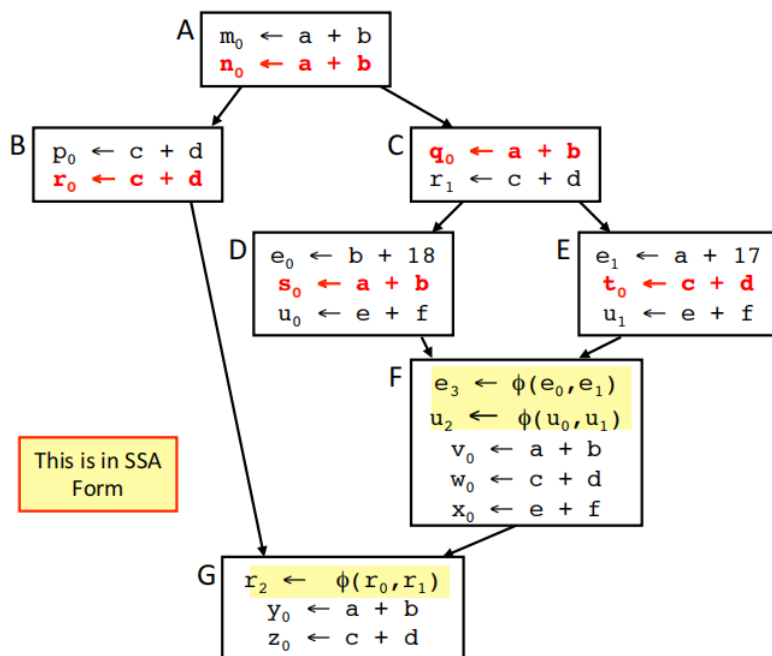


Oltre l'ottimizzazione regionale

Con l'uso di tecniche avanzate (*bells & whistles*), il *Superlocal Value Numbering* consente di individuare un numero maggiore di ridondanze con un costo aggiuntivo contenuto

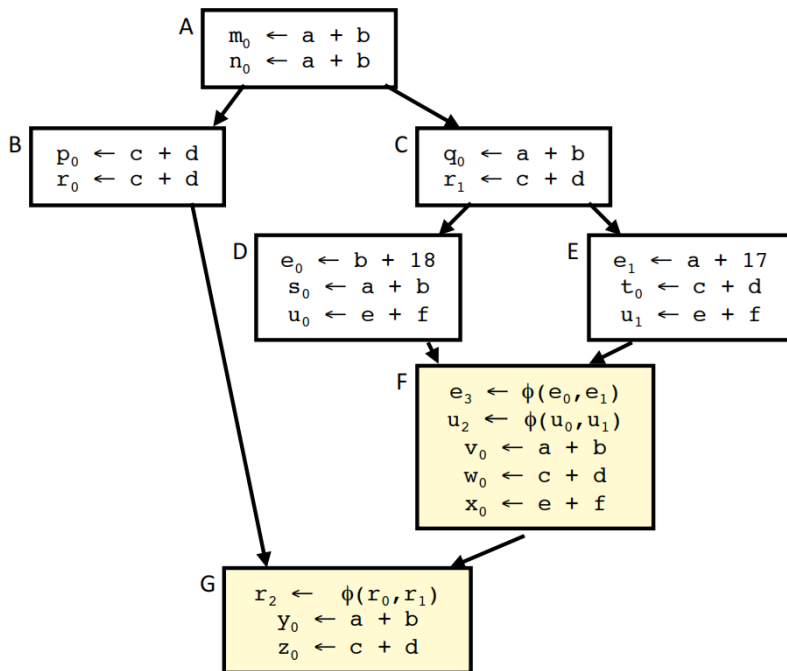


Tuttavia, questa tecnica non risolve i problemi relativi ai blocchi *F* e *G*.

Tecniche Superlocali

Alcuni metodi locali possono essere estesi, ma rimangono limitati nell'affrontare contesti più ampi.

Cosa Succede con Ambiti Più Ampi?



Il problema principale riguarda blocchi come F e G , che hanno caratteristiche specifiche:

- **Molteplici predecessori.** È necessario decidere quali informazioni rimangono valide in F e G .
- **Combinazione di informazioni.** Ad esempio, per il blocco G , bisogna combinare le informazioni provenienti da B e F . Questo processo può risultare costoso a causa della complessità della fusione degli stati.

Limiti del Superlocal Value Numbering

- **Fallback su ambiti limitati.** In presenza di incertezze, si ricorre a una gestione più semplice e locale, limitando l'efficacia sulle strutture più complesse.
- **Incapacità di "retrocedere".** Se il blocco C aggiunge valori a A , il sistema non è in grado di gestire correttamente questa dipendenza, creando ulteriori problemi nell'ottimizzazione.

Dominatori

In un grafo di flusso, un nodo x **domina** un nodo y se e solo se ogni percorso che parte dal nodo di ingresso del grafo di controllo e raggiunge y passa necessariamente per x . In particolare:

- Un nodo domina sempre sé stesso: x domina x .
- A ciascun nodo è associato un insieme di dominatori, indicato con $DOM(x)$.
- L'insieme dei dominatori di un nodo contiene almeno un elemento: $|DOM(x)| \geq 1$.

I dominatori trovano numerose applicazioni nell'analisi e nella trasformazione del codice, tra cui:

- Individuazione dei cicli (loop detection).
- Costruzione della forma SSA (*Static Single Assignment*).
- Supporto nelle decisioni per il movimento del codice (*code motion*).

Per ogni nodo x , esiste un nodo y in $DOM(x)$ che è il più vicino a x . Questo nodo è chiamato **dominatore immediato** (*Immediate Dominator*) di x e viene indicato come $IDOM(x)$. Inoltre:

- Se x è il nodo iniziale del grafo ($x = n_0$), allora x non ha un dominatore immediato, cioè $x \neq IDOM(x)$.
- Il dominatore immediato è un concetto cruciale per strutturare l'analisi dei grafi di flusso.

I dominatori possono migliorare il superlocal value numbering, ma ci sono alcune sfide da affrontare. Non abbiamo fornito supporto né per F né per G in situazioni con più predecessori, rendendo necessario decidere quali fatti siano validi in F e in G .

Per G , potrebbe essere necessario combinare B e F , ma la fusione degli stati risulta costosa. In questi casi, si può fare affidamento su ciò che è già noto. È possibile utilizzare una tabella generata da $IDOM(x)$ per iniziare con x , assegnando C per F e A per G . Questo approccio impone un ordine di applicazione basato sui dominatori.

Questa strategia porta all'introduzione della tecnica di Numerazione dei Dominatori (Dominador VN Technique, DVNT).

Dominator Value Numbering (DVNT)

L'algoritmo **Dominator Value Numbering (DVNT)** si basa sull'uso di un algoritmo superlocale applicato a blocchi di base estesi (Extended Basic Blocks, EBBs). Mantiene l'uso di tabelle hash con ambiti definiti e lo spazio dei nomi del Single Static Assignment (SSA).

Ogni nodo inizia con una tabella derivata dal suo IDOM, generalizzando così l'algoritmo superlocale. Tuttavia, i valori non fluiscono lungo gli edge di ritorno, ossia attorno ai cicli. Le tecniche di constant folding e le identità algebriche rimangono valide come in precedenza.

L'estensione del contesto consente di ottenere risultati potenzialmente migliori, combinando **LVN** (Local Value Numbering), **SVN** (Superlocal Value Numbering) e una solida base per gli EBBs che lo SVN potrebbe non considerare.

Vantaggi e Limiti della DVNT

L'approccio **DVNT** consente di individuare una maggiore quantità di ridondanze con un costo aggiuntivo minimo. Inoltre, mantiene un carattere online, garantendo un'analisi dinamica e incrementale.

Nonostante i suoi punti di forza, la DVNT presenta alcune limitazioni. Può mancare alcune opportunità di ottimizzazione, in particolare non riesce a gestire efficacemente le **Common Subexpression Eliminations (CSE)** o le costanti trasportate all'interno di loop.

Approfondimento sull'Analisi del Flusso di Dati

Il calcolo iterativo dei dominatori (*DOM*) è un esempio pratico di analisi del flusso di dati. Questa disciplina include tecniche utili per ragionare, in fase di compilazione, sul comportamento del flusso di valori durante l'esecuzione del programma. L'analisi del flusso di dati opera tipicamente su un grafo, dove i problemi locali risultano semplici all'interno di un blocco base, mentre i problemi globali sfruttano il grafo del flusso di controllo o derivati. Nei contesti interprocedurali si utilizza invece il grafo delle chiamate o sue varianti.

I problemi di flusso di dati sono formulati come un sistema di equazioni simultanee. A ciascun nodo e arco del grafo sono associati insiemi, e una delle tecniche di risoluzione più comuni è l'algoritmo iterativo. La soluzione desiderata è spesso una **meet over all paths (MOP)**, che risponde a domande come "Cosa è vero su ogni cammino dal nodo d'ingresso?" oppure "Questo evento può verificarsi su qualche cammino?". Questo tipo di analisi è strettamente legato al concetto di sicurezza.

Funzionamento dell'Algoritmo Iterativo

Il meccanismo di terminazione è garantito dal fatto che gli insiemi *DOM* iniziano come un insieme finito di nodi e si riducono in modo monotono. Questo porta inevitabilmente l'algoritmo a raggiungere un punto fisso in cui gli insiemi smettono di cambiare.

La correttezza si basa sulla dimostrazione che la soluzione al punto fisso coincide con la soluzione **MOP**. Sebbene la dimostrazione dettagliata non sia trattata in questa sede, sarà approfondita in lezioni successive.

In termini di efficienza, l'algoritmo round-robin non è particolarmente performante. L'ordine con cui vengono visitati i nodi nel grafo è un elemento chiave per migliorare la rapidità della convergenza verso la soluzione desiderata.

Calcolo dei Dominatori

Passo critico nella costruzione di SSA e DVNT

Un nodo n domina un nodo m se n si trova in ogni cammino che parte da n_0 e arriva a m . Ogni nodo domina sé stesso, mentre il dominatore immediato di n , indicato come $IDOM(n)$, è il dominatore più vicino a n .

Definizioni

La definizione formale dei dominatori è data dalle seguenti relazioni:

$DOM(n_0) = \{n_0\}$, ovvero il nodo iniziale domina solo sé stesso. Per ogni altro nodo n , si ha:

$$DOM(n) = \{n\} \cup \left(\bigcap_{p \in preds(n)} DOM(p) \right),$$

dove $preds(n)$ rappresenta l'insieme dei predecessori di n . Inizialmente, $DOM(n) = N$ per ogni nodo n diverso da n_0 .

Queste equazioni simultanee sui set definiscono un semplice problema nell'analisi del flusso di dati. Le equazioni ammettono una soluzione univoca al punto fisso, che può essere trovata rapidamente utilizzando un algoritmo iterativo.

Per convenzione, $IDOM(n) \neq n$, a meno che $n = n_0$.

Algoritmo Iterativo Round-Robin

Terminazione

L'algoritmo esegue cicli sui nodi del grafo e si arresta quando un ciclo non produce alcuna modifica negli insiemi dei dominatori.

Procedura

L'algoritmo inizia con l'inizializzazione degli insiemi DOM . Per il nodo iniziale b_0 , si assegna $DOM(b_0) \leftarrow \emptyset$. Per ogni altro nodo b_i , $DOM(b_i)$ è inizialmente considerato come l'insieme di tutti i nodi del grafo.

Durante le iterazioni, una variabile *change* viene impostata a vero. Finché *change* rimane vero, l'algoritmo continua a processare i nodi. Per ogni nodo b_i , si calcola un insieme temporaneo $TEMP$ come segue:

$$TEMP \leftarrow \{i\} \cup \left(\bigcap_{x \in pred(b_i)} DOM(x) \right).$$

Se $DOM(b_i)$ è diverso da $TEMP$, allora *change* viene impostato nuovamente a vero, e $DOM(b_i)$ viene aggiornato con il valore di $TEMP$.

Il processo continua fino a quando un ciclo completo non produce più cambiamenti negli insiemi DOM , segnando la convergenza al punto fisso.