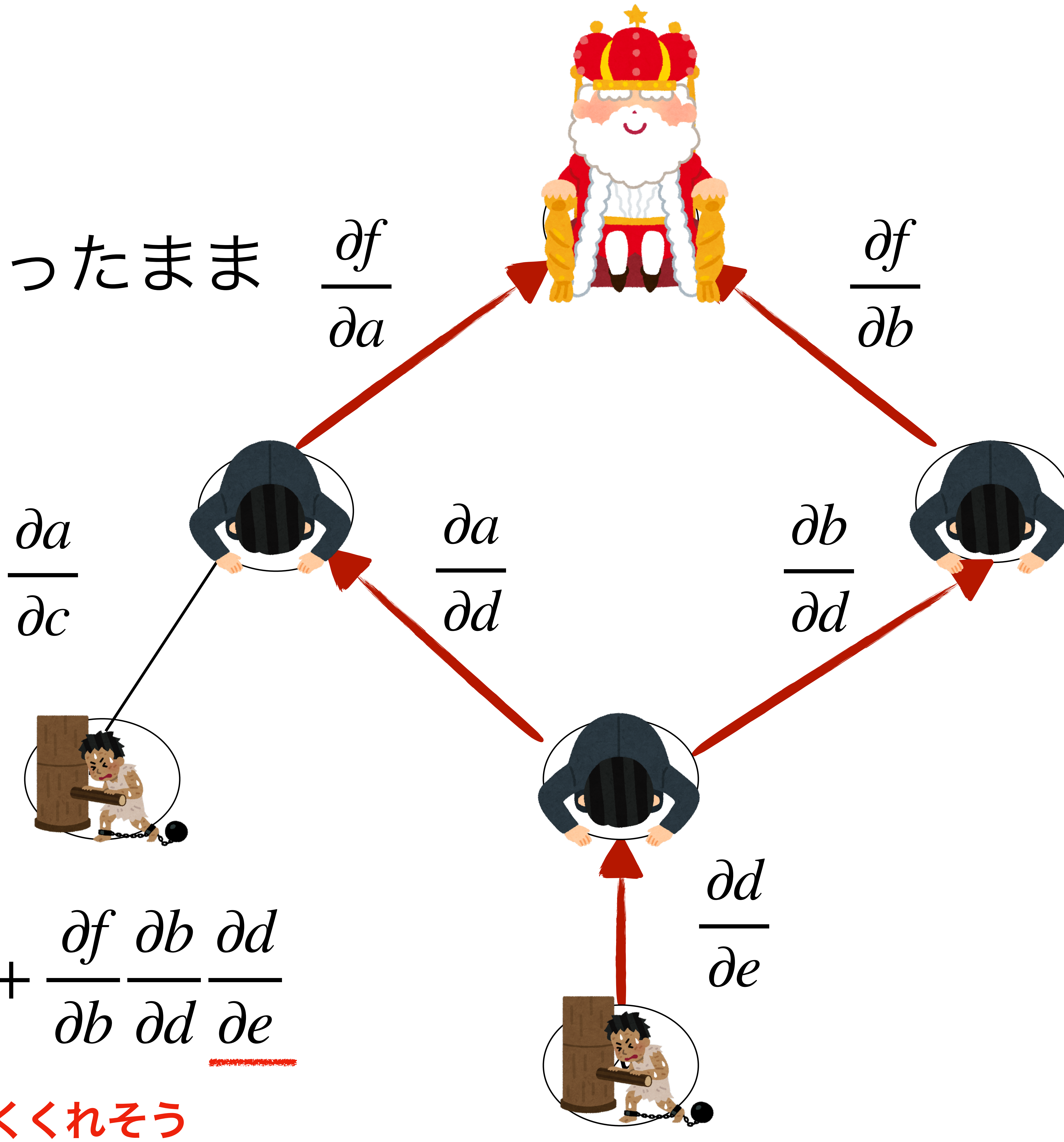
微分グラフの最適化

- 支配木を使った最適化をすると微分の情報を保ったまま微分グラフをコンパクトにできる

右図はイメージです

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial a}\frac{\partial a}{\partial d}\underline{\frac{\partial d}{\partial e}} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\underline{\frac{\partial d}{\partial e}}$$

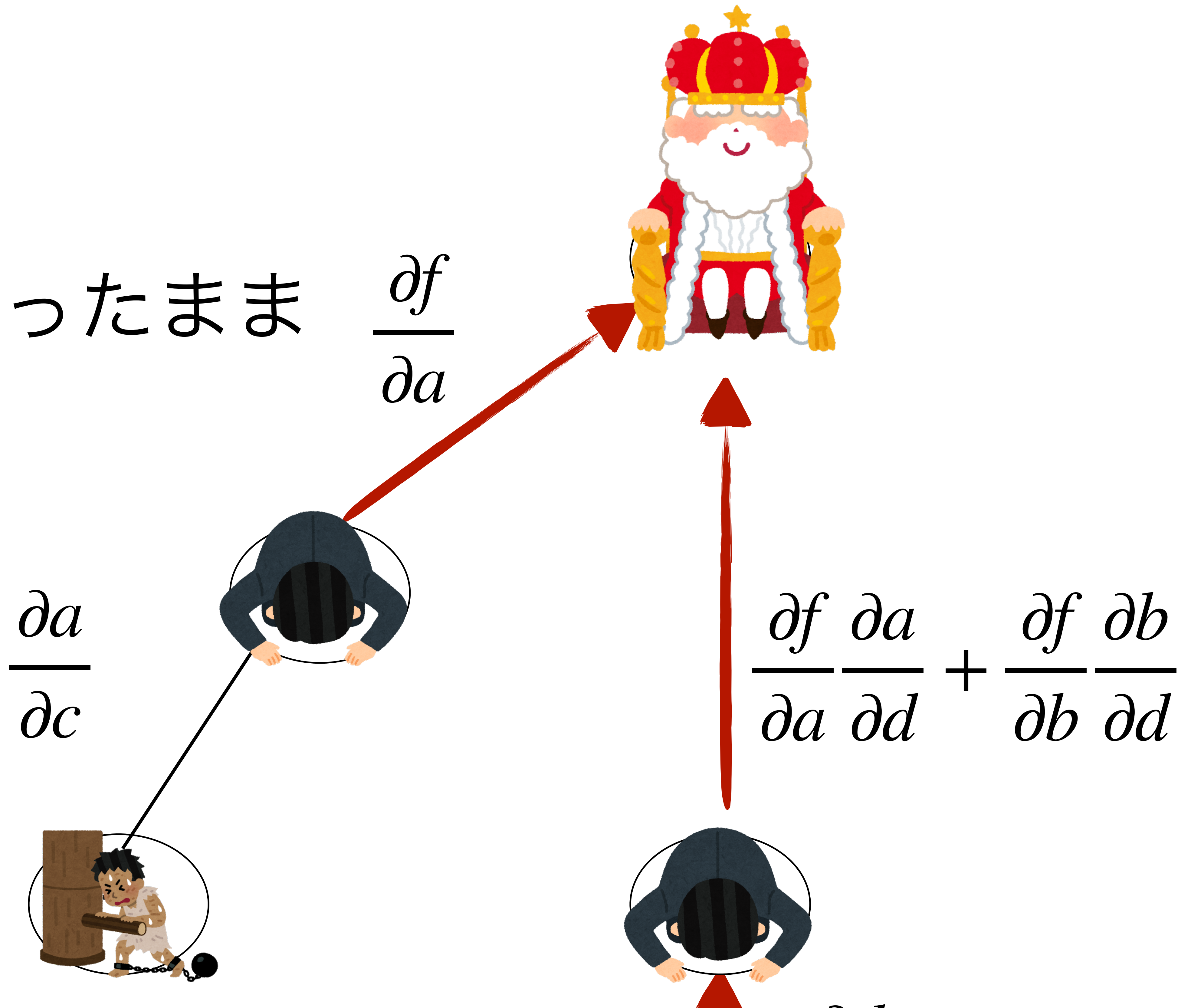
くくれそう

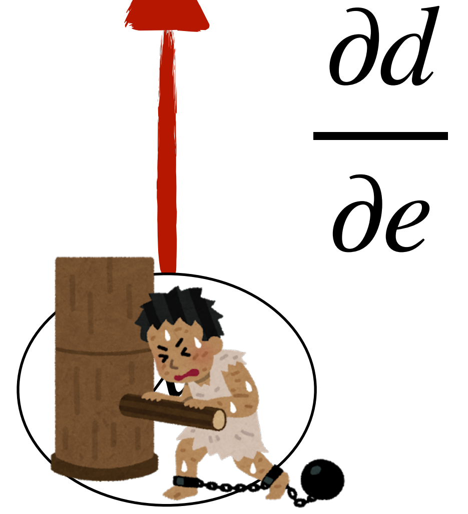


微分グラフの最適化

- 支配木を使った最適化をすると微分の情報を保ったまま微分グラフをコンパクトにできる

右図はイメージです



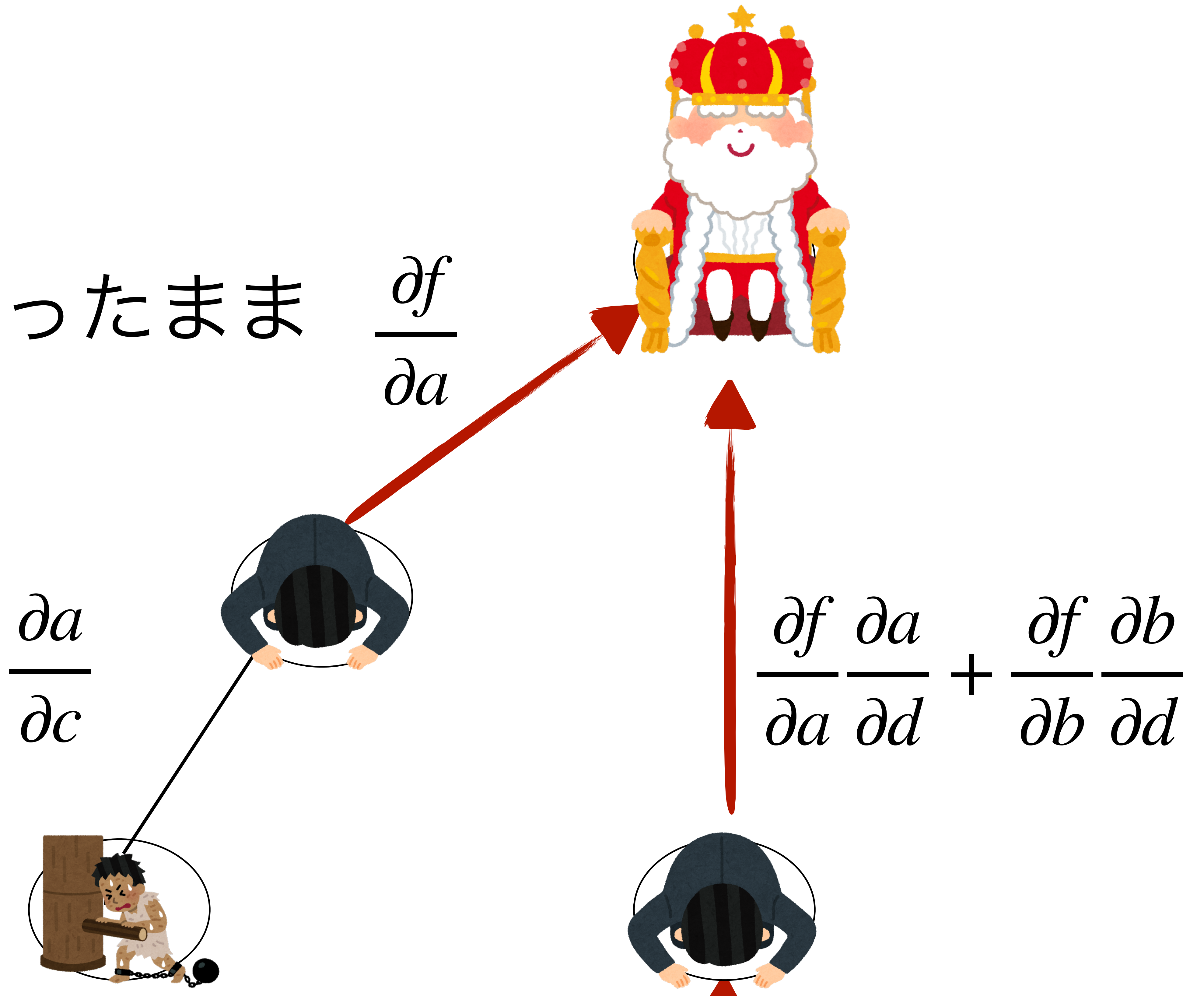


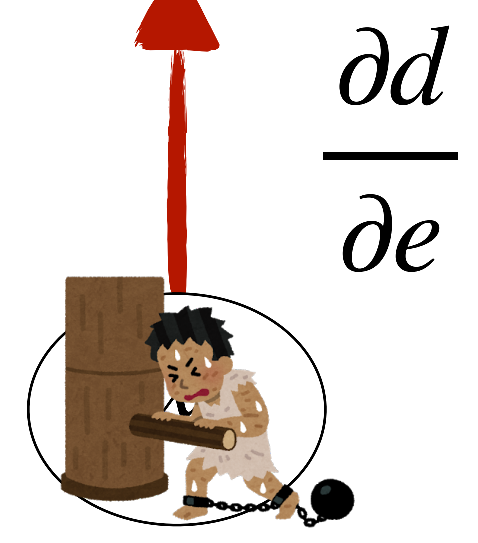
$$\frac{\partial f}{\partial x}=\underline{\left(\frac{\partial f}{\partial a}\frac{\partial a}{\partial d}+\frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\right)\frac{\partial d}{\partial e}}$$

乗算が減った！

微分グラフの最適化

- 支配木を使った最適化をすると微分の情報を保ったまま微分グラフをコンパクトにできる
  - 同じ式をたくさん評価するなら高速化が期待できる



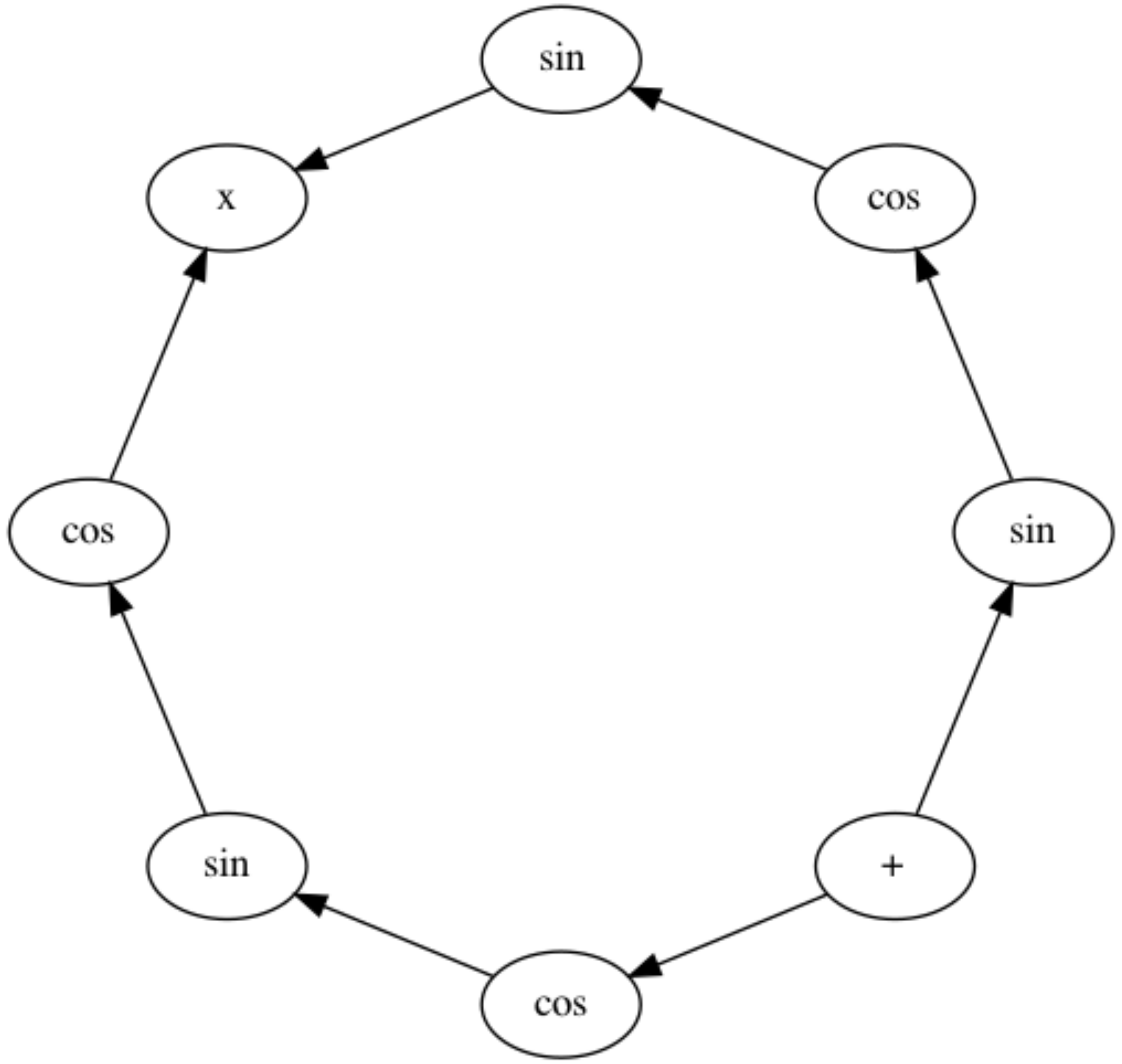


$$\frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial a}\frac{\partial a}{\partial d} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial d}\right)\frac{\partial d}{\partial e}$$

実験

- 3つの計算方法で比較
- DPしにくい例で微妙に早くなった

```
32
-0.00000000000000000000000000000000000000000000000000000000000000000006506295273620515
naive:6272 micro sec takes
naive:6272065 nano sec takes
-0.00000000000000000000000000000000000000000000000000000000000000000006506295273620503
dp:3869 micro sec takes
dp:3869179 nano sec takes
-0.00000000000000000000000000000000000000000000000000000000000000000006506295273620503
dom:3095 micro sec takes
dom:3095381 nano sec takes
32
-0.00000000000000000000000000000000000000000000000000000000000000000006506295273620515
naive:13149 micro sec takes
naive:13149971 nano sec takes
-0.00000000000000000000000000000000000000000000000000000000000000000006506295273620503
dp:5064 micro sec takes
dp:5064041 nano sec takes
-0.00000000000000000000000000000000000000000000000000000000000000000006506295273620503
dom:3768 micro sec takes
dom:3768377 nano sec takes
```



こういういかにもDPできなさそうなのを大きくした

実装よもやま

- 無駄に自作したParser Combinatorが魔境になった
  - https://bodil.lol/parser-combinators/
- Rustのくせに副作用を多用したのでとてもつらい

もくじ

- 計算機での微分
- 数式微分から自動微分
- 微分グラフの最適化と実験
- **その他参考**
- まとめ

その他参考

- 支配木を使った式DAGの最適化
  - 今回の最適化はこれ
- D言語で数式微分
- Automatic differentiation in machine learning: a survey
  - 自動微分の歴史とか参考にした
- 微分可能なプログラミング言語がある
  - Reverse modeの微分を組み込み関数として持つ言語の形式的定義

もくじ

- 計算機での微分
- 数式微分から自動微分
- 微分グラフの最適化と実験
- その他参考
- **まとめ**

まとめ

☀ (手動)微分を卒業できた