

# Chapitre 8: Optimisations

## INF7641 Compilation

Jean Privat

Université du Québec à Montréal

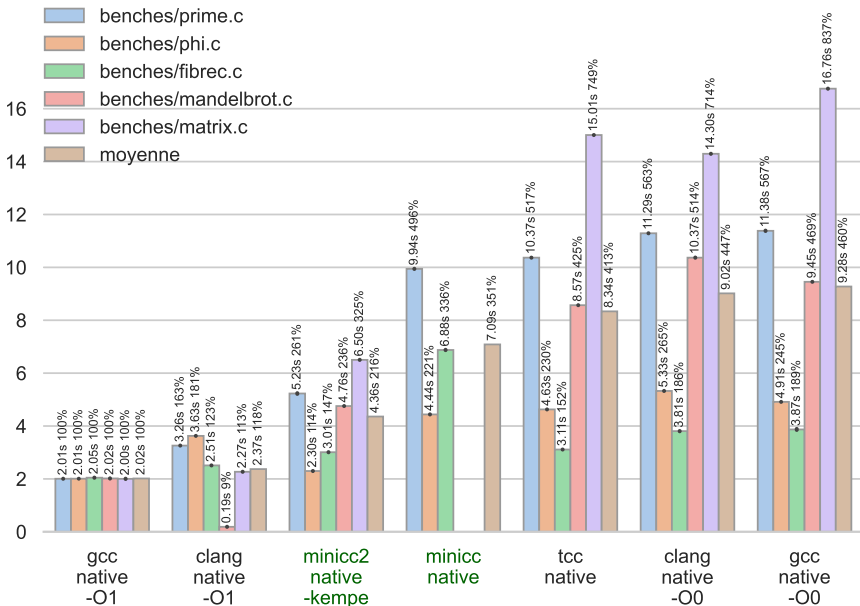
INF7641 Compilation  
v251



# Où on en est ?

- On a une représentation intermédiaire (IR)
- On sait faire quelques analyses (*dataflow*)

# MiniCC2 -O0



# Aujourd'hui ?

## IR $\rightarrow$ IR : MiniCC -O1

- Tant que pas satisfait: analyser & transformer IR

## Difficultés

- Représentation de l'information  
IR et résultats d'analyses
- Maintient/recalcul de l'information  
Transformer = changer le code,  
donc possiblement invalider des analyses

## Représentation de la satisfaction

- L'IR est plus joli ? Plus court ?

# Aujourd'hui ?

## IR $\rightarrow$ IR : MiniCC -O1

- Tant que pas satisfait: analyser & transformer IR

## Difficultés

- Représentation de l'information  
IR et résultats d'analyses
- Maintient/recalcul de l'information  
Transformer = changer le code,  
donc possiblement invalider des analyses

## Représentation de la satisfaction

- L'IR est plus joli ? Plus court ?
- Voire on mesure !

# Plan

- 1 *Single static assignment (SSA)*
- 2 Interlude : dominance
- 3 Transformer en SSA
- 4 Nos premières optimisations

# *Single static assignment (SSA)*

# Reachability analysis

## Chaînes *use-def/def-use*

- Utiles pour les optimisation
- Doit être recalculé souvent
- Information quadratique : defs  $\times$  uses

## SSA (*single static assignment*)

Situation idéale : une **seule affectation (définition)** par registre

- C'est **statique** car c'est dans le code
- Mais cette affectation peut être exécutée plusieurs fois

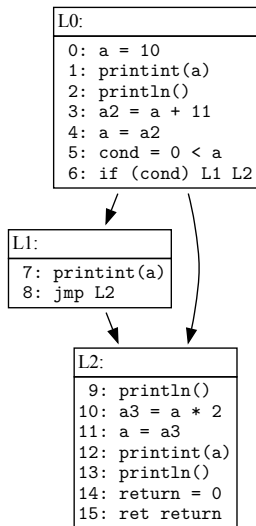
## Avantages

- Le *def* de ce registre est toujours connu (et unique)
- Les *uses* de ce registres sont **toutes** ses utilisations
- Pas besoin d'analyse *dataflow* pour les trouver
- Implémentation efficace possible pour mettre les uses en cache et accélérer les analyses et transformations



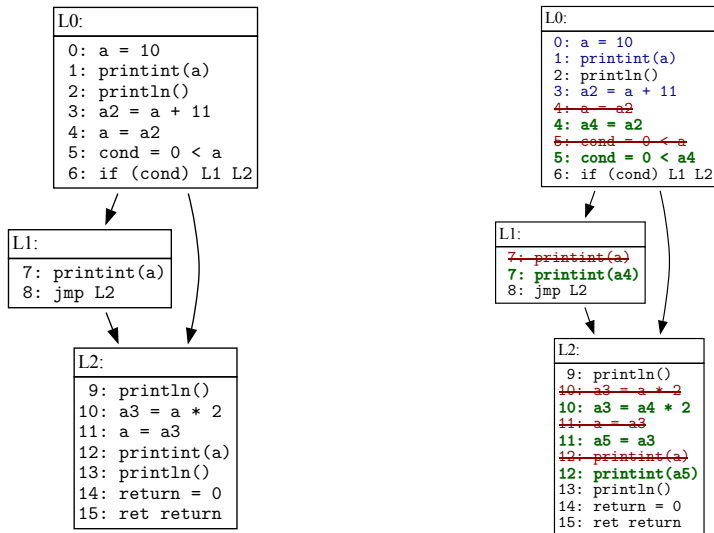
# Un seule définition par registre. Mise en œuvre ?

- Chaque affectation est associé à un nouveau registre
- → on renomme/numérote & maintient à jour les utilisations

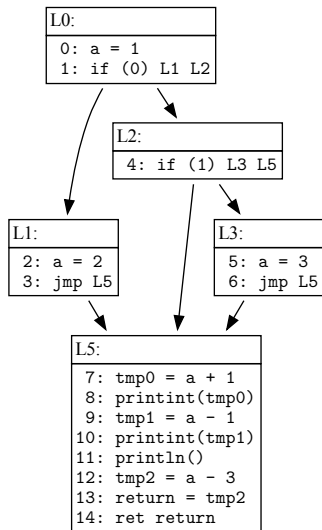


# Un seule définition par registre. Mise en œuvre ?

- Chaque affectation est associé à un nouveau registre
- → on renomme/numérote & maintient à jour les utilisations

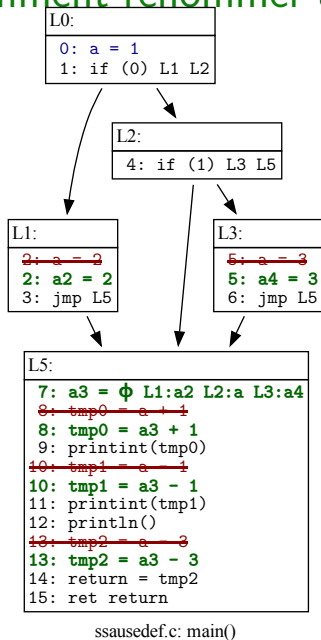
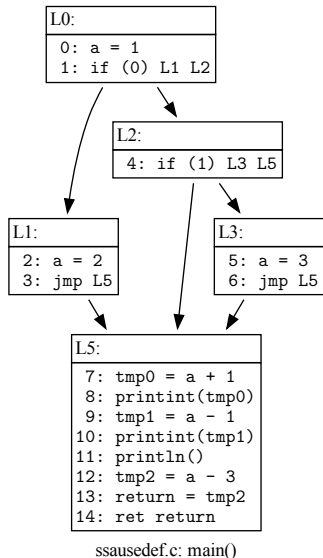


# Problème : les jonctions. Comment renommer a ?



ssausedef.c: main()

# Problème : les jonctions. Comment renommer a ?



# $\Phi$ (phi)

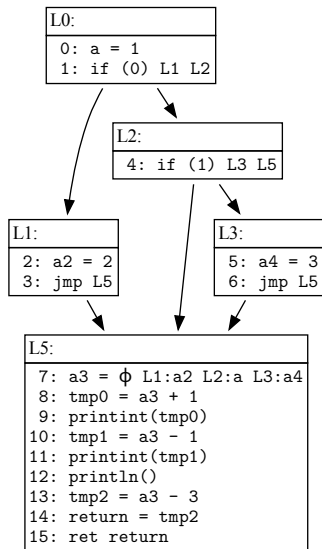
## Instruction $\Phi$

- Au début des blocs de base
- Combine les valeurs des différents arcs d'entrée
  - Chaque bloc d'entrée est associé à une opérande
- Dans un nouveau registre (car SSA)

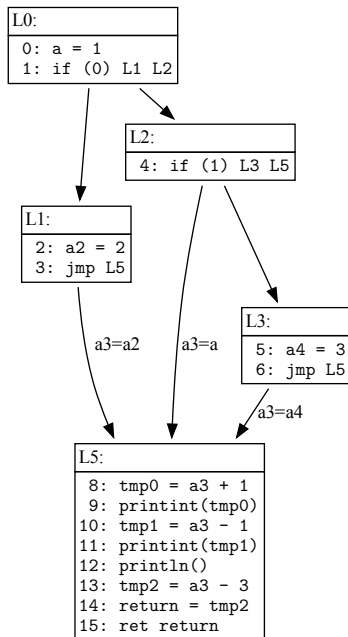
## Sémantique de $\Phi$

- Comme s'il y avait un `mv` sur chaque arc
- Attention, si plusieurs  $\Phi$ ,  
tous les `mv` sont à faire « en parallèle »

# Sémantique du $\Phi$



ssousedef.c: main()



# Avantage et inconvénients de $\Phi$

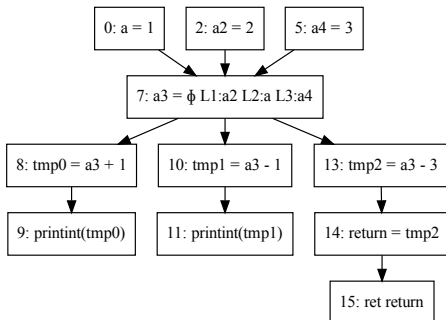
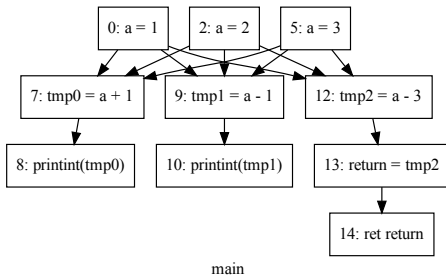
## Avantages

- $\Phi$  se comporte à peu près comme n'importe quelle instruction
  - un *dst*, des *src*
- La représentation est SSA
  - Simplifie les chaînes *use-def* et *def-use*

## Inconvénients

- Pénible à mettre
- Pénible à enlever
- Modifier le CFG  $\rightarrow$  maintenir la cohérence des  $\Phi$
- Plein de registres IR supplémentaires (l'allocateur doit être bon)

# Simplification use-def





# Transformer en SSA

Déterminer les blocs où les insérer les  $\Phi$

- Euh...

Renommer (numéroter) les registres

- Euh...

On est pas encore prêt

## Quitter SSA : Enlever les $\Phi$

Pour chaque bloc  $n$  du CFG avec des  $\Phi$

### Concrétiser les arcs entrants de $n$

- Soit un bloc prédécesseur  $p$  avec  $p \rightarrow n$
- On peut l'étendre si exclusif:  $\text{succ}(p) = \{n\}$
- Sinon, insérer un nouveau bloc  $p'$ 
  - $p \rightarrow p' \rightarrow n$ , avec  $\text{succ}(p') = \{n\}$

### Mov

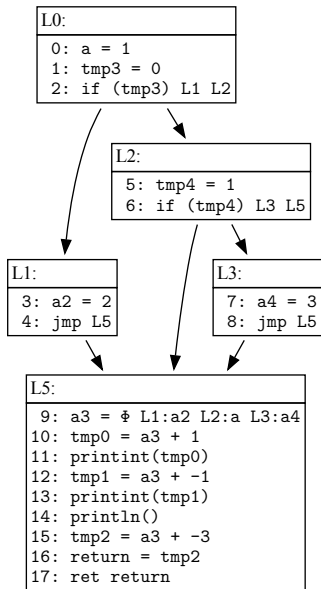
- Pour chaque  $x = \Phi(p : y, \dots)$
- Mettre un  $x=y$  à la fin de  $p$

### Sémantique parallèle

Attention, si plusieurs  $\Phi$

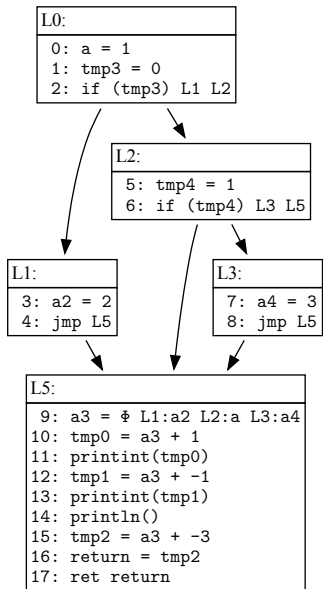
- Mettre les `mov` dans le bon ordre si dépendances
- Utiliser un registre temporaire supplémentaire si circularité
- $\rightarrow$  `PhiRemoval.java`

# Exercice 1 : Enlever le $\Phi$

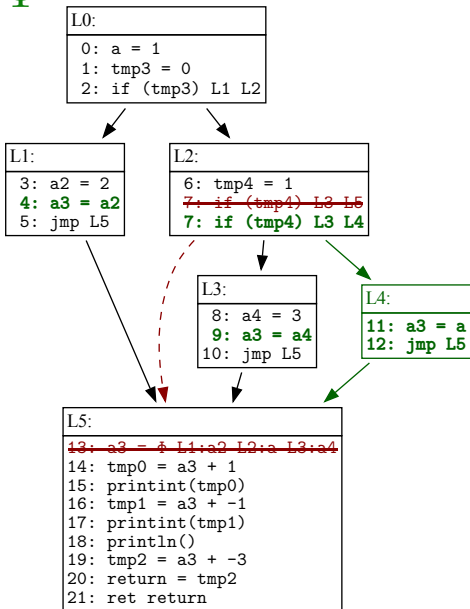


ssasedef.c: main()

# Exercice 1 : Enlever le $\Phi$

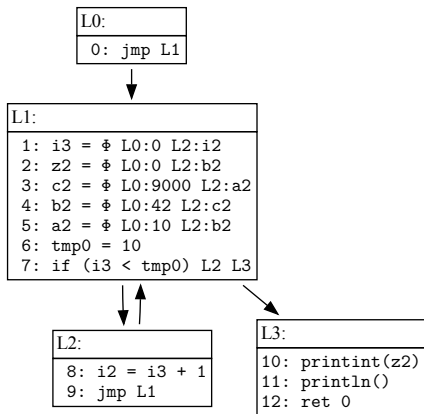


ssasedef.c: main()



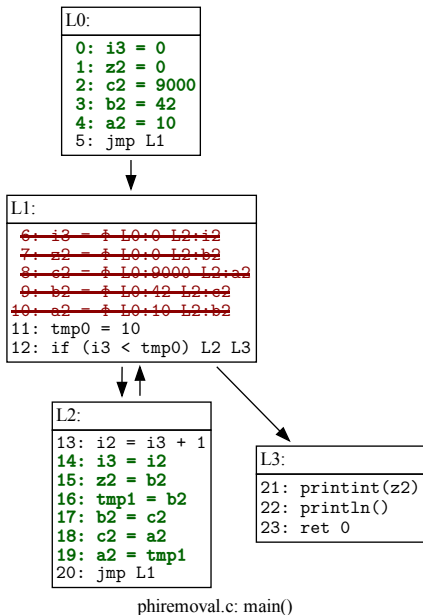
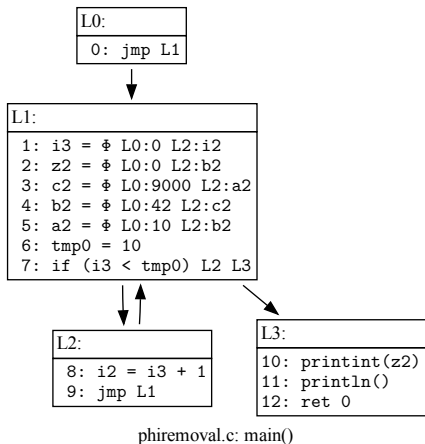
ssasedef.c: main()

## Exercice 2 : Enlever les $\Phi$



phiremoval.c: main()

## Exercice 2 : Enlever les $\Phi$



# Interlude : dominance

# Dominance

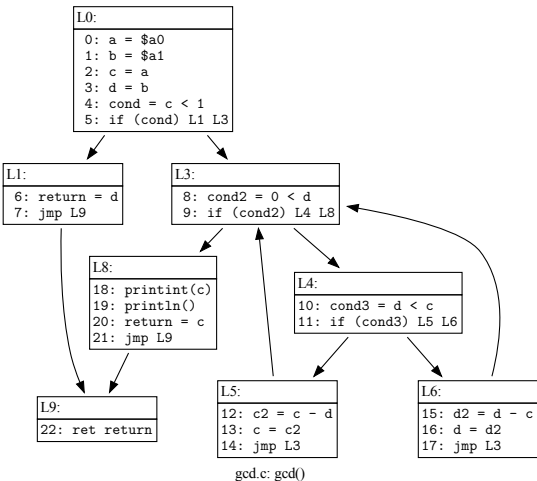
- Dans un CFG (départ  $n_0$ ), un nœud  $n_1$  **domine** un nœud  $n_2$
- (ou  $n_1$  est un **dominateur** de  $n_2$ )
- Si :  $n_1$  appartient à **tous** les chemins de  $n_0$  à  $n_2$

## Vocabulaire

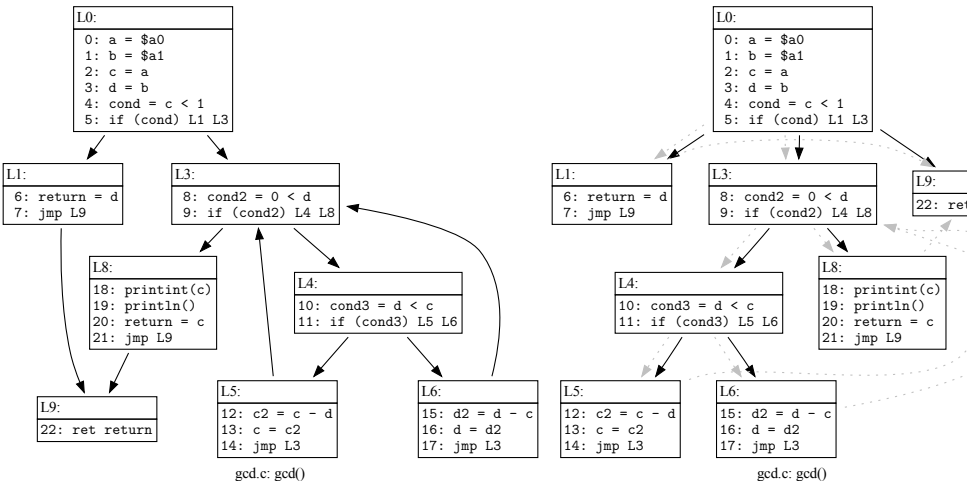
- dom (dominance) : relation d'ordre
  - réflexive, transitive, anti-associative
- sdom (domination stricte) : réduction réflexive de dom
  - $n_1 \text{ sdom } n_2 \Leftrightarrow n_1 \neq n_2 \wedge n_1 \text{ dom } n_2$
- idom (dominateur immédiat) : réduction transitive de sdom
  - $n_1 = \text{idom}(n_2) \Leftrightarrow n_1 \text{ sdom } n_2 \wedge \nexists n_3, n_1 \text{ sdom } s_3 \text{ sdom } n_2$
  - Tous les nœuds, ont **un seul** dominateur immédiat
  - Sauf celui de départ
  - $\rightarrow$  On peut construire un arbre de dominance



# Exercice : Calculer la dominance



# Exercice : Calculer la dominance



# Calcul de la dominance

## Plusieurs algorithmes existent

- $O(m\alpha(m, n))$  Lengauer&Tarjan, 1979  
où  $\alpha$  est l'inverse de la fonction d'Ackerman, croît très lentement
- $O(n^2)$  Cooper&Harvey&Kennedy, 2006

# Calcul de la dominance

## Plusieurs algorithmes existent

- $O(m\alpha(m, n))$  Lengauer&Tarjan, 1979  
où  $\alpha$  est l'inverse de la fonction d'Ackerman, croît très lentement
- $O(n^2)$  Cooper&Harvey&Kennedy, 2006

## Mais un simple dataflow fonctionne

- Chercher les dominateurs d'un bloc
- Ensemble :

# Calcul de la dominance

## Plusieurs algorithmes existent

- $O(m\alpha(m, n))$  Lengauer&Tarjan, 1979  
où  $\alpha$  est l'inverse de la fonction d'Ackerman, croît très lentement
- $O(n^2)$  Cooper&Harvey&Kennedy, 2006

## Mais un simple dataflow fonctionne

- Chercher les dominateurs d'un bloc
- Ensemble : les blocs
- Ensemble de chaque bloc : ses dominateurs
- Sens : en avant (le dominateur est **avant** le dominé)
- Fusion : intersection (sur **tous** les chemins)
- Gen: le bloc lui-même
- Kill: rien
- → `Dominance.java`

# Frontière de dominance

- Un nœud  $f$  appartient à la **frontière** de  $n$  si :
- Il existe un prédécesseur  $p$  de  $f$ , pour qui **tous les chemins** depuis l'entrée  $n_0$  passent par  $n$
- Et, il **existe un chemin** vers  $f$  depuis  $n_0$  qui ne passe pas par  $n$
- $\rightarrow f$  est hors mais à la limite de la zone de dominance de  $n$

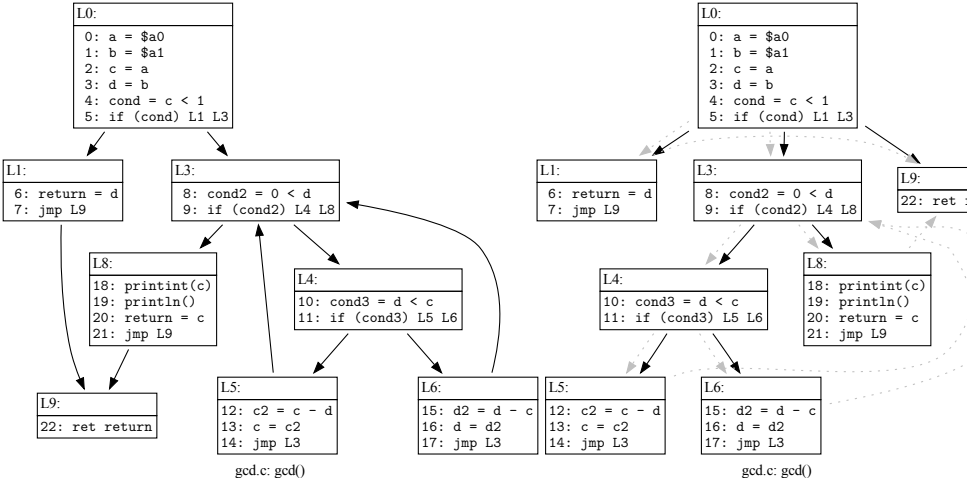
## Plus formellement ?

- $f \in \text{df}(n)$ 
  - $f$  pas strictement dominé par  $n$
  - $f$  successeur d'un sommet dominé par  $n$
- $\text{df}(n) = \{f \mid \neg(n \text{ sdom } f) \wedge \exists p \in \text{pred}(f), n \text{ dom } p\}$

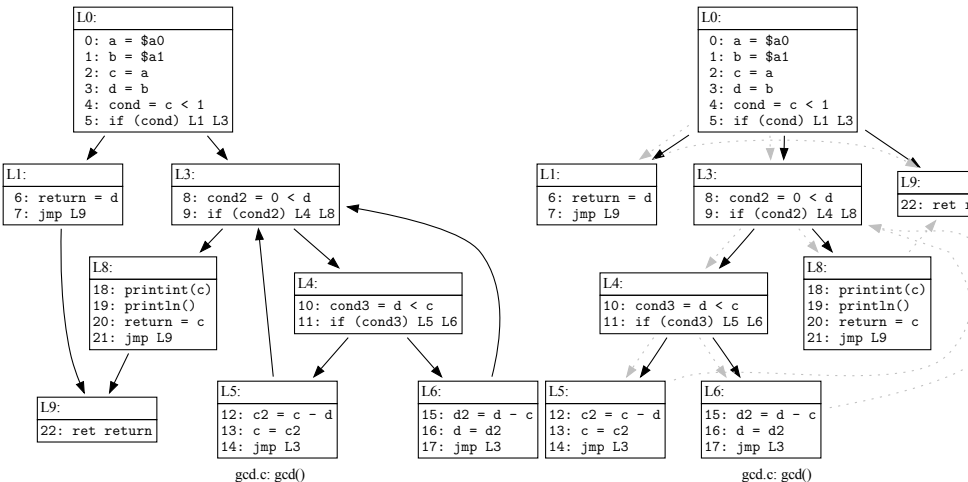
## Frontière étendue

- La frontière d'un ensemble de nœuds est l'union des frontières de ces nœuds
- $\text{df}(N) = \bigcup_{n \in N} \text{df}(n)$

# Exercice : déterminer les frontières



# Exercice : déterminer les frontières



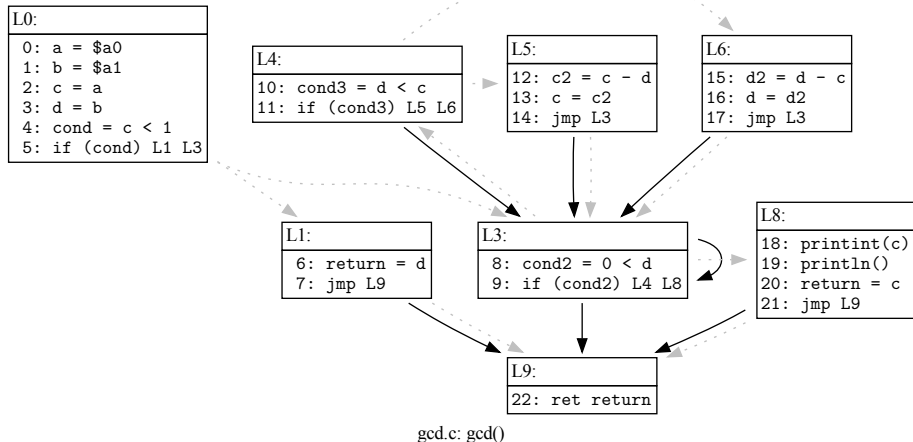
gcd.c: gcd()

gcd.c: gcd()

L0:[] L1:[L9] L3:[L3, L9] L4:[L3] L5:[L3] L6:[L3] L8:[L9] L9:[]



# Relation de frontière de dominance



# Transformer en SSA

# Placer les instructions $\Phi$

- Soit  $N$  l'ensemble des blocs qui ont une définition de  $r$
- Il faut placer un  $r = \Phi(\dots)$  à chaque nœuds de  $df(N)$
- Mais ! Ça ajoute une nouvelle définition de  $r$  !
- Donc on itère
  - $df_1(N) = df(N)$
  - $df_{i+1}(N) = df(N \cup df_i(N))$
  - Jusqu'au point fixe :  $df^+(N)$

# Algorithme d'insertion de $\Phi$

**pour** chaque registre  $r$  **faire**

Liste  $\leftarrow \emptyset$  ;

**pour** chaque nœud  $n$  affectant  $r$  **faire**

└ Ajouter  $n$  à Liste ;

**tant que** Liste  $\neq \emptyset$  **faire**

└ Retirer  $n$  de Liste ;

└ **pour** chaque  $f \in df(n)$  **faire**

└└ Ajouter une  $\Phi$  dans  $f$  pour  $r$  (sauf si déjà un) ;

└└ Ajouter  $f$  à Liste (sauf si on a déjà traité  $f$ ) ;

# Renommage des registres

- Pour chaque registre, on conserve une pile de noms
  - L'historique des noms
- On visite en profondeur le CFG
  - Remplacer les opérandes par les noms les plus récents
  - Créer et empiler des nouveaux noms pour les résultats
  - Mettre à jour les entrées des  $\Phi$

**pour** *chaque registre*  $r$  **faire**

$\lfloor$  pile[ $r$ ] = nouvelle pile vide;

renommer(nœud d'entrée du CFG) ;

**fun** *renommer*( $n$ : Nœud)

**pour** *chaque instruction  $i$  de  $n$*  **faire**

**pour** *chaque opérande  $r$  de  $i$*  **faire**

└ Remplacer  $r$  par tête(pile[ $r$ ]) ;

**pour** *le résultat  $r$  de  $i$*  **faire**

└ Créer un nouveau nom  $r'$  ;

└ Empiler  $r'$  dans pile[ $r$ ] ;

└ Remplacer  $r$  par  $r'$  ;

**pour** *chaque successeur  $s$  de  $n$*  **faire**

**pour** *chaque instruction  $r = \Phi(\dots)$  de  $s$*  **faire**

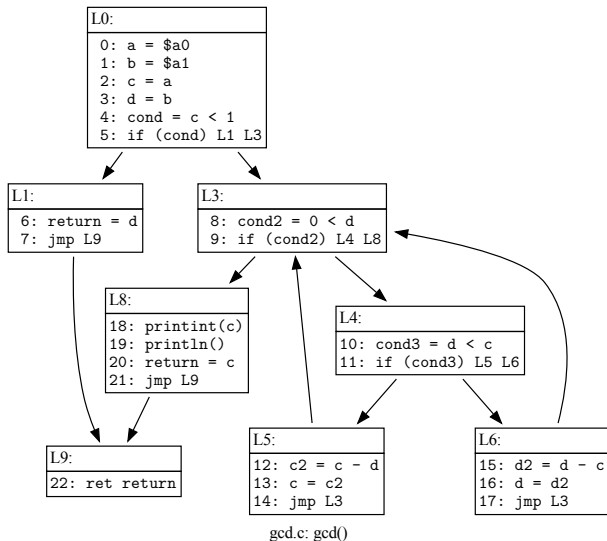
└ Associer tête(pile[ $r$ ]) pour l'arc  $n \rightarrow s$  ;

**si**  $s$  *pas encore visité* **alors**

└ *renommer*( $s$ ) ;

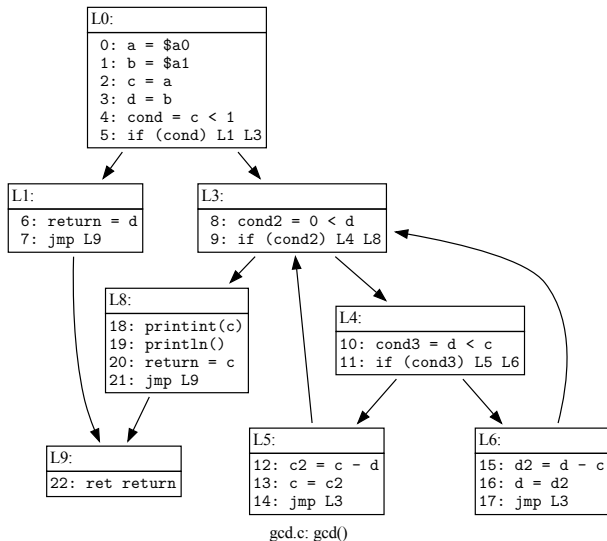
└ Dépiler (restaurer) les piles ;

# Exercice : Renommer c, d et return



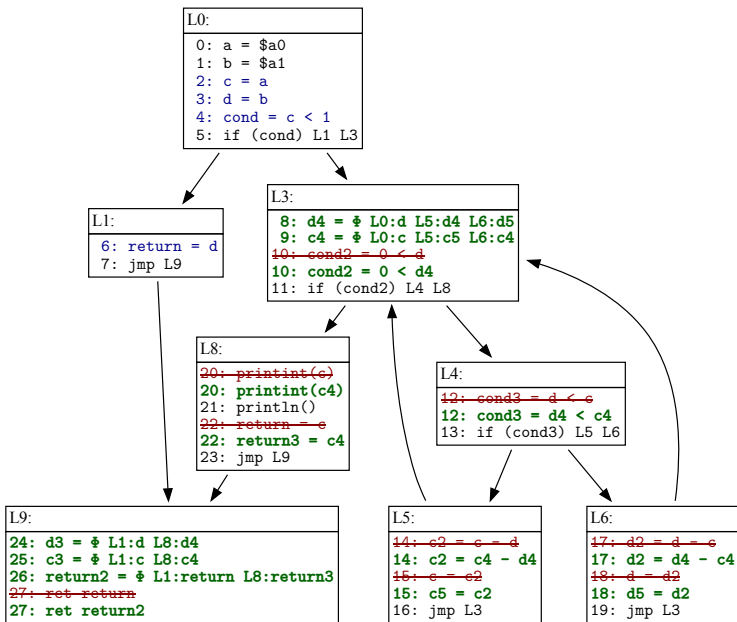
Où sont les  $\Phi$  ?

# Exercice : Renommer c, d et return



Où sont les  $\Phi$  ? Tous dans L3 et L9





gcd.c: gcd()

# SSA: résumé

- Une **seule définition** par registre
- Instruction **magique**  $\Phi$  pour combiner les valeurs possibles
- La définition **domine** les utilisations
  - (sauf, bien sûr, les opérandes des nœuds  $\Phi$ )
- $\rightarrow$  `SSA.java`

# Nos premières optimisations

# Quelques optimisations faciles

## Déjà faciles

- *Constant folding* :  $x=1+2 \rightarrow x=3$
- *Algebraic simplification* :  $x*1 \rightarrow x$
- *Strength reduction* :  $y=x*8 \rightarrow y=x<<3$

## Plus faciles avec SSA

- *Constant propagation* :  $y=1; z=x+y \rightarrow z=x+1$
- *Copy propagation* :  $y=x; z=y+1 \rightarrow y=x; z=x+1$
- *Algebraic combinaison* :  $y=x+1; z=y+1 \rightarrow y=x+1; z=x+2$
- *Common subexpression elimination*  
 $x=b*c+d; y=b*c+e \rightarrow t=b*c; x=t+d; y=t+e$
- *Dead code elimination* : supprimer variables inutilisées

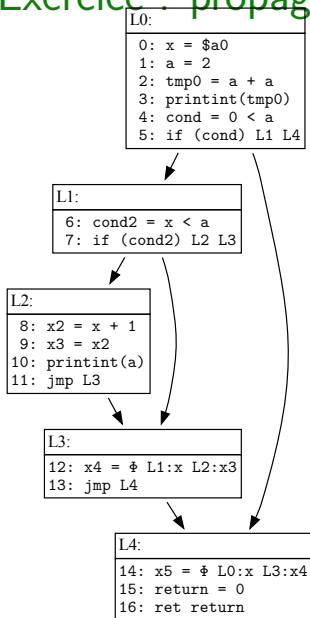
# Sparse Conditional Constant Propagation (SCCP)

- [Wegman&Zadeck, 1991](#)
- Analyse *dataflow* « moderne »
- Propage les valeurs et calculs constantes
- Élimine les branches conditionnelles mortes
- Profite de SSA
- Rapide

Alternative : à la main, plus lent et moins efficace

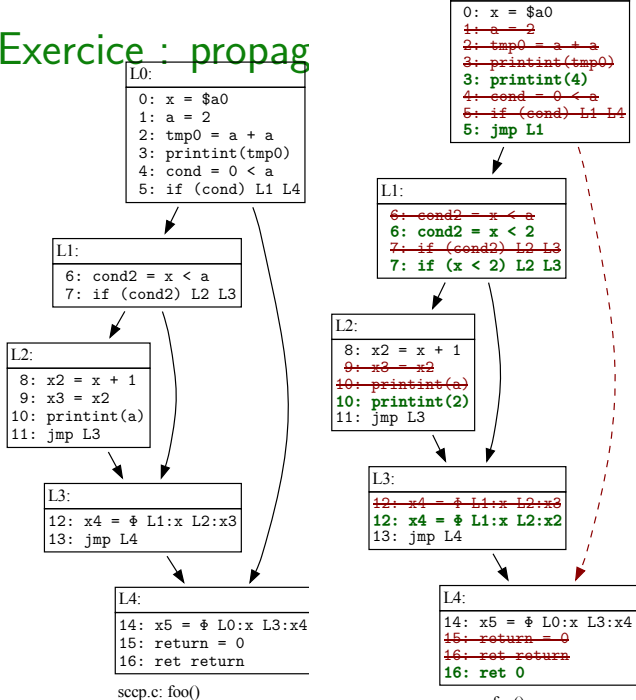
- Itérer, simplifier, propager
- Voir `ConstantPropagation.java` et `DeadCodeElimination.java`

# Exercice : propager les constantes



sccp.c: foo()

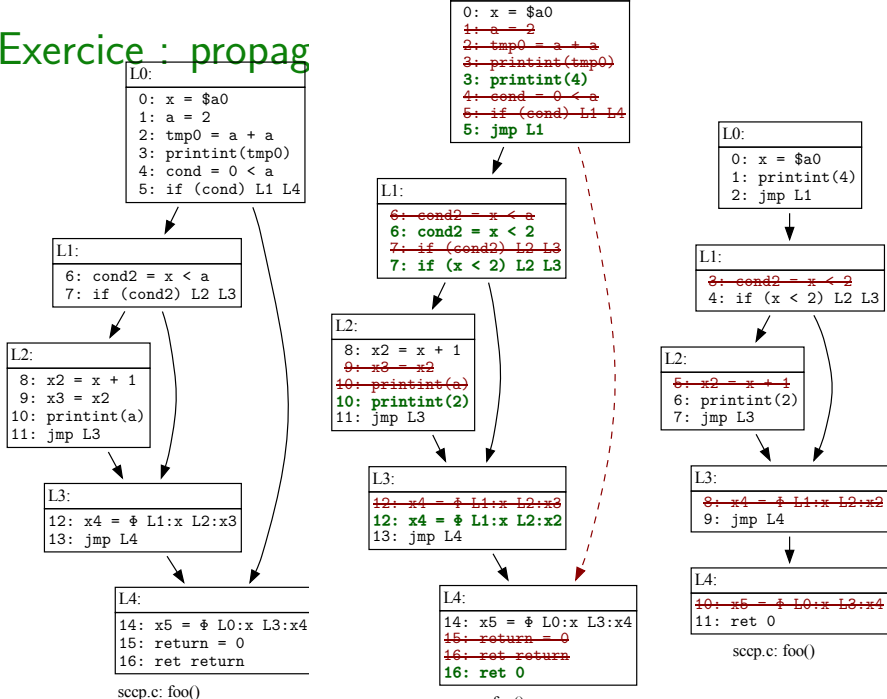
# Exercice : propag



sccp.c: foo()

sccp.c: foo()

# Exercice : propag



sccp.c: foo()

sccp.c: foo()

sccp.c: foo()



# Simplifications du CFG

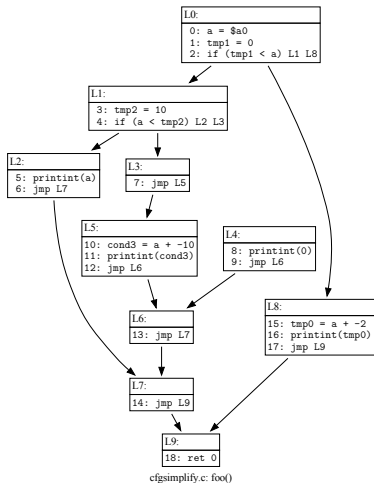
- Supprimer bloc sans prédécesseur
  - → Voir `CFGPrune.java`
- Fusionner bloc avec son successeur unique
  - Si ce successeur a un seul prédécesseur
  - → Voir `CFGMerge.java`
- Supprimer bloc qui contient juste un `Jump`
  - → Voir `CFGFoldJump.java`
- Supprimer  $\Phi$  des blocs avec un seul prédécesseur
  - → Voir `CFGPrune.java` aussi

Note: Pénible avec SSA, car il faut maintenir les  $\Phi$

# Exercice : Simplifier le CFG

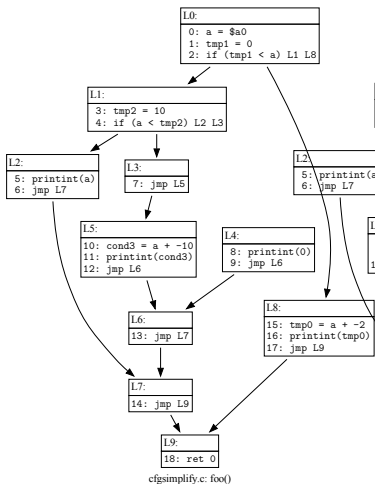
original

prune

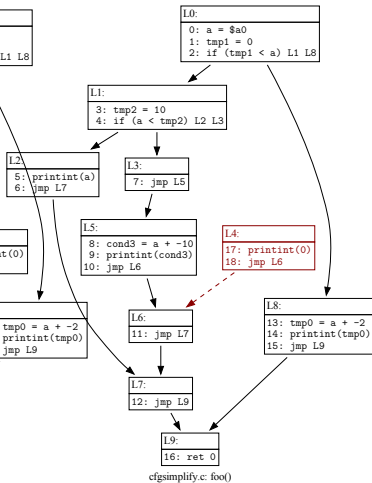


# Exercice : Simplifier le CFG

original



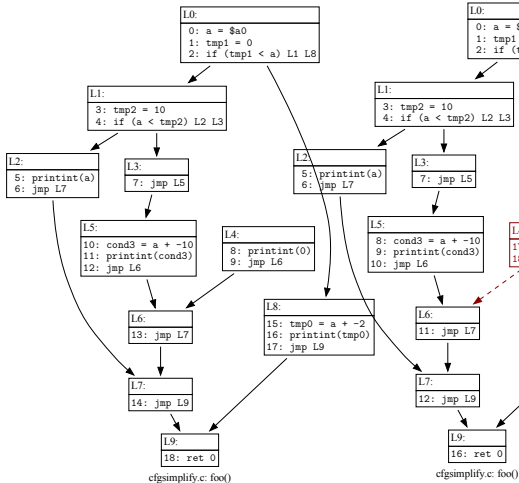
prune



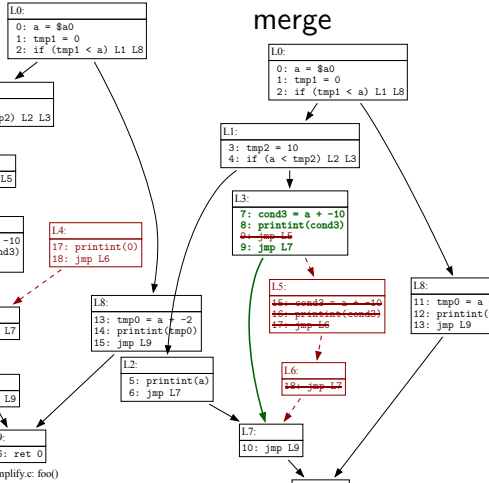
merge

# Exercice : Simplifier le CFG

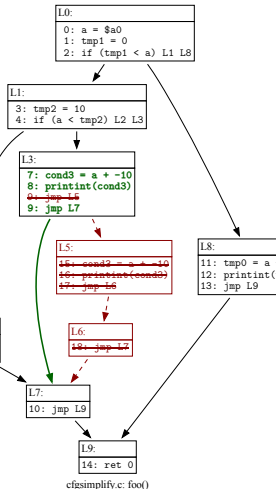
original



prune



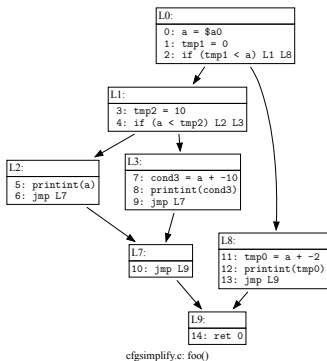
merge



# Exercice : Simplifier le CFG

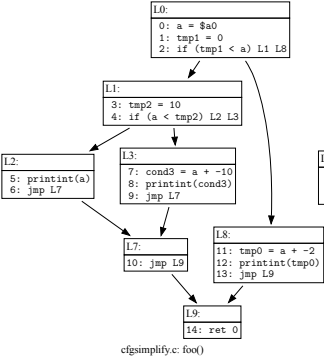
post-merge

suppr. jump

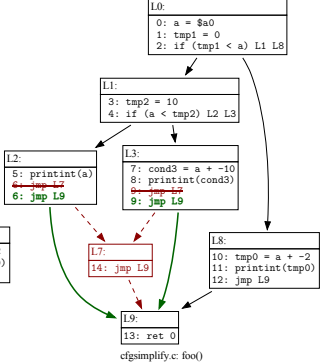


# Exercice : Simplifier le CFG

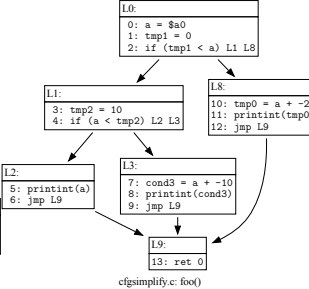
post-merge



suppr. jump



final



# Common subexpression elimination, CSE

## Expressions disponibles (*available expressions*)

- Analyse dataflow
- Détermine les expressions déjà calculées
- Déjà vu

## Numérotation des valeurs (*value numbering*)

Griggs&Cooper&Simpson,1997

- Technique alternative
- S'intéresse aux valeurs plutôt qu'aux formes syntaxiques
- Basée sur SSA

## Gain et pertes : CSE

# Common subexpression elimination, CSE

## Expressions disponibles (*available expressions*)

- Analyse dataflow
- Détermine les expressions déjà calculées
- Déjà vu

## Numérotation des valeurs (*value numbering*)

Griggs&Cooper&Simpson,1997

- Technique alternative
- S'intéresse aux valeurs plutôt qu'aux formes syntaxiques
- Basée sur SSA

## Gain et pertes : CSE

- Réduit la taille du code
- Réduit le temps d'exécution
- Un registre à faire vivre plus longtemps



# Élimination de code mort

## Code mort (on y revient)

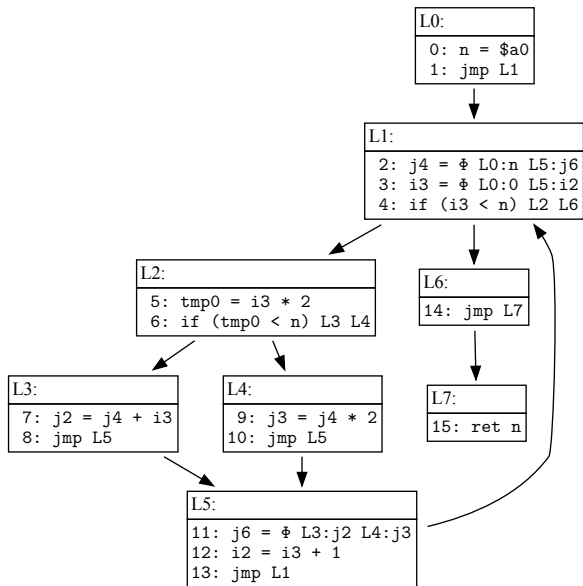
- Registre inutilisé en opérande
- Supprimer les opérations **sans effet de bord** (pures) qui le calcule
  - `mov` ? arithmétique ? `return` ? `jmp` ? `if` ? `call` ?

# Élimination de code mort

## Code mort (on y revient)

- Registre inutilisé en opérande
- Supprimer les opérations **sans effet de bord** (pures) qui le calcule
  - `mov` ? arithmétique ? `return` ? `jmp` ? `if` ? `call` ?
  - `mov` et arithmétique : pures
  - `return` et `call` : effet de bord
  - `jmp` et `if` : gardent le CFG cohérent
- Et répéter jusqu'à stabilité
- → `DeadCodeElimination.java`

# Exercice : éliminer le code mort



dce.c: foo()

# Un peu simpliste ?

- Un registre peut être utilisé mais pas très utile quand même
- Dépendances entre registres inutiles
  - `x = y+5; y = x-5;` (où `x` et `y` pas utilisés par ailleurs)
- Boucles et conditions inutiles
  - `while(i<10) { i=i+1; }` (où `i` pas utilisé par ailleurs)

# Dépendance de contrôle (*control dependency*)

## Définition

- Un bloc  $n$  a une **dépendance de contrôle** envers le bloc  $f$  si
- Il existe un successeur  $s$  de  $f$ , pour qui tous **les chemins** vers la sortie  $n_\omega$  passent par  $n$
- Et, il **existe un chemin** de  $f$  vers  $n_\omega$  qui ne passe pas par  $n$
- $\rightarrow n$  contrôle l'exécutabilité de  $f$

... ça nous rappelle quelque chose, non?

# Dépendance de contrôle (*control dependency*)

## Définition

- Un bloc  $n$  a une **dépendance de contrôle** envers le bloc  $f$  si
- Il existe un successeur  $s$  de  $f$ , pour qui tous **les chemins** vers la sortie  $n_\omega$  passent par  $n$
- Et, il **existe un chemin** de  $f$  vers  $n_\omega$  qui ne passe pas par  $n$
- $\rightarrow n$  contrôle l'exécutabilité de  $f$

... ça nous rappelle quelque chose, non?

# Post-dominance

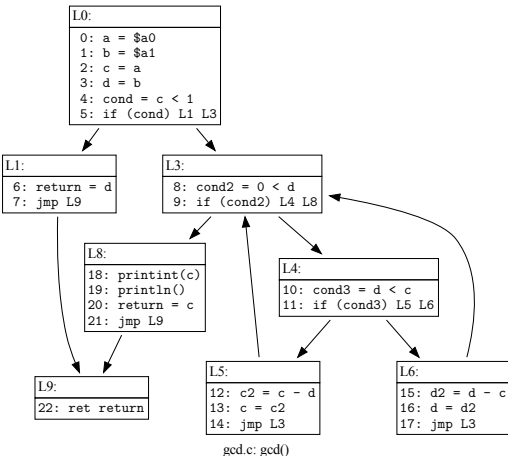
## Notion duale de la dominance

- Dans un CFG (sortie  $n_\omega$ ), un nœud  $n_1$  **post-domine** un nœud  $n_2$
- (ou  $n_1$  est un **post-dominateur** de  $n_2$ )
- Si :  $n_1$  appartient à **tous** les chemins de  $n_2$  à  $n_\omega$ .

## En fait, c'est pareil mais à l'envers

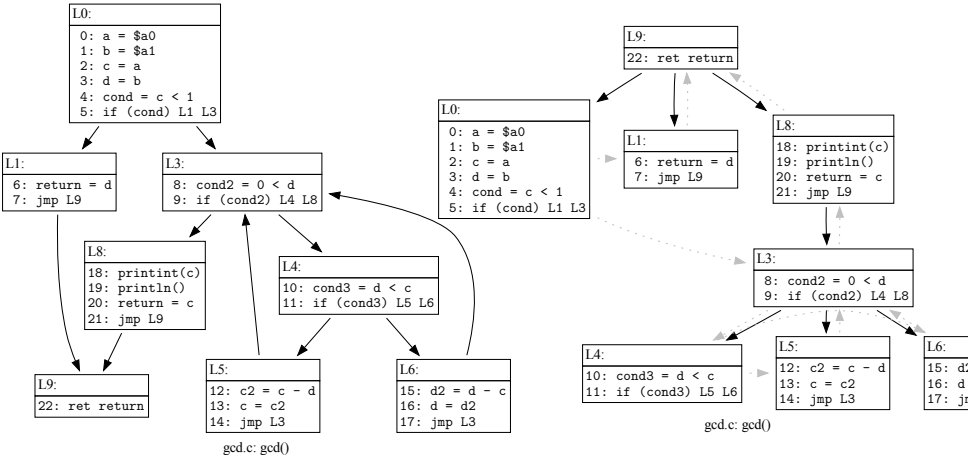
- On travaille sur le CFG dont on a inversé les arcs
- Ou on fait le *dataflow* en mode **arrière**
- Attention: inverser l'arbre de dominance ne suffit pas
- relation d'ordre (pdom), version stricte (spdom), version arbre (ipdom)
- et une frontière de post-dominance (fpd) !
- → `PostDominance.java`

# Exercice : Calculer la post-dominance





# Exercice : Calculer la post-dominance



# Frontière de post-dominance

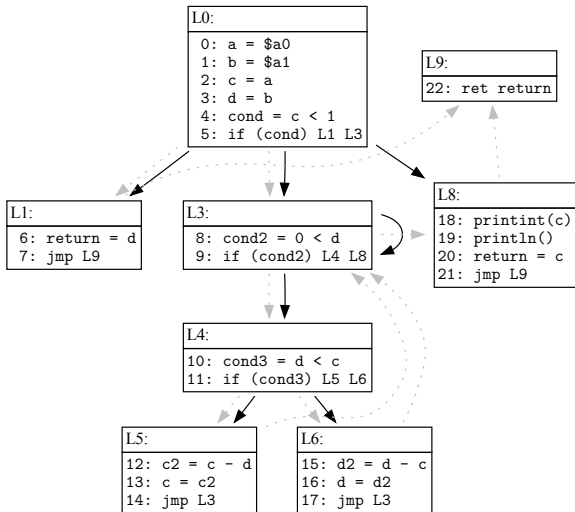
- Un bloc  $n$  a une **dépendance de contrôle** envers le bloc  $f$  si
- $f$  appartient à la **frontière de post-dominance** de  $n$
- $\rightarrow$  Il existe  $s$ ,  $f \rightarrow s$  tel que  $n$  pdom  $s$  et  $\neg(n \text{ spdom } f)$

## Exercice : postes frontières

- Diapo précédente

# Exercice : postes frontières

L0:[] L1:[L0] L3:[L0, L3] L4:[L3] L5:[L4] L6:[L4] L8:[L0] L9:[]



gcd.c: gcd()

# Élimination agressive de code mort (ADCE)

Les registres sont **inutiles** (morts), sauf preuve du contraire

## Sont **utiles**

- Les opérandes et les blocs des opérations utiles
- `call`, `ret`
- Les opérations qui calculent des registres **utiles**
- Les `if` de frontière de post-domination d'un bloc utile
  - Les aiguillages doivent fonctionner...
- Les blocs entrants d'un  $\Phi$  utile
  - Car les `mov` du  $\Phi$  *sont sur* les arcs entrants

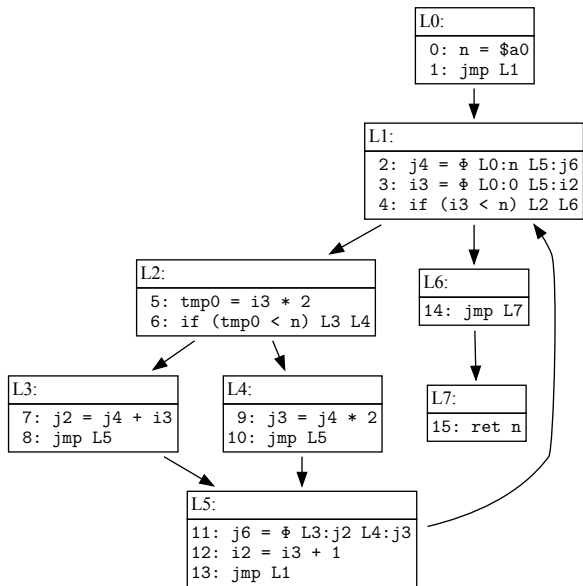
## Supprimer l'inutile

- Tout ce qui n'est pas marqué utile
- On garde quand même les `jmp`
  - La simplification du CFG s'en occupera
- Un `if` inutile est remplacé par un `jmp`

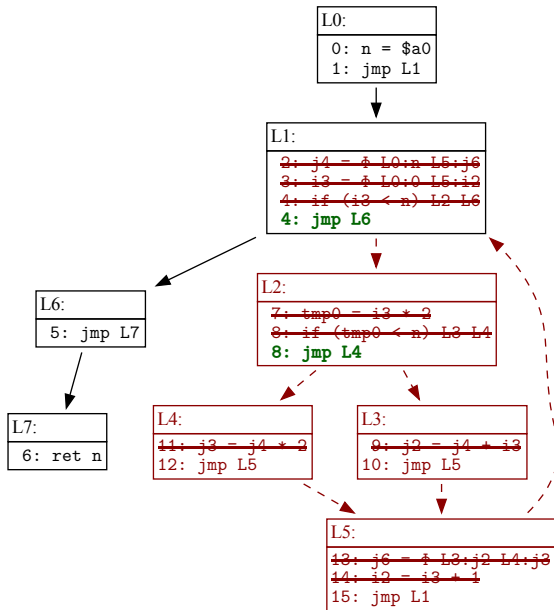
# Implémentation

- Une simple liste de travail pour ramasser les trucs utiles
- Une itération pour éliminer
- → `AggressiveDeadCodeElimination.java`

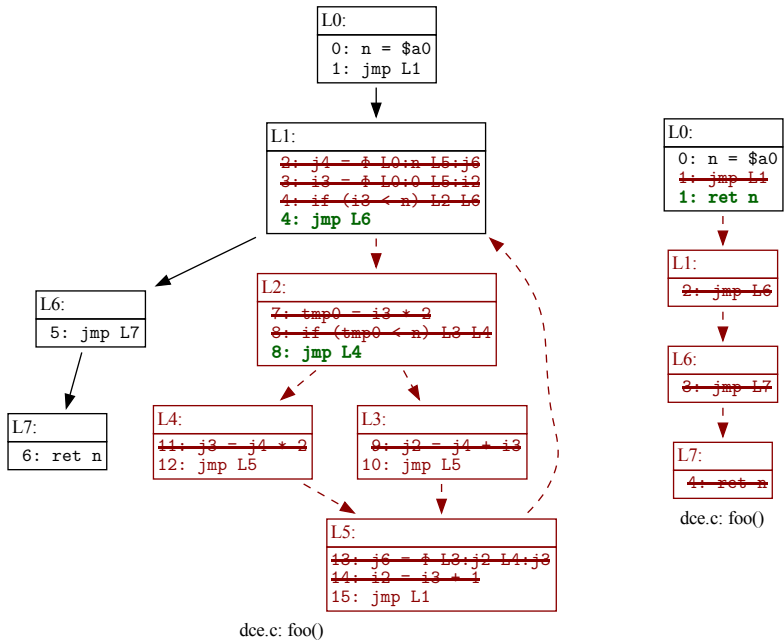
# Exercice : ADCE



dce.c: foo()



dce.c: foo()





# Remontée de code (*code hoisting*)

## *Very busy expression* (déjà vu)

- Quelles opérations seront nécessairement faites plus tard ?
- ... ou vraisemblablement ? (on prend un chance!)
- → Les remonter

## Invariants de boucle

- Sortie du code des boucles
- Une autre fois...

## Gains et pertes

# Remontée de code (*code hoisting*)

## *Very busy expression* (déjà vu)

- Quelles opérations seront nécessairement faites plus tard ?
- ... ou vraisemblablement ? (on prend un chance!)
- → Les remonter

## Invariants de boucle

- Sortie du code des boucles
- Une autre fois...

## Gains et pertes

- Réduit la taille du code
- Réduit éventuellement le temps d'exécution
- Les opérandes deviennent peut-être inutiles plus tôt
- Un registre à faire vivre plus longtemps

# Remat rialisation (*rematerialization, remat*)

- Contre-mesure !
- Recalculer une valeur plut t que la stocker dans un registre

Pourquoi ?

# Remat rialisation (*rematerialization, remat*)

- Contre-mesure !
- Recalculer une valeur plut t que la stocker dans un registre

## Pourquoi ?

- Un recalcul est moins cher qu'un registre *spill *

## Contraintes

- Calculs : les op randes doivent  tre encore disponibles
- Les initialisations   un imm diat sont de bons candidats
- Doit  tre int gr  avec l'allocation de registres

# Code sinking

- Descendre une instruction à un bloc successeur
- Facile en SSA !

## Gains

# Code sinking

- Descendre une instruction à un bloc successeur
- Facile en SSA !

## Gains

- Un registre à faire survivre moins longtemps
- L'instruction sera peut-être même pas exécutée

## Pertes

- Augmentation de la durée de vie des opérandes
- Déplacement par erreur dans un boucle
  - Calcul exécuté plusieurs fois pour rien
- Pression sur le parallélisme processeur
  - Définition juste avant les utilisations, pas forcément optimal

# Ordonnancement des instructions (*instruction scheduling*)

Objectif 1 : Réduire la pression sur les registres

Objectif 2 : Profiter du parallélisme d'instructions des processeurs modernes : **pipelines**

- → Prévenir les **bulles** (*bubble* ou *pipeline stall*)
- → Réordonner les instructions

## Contraintes

- Besoin de connaître les détails micro-architecturaux (très dépendant de la plateforme)
  - Souvent par rétro-ingénierie : on mesure en boîte-noire et on construit des modèles plus ou moins représentatifs
  - **-march** et **-mtune**
- À faire vers la fin du processus de compilation

# Peephole (à la lorgnette? au judas?)

McKeeman, 1965

## Cherche et substitue des patterns

- Localement
- Sur un petit nombre d'instructions
- Très nombreuses et spécifiques (plus de 1000 chez LLVM)

## Exemples indépendant de la plateforme

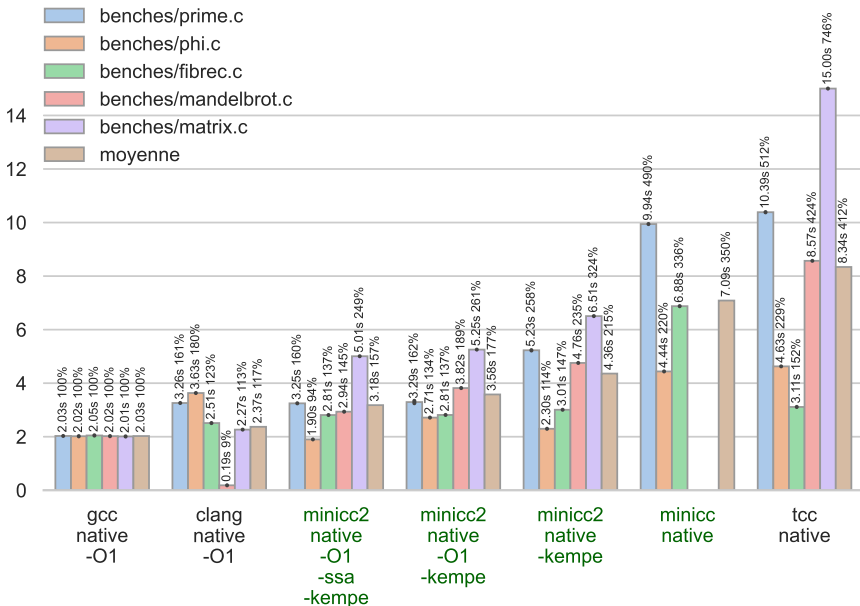
- *Constant folding* ( $x=40+2 \rightarrow x=42$ )
- *Strength reduction* ( $x=y*8 \rightarrow x=y<<3$ )

## Exemples Dépendant de la plateforme

- Très utile en CISC : beaucoup d'instructions qui font plusieurs trucs d'un coup
- $a=a+1 \rightarrow \text{INC } a$
- $b=a*8; d=b+c; e=d+42 \rightarrow \text{LEA } e, [c + a*8 + 42]$



# MiniCC2 -O1



# La prochaine fois

- Les boucles