

# モンテカルロ木探索

AI開発演習II



# モンテカルロ木探索の全体像

---

次に打つべき手は何か？

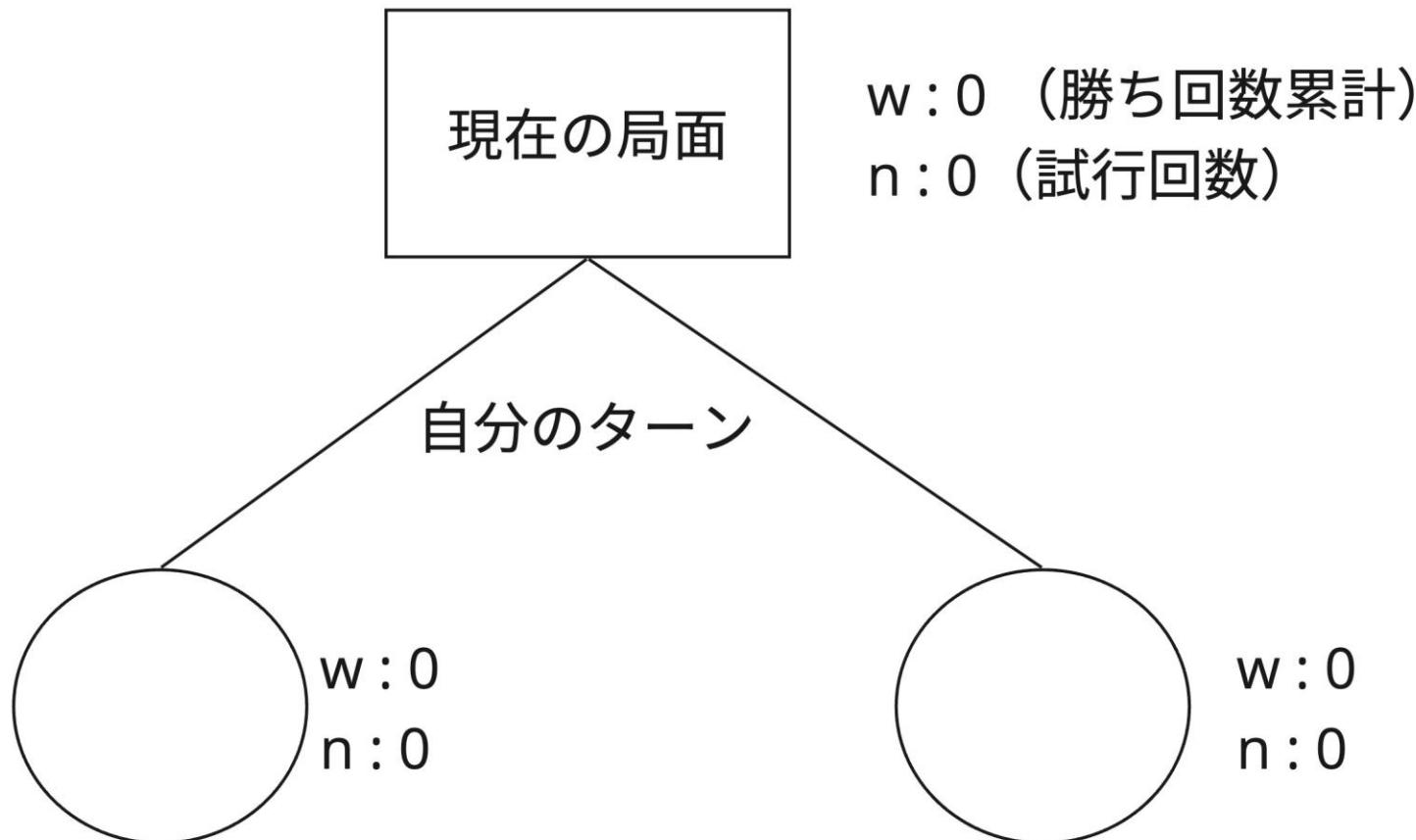
それを答えるためにモンテカルロ木を作成して判断する。

以下の四つのステップを繰り返し行います

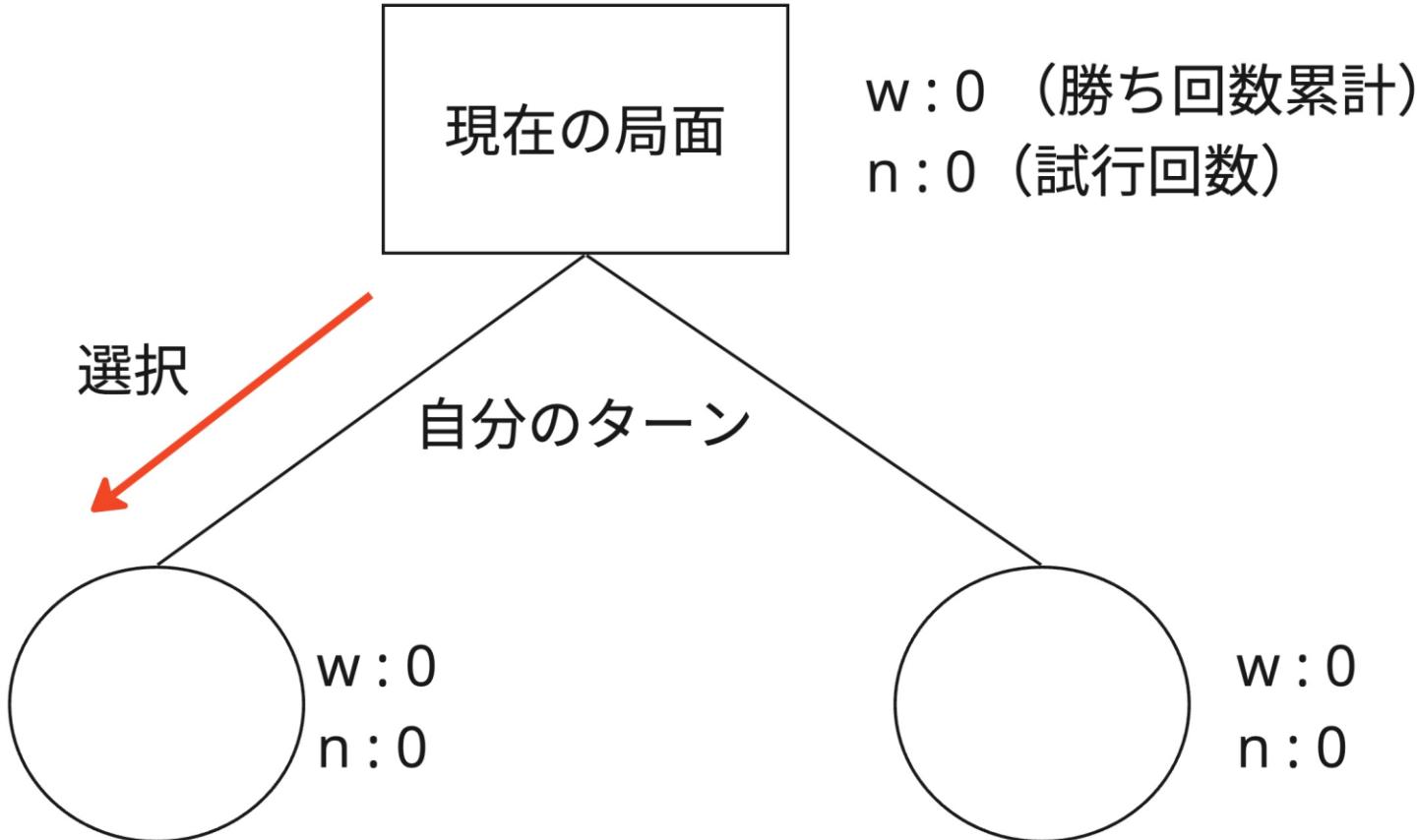
1. 「選択」
  - どの手を計算するか選択
2. 「評価」
  - 選択した手からゲーム終了までランダムに打つ（プレイアウト）
3. 「更新」
  - 勝ち回数と試行回数の更新
4. 「展開」
  - プレイアウトの回数が閾値に達する場合、次の合法手を木に追加

# 初期状態

---

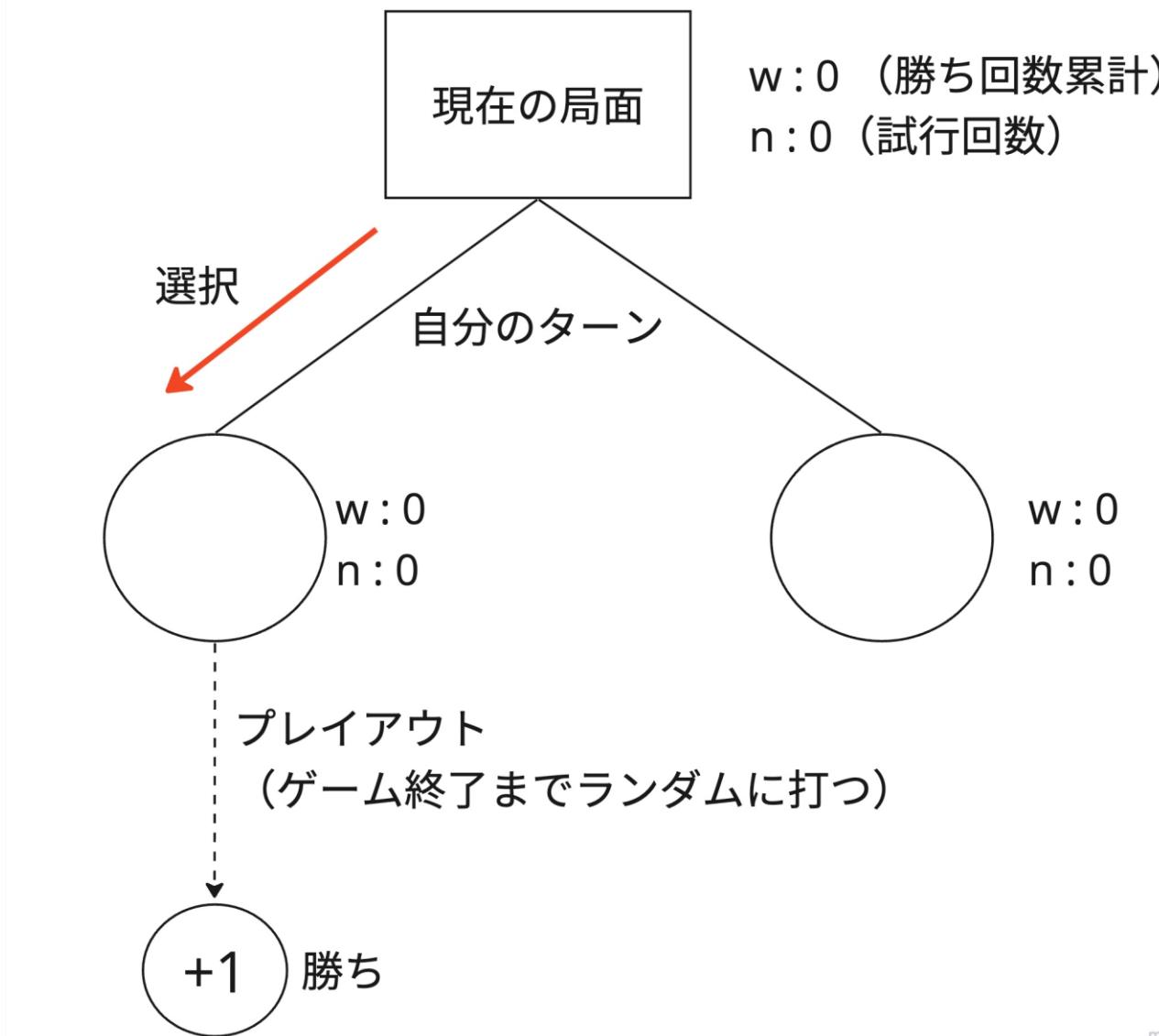


# 1回目の選択



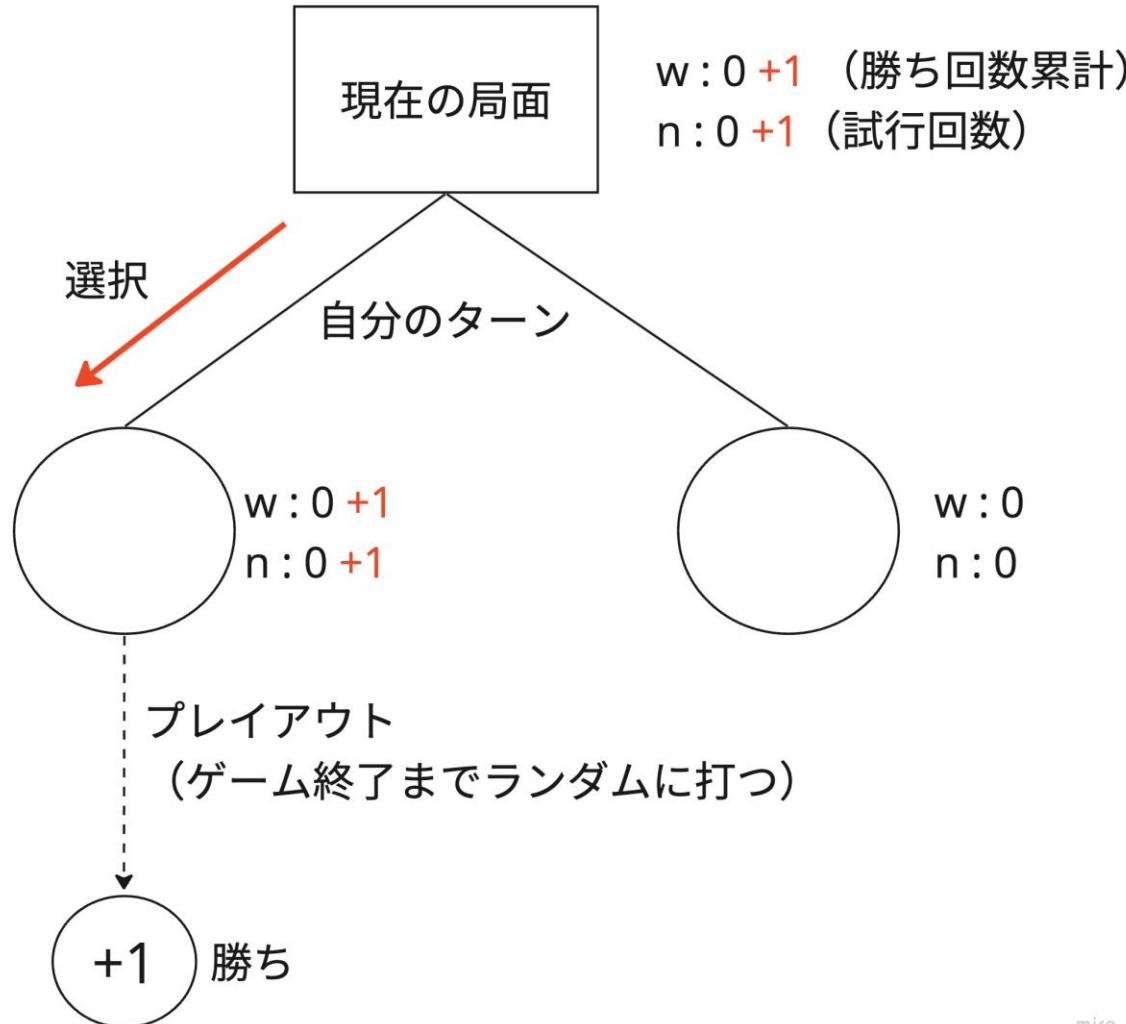
両方とも選択されたことが無いため、  
ランダムに選ぶ

# 1回目の評価

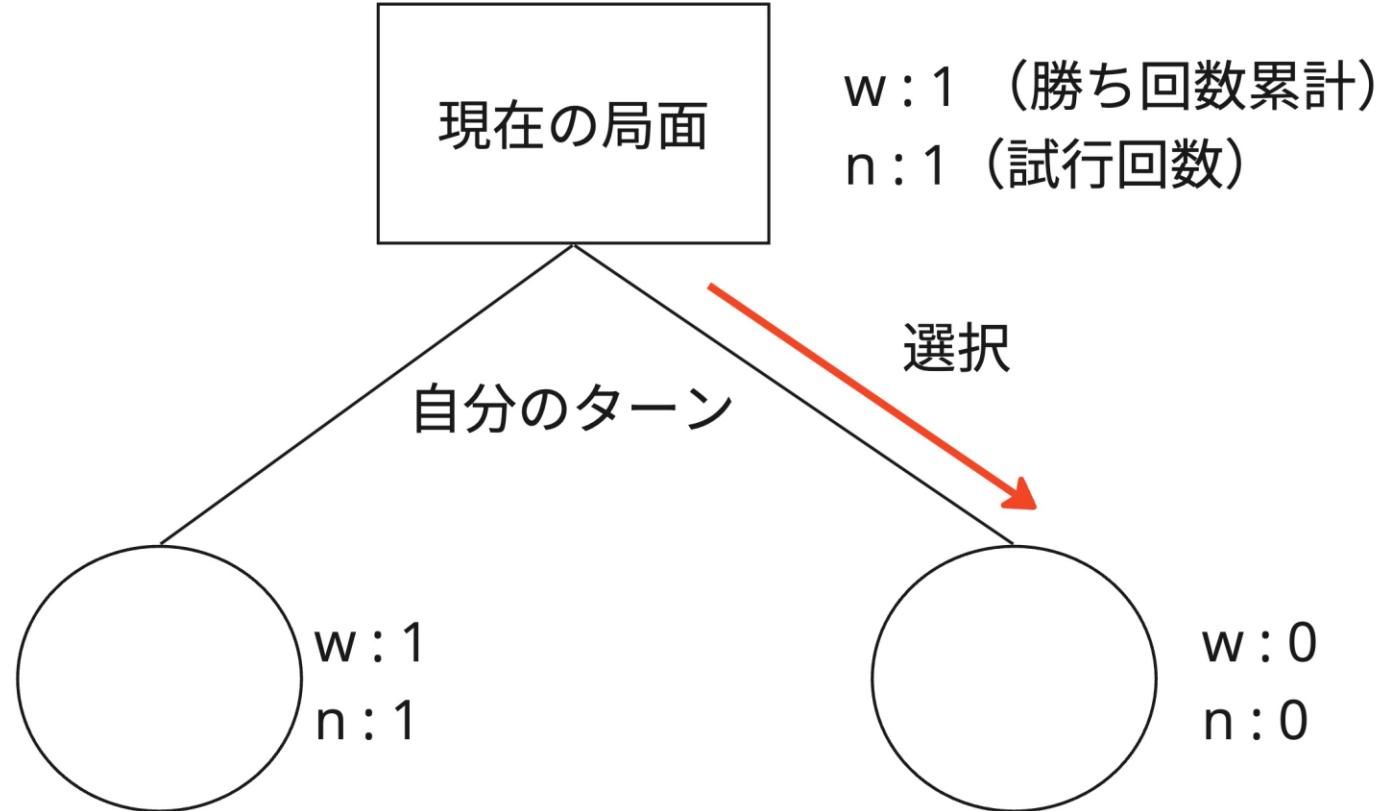


# 1回目の更新

$n$  (試行回数) が10に達するまで  
「展開」は行わないことにする

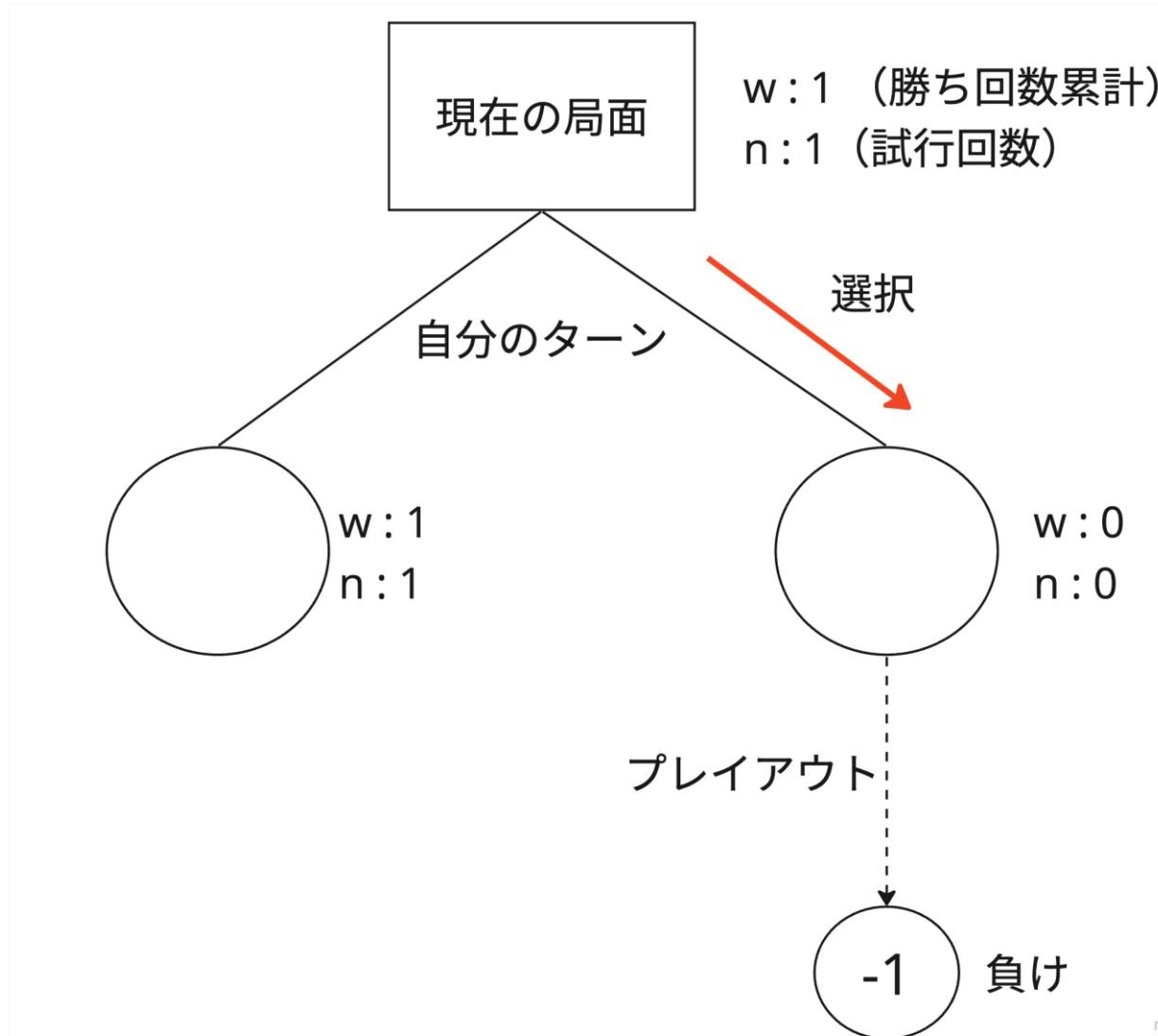


# 2回目の選択

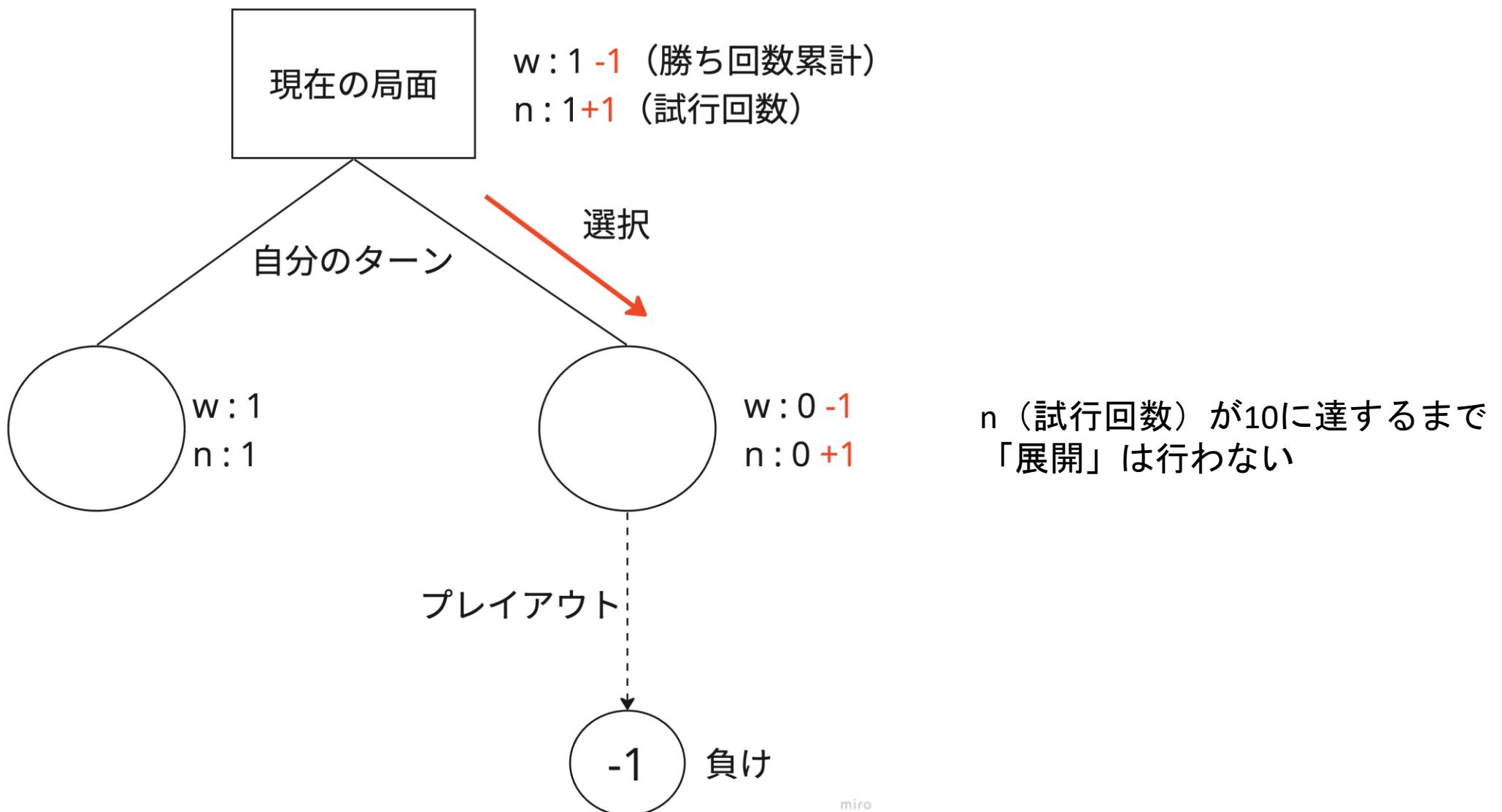


右のパターンがまだ選択されていないので、  
右を選ぶ

## 2回目の評価

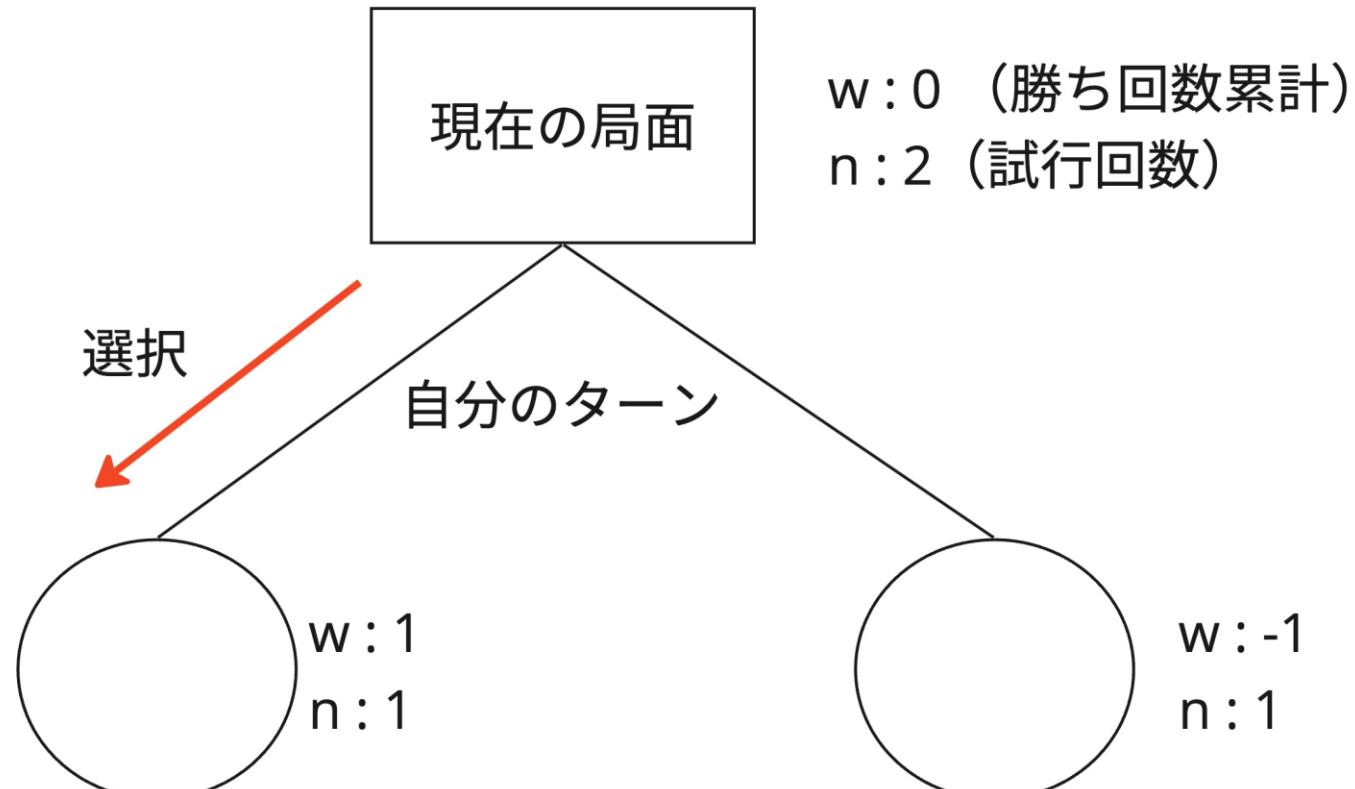


## 2回目の更新



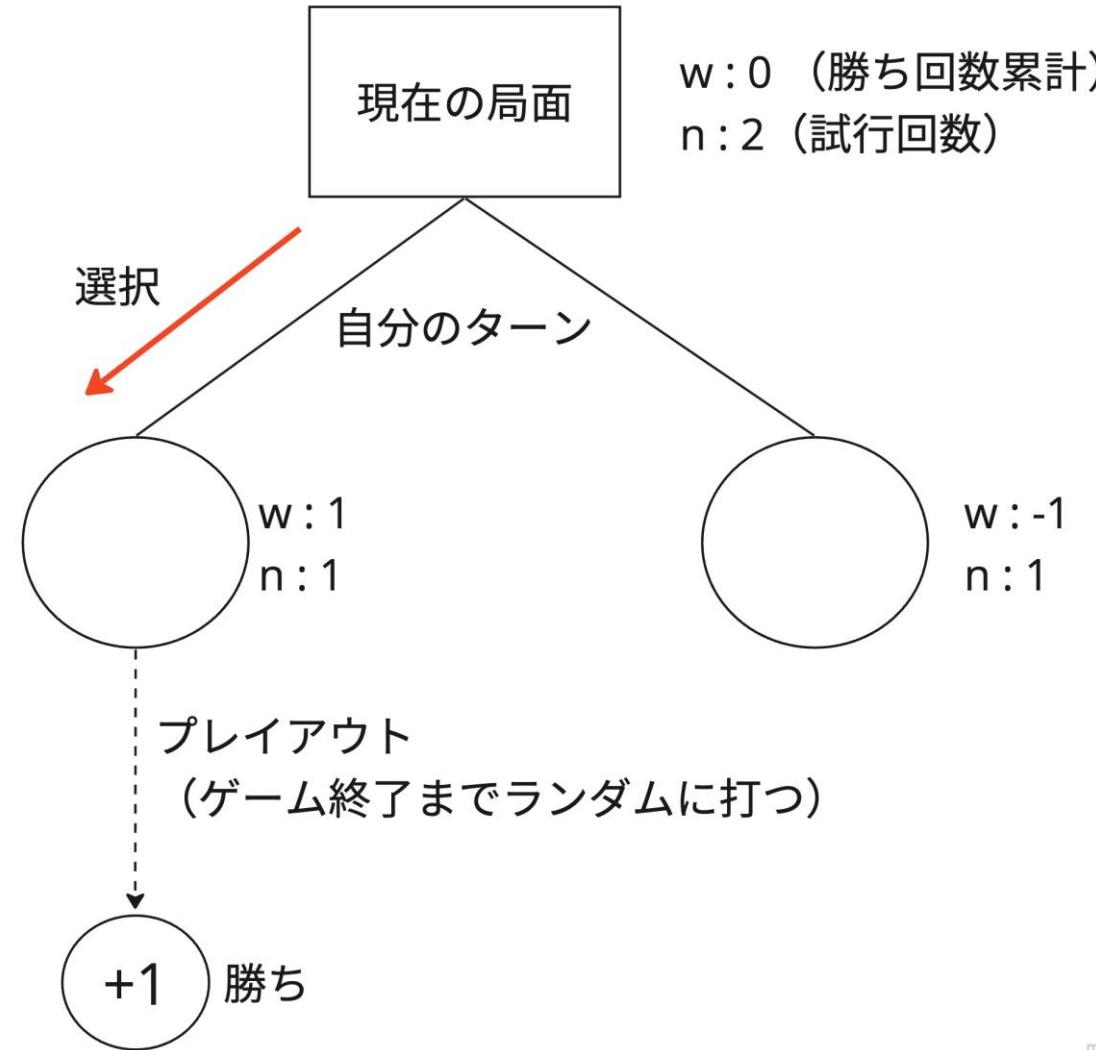
# 3回目の選択

---



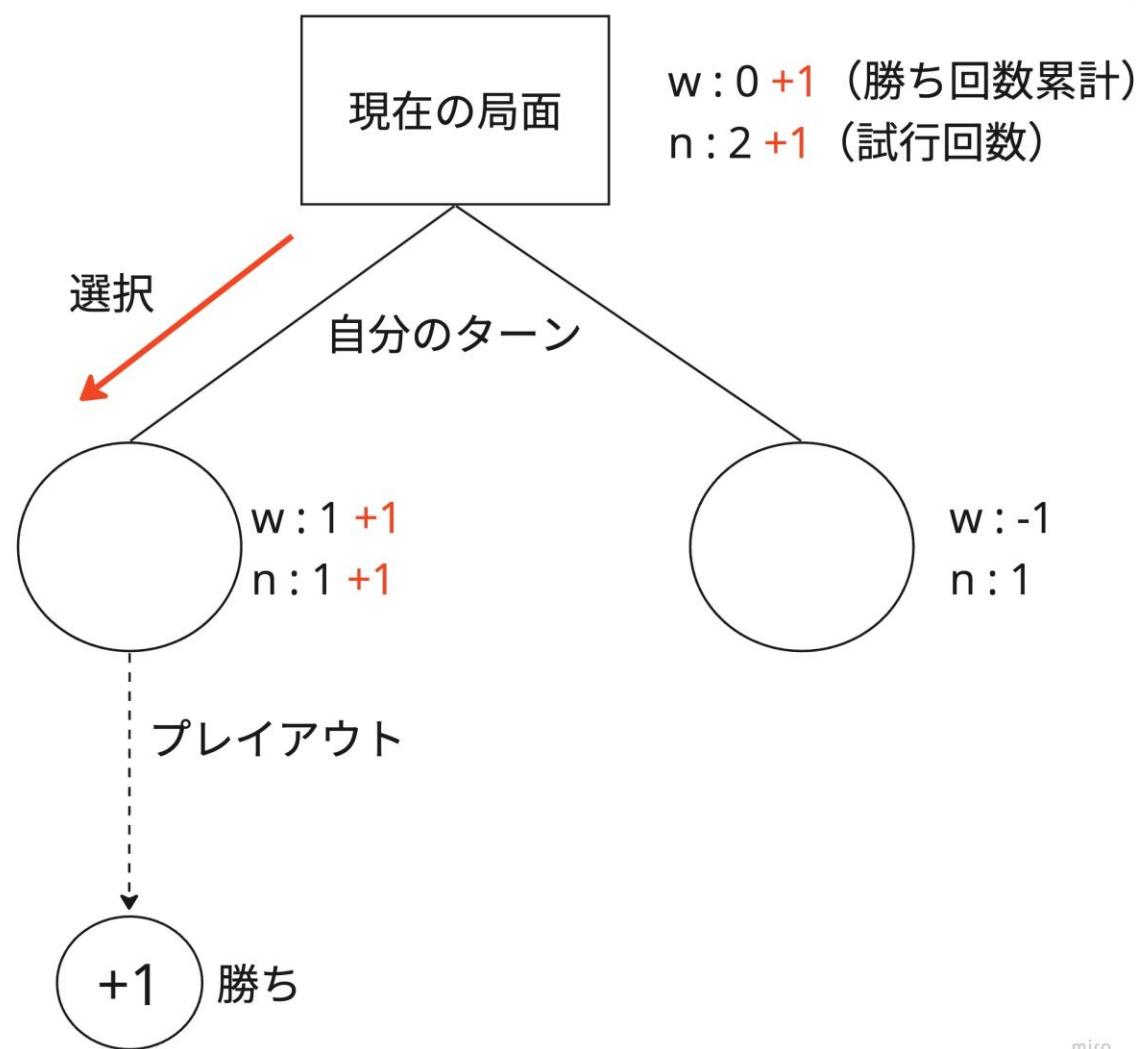
勝率が高い手を選択する  
(具体的な計算方法は後ほど)

# 3回目の評価



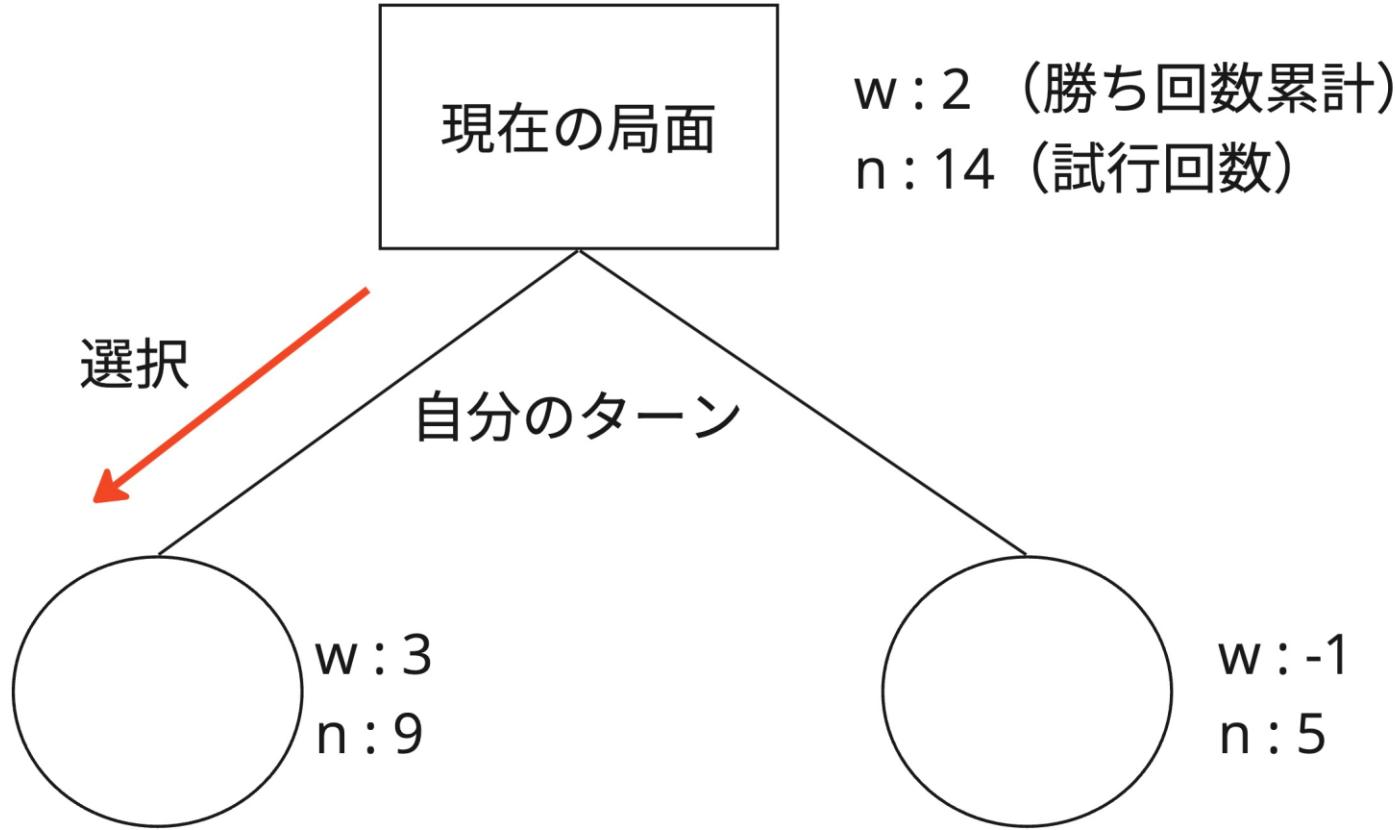
# 3回目の更新

$n$  (試行回数) が10に達するまで  
「展開」は行わない



# 15回目の選択

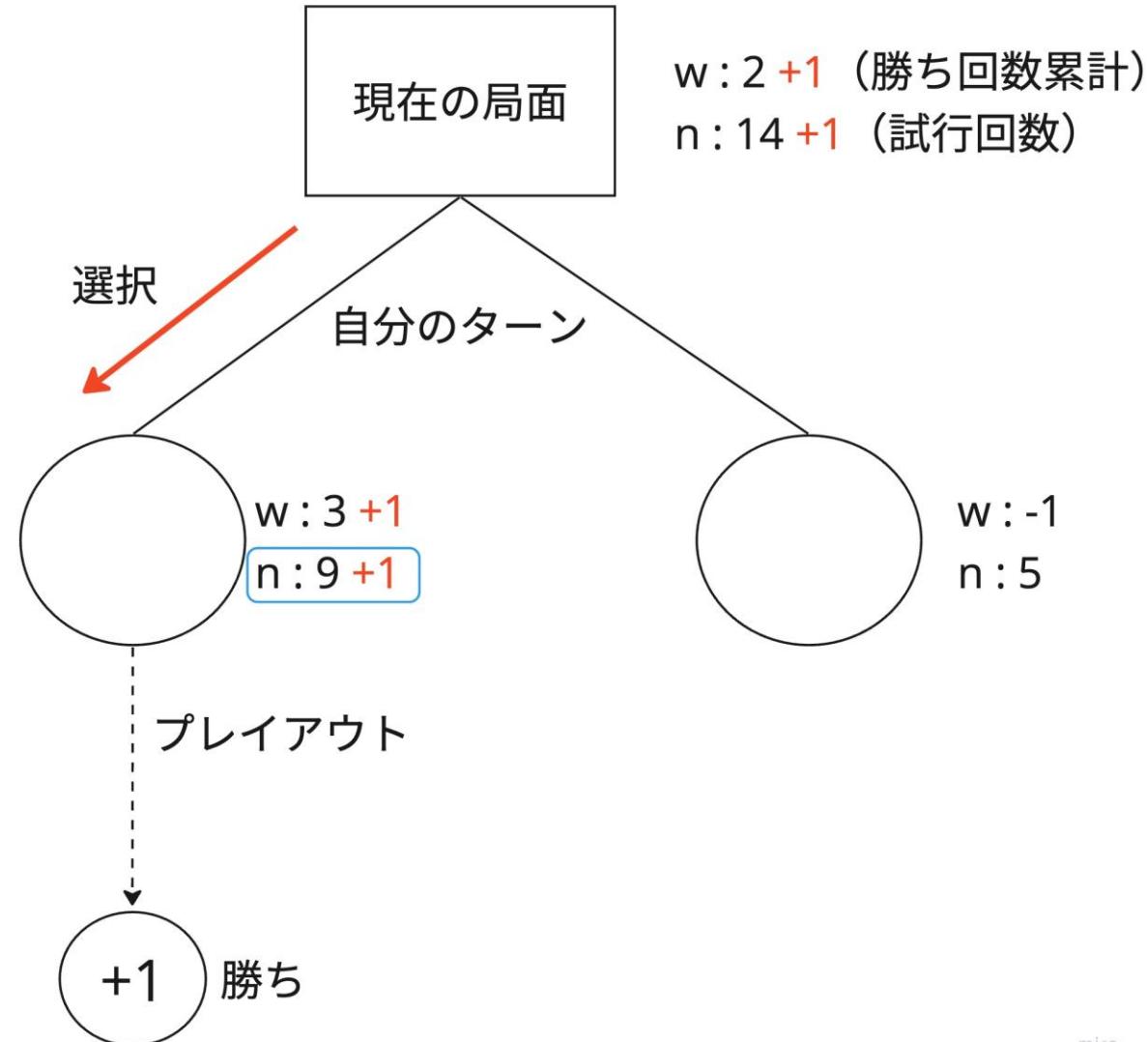
---



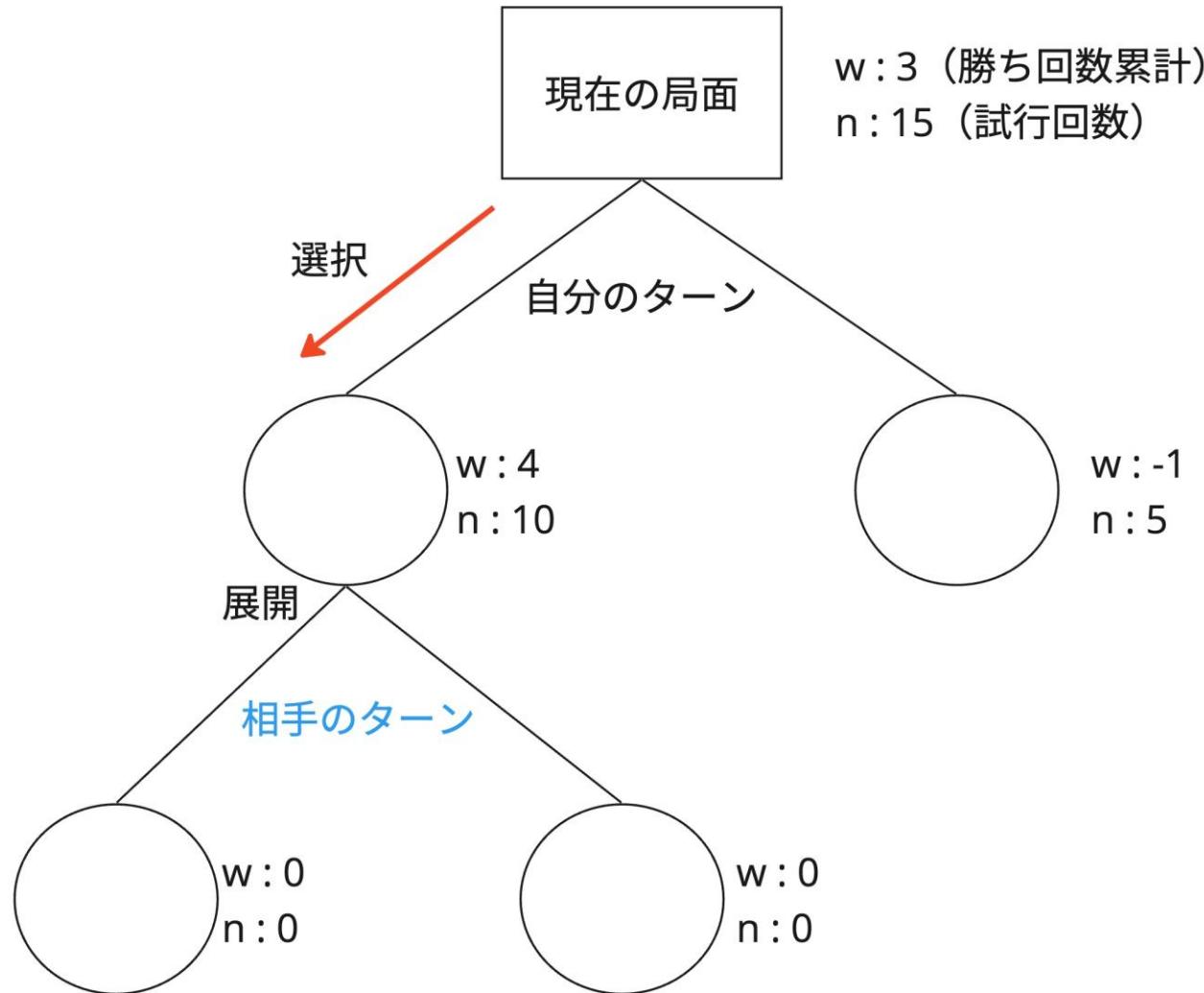
勝率が高い手を選択する  
(具体的な計算方法は後ほど)

# 15回目の評価

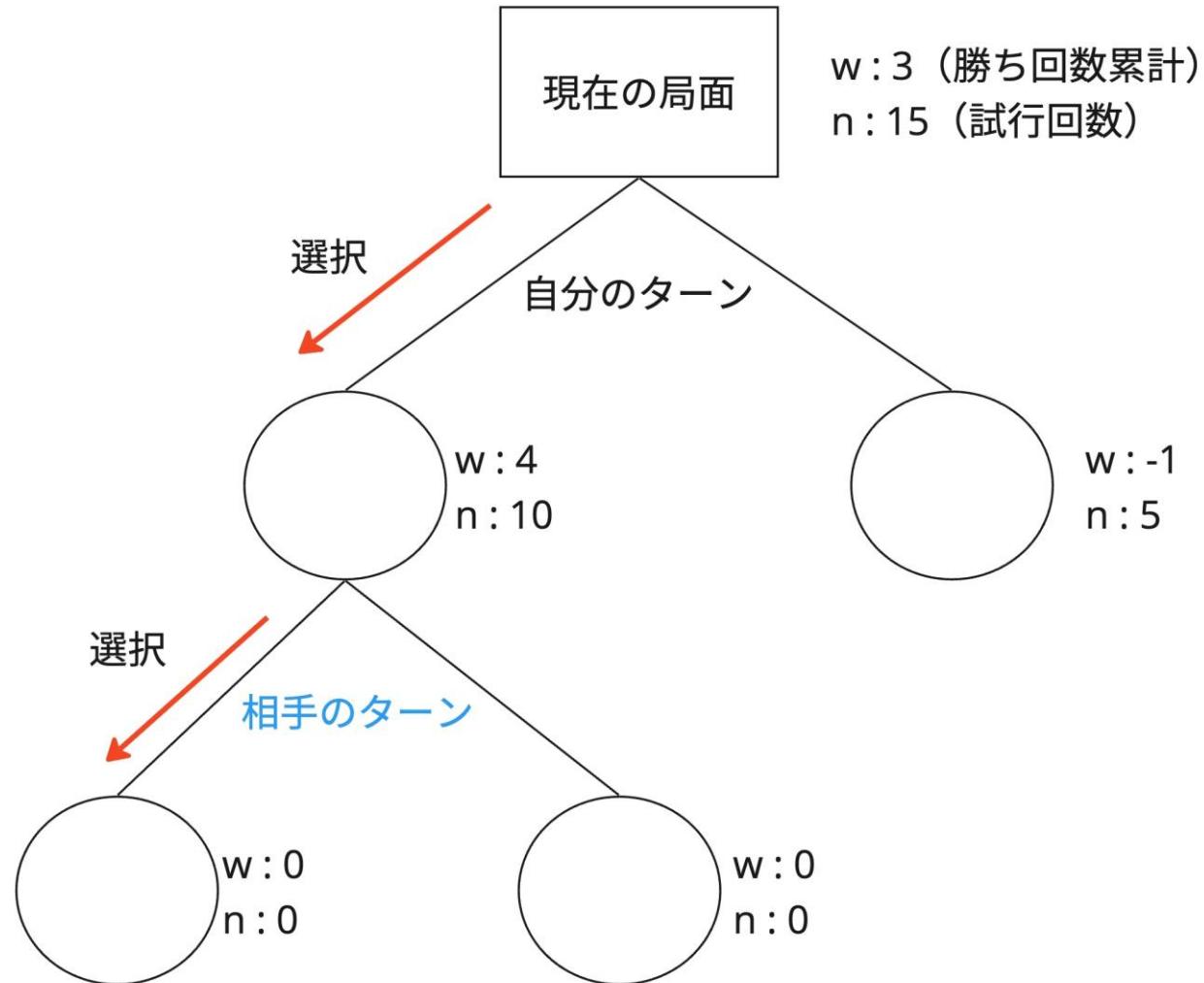
n（試行回数）が10に達したので  
「展開」を行う



# 15回目で初めての展開

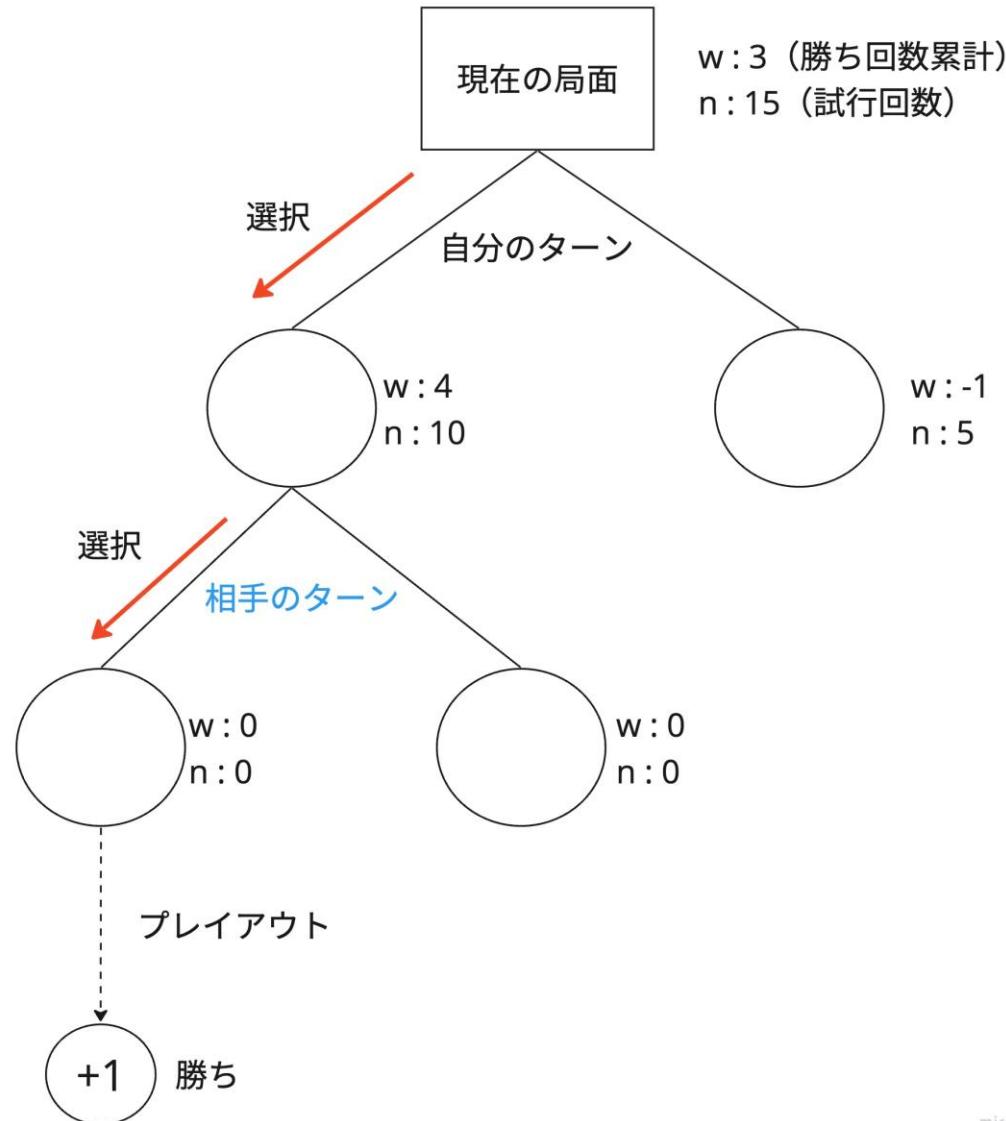


# 16回目の選択

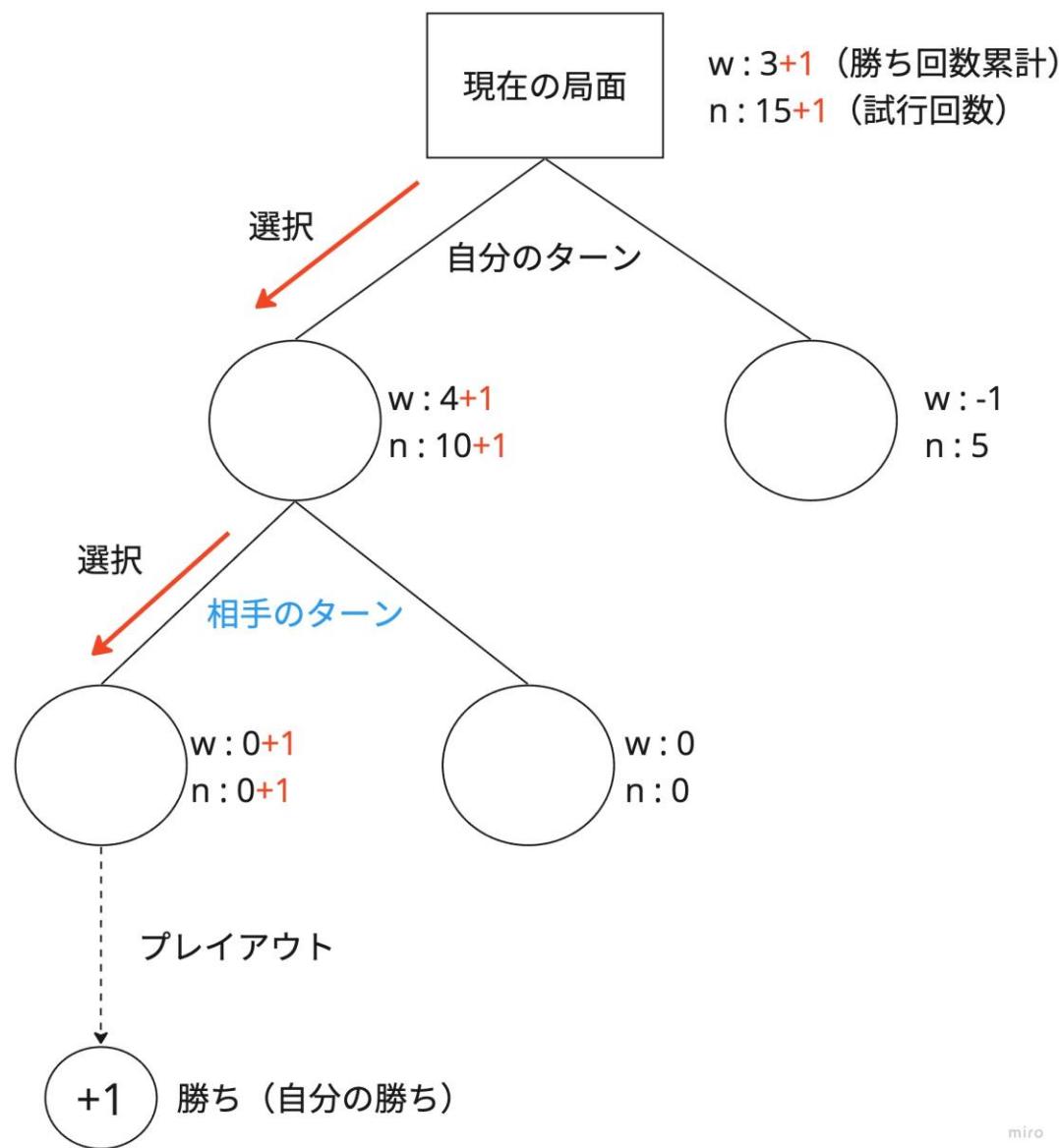


両方とも選択されたことが無いため、ランダムに選ぶ

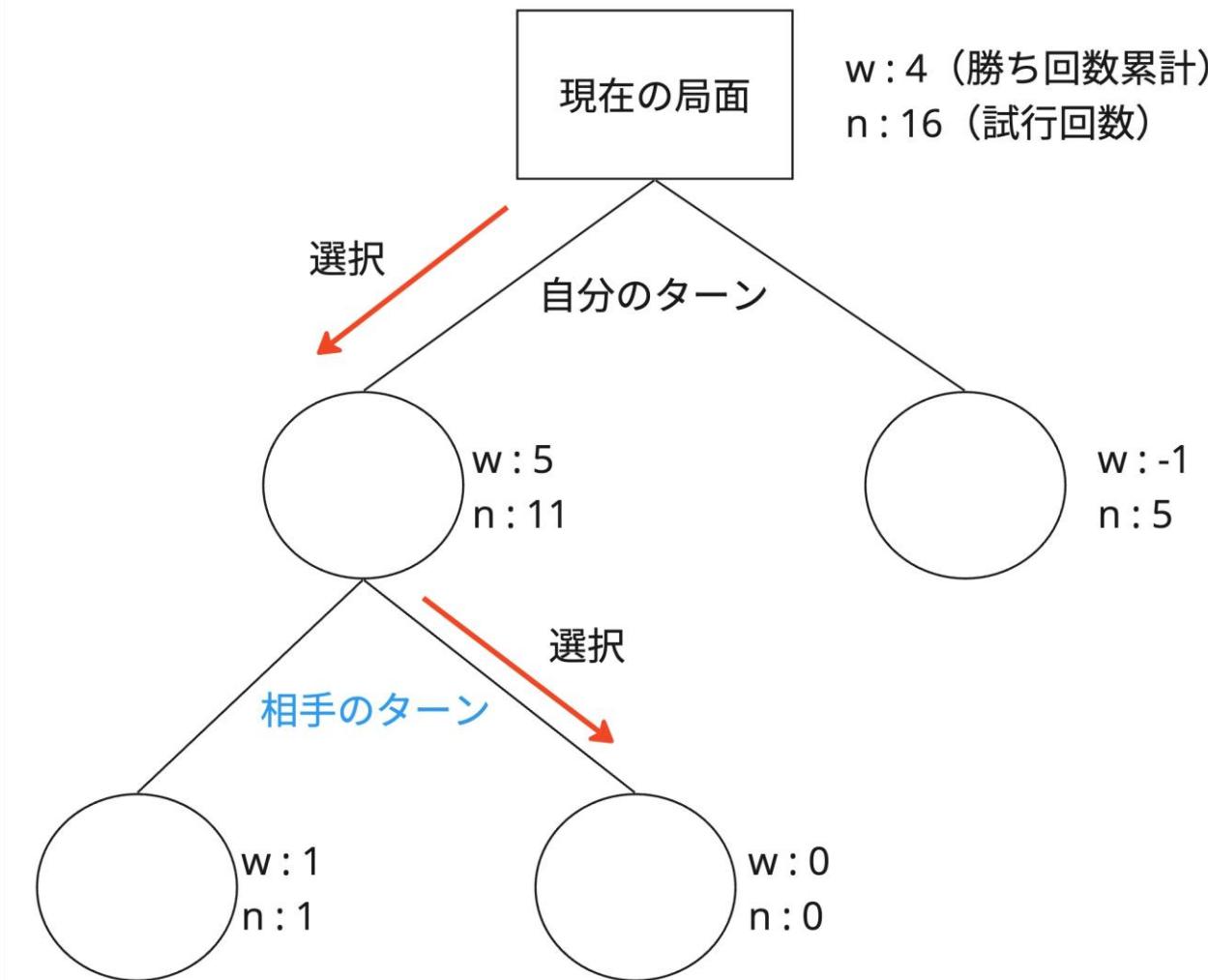
# 16回目の評価



# 16回目の更新

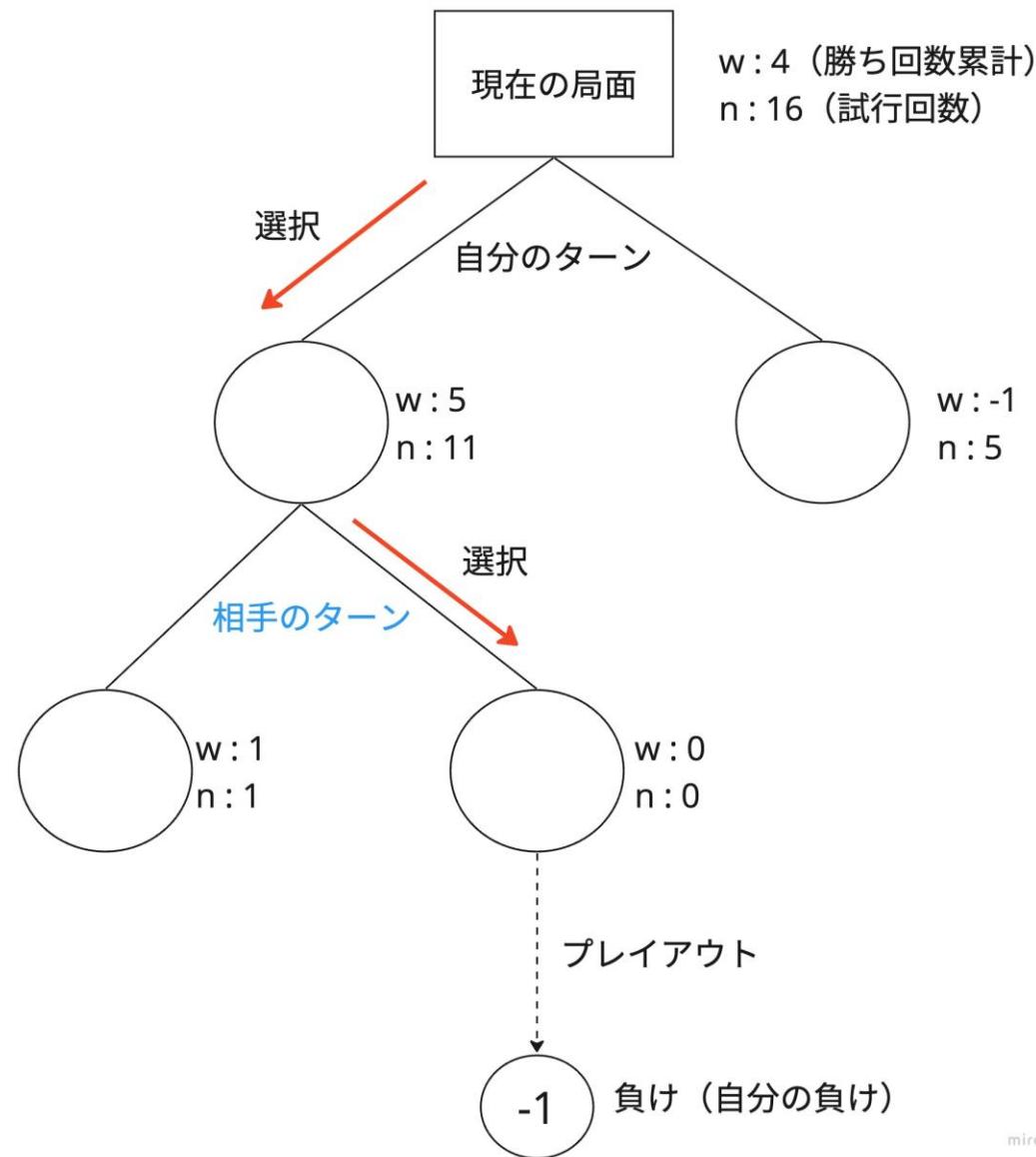


# 17回目の選択

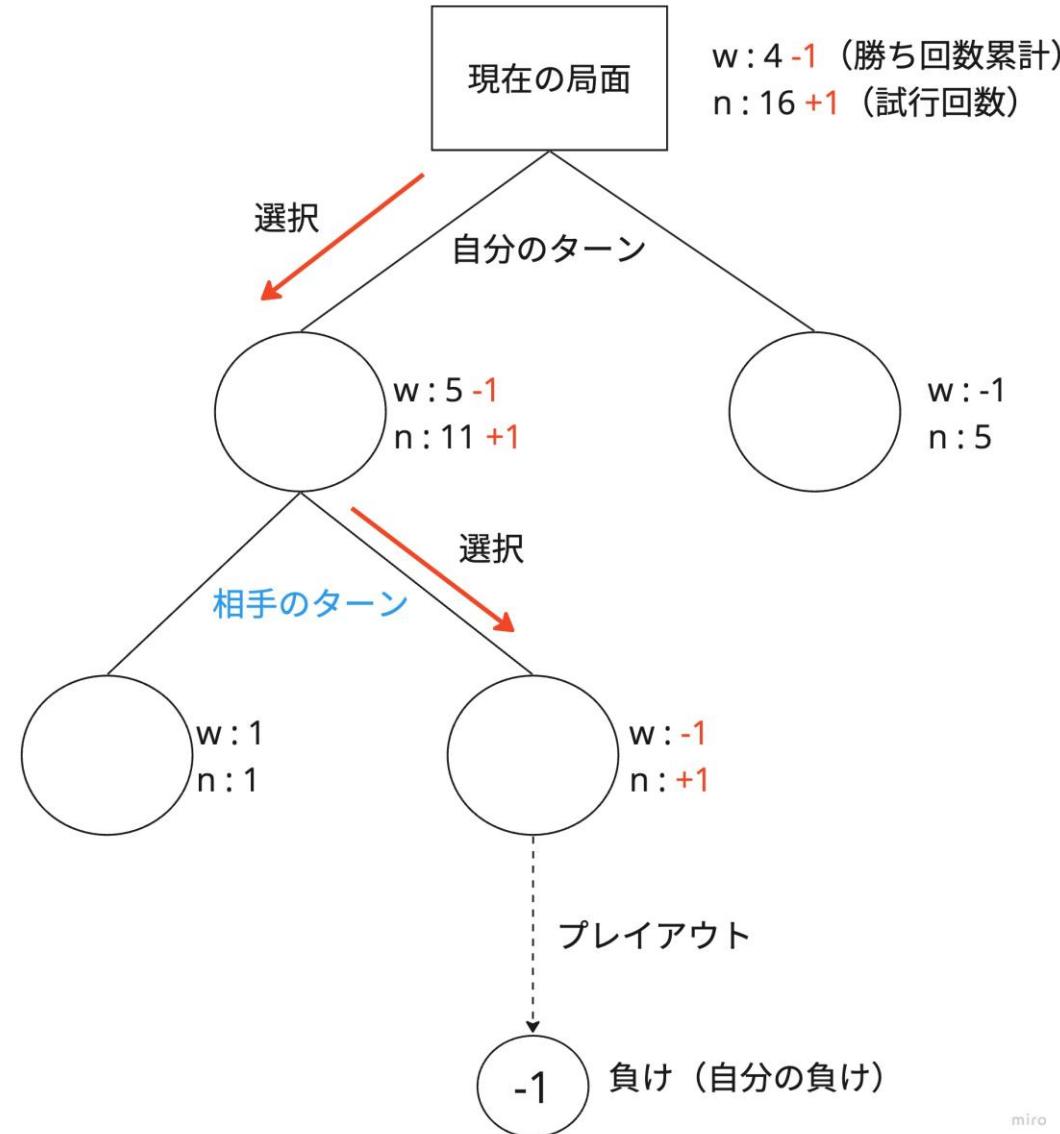


右のパターンが選択されていないため、右を選ぶ

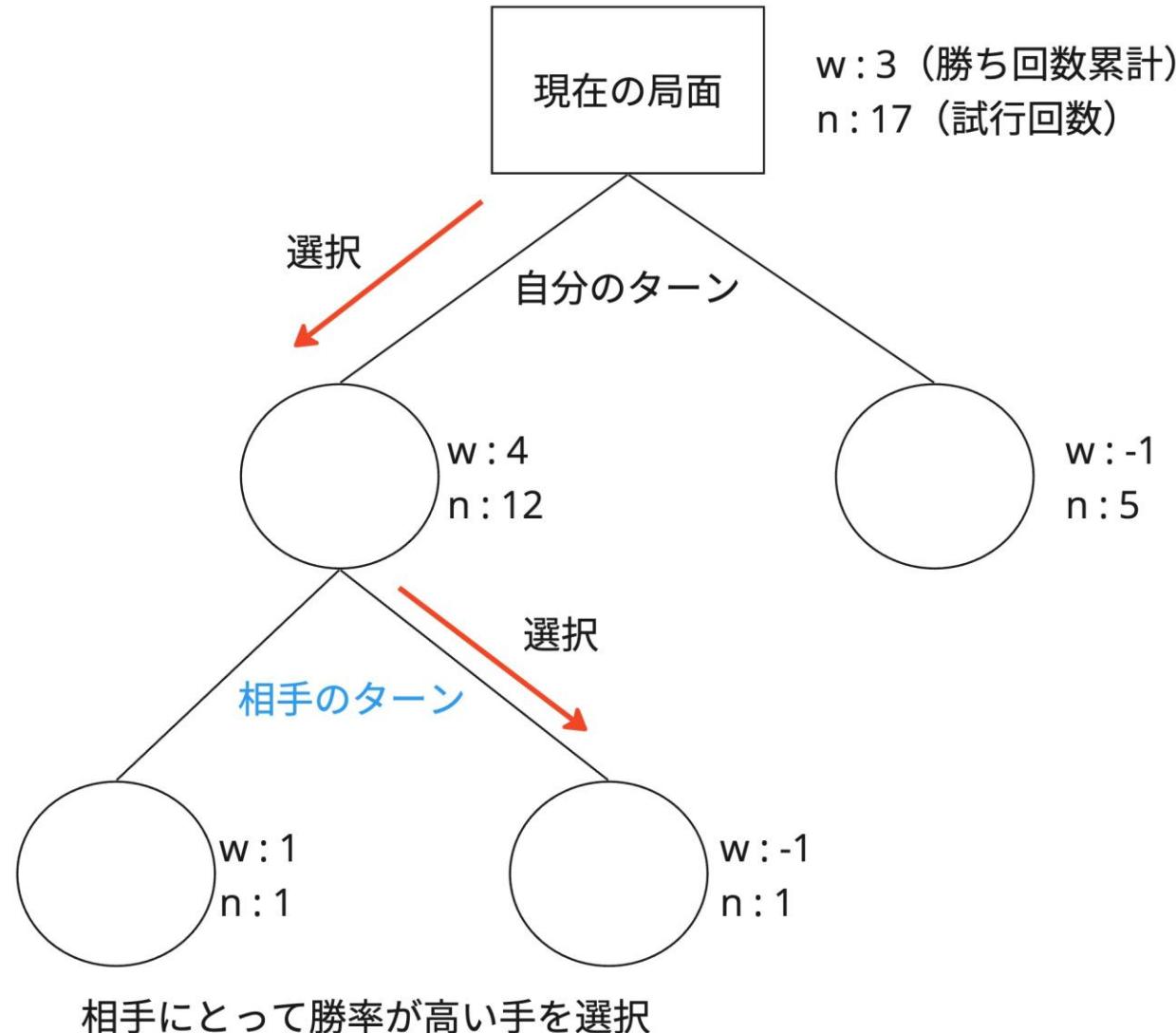
# 18回目の評価



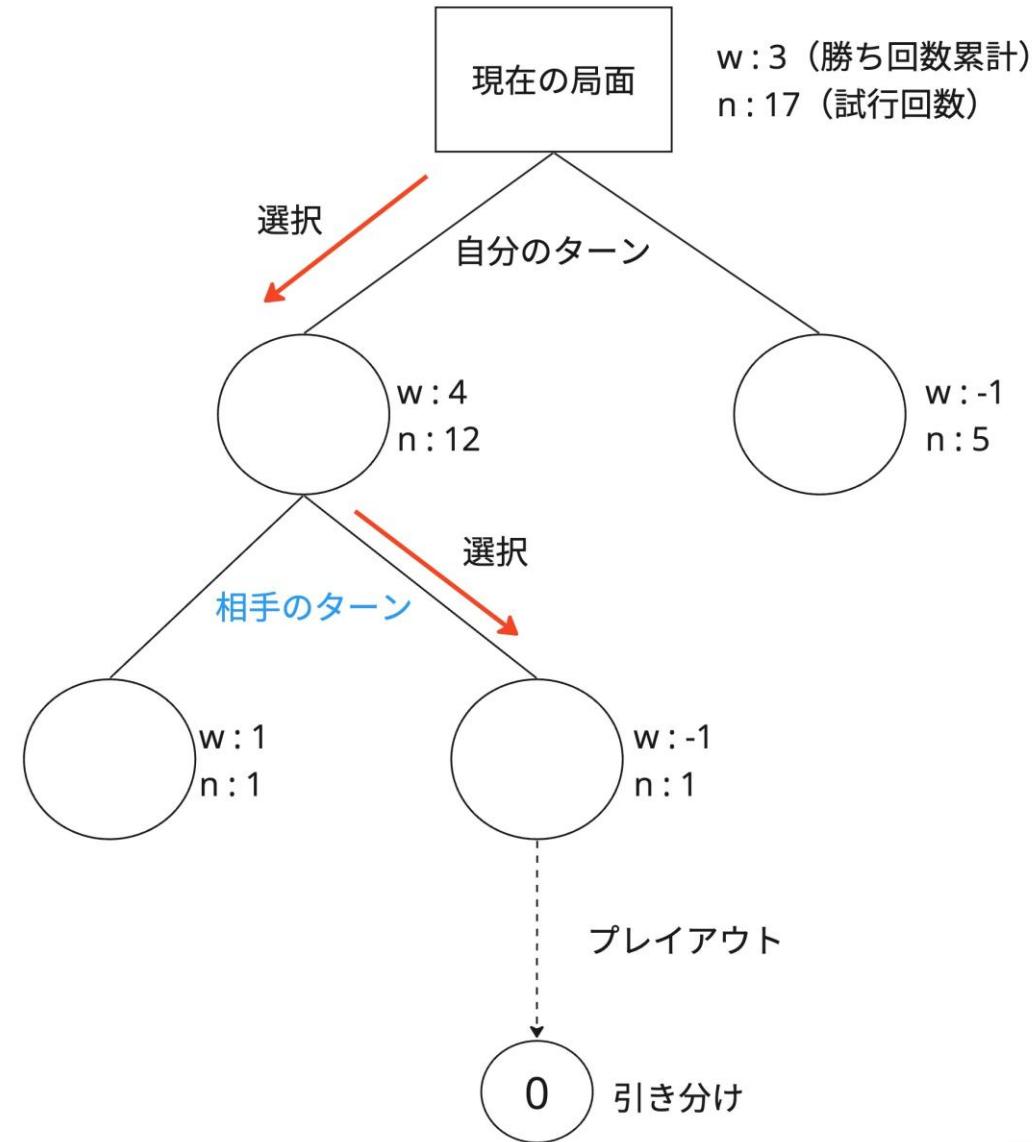
# 18回目の更新



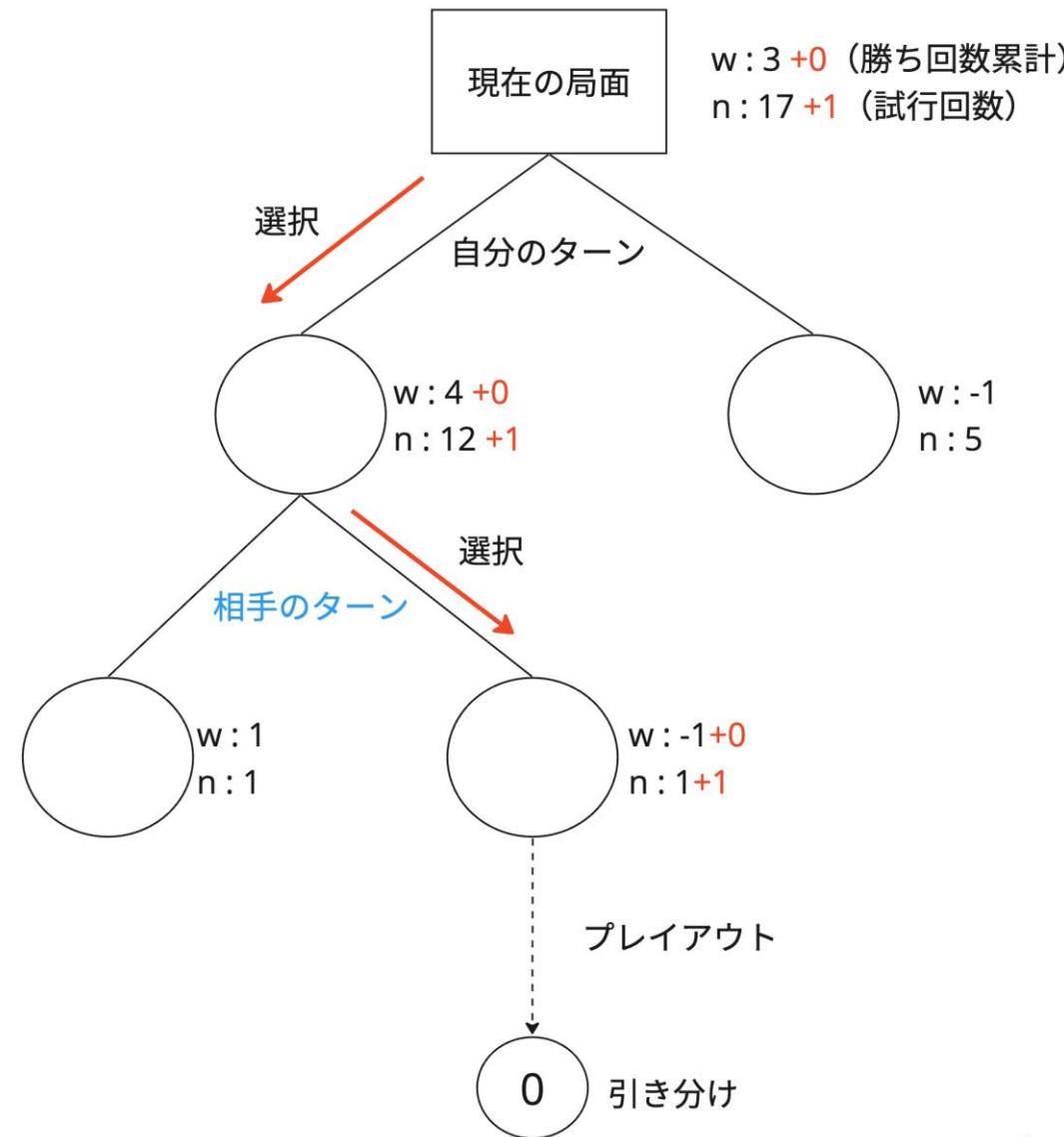
# 19回目の選択



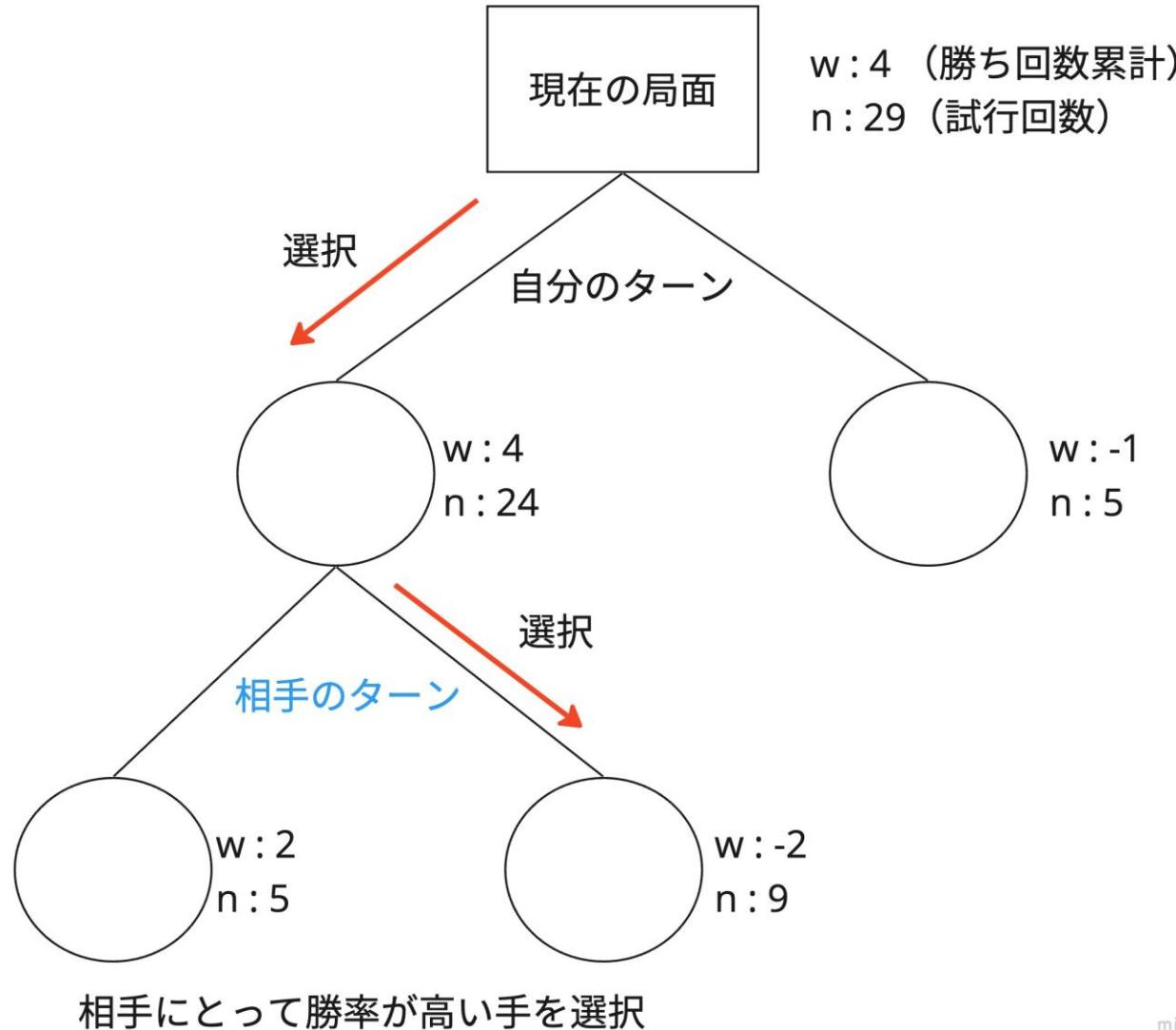
# 19回目の評価



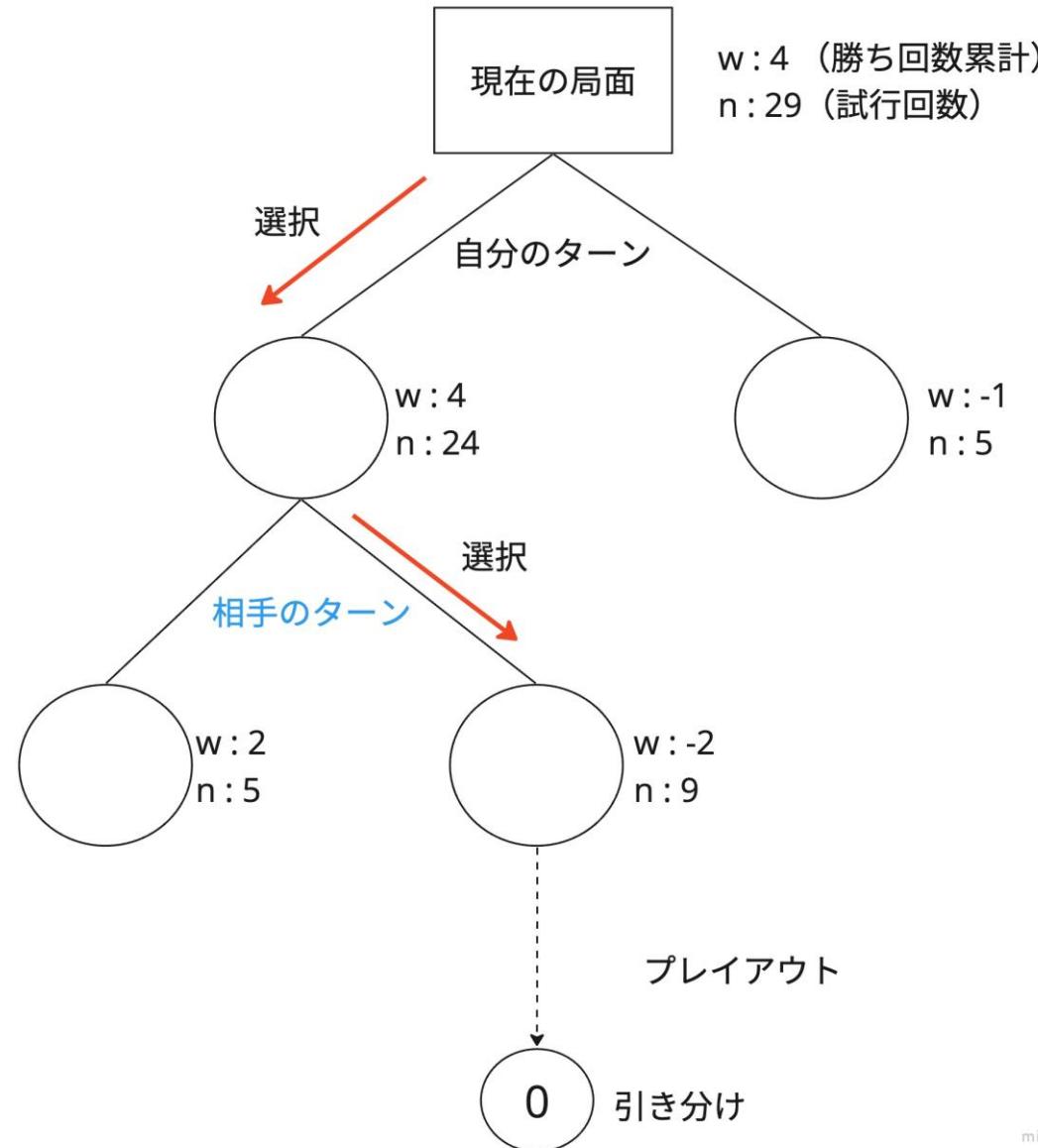
# 20回目の更新



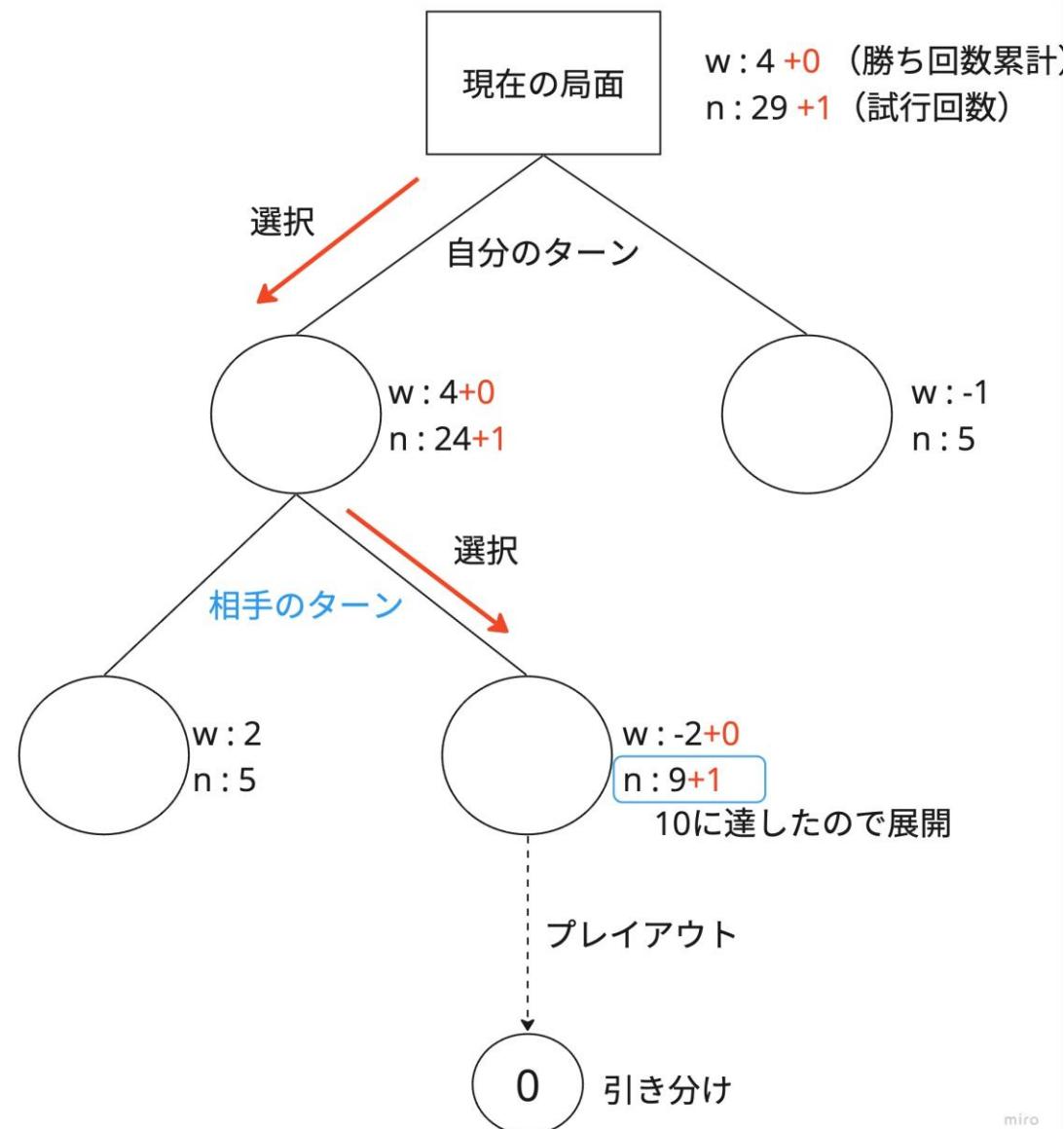
# 30回目の選択



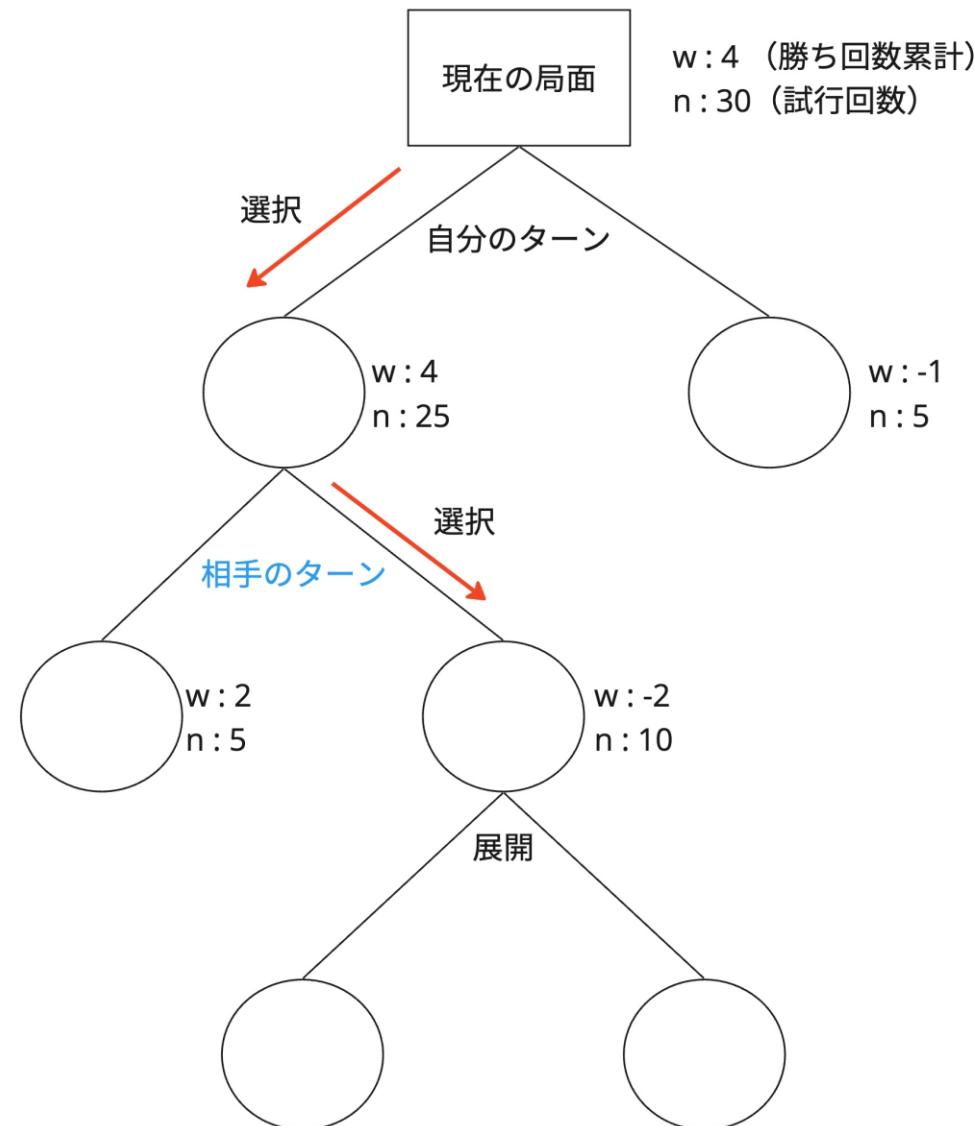
# 30回目の評価



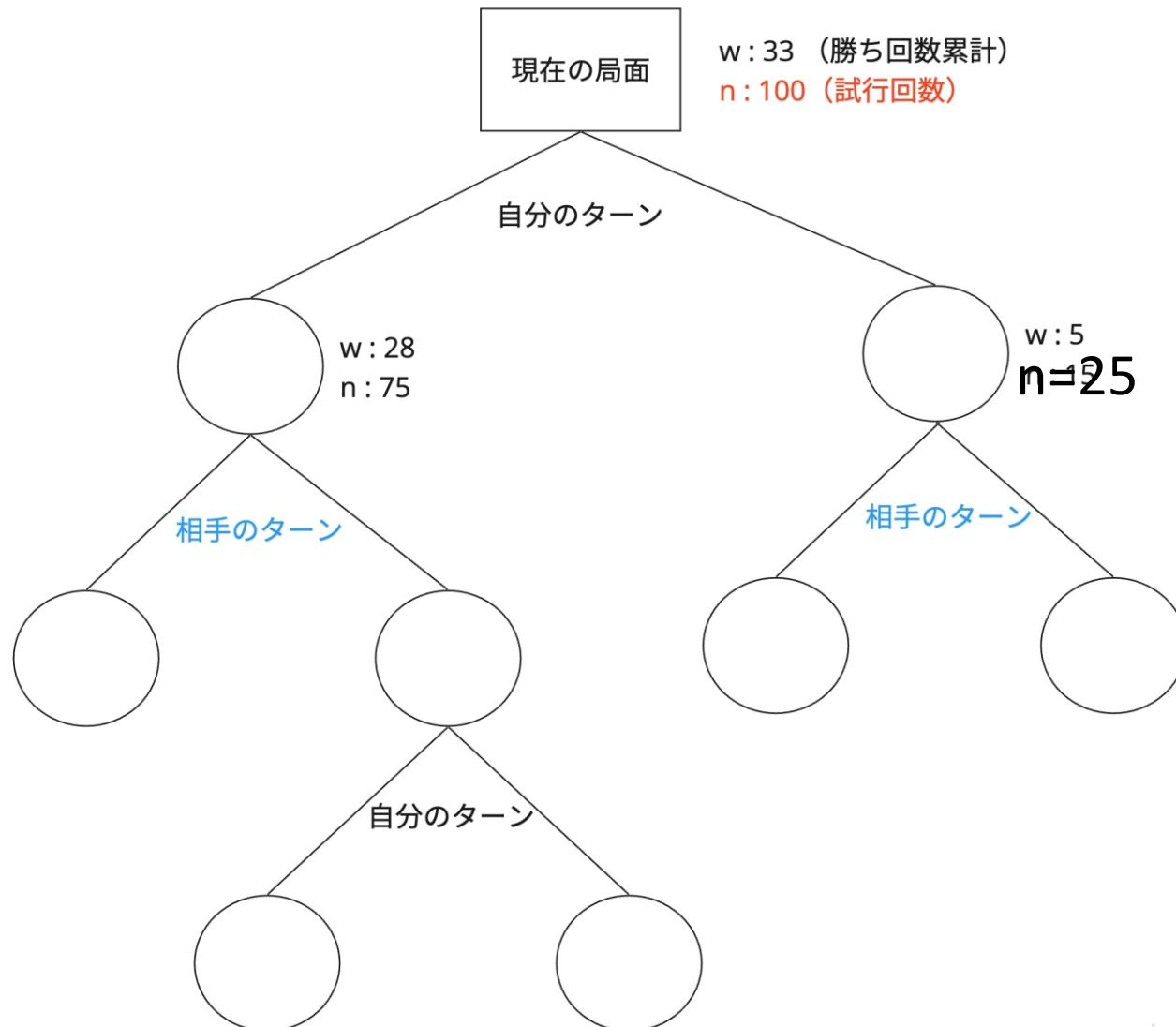
# 30回目の更新



# 第2回目の展開

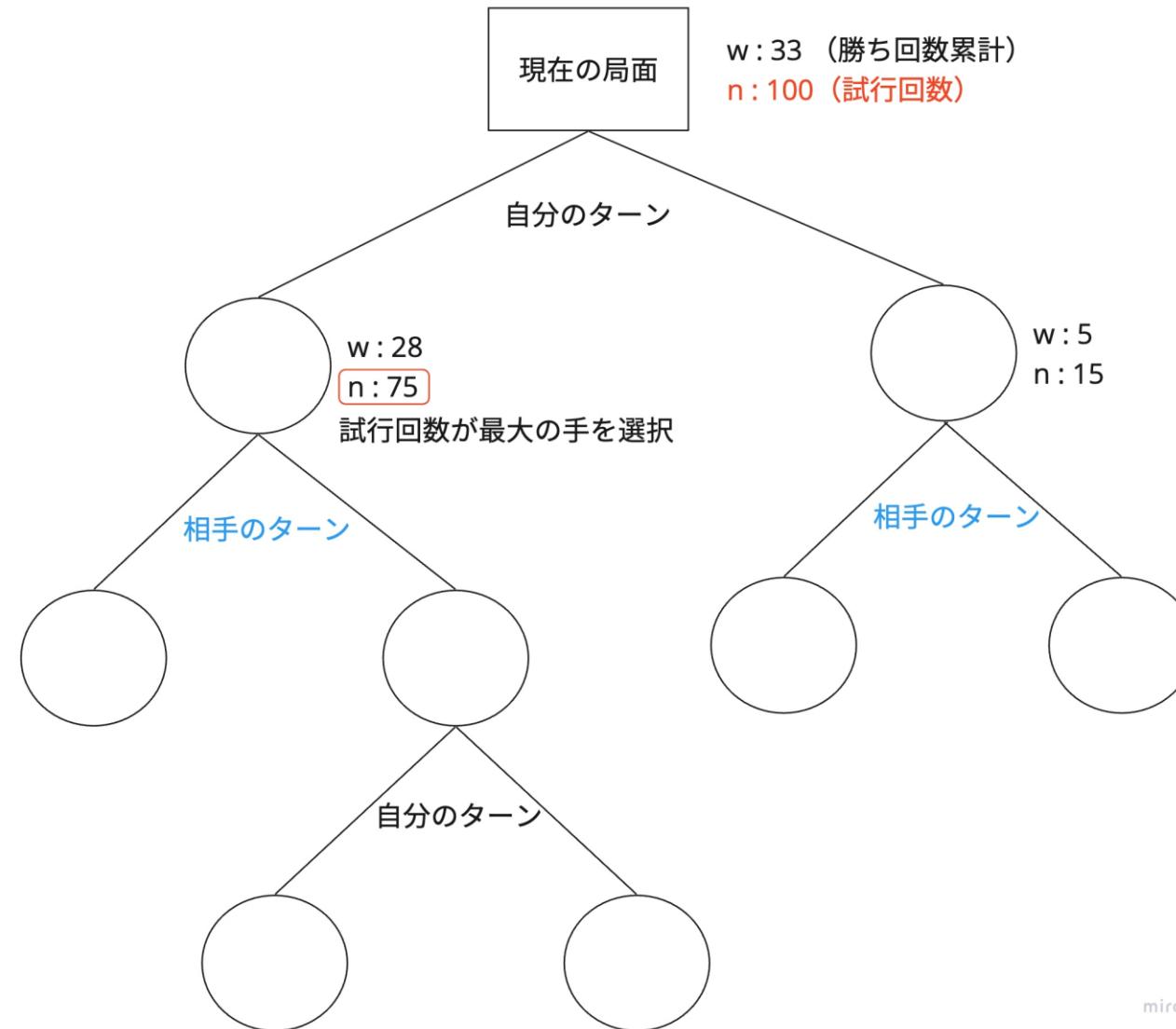


# 作成終了



- 試行回数が100に達したら作成プロセスを停止する  
※閾値の設定は自由

# 次の手の決定



End !

# 「選択」の理論

---

自分のターンの時、UCB1が最大になる手を選択する

$$UCB1 = \frac{w}{n} + \sqrt{\frac{2\log(t)}{n}}$$

$w$  = 勝ち回数  
 $n$  = 試行回数  
 $t$  = 試行回数合計

# 「選択」の理論

---

相手のターンの時は相手にとって最善の手を打ちたいので

$$UCB1 = \frac{-w}{n} + \sqrt{\frac{2\log(t)}{n}}$$

$w$  = 自分の勝ち回数

$n$  = 試行回数

$t$  = 試行回数合計

# 全体のアルゴリズムのまとめ

