# ENGINEERING A COMPILER

9.3.4 - 9.3.6

山本 航平

SSA (Static Single-Assignment) 形式のルール:

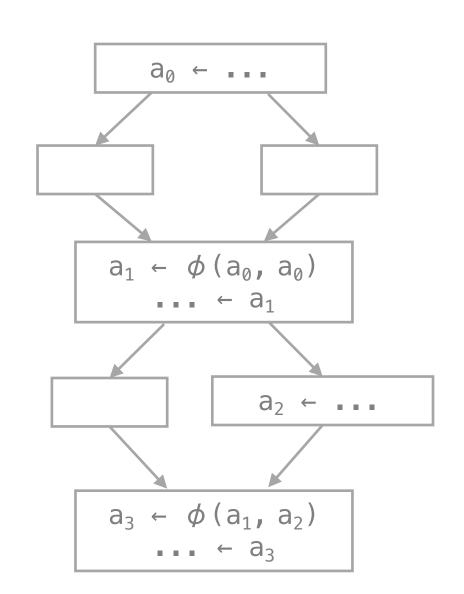
- 1. 手続き内の各計算は一意の名前を定義する
- 2. 手続き内の各使用は単一の名前を参照する

同じ変数には静的に一度しか代入されない



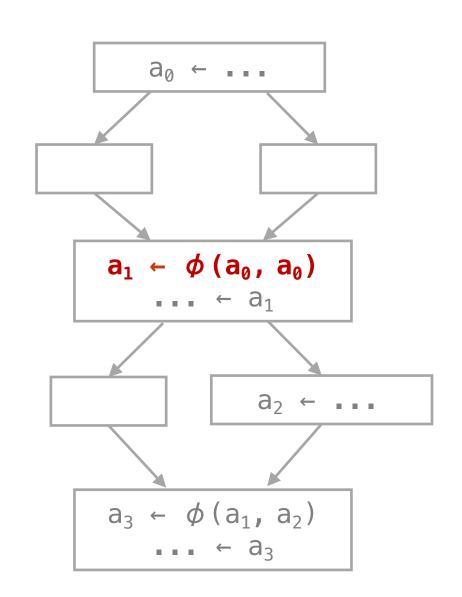
### 9.3.1: A Naive for Building SSA Form

- ・CFG の合流点すべてに  $\phi$ 関数を挿入する
  - → 無駄な φ 関数が多く挿入されるため 必要な箇所だけ挿入したい



### 9.3.1: A Naive for Building SSA Form

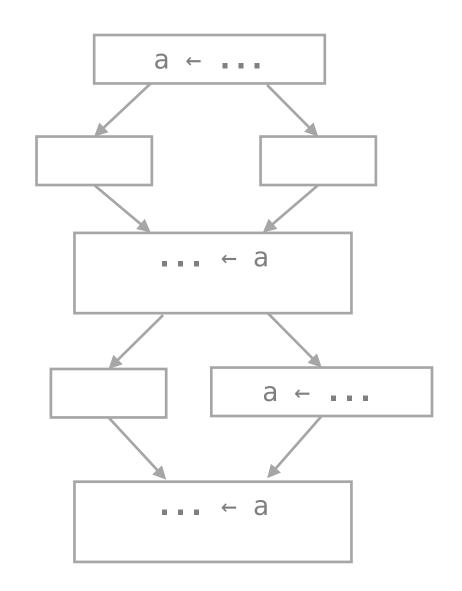
- ・CFG の合流点すべてに  $\phi$ 関数を挿入する
  - → 無駄なφ関数が多く挿入されるため 必要な箇所だけ挿入したい



支配関係を考慮した SSA を構築する

#### 9.3.2 : Dominance Frontiers

CFG の支配関係から, φ関数が必要な基本ブロックを特定する

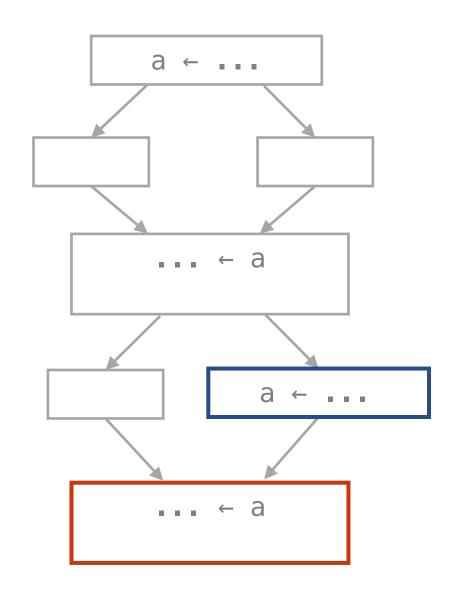


支配関係を考慮した SSA を構築する

#### 9.3.2 : Dominance Frontiers

CFG の支配関係から, φ関数が必要な基本ブロックを特定する

例: で定義された変数はで φ 関数が必要



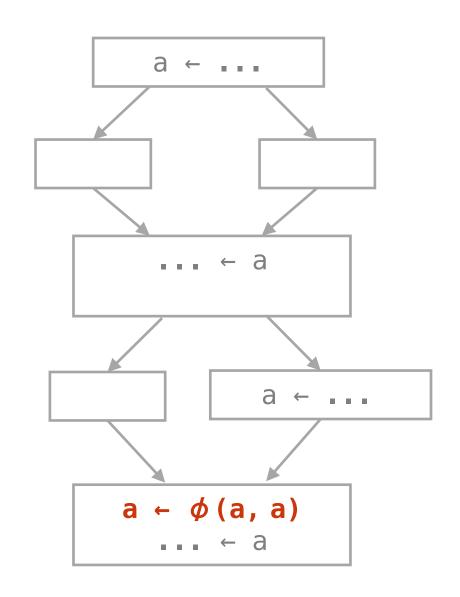
支配関係を考慮した SSA を構築する

#### 9.3.2 : Dominance Frontiers

CFG の支配関係から, φ関数が必要な基本ブロックを特定する

### 9.3.3 : Placing $\phi$ -Functions

・DF集合に基づいて  $\phi$ 関数を挿入

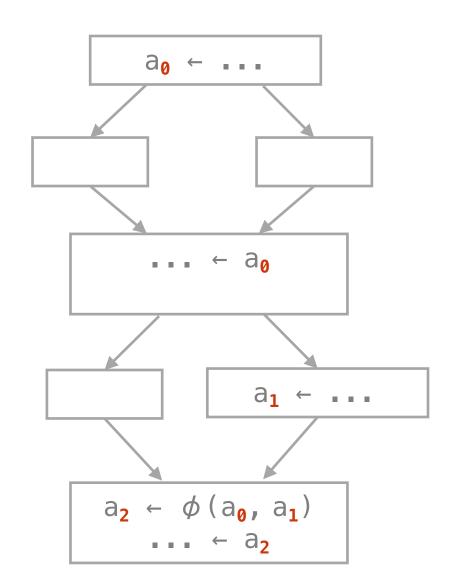


### 今回やること

支配関係を考慮した SSA を構築する

9.3.4 : Renaming

・変数名の変更



### 今回やること

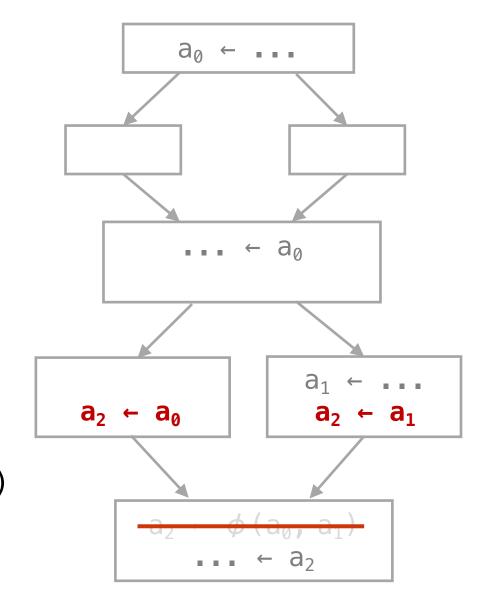
支配関係を考慮した SSA を構築する

9.3.4 : Renaming

・変数名の変更

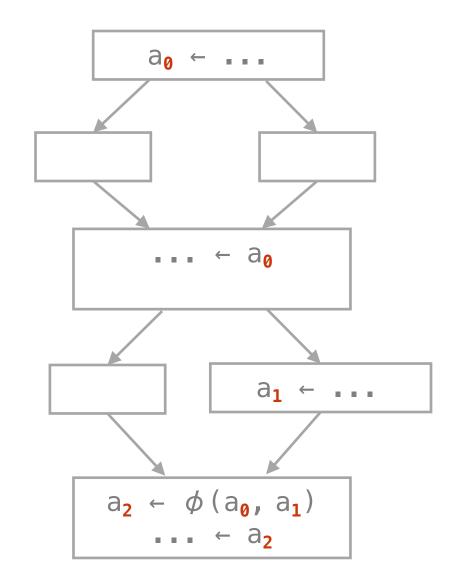
9.3.5: Translation out of SSA Form

- ・SSA を通常の形式に変換
- ・ $\phi$ 関数のない形にコードを変換 (コンピュータは $\phi$ 関数をそのまま実行できない)

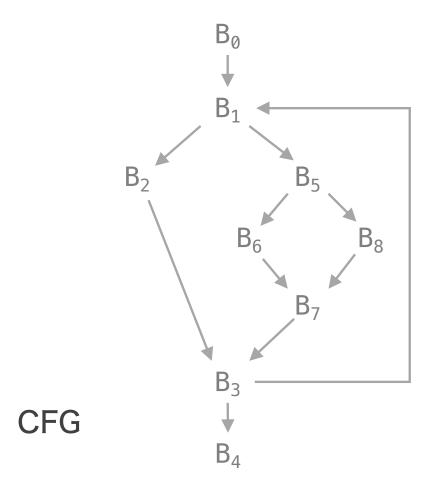


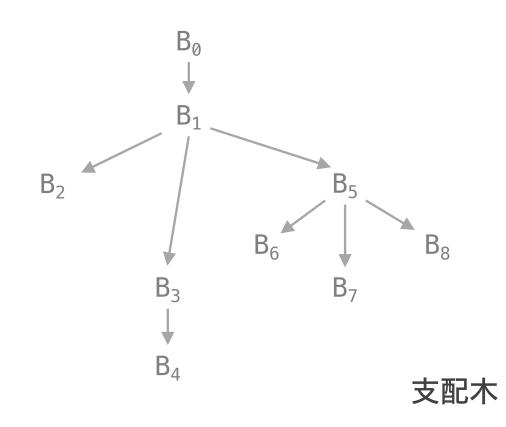
### 9.3.4 Renaming

- ・変数名の変更
- ・元々の変数名をベース名として0 → 1 → … と添え字をつける



・支配木を preorder で走査





- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

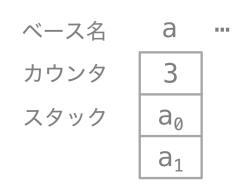
- ・定義された時にカウンタの値を添え字に
- ・値は単調増加

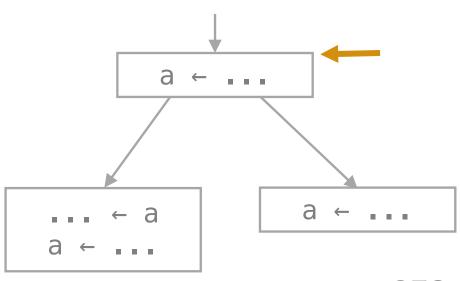
- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加



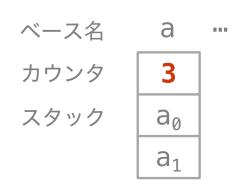


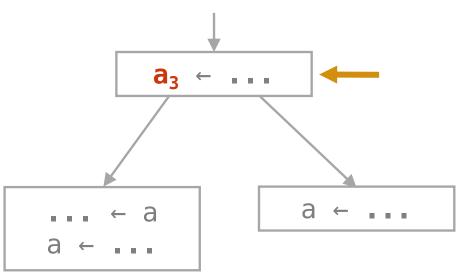
- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加



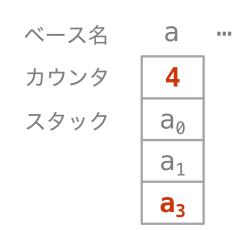


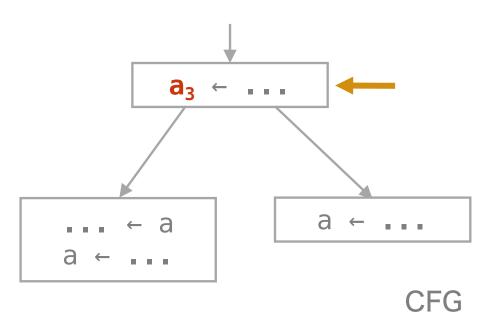
- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加



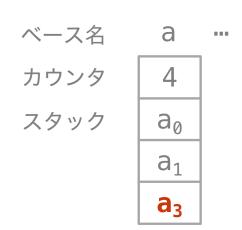


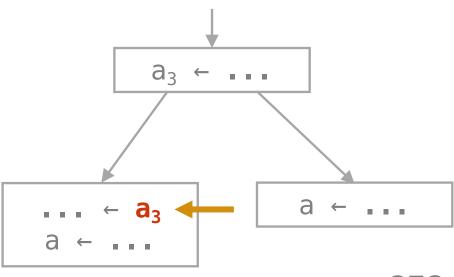
- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加





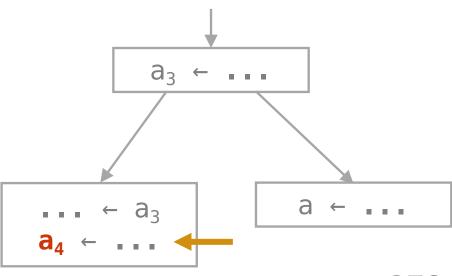
- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加



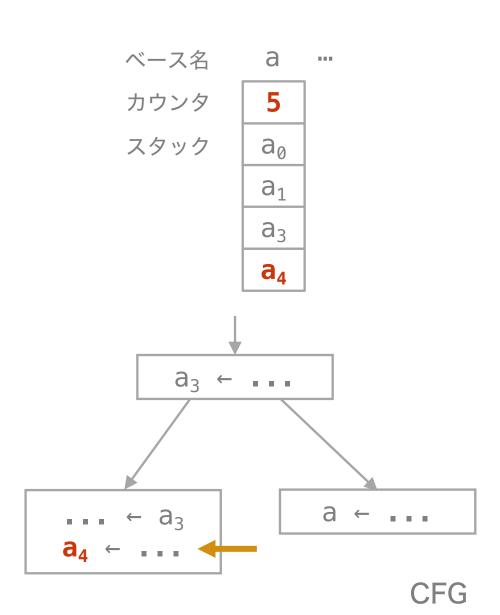


- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加

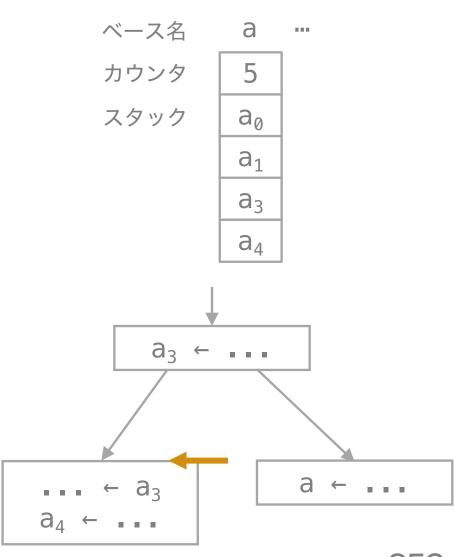


- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加

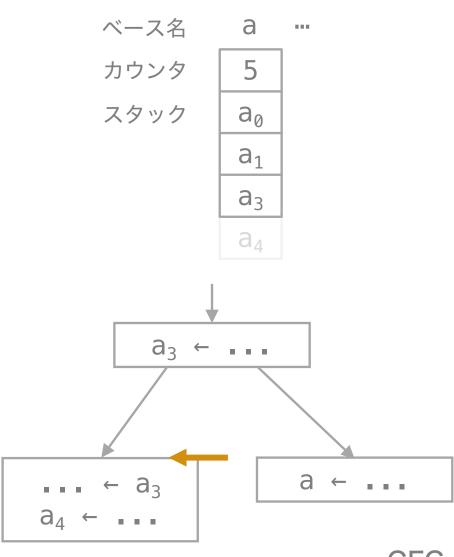


- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加

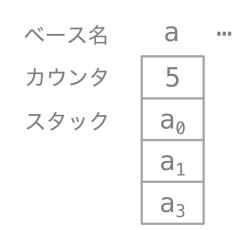


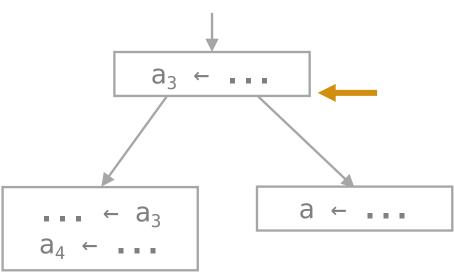
- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加





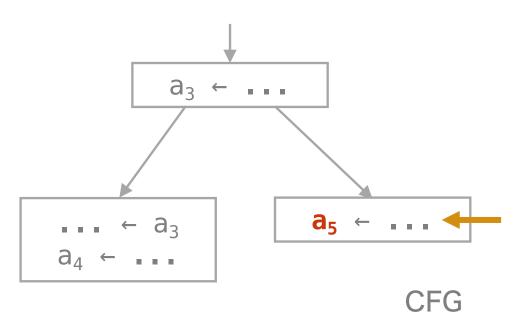
- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### スタック

- ・top が現在使用可能な添え字を表す
- ・新しく定義されたら添え字を push
- ・BB を探索し終えたら BB 内で定義された添え字を pop

- ・定義された時にカウンタの値を添え字に
- ・値は単調増加

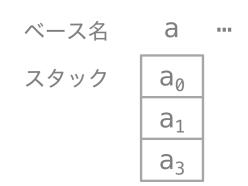


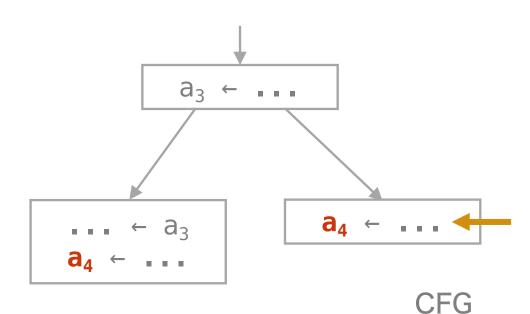


- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用

#### カウンタ

→ カウンタがなかったら同じ添え字で 定義される可能性

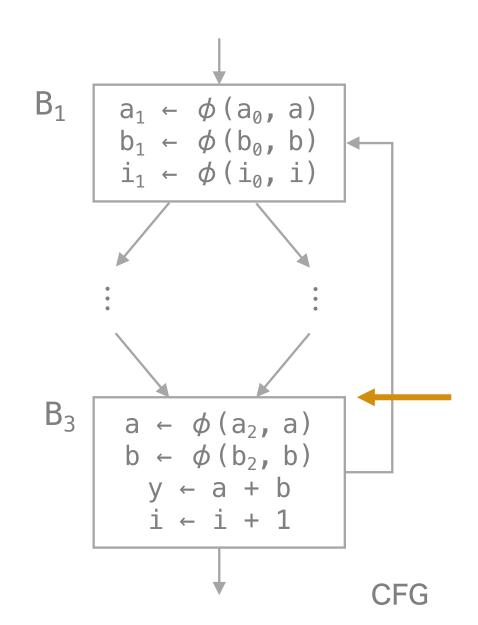




- ・支配木を preorder で走査
- ・各ベース名ごとにスタックとカウンタを使用
- ・各基本ブロックでの操作
  - 1. φ関数の定義を SSA 名に
  - 2. BB 内の各命令の使用と定義を SSA 名に
  - 3. CFG 上で後続の BB の  $\phi$ 関数のパラメータを現在の SSA 名に変更
  - 4. 再帰的に支配木上で後続の BB に移動
  - 5. BB 内で定義された SSA 名をスタックから pop

Rename(B<sub>3</sub>):

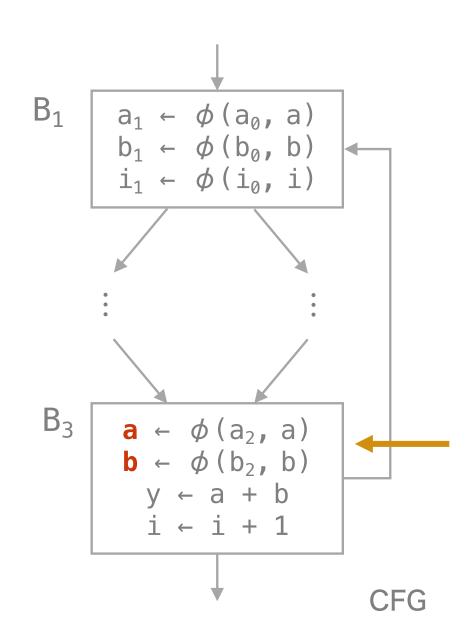
| ベース名 | а              | b              | i              |
|------|----------------|----------------|----------------|
| カウンタ | 3              | 3              | 2              |
| スタック | a <sub>0</sub> | b <sub>0</sub> | i <sub>0</sub> |
|      | a <sub>1</sub> | b <sub>1</sub> | $i_1$          |
|      | a <sub>2</sub> |                |                |



Rename  $(B_3)$ :

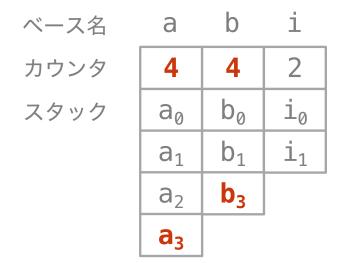
1. *ϕ* 関数の定義を SSA 名に

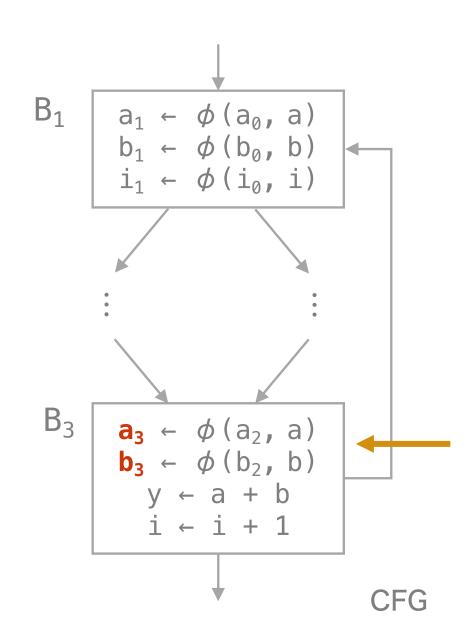
Q. a と b の添え字は?



### Rename $(B_3)$ :

1. *ϕ* 関数の定義を SSA 名に

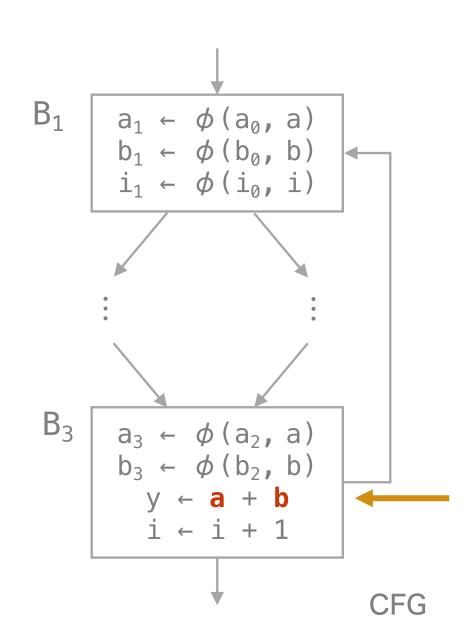




Rename  $(B_3)$ :

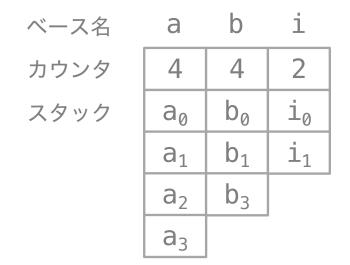
2. BB 内の各命令の使用と定義を SSA 名に

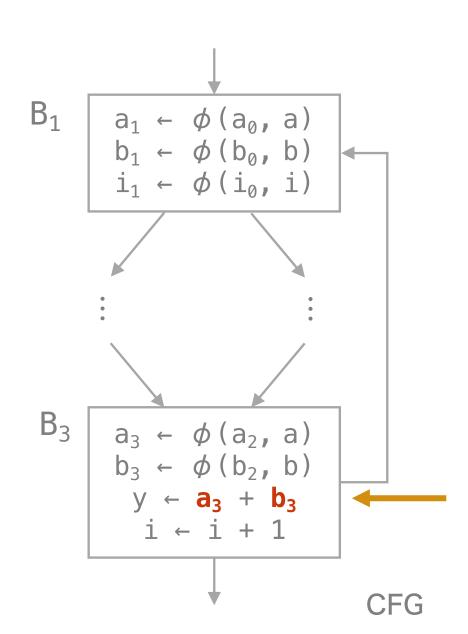
Q. a と b の添え字は?



Rename(B<sub>3</sub>):

2. BB 内の各命令の使用と定義を SSA 名に

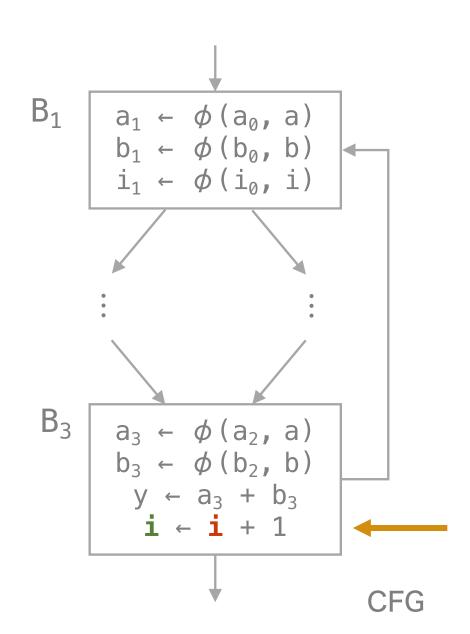




Rename  $(B_3)$ :

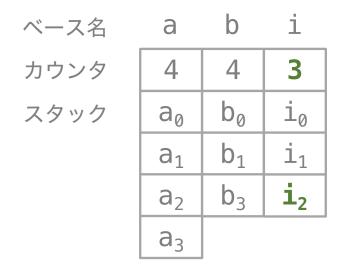
2. BB 内の各命令の使用と定義を SSA 名に

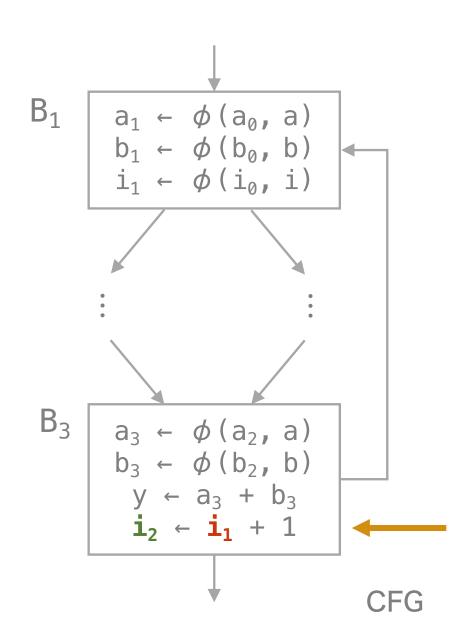
Q. i と i の添え字は?(使用 → 定義の順で SSA 名にする)



Rename(B<sub>3</sub>):

2. BB 内の各命令の使用と定義を SSA 名に

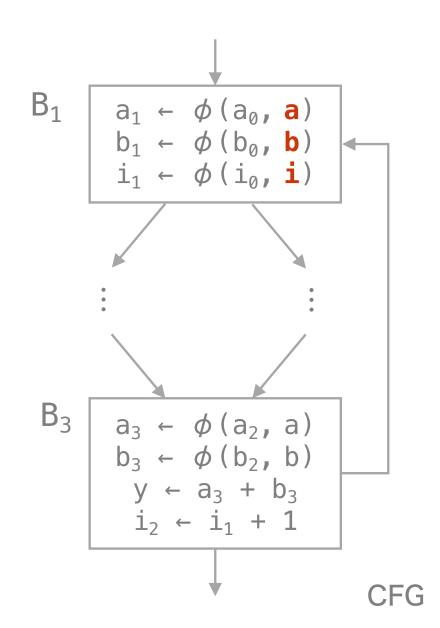




Rename(B<sub>3</sub>):

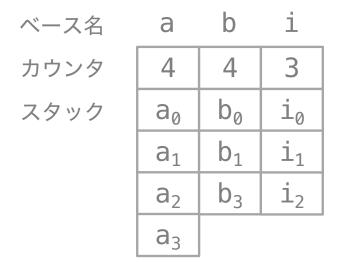
3. CFG 上で後続の BB の  $\phi$ 関数のパラメータを現在の SSA 名に変更

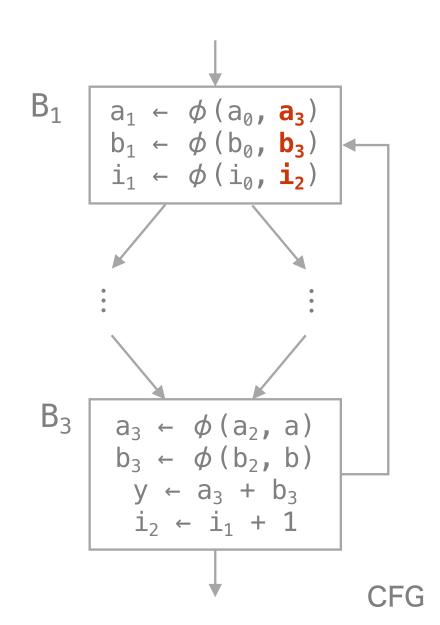
Q. a と b と i の添え字は?



### Rename $(B_3)$ :

3. CFG 上で後続の BB の  $\phi$ 関数のパラメータを現在の SSA 名に変更

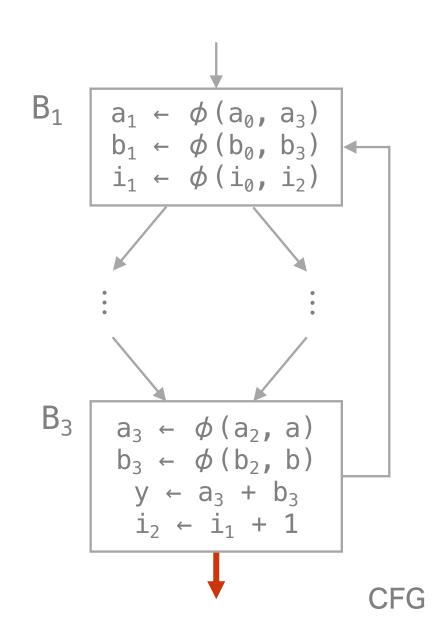




### Rename $(B_3)$ :

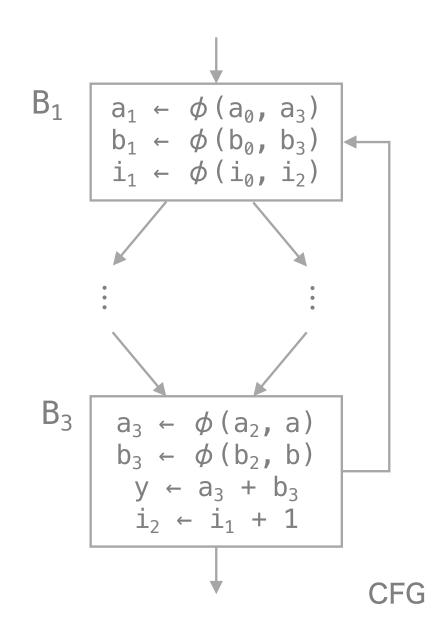
4. 再帰的に支配木上で後続の BB に移動

| ベース名 | а                     | b              | i              |
|------|-----------------------|----------------|----------------|
| カウンタ | 4                     | 4              | 3              |
| スタック | a <sub>0</sub>        | b <sub>0</sub> | i <sub>0</sub> |
|      | $a_1$                 | b <sub>1</sub> | i <sub>1</sub> |
|      | a <sub>2</sub>        | b <sub>3</sub> | $i_2$          |
|      | <b>a</b> <sub>3</sub> |                |                |



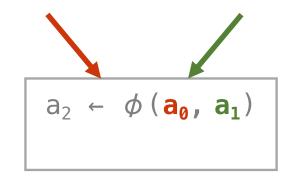
Rename  $(B_3)$ :

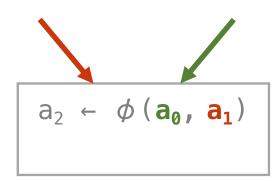
5. BB 内で定義された SSA 名をスタックから pop



### **ø関数のパラメータ**

- 3. CFG 上で後続の BB の  $\phi$ 関数のパラメータを現在の SSA 名に変更
  - → どのパラメータを変更するか知る必要がある
- ・CFG の実装と SSA の構築で一貫したルールを定める
  - → CFG エッジがリストで保存されているなら, その順序で引数を決定する
- ・教科書では図の CFG エッジの 左  $\rightarrow$  右が,  $\phi$  関数の引数の 左  $\rightarrow$  右 に対応





# A Final Improvement

・スタックには最新の名前だけ push すれば良い ex.)  $a_1$  はスタックに乗せる必要なし

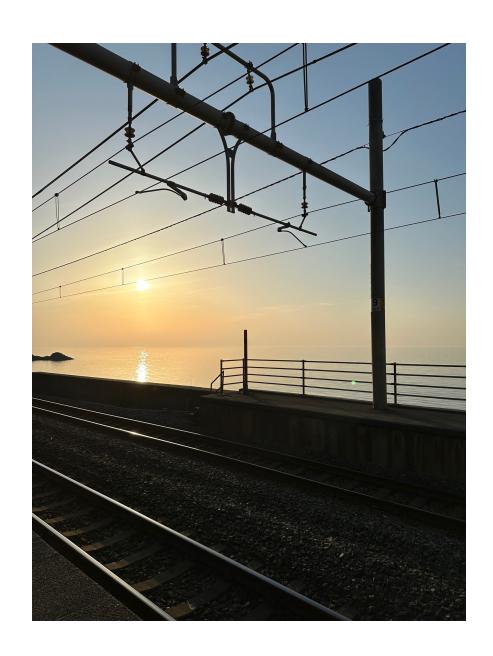
$$a_{1} \leftarrow \phi(a_{0}, a_{3})$$

$$b_{1} \leftarrow \phi(b_{0}, b_{3})$$

$$i_{1} \leftarrow \phi(i_{0}, i_{2})$$

$$a_{2} \leftarrow \cdots$$

- ・ブロック内で定義されたベース名1つにつき push と pop 1回ずつ
  - → スタック操作の時間削減
  - → スタックのスペース削減・オーバーフロー回避(スタックの深さは必ず支配木の深さ以下になる)

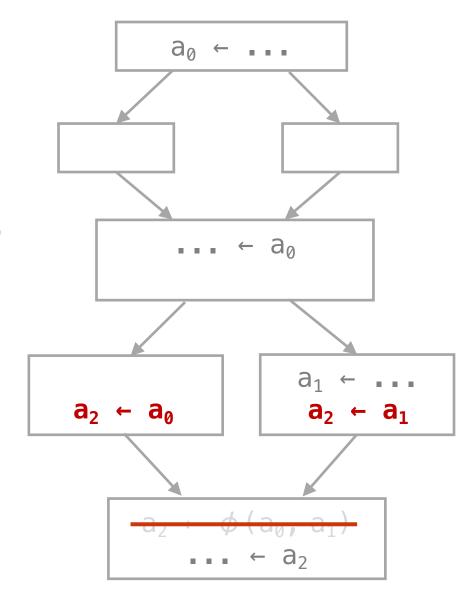


新潟県 青梅川駅

### 9.3.5 Translation out of SSA Form

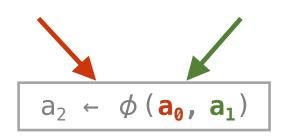
#### SSA 逆変換:

- ・SSA を通常の形式に変換
- ・ $\phi$ 関数のない形にコードを変換 (コンピュータは $\phi$ 関数をそのまま実行できない)
  - ightarrow  $\phi$  関数のセマンティクスを満たすように  $\phi$  関数をコピー演算に置き換えたい

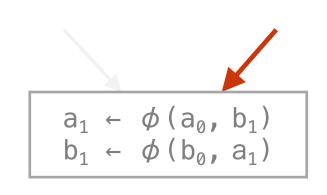


### φ関数のセマンティクス

1. どのエッジから来たかによって値が選択される



2. 同じブロック内の  $\phi$ 関数は **並列に** 計算される





 $b_1$  の値を  $a_1$  に 元々の  $a_1$  の値を  $b_1$  に ( $a_1$  と  $b_1$  の値がスワップ)



 $b_1$  の値を  $a_1$  に その  $a_1$  の値を  $b_1$  に

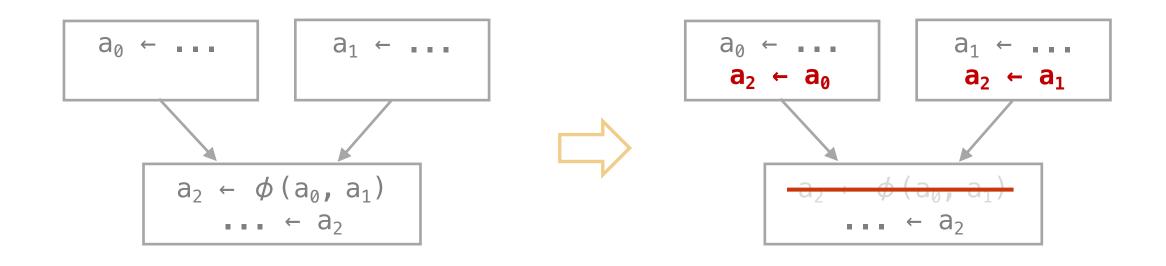
### 9.3.5 Translation out of SSA Form

#### <u>説明の流れ</u>:

- 1. 素朴な SSA 逆変換
- 2. 素朴な SSA 逆変換の 2つの問題点
- 3. 問題点を解決する SSA 逆変換

### 素朴な SSA 逆変換

・CFG の前のノードに、適切な $\phi$ 関数の引数を $\phi$ 関数で定義された変数に コピーする操作を挿入



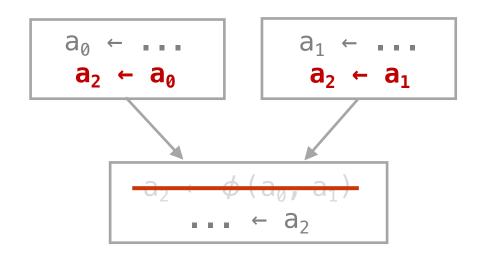
### 素朴な SSA 逆変換

・CFG の前のノードに、適切な $\phi$ 関数の引数を $\phi$ 関数で定義された変数に コピーする操作を挿入

#### 問題点:

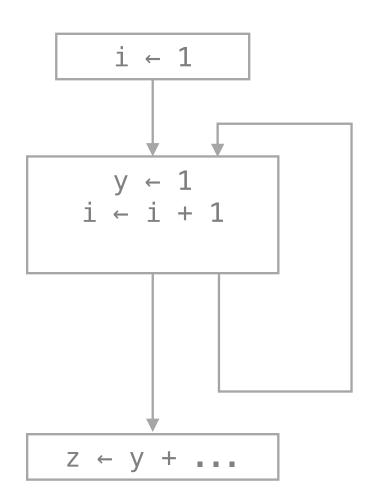
正しくないコードを生成する可能性がある

- 1. The Lost-Copy Problem
- 2. The Swap Problem

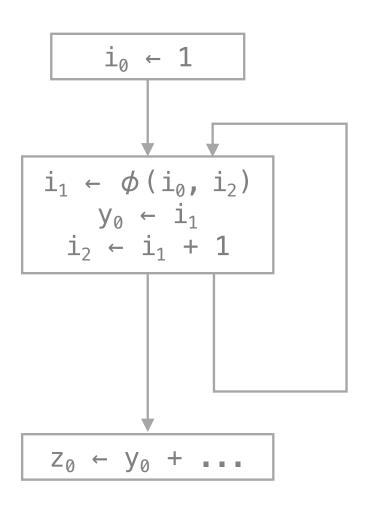


#### 元々のコード

- ・ループ内で i をインクリメントする
- ・ループ後の z の計算にはi の最後から2番目の値が使われる



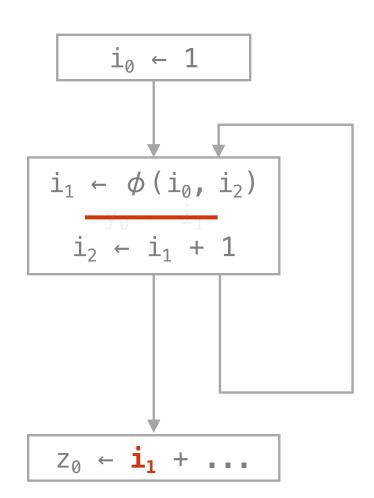
**Pruned SSA** 



#### **After Copy Folding**

#### Copy Folding (コピー畳み込み):

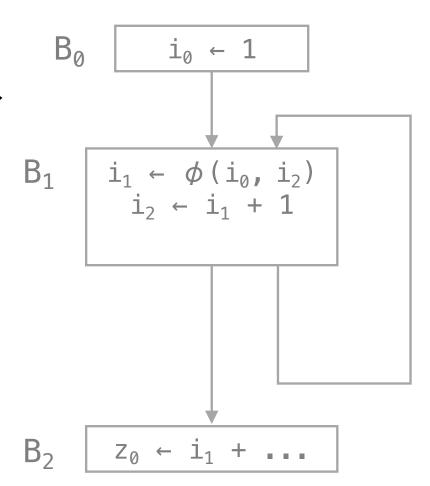
source と destination の名前を変更することで 不必要なコピー操作を削除する最適化



#### 素朴な SSA 逆変換

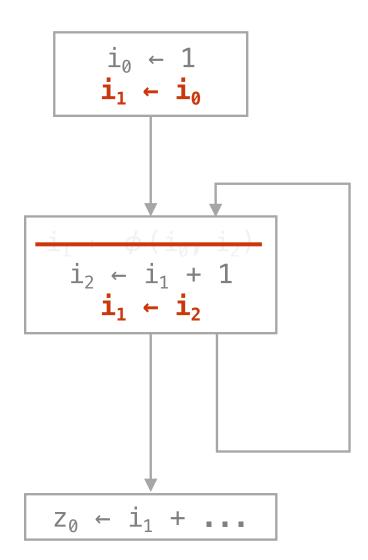
 $\phi$ 関数の前の BB に適切なコピー操作を挿入

Q. どこに何が挿入される?



#### 素朴な SSA 逆変換

 $\phi$ 関数の前の BB に適切なコピー操作を挿入



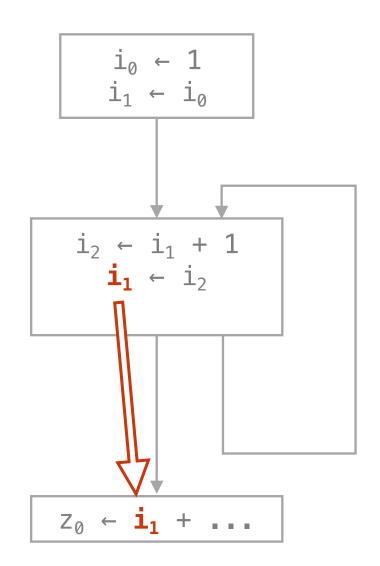
#### 素朴な SSA 逆変換

φ関数の前の BB に適切なコピー操作を挿入

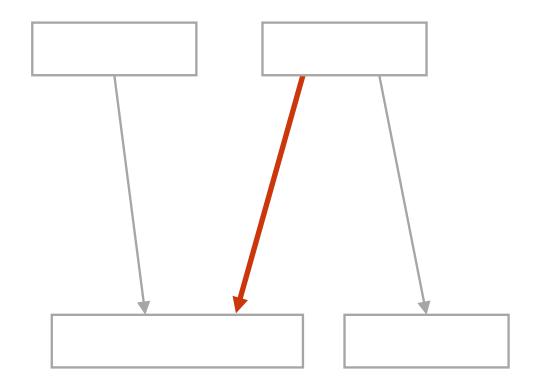
#### <u>元々のコード</u>

- ・ループ内で i をインクリメントする
- ・ループ後の z の計算にはi の最後から2番目の値が使われる

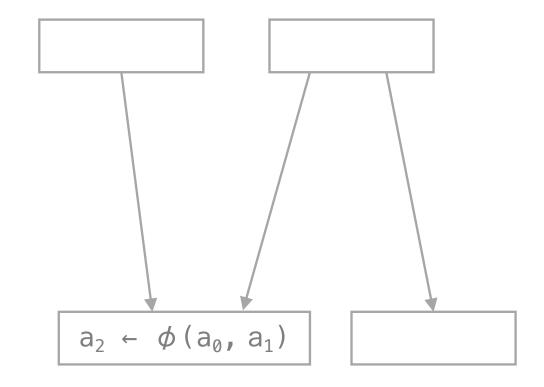
→ 元々のコードとは違うコードが生成



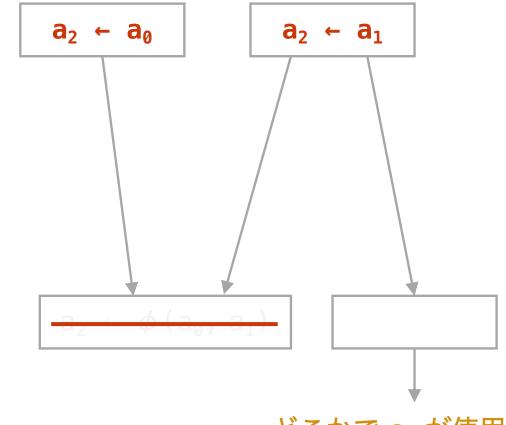
・始点のノードが複数の子を持ち、終点のノードが複数の親を持つ辺



- ・始点のノードが複数の子を持ち、終点のノードが複数の親を持つ辺
- ・始点のノードへのコピー操作の挿入は live な変数の値を変更する可能性がある

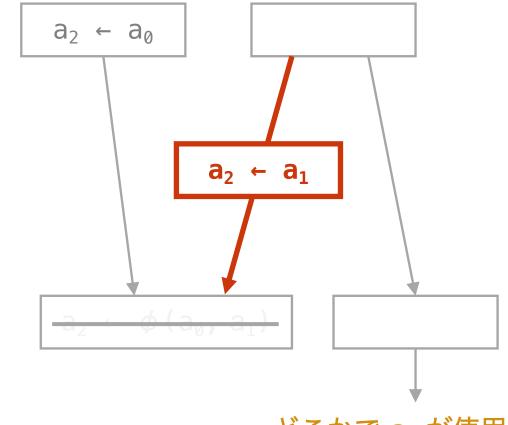


- ・始点のノードが複数の子を持ち、終点のノードが複数の親を持つ辺
- ・始点のノードへのコピー操作の挿入は live な変数の値を変更する可能性がある



どこかで **a**<sub>2</sub> が使用 される可能性

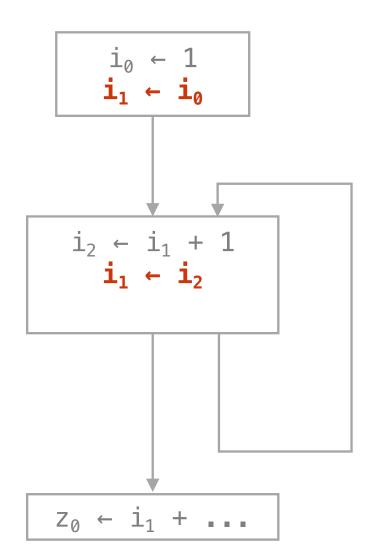
- ・始点のノードが複数の子を持ち、終点のノードが複数の親を持つ辺
- ・始点のノードへのコピー操作の挿入は live な変数の値を変更する可能性がある
- ・クリティカルエッジを分割できればこの問題は解決する



どこかで **a**<sub>2</sub> が使用 される可能性

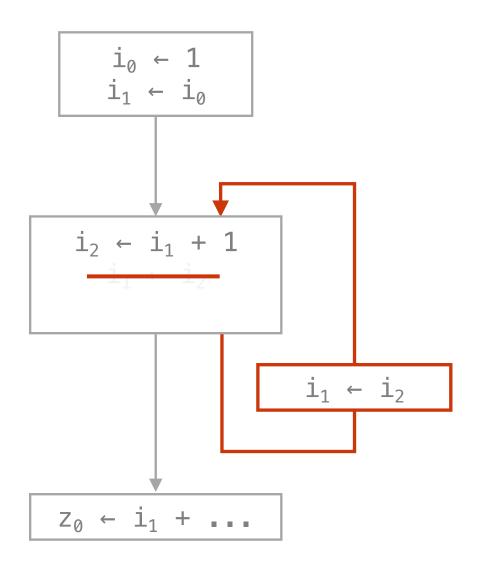
#### 素朴な SSA 逆変換

 $\phi$ 関数の前の BB に適切なコピー操作を挿入



#### クリティカルエッジの分割

元々のコードと同じ意味のコードが生成



### クリティカルエッジの分割

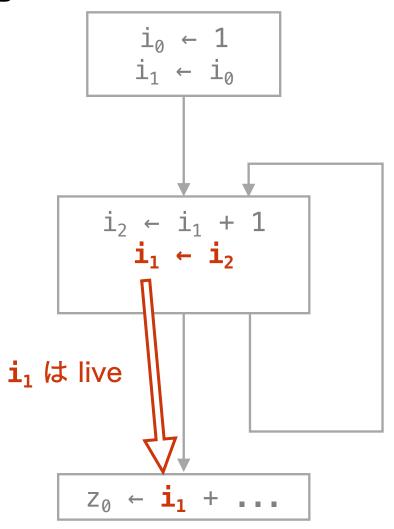
- ・すべてのクリティカルエッジを分割できれば、素朴な方法でも 正しいコードを生成する
- ·CFG 上のクリティカルエッジを分割できない・すべきでない状況

#### 素朴な SSA 逆変換

φ関数の前の BB に適切なコピー操作を挿入

#### クリティカルエッジを分割できない場合

- ・挿入ポイントでコピーのターゲットが 生きているか確認
- ・生きていれば新しい名前を導入

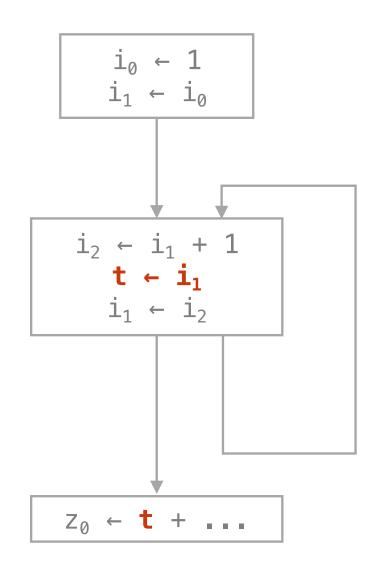


#### 素朴な SSA 逆変換

φ関数の前の BB に適切なコピー操作を挿入

#### クリティカルエッジを分割できない場合

- ・挿入ポイントでコピーのターゲットが 生きているか確認
- ・生きていれば新しい名前を導入



#### <u>φ関数のセマンティクス</u>

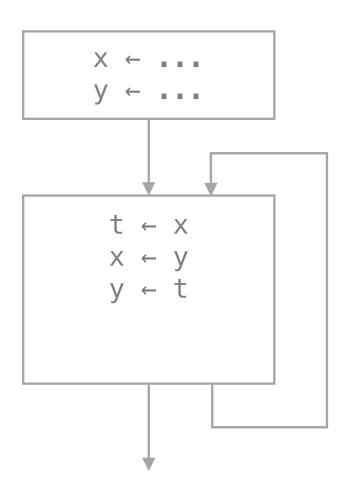
- 1. どのエッジから来たかによって値が選択される
- 2. 同じブロック内の  $\phi$ 関数は **並列に** 計算される

#### 素朴な SSA 逆変換の問題点

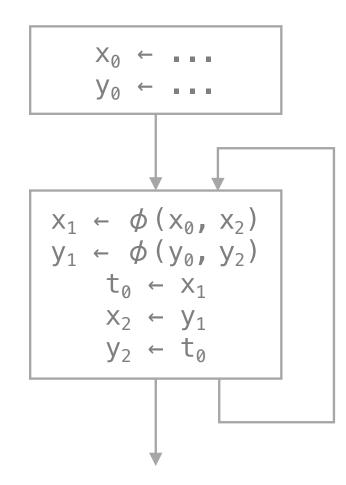
素朴な SSA 逆変換は 並列な Ø 関数 を 逐次コピー操作 に置き換える

#### 元々のコード

·x と y の値をスワップする



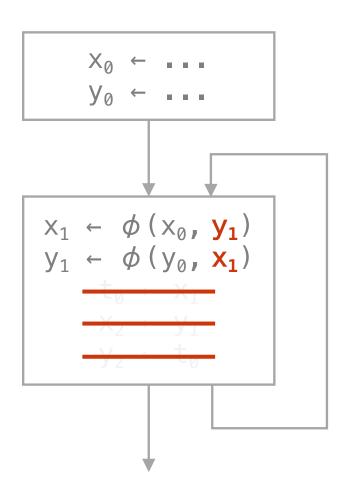
#### **Pruned SSA**



#### **After Copy Folding**

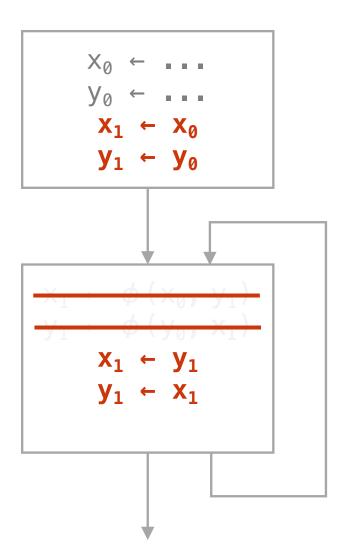
#### Copy Folding (コピー畳み込み):

source と destination の名前を変更することで 不必要なコピー操作を削除する最適化



#### 素朴な SSA 逆変換

 $\phi$ 関数の前の BB に適切なコピー操作を挿入

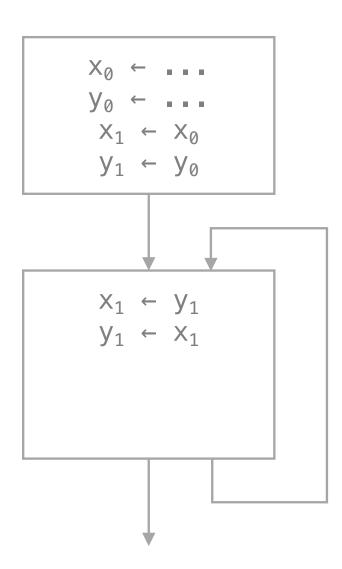


#### 素朴な SSA 逆変換

φ関数の前の BB に適切なコピー操作を挿入

#### 元々のコード

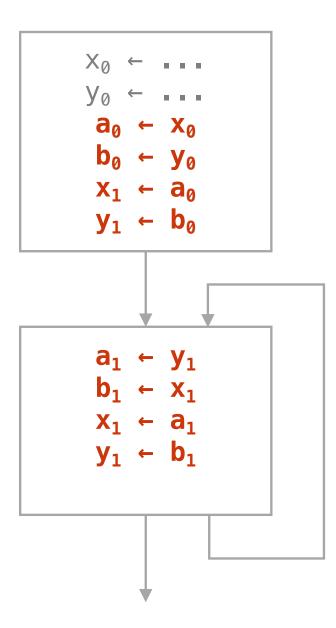
- ·x と y の値をスワップする
  - → 元々のコードとは違うコードが生成



#### 単純な解決法

 $\phi$  関数の各引数を一時的な変数にコピーする

→ 必要なコピー操作の数が2倍

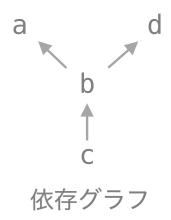


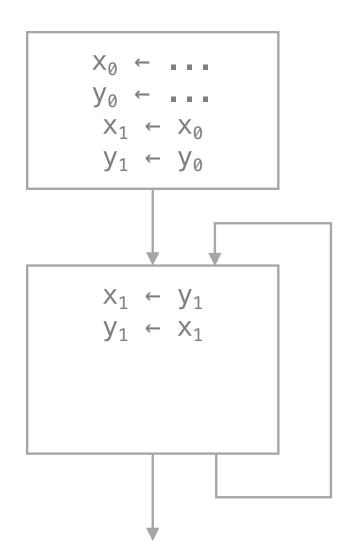
#### 依存グラフを用いた解決法

・余分なコピーを減らすことができる

#### Dependence graph(依存グラフ):

定義から使用に値の流れを表したグラフ

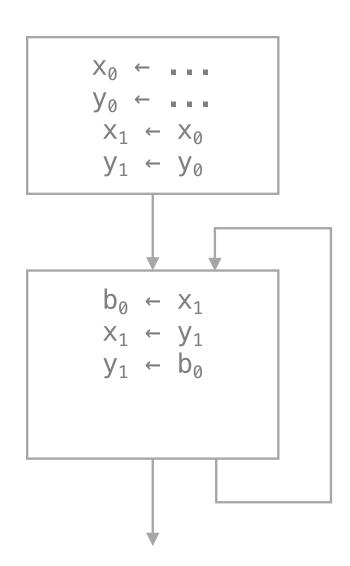




#### 依存グラフを用いた解決法

- ・余分なコピーを減らすことができる
- ・依存グラフにサイクルが含まれるときサイクルをなくすように新しい変数を導入



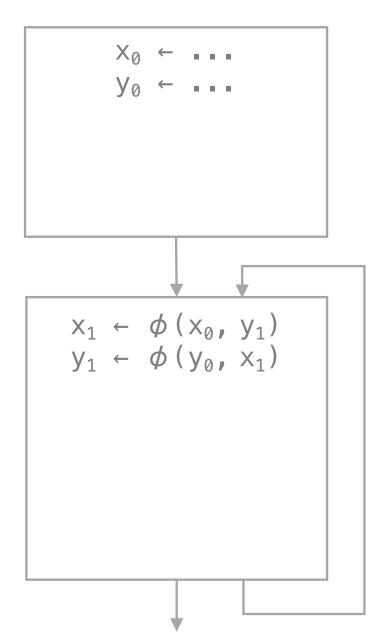


### A Unified Approach to Out-of-SSA Translation

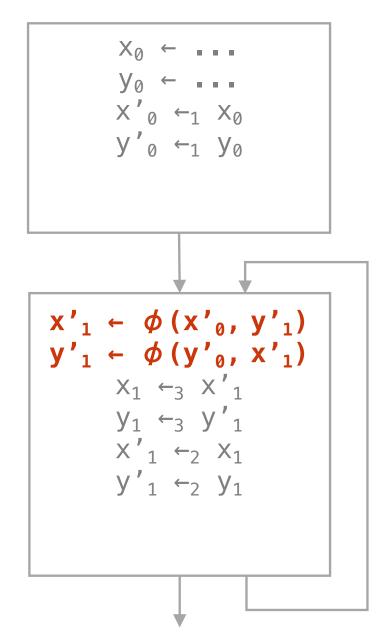
- ・素朴な SSA 逆変換とその問題点を2つ見てきた
  - The Lost-Copy Problem
  - The Swap Problem
- ・前に見た解決法は美しくないので、統一された SSA 逆変換の アルゴリズムを説明する

#### Pruned SSA Form, Copies Folded

・この例も x と y の値をスワップするプログラム



· φ関数の名前空間を分離する

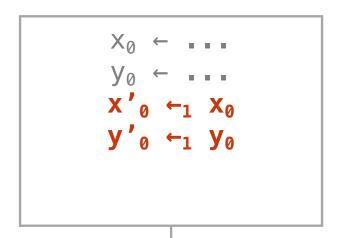


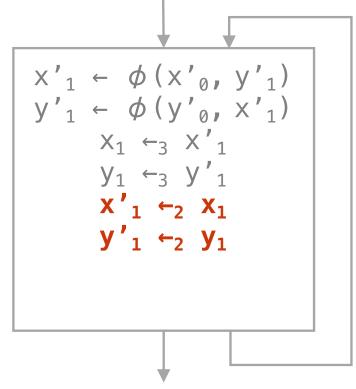
- · φ関数の名前空間を分離する
  - ・引数をコードの名前空間と対応させる
    - $\rightarrow \phi$  関数の前の BB の最後にコピー操作を追加

#### 並列コピーグループ: ←i

同じグループの操作は並列に実行される

→ 命令順を入れ替えても結果は同じ



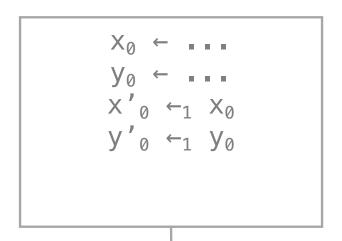


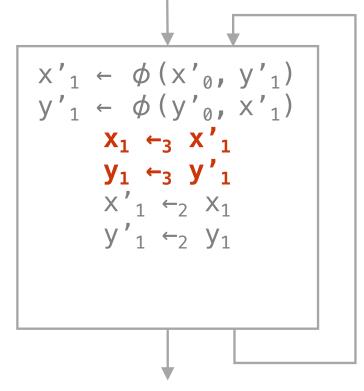
- · φ関数の名前空間を分離する
  - ・引数をコードの名前空間と対応させる
    - → φ関数の前の BB の最後にコピー操作を追加
  - · φ関数の定義をコードの名前空間と対応させる

#### 並列コピーグループ: ←1

同じグループの操作は並列に実行される

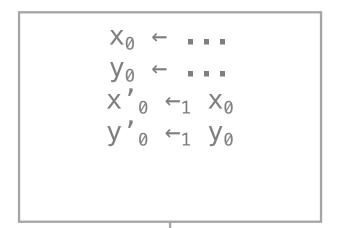
→ 命令順を入れ替えても結果は同じ

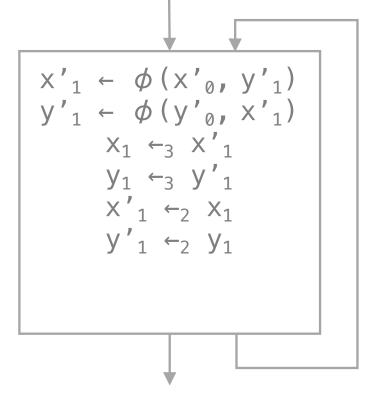




### Phase One

- · φ関数の名前空間を分離する
  - · ø 関数で使われる変数を外部と分離
  - ・ φ 関数の外部 (コピー操作) で並列実行の 影響を表せる



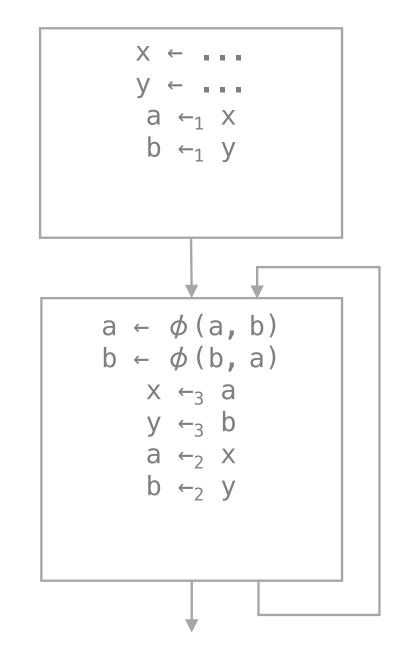


### Phase One

- · φ関数の名前空間を分離する

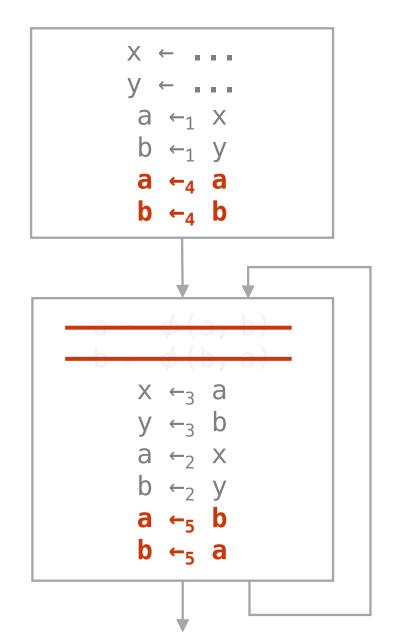
  - ・ φ 関数の外部 (コピー操作) で並列実行の 影響を表せる

・(説明の都合上?) 元のコードと意味が同じになるため、 変数を改名し、変数の添え字を外す



### Phase Two

- ・素朴な手法のように、 $\phi$ 関数を削除し、
  - コピー操作を追加
  - $\rightarrow \phi$  関数の並列セマンティクスを表すため、 並列コピー操作を挿入

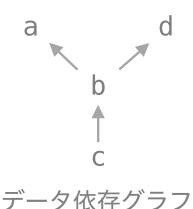


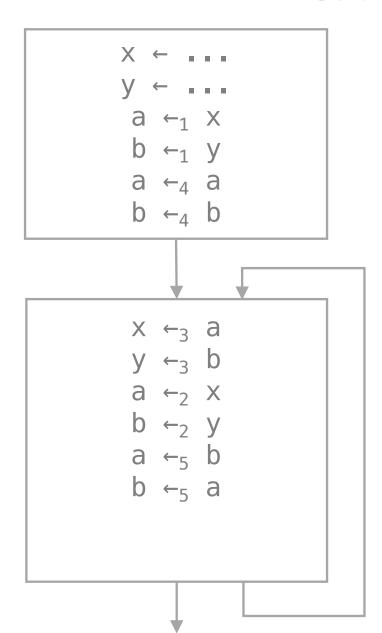
- ・並列コピーグループを意味が等しくなるように 逐次コピー操作に変更
  - → 並列コピーグループのデータ依存グラフを作成

### Data-dependence graph (データ依存グラフ)

定義から使用に値の流れを表したグラフ



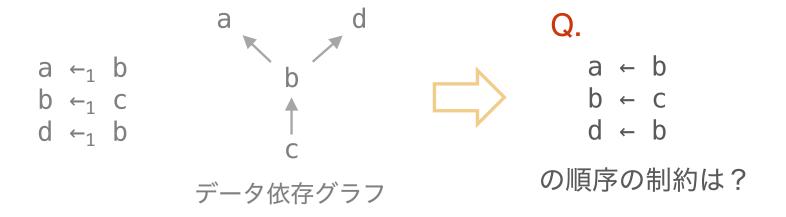


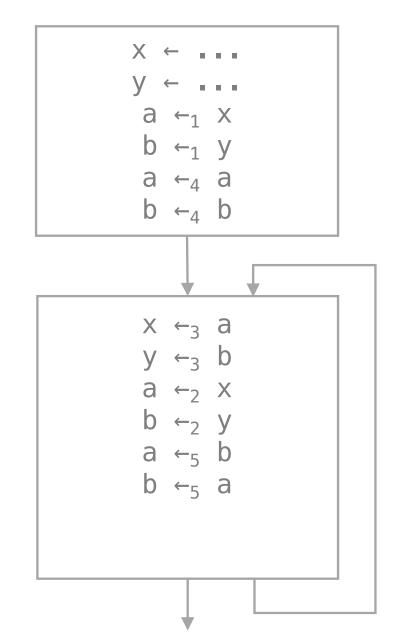


- ・並列コピーグループを意味が等しくなるように 逐次コピー操作に変更
  - → 並列コピーグループのデータ依存グラフを作成

#### 依存グラフにサイクルがないとき

グラフが表す順序で逐次コピー操作に置き換える

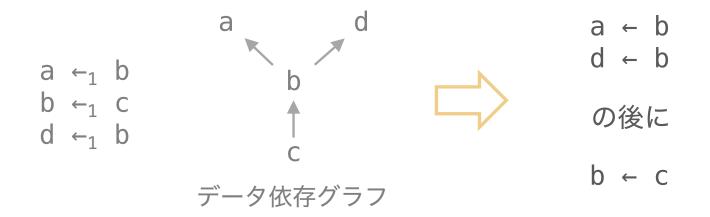


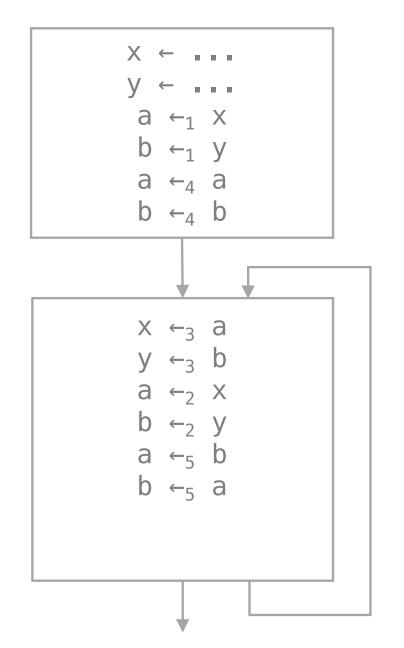


- ・並列コピーグループを意味が等しくなるように 逐次コピー操作に変更
  - → 並列コピーグループのデータ依存グラフを作成

#### 依存グラフにサイクルがないとき

グラフが表す順序で逐次コピー操作に置き換える



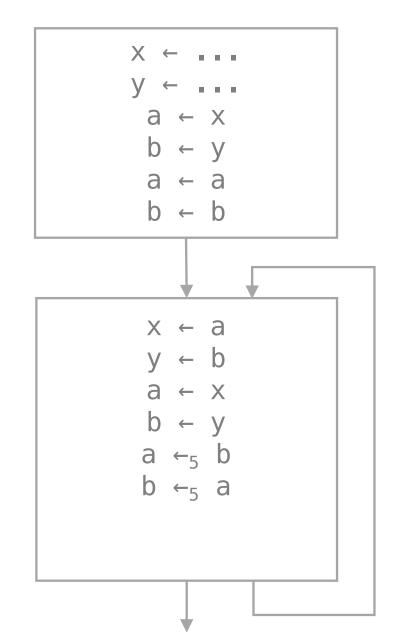


- ・並列コピーグループを意味が等しくなるように 逐次コピー操作に変更
  - → 並列コピーグループのデータ依存グラフを作成

#### 依存グラフにサイクルがあるとき

サイクルをなくすようにコピー操作に書き換える

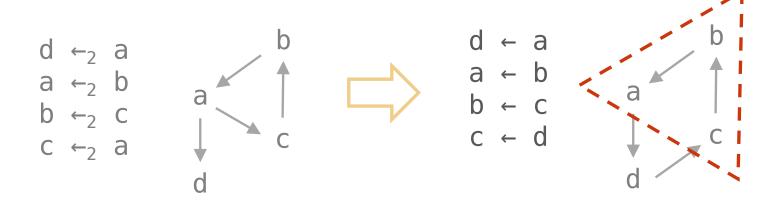


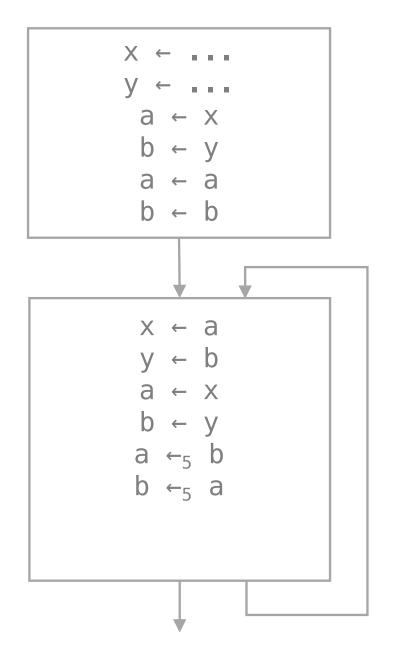


- ・並列コピーグループを意味が等しくなるように 逐次コピー操作に変更
  - → 並列コピーグループのデータ依存グラフを作成

#### 依存グラフにサイクルがあるとき

サイクルがあっても新しい変数を追加しなくて良い場合も



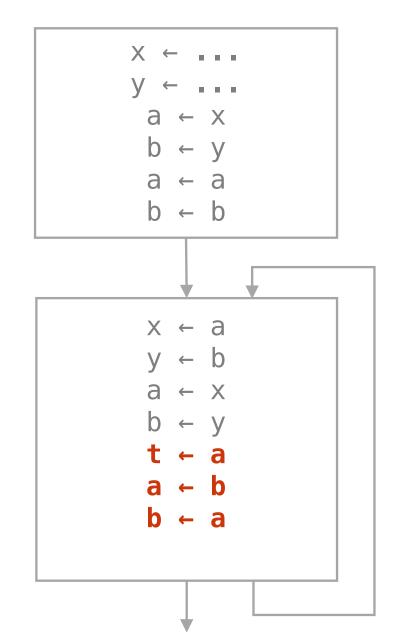


- ・並列コピーグループを意味が等しくなるように 逐次コピー操作に変更
  - → 並列コピーグループのデータ依存グラフを作成

#### 依存グラフにサイクルがあるとき

サイクルをなくすようにコピー操作に書き換える

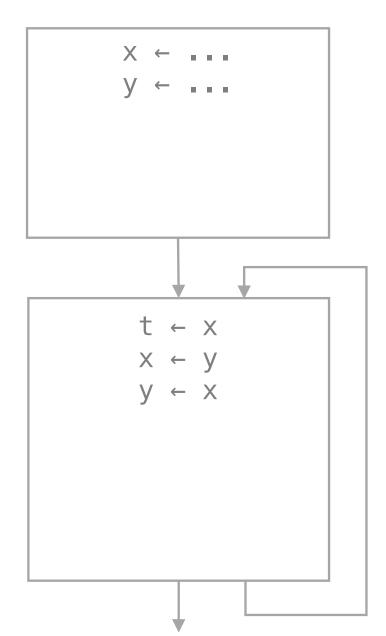




- ・並列コピーグループを意味が等しくなるように 逐次コピー操作に変更
  - → 並列コピーグループのデータ依存グラフを作成

### Copy folding 後

・元々のコードの意味を保持しつつ SSA ではない形式に戻すことができた





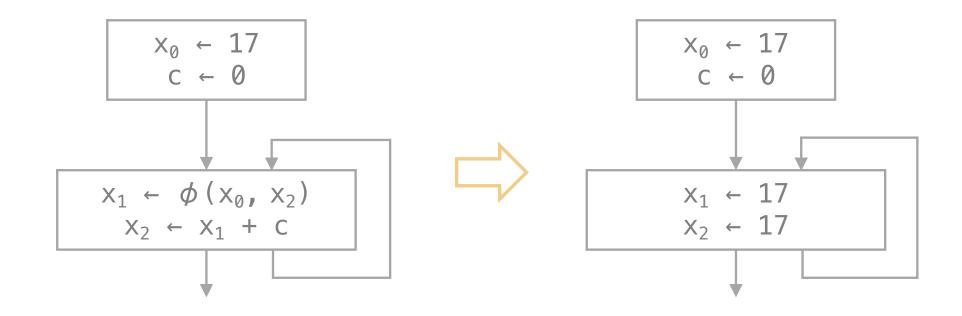
いわき vs 山口 at いわき

# 9.3.6 Using SSA Form

・SSA 形式を使うことで解析や最適化の質を上げることができる

### Sparce Simple Constant Propagation (SSCP)

定数伝播問題



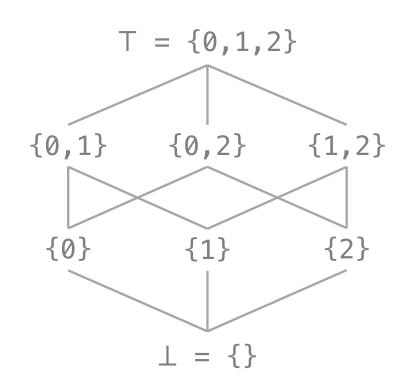
順序集合  $(L, \geq)$  が下半束 (meet-semilattice):

集合の任意の2元 x,y に対して、 それらの下限  $x \wedge y$  が存在する

meet operator ∧

$$\forall a, b, c \in L$$

- 1. 幂等性:  $a \wedge a = a$
- 2. 可換性:  $a \wedge b = b \wedge a$
- 3. 結合性:  $a \wedge (b \wedge c) = (a \wedge b) \wedge c$



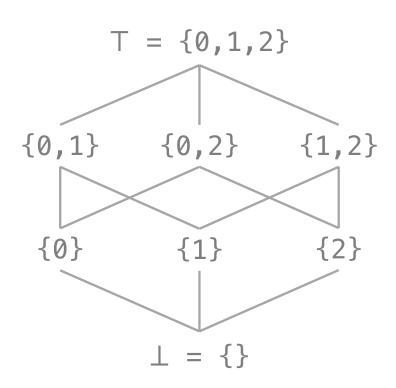
順序集合  $(L, \geq)$  が下半束 (meet-semilattice):

集合の任意の2元 x,y に対して、 それらの下限  $x \wedge y$  が存在する

meet operator ∧

$$a > b \iff a \land b = b$$

$$a > b \iff a \ge b \text{ and } a \ne b$$



順序集合  $(L, \geq)$  が下半束 (meet-semilattice):

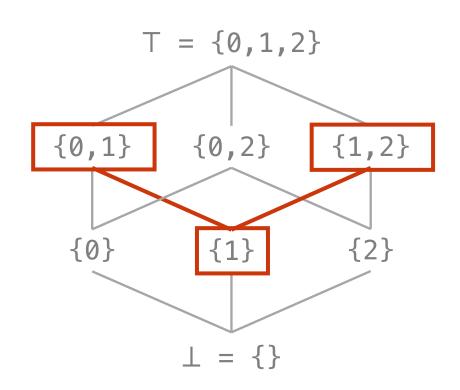
集合の任意の2元 x,y に対して、 それらの下限  $x \wedge y$  が存在する

L: {0,1,2} の冪集合

二項関係: ⊆

meet operator: n

 $\{0,1\} \cap \{1,2\} = \{1\}$ 



順序集合  $(L, \geq)$  が下半束 (meet-semilattice):

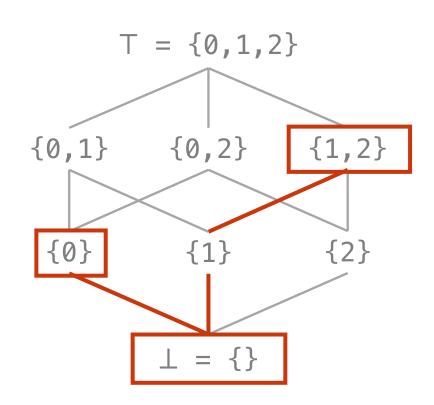
集合の任意の2元 x,y に対して、 それらの下限  $x \wedge y$  が存在する

L: {0,1,2} の冪集合

二項関係: ⊆

meet operator: n

 $\{0\} \cap \{1,2\} = \{\}$ 



順序集合  $(L, \geq)$  が下半束 (meet-semilattice):

集合の任意の2元 x,y に対して、

それらの下限  $x \wedge y$  が存在する

bottom: ⊥

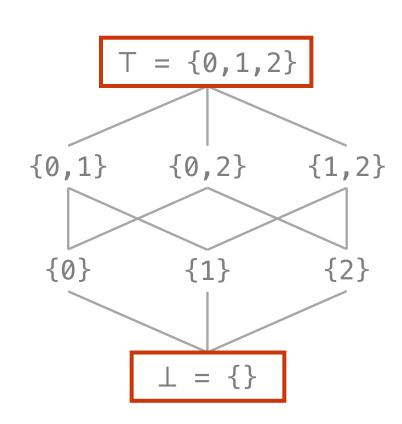
→ すべての要素の下限

 $\forall a \in L, a \land \bot = \bot, \text{ and } \forall a \in L, a \ge \bot$ 

top: ⊤

→ すべての要素の上限(なくてもよい)

 $\forall a \in L, a \land \top = a, \text{ and } \forall a \in L, \top \geq a$ 



順序集合  $(L, \geq)$  が下半束 (meet-semilattice):

集合の任意の2元 x,y に対して、 それらの下限  $x \wedge y$  が存在する

L: 自然数の集合

二項関係: 通常の数の大小関係

Q. meet operator は?

$$a \ge b \iff a \land b = b$$

順序集合  $(L, \geq)$  が下半束 (meet-semilattice):

集合の任意の2元 x,y に対して、 それらの下限  $x \wedge y$  が存在する

L: 自然数の集合

二項関係: 通常の数の大小関係

meet operator: min

$$a \ge b \iff a \land b = b$$

## **SSCP**

- ・定数伝播のアルゴリズム
- · 」, T, 定数値の無限集合

T: 定数かどうか分からない

→ その値に関する情報を今後発見するかも

⊥: 定数ではない

