Liveness Analysis and Register Allocation

Cosmin E. Oancea
cosmin.oancea@diku.dk
Modified by Marco Valtorta (UofSC) for CSCE 531 Spring 2020

Department of Computer Science (DIKU)
University of Copenhagen

February 2018 IPS Lecture Slides
Structure of a Compiler

Program text
↓
Lexical analysis
↓
Symbol sequence
↓
Syntax analysis
↓
Syntax tree
↓
Type Checking
↓
Syntax tree
↓
Intermediate code generation

Intermediary code:
→
Intermediate code

Binary machine code
↑
Assembly and linking
↑
Ditto with named registers
↑
Register allocation
↑
Symbolic machine code
↑
Machine code generation
↑
Intermediate code
1. Problem Statement and Intuition

2. Liveness-Analysis Preliminaries: Succ, Gen and Kill Sets

3. Liveness Analysis: Equations, Fix-Point Iteration and Interference

4. Register-Allocation via Coloring: Interference Graph & Intuitive Alg

5. Register-Allocation via Coloring: Improved Algorithm with Spilling
Problem Statement

Processors have a limited number of registers:

- **X86**: 8 (integer) registers,
- **ARM**: 16 (integer) registers,
- **MIPS**: 31 (integer) registers.

In addition, 3 – 4 special-purpose registers (can’t hold variables).

Solution:

- Whenever possible, let several variables share the same register,
- If there are still variables that cannot be mapped to a register, store them in memory.
Where to Implement Register Allocation?

Two possibilities: at IL or at machine-language level. Pro/Cons?
Where to Implement Register Allocation?

Two possibilities: at IL or at machine-language level. Pro/Cons?

- **IL Level:**
  - Can be shared between multiple architectures (parameterized on the number of registers).
  - Translation to machine code can introduce/remove intermediate results.

- **Machine-Code Level:**
  - Accurate, near-optimal mapping.
  - Implemented for every architecture, no code reuse.

We show register allocation at IL level. Similar for machine code.
Register-Allocation Scope

- Code Sequence Without Jumps:
  - Simple.
  - A variable is saved to memory when jumps occur.

- Procedure/Function Level:
  - Variables can still be in registers even across jumps.
  - A bit more complicated.
  - Variables saved to memory before function calls.

- Module/Program Level:
  - Sometimes variables can still be hold in registers across function calls (but not always: recursion).
  - More complicated alg of higher time complexity.

Most compilers implement register allocation at function level.
When Can Two Variables Share a Register?

Intuition: Two vars can share a register if the two variables do not have overlapping periods of use.

Period of Use: From var’s assignment to the last use of the assigned value. A variable can have several periods of use (live ranges).

Liveness: If a variable’s value may be used on the continuation of an execution path passing through program point PP, then the variable is live at PP. Otherwise: dead at PP.
When Can Two Variables Share a Register?

With the code below, can variables \( a \) and \( c \) share the same register?

\[
\begin{align*}
a & := 1 \\
c & := a + 1 \\
a & := c + 3 \\
a & := a + 2
\end{align*}
\]

(a) TRUE  
(b) FALSE
1. Problem Statement and Intuition

2. Liveness-Analysis Preliminaries: Succ, Gen and Kill Sets

3. Liveness Analysis: Equations, Fix-Point Iteration and Interference

4. Register-Allocation via Coloring: Interference Graph & Intuitive Alg

5. Register-Allocation via Coloring: Improved Algorithm with Spilling
Prioritized Rules for Liveness

1) If a variable, \( \text{VAR} \), is used, i.e., its value, in an instruction, \( I \), then \( \text{VAR} \) is live at the entry of \( I \).

2) If \( \text{VAR} \) is assigned a value in instruction \( I \) (and 1 does not apply) then \( \text{VAR} \) is dead at the entry of \( I \).

3) If \( \text{VAR} \) is live at the end of instruction \( I \) then it is live at the entry of \( I \) (unless 2 applies).

4) A \( \text{VAR} \) is live at the end of instruction \( I \) \( \Leftrightarrow \) \( \text{VAR} \) is live at the entry of any instructions that may be executed immediately after \( I \), i.e., immediate successors of \( I \).
Liveness-Analysis Concepts

We number program instructions from 1 to $n$.
For each instruction we define the following sets:

$succ[i]$: The instructions (numbers) that can possibly be executed immediately after instruction (numbered) $i$.

$gen[i]$: The set of variables whose values are read by instruction $i$.

$kill[i]$: The set of variables that are overwritten by instruction $i$.

$in[i]$: The set of variables that are live at the entry of instruction $i$.

$out[i]$: The set of variables that are live at the end of instruction $i$.

In the end, what we need is $out[i]$ for all instructions.
Immediate Successors

- \( \text{succ}[i] = \{i + 1\} \) unless instruction \( i \) is a GOTO, an IF-THEN-ELSE, or the last instruction of the program.

- \( \text{succ}[i] = \{j\} \), if instruction \( i \) is: GOTO \( l \)
  and instruction \( j \) is: LABEL \( l \).

- \( \text{succ}[i] = \{j, k\} \), if instruction \( i \) is IF \( c \) THEN \( l_1 \) ELSE \( l_2 \),
  instruction \( j \) is LABEL \( l_1 \), and instruction \( k \) is LABEL \( l_2 \).

- If \( n \) denotes the last instruction of the program, and \( n \) is not a GOTO or an IF-THEN-ELSE instruction, then \( \text{succ}[n] = \emptyset \).

Note: Programs always exit by executing a RETURN instruction.
## Rules for Constructing \textit{gen} and \textit{kill} Sets

Below $k$ denotes a constant (value), $M[...]$ denotes memory access.

<table>
<thead>
<tr>
<th>Instruction $i$</th>
<th>$\text{gen}[i]$</th>
<th>$\text{kill}[i]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LABEL $l$</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>$x := y$</td>
<td>${y}$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$x := k$</td>
<td>$\emptyset$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$x := \text{unop } y$</td>
<td>${y}$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$x := \text{unop } k$</td>
<td>$\emptyset$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$x := y \text{ binop } z$</td>
<td>${y, z}$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$x := y \text{ binop } k$</td>
<td>${y}$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$x := M[y]$</td>
<td>${y}$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$x := M[k]$</td>
<td>$\emptyset$</td>
<td>${x}$</td>
</tr>
<tr>
<td>$M[x] := y$</td>
<td>${x, y}$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>$M[k] := y$</td>
<td>${y}$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>GOTO $l$</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>IF $x \text{ relop } y$ THEN $l_t$ ELSE $l_f$</td>
<td>${x, y}$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>$x := \text{CALL } f(args)$</td>
<td>$\text{args}$</td>
<td>${x}$</td>
</tr>
</tbody>
</table>
Gen & Kill Sets Multiple-Choice

The *kill* and *gen* sets of instruction \( x := a + x \) are:

(A) \( \text{kill} = \emptyset, \quad \text{gen} = \{a\} \)
(B) \( \text{kill} = \{x\}, \quad \text{gen} = \{a, x\} \)
(C) \( \text{kill} = \{a, x\}, \quad \text{gen} = \{x\} \)
(D) \( \text{kill} = \{a\}, \quad \text{gen} = \emptyset \)
(E) \( \text{kill} = \{x\}, \quad \text{gen} = \{x\} \)
Gen & Kill Sets Multiple-Choice

The *kill* and *gen* sets of instruction $M[i] := i + a$ are:

(A) $kill = \emptyset$, $gen = \{a\}$
(B) $kill = \{i\}$, $gen = \{a, i\}$
(C) $kill = \{a, i\}$, $gen = \{i\}$
(D) $kill = \{i\}$, $gen = \{a\}$
(E) $kill = \emptyset$, $gen = \{i, a\}$
Successors Multiple-Choice Question

With the code below, which of the following statements is TRUE?

1. \( x := 0 \)
2. \( \text{IF } x = 0 \text{ THEN lab1 ELSE lab2} \)
3. Label lab1:
4. \( x := 3 \)
5. \( \text{GOTO lab3} \)
6. Label lab2:
7. \( x := 4 \)
8. Label lab3:

(A) \( \text{succ}[2] = \{3, 6, 8\} \) and \( \text{succ}[4] = \{5\} \)
(B) \( \text{succ}[2] = \{3\} \) and \( \text{succ}[5] = \{6\} \)
(C) \( \text{succ}[2] = \{3\} \) and \( \text{succ}[1] = \{2\} \)
(D) \( \text{succ}[2] = \{3, 6\} \) and \( \text{succ}[5] = \{8\} \)
(E) \( \text{succ}[3] = \{2\} \) and \( \text{succ}[6] = \{2\} \)
1. Problem Statement and Intuition

2. Liveness-Analysis Preliminaries: \textit{Succ}, \textit{Gen} and \textit{Kill} Sets

3. Liveness Analysis: Equations, Fix-Point Iteration and Interference

4. Register-Allocation via Coloring: Interference Graph & Intuitive Alg

5. Register-Allocation via Coloring: Improved Algorithm with Spilling
Data-Flow Equation for Liveness Analysis

4) A VAR is live at the end of instruction I ⇔ VAR is live at the entry of any instructions that may be executed immediately after I, i.e., immediate successors of I.

The CORRECT Equation for \( out[i] \) is:

(A) \( out[i] = in[i] \)

(B) \( out[i] = \bigcup_{j \in \text{succ}[i]} in[j] \)

(C) \( out[i] = gen[i] \cup (\bigcup_{j \in \text{succ}[i]} in[j]) \)

(D) \( out[i] = in[i] \setminus (\bigcup_{j \in \text{succ}[i]} in[j]) \)

(E) \( out[i] = (gen[i] \cup out[i]) \setminus \text{kill}[i] \)

Recall that \( \setminus \) indicates set difference, i.e. \( A \setminus B \) is the set of elements of A that are not in B.
Data-Flow Equation for Liveness Analysis

1) If a variable, $\text{VAR}$, is used, i.e., its value, in an instruction, $I$, then $\text{VAR}$ is live at the entry of $I$.

2) If $\text{VAR}$ is assigned a value in instruction $I$ (and 1 does not apply) then $\text{VAR}$ is dead at the entry of $I$.

3) If $\text{VAR}$ is live at the end of instruction $I$ then it is live at the entry of $I$ (unless 2 applies).

The CORRECT Equation for $\text{in}[i]$ is:

(A) $\text{in}[i] = \text{gen}[i] \cup \text{out}[i]$
(B) $\text{in}[i] = \text{gen}[i] \setminus \text{kill}[i]$
(C) $\text{in}[i] = \text{gen}[i] \cup (\text{out}[i] \setminus \text{kill}[i])$
(D) $\text{in}[i] = (\text{gen}[i] \setminus \text{kill}[i]) \cup \text{out}[i]$
(E) $\text{in}[i] = (\text{gen}[i] \cup \text{out}[i]) \setminus \text{kill}[i]$
Data-Flow Equations for Liveness Analysis

\[ \text{in}[i] = \text{gen}[i] \cup (\text{out}[i] \setminus \text{kill}[i]) \] (1)

\[ \text{out}[i] = \bigcup_{j \in \text{succ}[i]} \text{in}[j] \] (2)

Exception: If \( \text{succ}[i] = \emptyset \), then \( \text{out}[i] \) is the set of variables that appear in the function’s result.

The (recursive) equations are solved by iterating to a fix point: \( \text{in}[i] \) and \( \text{out}[i] \) are initialized to \( \emptyset \), and iterate until no changes occur.

Why does it converge?

For fast(er) convergence: compute \( \text{out}[i] \) before \( \text{in}[i] \) and \( \text{in}[i + 1] \) before \( \text{out}[i] \), respectively (i.e., backward flow analysis).
Imperative-Fibonacci Example

\[
\text{fibo}(n) \begin{align*}
1: & \quad a := 0 \\
2: & \quad b := 1 \\
3: & \quad z := 0 \\
4: & \quad \text{LABEL loop} \\
5: & \quad \text{IF } n = z \text{ THEN end ELSE body} \\
6: & \quad \text{LABEL body} \\
7: & \quad t := a + b \\
8: & \quad a := b \\
9: & \quad b := t \\
10: & \quad n := n - 1 \\
11: & \quad z := 0 \\
12: & \quad \text{GOTO loop} \\
13: & \quad \text{LABEL end}
\end{align*}
\]

\begin{array}{|c|c|c|c|}
\hline
i & succ[i] & gen[i] & kill[i] \\
\hline
1 & 2 & & a \\
2 & 3 & & b \\
3 & 4 & & z \\
4 & 5 & & \\
5 & 6, 13 & n, z & \\
6 & 7 & & \\
7 & 8 & a, b & t \\
8 & 9 & b & a \\
9 & 10 & t & b \\
10 & 11 & n & n \\
11 & 12 & & z \\
12 & 4 & & \\
13 & & & \\
\hline
\end{array}

We omitted RETURN a. Means \(out[i] = \{a\}\) result used after fct call.
Imperative-Fibonacci Example

\[
fibo(n)1: \quad a := 0 \\
2: \quad b := 1 \\
3: \quad z := 0 \\
4: \quad \text{LABEL loop} \\
5: \quad \text{IF } n = z \text{ THEN end ELSE body} \\
6: \quad \text{LABEL body} \\
7: \quad t := a + b \\
8: \quad a := b \\
9: \quad b := t \\
10: \quad n := n - 1 \\
11: \quad z := 0 \\
12: \quad \text{GOTO loop} \\
13: \quad \text{LABEL end}
\]

<table>
<thead>
<tr>
<th></th>
<th>succ[i]</th>
<th>gen[i]</th>
<th>kill[i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td></td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td></td>
<td>b</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td></td>
<td>z</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6, 13</td>
<td>n, z</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>a, b</td>
<td>t</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>t</td>
<td>b</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td></td>
<td>z</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We omitted \text{RETURN a}. Means \text{out[i]} = \{a\} result used after fct call.
Fix-Point Iteration for the Fibonacci Example

Use backwards evaluation order: out[14], in[14], ..., out[1], in[1].

<table>
<thead>
<tr>
<th>i</th>
<th>Initial out[i]</th>
<th>Initial in[i]</th>
<th>Iteration 1 out[i]</th>
<th>Iteration 1 in[i]</th>
<th>Iteration 2 out[i]</th>
<th>Iteration 2 in[i]</th>
<th>Iteration 3 out[i]</th>
<th>Iteration 3 in[i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n, a</td>
<td>n</td>
<td>n, a</td>
<td>n</td>
<td>n, a</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>n, a, b</td>
<td>n, a</td>
<td>n, a, b</td>
<td>n, a</td>
<td>n, a, b</td>
<td>n, a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>n, z, a, b</td>
<td>n, a, b</td>
<td>n, z, a, b</td>
<td>n, a, b</td>
<td>n, z, a, b</td>
<td>n, a, b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>a, b, n</td>
<td>n, z, a, b</td>
<td>a, b, n</td>
<td>n, z, a, b</td>
<td>a, b, n</td>
<td>n, z, a, b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>a, b, n</td>
<td>a, b, n</td>
<td>a, b, n</td>
<td>a, b, n</td>
<td>a, b, n</td>
<td>a, b, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>b, t, n</td>
<td>a, b, n</td>
<td>b, t, n</td>
<td>a, b, n</td>
<td>b, t, n</td>
<td>a, b, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>t, n</td>
<td>b, t, n</td>
<td>t, n, a</td>
<td>b, t, n</td>
<td>t, n, a</td>
<td>b, t, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>n</td>
<td>t, n</td>
<td>n, a, b</td>
<td>t, n, a</td>
<td>n, a, b</td>
<td>t, n, a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>n</td>
<td>n, a, b</td>
<td>n, a, b</td>
<td>n, a, b</td>
<td>n, a, b</td>
<td>n, a, b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>n, z, a, b</td>
<td>n, a, b</td>
<td>n, z, a, b</td>
<td>n, a, b</td>
<td>n, z, a, b</td>
<td>n, a, b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td>n, z, a, b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Usually less than 5 iterations.
More Multiple Choice Questions

If a formal parameter $p$ is NOT LIVE at the entry point of a function, then it means that:

(A) the original value of $p$ is never used, e.g., $p$ is redefined before being used or is never used.

(B) parameter $p$ may be used before being written/updated.

(C) parameter $p$ is only written inside the function and never read

(D) parameter $p$ is only read inside the function and never written

(E) parameter $p$ is read-and-written in all instructions in which it appears, i.e., $p:=p+x$
More Multiple Choice Questions

If a variable $a$, which is NOT a formal argument, is LIVE at the entry point of a function, then it means that:

(A) $a$ is only read inside the function
(B) $a$ may be used without being initialized
(C) $a$ is only written inside the function
(D) $a$ will certainly be used before being initialized
(E) $a$ is read-and-written in all instructions in which it appears, i.e., $a := a + x$
1. Problem Statement and Intuition

2. Liveness-Analysis Preliminaries: Succ, Gen and Kill Sets

3. Liveness Analysis: Equations, Fix-Point Iteration and Interference

4. Register-Allocation via Coloring: Interference Graph & Intuitive Alg

5. Register-Allocation via Coloring: Improved Algorithm with Spilling
Interference

Definition: Variable $x$ interferes with variable $y$, if there is an instruction numbered $i$ such that:

1. $x \in \text{kill}[i]$ and
2. $y \in \text{out}[i]$ and
3. $x \neq y$ and
4. If instruction $i$ is $x := y$ then $x$ does not interferes with $y$ (but it interferes with any other variable in $\text{out}[i]$)

Two variables can share the same register iff they do not interfere with each other!
## Interference for the Fibonacci Example

<table>
<thead>
<tr>
<th>Instruction</th>
<th>Left-hand side</th>
<th>Interferes with</th>
<th>out(i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 : $a := 0$</td>
<td>$a$</td>
<td>$n$</td>
<td>$n, a$</td>
</tr>
<tr>
<td>2 : $b := 1$</td>
<td>$b$</td>
<td>$n, a$</td>
<td>$n, a, b$</td>
</tr>
<tr>
<td>3 : $z := 0$</td>
<td>$z$</td>
<td>$n, a, b$</td>
<td>$n, z, a, b$</td>
</tr>
<tr>
<td>7 : $t := a + b$</td>
<td>$t$</td>
<td>$b, n$</td>
<td>$a, b, n$</td>
</tr>
<tr>
<td>8 : $a := b$</td>
<td>$a$</td>
<td>$t, n$</td>
<td>$t, b, n$</td>
</tr>
<tr>
<td>9 : $b := t$</td>
<td>$b$</td>
<td>$n, a$</td>
<td>$t, n, a$</td>
</tr>
<tr>
<td>10 : $n := n - 1$</td>
<td>$n$</td>
<td>$a, b$</td>
<td>$n, a, b$</td>
</tr>
<tr>
<td>11 : $z := 0$</td>
<td>$z$</td>
<td>$n, a, b$</td>
<td>$n, z, a, b$</td>
</tr>
</tbody>
</table>

Since interference is a symmetric and non-reflexive relation, we can draw interference as a (undirected) graph:

![Interference Graph](image-url)
Register Allocation By Graph Coloring

Two variables connected by an edge in the interference graph cannot share a register!

Idea: Associate variables with register numbers such that:

1. Two variables connected by an edge receive different numbers.
2. Numbers represent the (limited number of) hardware registers.

Equivalent to graph-coloring problem: color each node with one of \( n \) (available) colors, such that any two neighbors are colored differently.

Since graph coloring is NP complete, we use a heuristic method that gives good results in most cases.

Idea: a node with less-than-\( n \) neighbors can always be colored. Eliminate such nodes from the graph and solve recursively!
Coloring The Graph With Four Colors

z and t have only three neighbors so they can wait.
The remaining three nodes can now be given different colors!
Coloring The Graph With Four Colors

$z$ and $t$ can now be given a different color!
Coloring The Graph With Four Colors

But what if we only have three colors (registers) available?
1. Problem Statement and Intuition

2. Liveness-Analysis Preliminaries: $Succ$, $Gen$ and $Kill$ Sets

3. Liveness Analysis: Equations, Fix-Point Iteration and Interference

4. Register-Allocation via Coloring: Interference Graph & Intuitive Alg

5. Register-Allocation via Coloring: Improved Algorithm with Spilling
Improved Algorithm

Initialization: Start with an empty stack.

Simplify: 1) If there is a node with less than \( n \) edges (neighbors): (i) place it on the stack together with the list of edges, and (ii) remove it and its edges from the graph.
2. If there is no node with less than \( n \) neighbors, pick any node and do as above.
3. Continue until the graph is empty. If so go to select.

Select: 1. Take a node and its neighbor list from the stack.
2. If possible, color it differently than its neighbor’s.
3. If not possible, select the node for spilling (fails).
4. Repeat until stack is empty.

The quality of the result depends on (i) how to chose a node in simplify, and (ii) how to chose a color in select.
Example: Coloring the Graph with Three Colors

No node has < 3 neighbors, hence choose arbitrarily, say $z$.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z$</td>
<td>$a, b, n$</td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

There are still no nodes with $< 3$ neighbors, hence we chose $a$.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>$b, n, t$</td>
<td></td>
</tr>
<tr>
<td>$z$</td>
<td>$a, b, n$</td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

$b$ has two neighbors, so we choose it.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b$</td>
<td>$t, n$</td>
<td></td>
</tr>
<tr>
<td>$a$</td>
<td>$b, n, t$</td>
<td></td>
</tr>
<tr>
<td>$z$</td>
<td>$a, b, n$</td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

Finally, choose $t$ and $n$.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$n$</td>
<td></td>
</tr>
<tr>
<td>$t$</td>
<td>$n$, $t$</td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>$t$, $n$</td>
<td></td>
</tr>
<tr>
<td>$a$</td>
<td>$b$, $n$, $t$</td>
<td></td>
</tr>
<tr>
<td>$z$</td>
<td>$a$, $b$, $n$</td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

$n$ has no neighbors so we can choose 1.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$n$</td>
<td>1</td>
</tr>
<tr>
<td>$t$</td>
<td>$n$</td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>$t, n$</td>
<td></td>
</tr>
<tr>
<td>$a$</td>
<td>$b, n, t$</td>
<td></td>
</tr>
<tr>
<td>$z$</td>
<td>$a, b, n$</td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

t only has \( n \) as neighbor, so we can color it with 2.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( n )</td>
<td>1</td>
</tr>
<tr>
<td>( t )</td>
<td>( t, n )</td>
<td>2</td>
</tr>
<tr>
<td>( b )</td>
<td>( b, n, t )</td>
<td></td>
</tr>
<tr>
<td>( a )</td>
<td>( a, b, n )</td>
<td></td>
</tr>
<tr>
<td>( z )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

$b$ has $t$ and $n$ as neighbors, hence we can color it with 3.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$t$</td>
<td>$n$</td>
<td>2</td>
</tr>
<tr>
<td>$b$</td>
<td>$t, n$</td>
<td>3</td>
</tr>
<tr>
<td>$a$</td>
<td>$b, n, t$</td>
<td></td>
</tr>
<tr>
<td>$z$</td>
<td>$a, b, n$</td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

$\text{a}$ has three differently-colored neighbors, so it is marked as $\text{spill}$.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$n$</td>
<td>1</td>
</tr>
<tr>
<td>$t$</td>
<td>$n$</td>
<td>2</td>
</tr>
<tr>
<td>$b$</td>
<td>$t, n$</td>
<td>3</td>
</tr>
<tr>
<td>$a$</td>
<td>$b, n, t$</td>
<td>spill</td>
</tr>
<tr>
<td>$z$</td>
<td>$a, b, n$</td>
<td></td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

z has colors 1 and 3 as neighbors, hence we can color it with 2.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>n</td>
<td>1</td>
</tr>
<tr>
<td>t</td>
<td>n</td>
<td>2</td>
</tr>
<tr>
<td>b</td>
<td>t, n</td>
<td>3</td>
</tr>
<tr>
<td>a</td>
<td>b, n, t</td>
<td>spill</td>
</tr>
<tr>
<td>z</td>
<td>a, b, n</td>
<td>2</td>
</tr>
</tbody>
</table>
Example: Coloring the Graph with Three Colors

We are now finished, but we need to *spill* $a$.

<table>
<thead>
<tr>
<th>Node</th>
<th>Neighbours</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$n$</td>
<td>1</td>
</tr>
<tr>
<td>$t$</td>
<td>$n$</td>
<td>2</td>
</tr>
<tr>
<td>$b$</td>
<td>$t, n$</td>
<td>3</td>
</tr>
<tr>
<td>$a$</td>
<td>$b, n, t$</td>
<td>spill</td>
</tr>
<tr>
<td>$z$</td>
<td>$a, b, n$</td>
<td>2</td>
</tr>
</tbody>
</table>
Spilling

Spilling means that some variables will reside in memory (except for brief periods). For each spilled variable:

1) Select a memory address $addr_x$, where the value of $x$ will reside.
2) If instruction $i$ uses $x$, then rename it locally to $x_i$.
3) Before an instruction $i$, which reads $x_i$, insert $x_i := M[addr_x]$.
4) After an instruction $i$, which updates $x_i$, insert $M[addr_x] := x_i$.
5) If $x$ is alive at the beginning of the function/program, insert $M[addr_x] := x$ before the first instruction of the function.
6) If $x$ is live at the end of the program/function, insert $x := M[addr_x]$ after the last instruction of the function.

Finally, perform liveness analysis and register allocation again.
Spilling Example

1: \( a_1 := 0 \)
   \( M[address_a] := a_1 \)
2: \( b := 1 \)
3: \( z := 0 \)
4: LABEL loop
5: IF \( n = z \) THEN end ELSE body
6: LABEL body
   \( a_7 := M[address_a] \)
7: \( t := a_7 + b \)
8: \( a_8 := b \)
   \( M[address_a] := a_8 \)
9: \( b := t \)
10: \( n := n - 1 \)
11: \( z := 0 \)
12: GOTO loop
13: LABEL end
   \( a := M[address_a] \)
After Spilling, Coloring Succeeds!

Diagram:

- Nodes: $a$, $z$, $t$, $n$, $a_1$, $a_3$, $a_7$, $b$
- Edges: Connections between the nodes as shown in the diagram.
Heuristics

For **Simplify**: when choosing a node with $\geq n$ neighbors:

- Choose the node with fewest neighbors, which is more likely to be colorable, or
- Choose a node with many neighbors, each of them having close to $n$ neighbors, i.e., spilling this node would allow the coloring of its neighbors.

For **Select**: when choosing a color:

- Choose colors that have already been used.
- If instructions such as $x := y$ exist, color $x$ and $y$ with the same color, i.e., eliminate this instruction.