## CSE211: Compiler Design

 Oct. 22, 2020- Topic: Local value numbering continued and data flow analysis
- Questions:

Questions/comments about homework 1?

What are some difficult programs for local value numbering?


## Announcements

- Homework 2 released! Have a look but don't panic
- Remember, due dates pushed back 1 week
- Part 1 should be possible after today's lecture
- The theory for Part 2 is in lecture. We will go over code next lecture.



## CSE211: Compiler Design

 Oct. 22, 2020- Topic: Local value numbering continued and data flow analysis
- Questions:

Questions/comments about homework 1?

What are some difficult programs for local value numbering?


## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\longrightarrow \begin{aligned}
& \mathrm{a} 2=\mathrm{b} 0+\mathrm{c} 1 ; \\
& \mathrm{b} 4=\mathrm{a} 2-\mathrm{d} 3 ; \\
& \mathrm{c} 5=\mathrm{b} 4+\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

$$
\begin{gathered}
\mathrm{H} \\
\}
\end{gathered}=\{
$$

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\longrightarrow \begin{aligned}
& \mathrm{a} 2=\mathrm{b} 0+\mathrm{c} 1 ; \\
& \mathrm{b} 4=\mathrm{a} 2-\mathrm{d} ; \\
& \mathrm{c} 5=\mathrm{b} 4+\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

$$
\begin{aligned}
& H=\left\{{ }^{H} b 0+c 1 ": " a 2 ",\right. \\
& \}
\end{aligned}
$$

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{b} 0+\mathrm{c} 1 ; \\
& \mathrm{b} 4=\mathrm{a} 2-\mathrm{d} 3 ; \\
& \mathrm{c} 5=\mathrm{b} 4+\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

$$
\begin{gathered}
\mathrm{H}=\left\{{ }^{\prime} \mathrm{b0}+\mathrm{c} 1 ": " \mathrm{a} 2 ",\right.
\end{gathered}
$$

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\begin{aligned}
\mathrm{a} 2 & =\mathrm{b} 0+\mathrm{c} 1 ; \\
\mathrm{b} 4 & =\mathrm{a} 2-\mathrm{d} 3 ; \\
\mathrm{c} 5 & =\mathrm{b} 4+\mathrm{c} 1 ; \\
\mathrm{d} 6 & =\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{b} 0+\mathrm{c} 1 ; \\
& \mathrm{b} 4=\mathrm{a} 2-\mathrm{d} 3 ; \\
& \mathrm{c} 5=\mathrm{b} 4+\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

$\begin{array}{ll}\mathrm{H}=\left\{\begin{array}{c}\text { "b0 }+\mathrm{c} 1 ": ~ " a 2 ", ~\end{array}\right. \\ & \text { "a2-d3": "b4", }\end{array}$

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{b} 0+\mathrm{c} 1 ; \\
& \mathrm{b} 4=\mathrm{a} 2-\mathrm{d} 3 ; \\
& \mathrm{c} 5=\mathrm{b} 4+\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

$H=\{$

$$
" \mathrm{~b} 0+\mathrm{c} 1 ": " \mathrm{a} 2 "
$$

"a2 - d3" : "b4"

$$
\}
$$

mismatch due to numberings!

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\begin{aligned}
\mathrm{a} 2 & =\mathrm{b} 0+\mathrm{c} 1 ; \\
\mathrm{b} 4 & =\mathrm{a} 2-\mathrm{d} 3 ; \\
\mathrm{c} 5 & =\mathrm{b} 4+\mathrm{c} 1 ; \\
\mathrm{d} 6 & =\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{H}=\{ \\
& \text { "b0 + c1" : "a2", } \\
& \text { "a2 - d3" : "b4", } \\
& \text { "b4 + c1" : "c5", } \\
& \text { \} }
\end{aligned}
$$

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{b} 0+\mathrm{c} 1 ; \\
& \mathrm{b} 4=\mathrm{a} 2-\mathrm{d} 3 ; \\
& \mathrm{c} 5=\mathrm{b} 4+\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2-\mathrm{d} 3 ;
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{H}=\{ \\
& \text { "b0 + c1" : "a2", } \\
& \text { "a2 - d3" : "b4", } \\
& \text { "b4 + c1" : "c5", } \\
& \text { \} }
\end{aligned}
$$

## Local Value Numbering

- Algorithm: Now that variables are numbered
- Iterate sequentially through instructions. Keep a hash table of the rhs (numbered variables and operation) mapped to their Ihs.
- At each step, check to see if the rhs has already been computed.

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{b} 0+\mathrm{c} 1 ; \\
& \mathrm{b} 4=\mathrm{a} 2-\mathrm{d} 3 ; \\
& \mathrm{c} 5=\mathrm{b} 4+\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{b} 4 ;
\end{aligned}
$$

$\mathrm{H}=\{$
"b0 + c1" : "a2",
"a2 - d3" : "b4",
"b4 + c1" : "c5",

## Adding Commutativity

- Certain operators are commutative (e.g. ADD and MULTIPLY)
- In this case, the analysis should consider a deterministic order of operands.
- You can use variable numbers or lexigraphical order


## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order

$$
\longrightarrow \begin{aligned}
& \mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ; \\
& \mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 ; \\
& \mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2 * \mathrm{~d} 3 ;
\end{aligned}
$$

$$
\underset{\}}{\mathrm{H}}=\{
$$

## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order

$$
\longrightarrow \begin{aligned}
& \mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ; \\
& \mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 ; \\
& \mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2 * \mathrm{~d} 3 ;
\end{aligned}
$$

```
H = {
    "c1 - b0" : "a2",
}
```


## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order

$$
\begin{array}{|ll|}
\hline \mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ; \\
\mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 \\
\mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ; \\
\mathrm{d} 6=\mathrm{a} 2 * \mathrm{~d} 3 ;
\end{array} \quad \mathrm{H}=\left\{{ }^{\prime} \mathrm{c} 1-\mathrm{b} 0 ": " \mathrm{a} 2 ",\right.
$$

## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order
re-ordered because a2 < d3 lexigraphically

$\longrightarrow$| $\mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ;$ |
| :--- |
| $\mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 ;$ |
| $\mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ;$ |
| $\mathrm{d} 6=\mathrm{a} 2 * \mathrm{~d} 3 ;$ |

$$
\begin{aligned}
& \mathrm{H}=\{ \\
& \text { "c1 - b0" : "a2", } \\
& \text { "a2 * d3" : "f4", } \\
& \text { \} }
\end{aligned}
$$

## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ; \\
& \mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 ; \\
& \mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2 * \mathrm{~d} 3 ;
\end{aligned}
$$

```
H={
"c1 - b0" : "a2",
"a2 * d3" : "f4",
}
```


## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ; \\
& \mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 ; \\
& \mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2 * \mathrm{~d} 3 ;
\end{aligned}
$$

$$
\}
$$

$$
\begin{aligned}
& \text { H = \{ } \\
& \text { "c1 - b0" : "a2", } \\
& \text { "a2 * d3" : "f4", } \\
& \text { "b0 - c1" : "c5", }
\end{aligned}
$$

## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ; \\
& \mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 ; \\
& \mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{a} 2 * \mathrm{~d} 3 ;
\end{aligned}
$$

```
H = {
    "c1 - b0" : "a2",
    "a2 * d3" : "f4",
    "b0 - c1" : "c5",
}
```


## Local Value Numbering

- Algorithm optimization: for commutative operations, re-order operands into a deterministic order

$$
\begin{aligned}
& \mathrm{a} 2=\mathrm{c} 1-\mathrm{b} 0 ; \\
& \mathrm{f} 4=\mathrm{d} 3 * \mathrm{a} 2 ; \\
& \mathrm{c} 5=\mathrm{b} 0-\mathrm{c} 1 ; \\
& \mathrm{d} 6=\mathrm{f} 4 ;
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{H}=\{ & \\
& \text { "c1-b0": "a2", } \\
& " \mathrm{a} 2 * \mathrm{~d} " \text { " : "f4", } \\
& \text { "b0-c1": "c5", }
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- We've assumed we have access to an unlimited number of virtual registers.
- In some cases we may not be able to add virtual registers
- If an expensive register allocation pass has already occurred.
- We need to give back a program such that variables without numbers is still valid.


## Local Value Numbering w/out adding registers

- Example:



## Local Value Numbering w/out adding registers

- Solutions?

$$
\begin{aligned}
& \mathrm{a}=\mathrm{x}+\mathrm{y} ; \\
& \mathrm{a}=\mathrm{z} ; \\
& \mathrm{b}=\mathrm{x}+\mathrm{y} ;
\end{aligned} \xrightarrow{\text { numbering }} \begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& a=x+y ; \\
& a=z ; \\
& b=x+y ; \\
& c=x+y ;
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number
$a=x+y ;$
$a=z i$
$b=x+y i$
$c=x+y i$

We cannot optimize the first line, but we can optimize the second

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& a=x+y ; \\
& a=z ; \\
& b=x+y ; \\
& c=x+y ;
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

```
Current_val = {
}
```

$\longrightarrow$| $\mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ;$ |
| :--- |
| $\mathrm{a} 5=\mathrm{z} 4 ;$ |
| $\mathrm{b} 6=\mathrm{x}=\mathrm{y} 2 ;$ |
| $\mathrm{c} 7=\mathrm{x}=\mathrm{y}=;$ |

$$
\begin{aligned}
& H=\{ \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 3, } \\
& \} \\
& \text { H = \{ "x1 + y2" : "a3", } \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 3, } \\
& \} \\
& \text { H = \{ "x1 + y2" : "a3", } \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 5, } \\
& \} \\
& \text { H = \{ "x1 + y2" : "a3", } \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 5, } \\
& \} \\
& \begin{array}{l}
\text { H = \{ "x1 + y2" : "a3", } \\
\}
\end{array}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 5, } \\
& \} \\
& \mathrm{H}=\{\text { "x1 + y2" : "a3", } \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\longrightarrow \quad \begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x}=\mathrm{y}+ \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ } \\
& \qquad \begin{array}{l}
" a ": 5, \\
\text { "b" : 6 }
\end{array} \\
& \begin{array}{l}
\mathrm{H}=\{ \\
\}
\end{array} \quad \mathrm{x} 1+\mathrm{y} 2 ": " \mathrm{~b} 6 ",
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 5, } \\
& \qquad \begin{array}{l}
" \mathrm{b"} \text { : } 6
\end{array} \\
& \} \\
& \mathrm{H}=\{ \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 5, } \\
& \qquad \begin{array}{l}
\text { "b" : 6 }
\end{array} \\
& \} \\
& \text { H = \{ "x1 + y2" : "b6", }
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Keep another hash table to keep the current variable number

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 5=\mathrm{z} 4 ; \\
& \mathrm{b} 6=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{c} 7=\mathrm{b} 6 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ "a" : 5, } \\
& \qquad \begin{array}{l}
" \mathrm{b"} \text { : } 6
\end{array} \\
& \} \\
& \mathrm{H}=\{ \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values


## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

$$
\begin{aligned}
& \mathrm{a}=\mathrm{x}+\mathrm{y} ; \\
& \mathrm{b}=\mathrm{x}+\mathrm{yi} \\
& \mathrm{a}=\mathrm{zi} \\
& \mathrm{c}=\mathrm{x}+\mathrm{y} ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ } \\
& \} \\
& \begin{array}{l}
\mathrm{H}=\{ \\
\}
\end{array}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

```
Current_val = {
}
```

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{b} 4=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{a} 6=\mathrm{z} 5 ; \\
& \mathrm{c} 7=\mathrm{x}=\mathrm{y} 2 ;
\end{aligned}
$$

$$
\begin{aligned}
& H=\{ \\
& \}
\end{aligned}
$$

## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

```
Current_val = {
                                    "a" : 6,
}
H= {
}
```

$\mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ;$
$\mathrm{b} 4=\mathrm{a} 3 ;$
$\mathrm{a} 6=\mathrm{z} ;$
$\mathrm{c} 7=\mathrm{x}=\mathrm{y}+$

## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

```
Current_val = {
                                    "a" : 6,
}
```

```
H = { \
```

H = { \
}

```
}
```

```
a3 = x1 + y2;
```

a3 = x1 + y2;
b4 = a3;
b4 = a3;
a6 = z5;
a6 = z5;
c7 = x1 + y2;

```
c7 = x1 + y2;
```

but we could have replaced it with b4!

## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

```
Current_val = {
    "a" : 3,
}
```

| rewind to this point | $\begin{aligned} \mathrm{a} 3 & =\mathrm{x} 1+\mathrm{y} 2 ; \\ \mathrm{b} 4 & =\mathrm{x} 1+\mathrm{y} 2 ; \\ \mathrm{a} 6 & =\mathrm{z} 5 ; \\ \mathrm{c} 7 & =\mathrm{x} 1+\mathrm{y} 2 ; \end{aligned}$ |
| :---: | :---: |

```
H = { {
}
```


## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

```
Current_val = {
        "a" : 3,
}
H= {
}
```

$\mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ;$
$\mathrm{b} 4=\mathrm{a} 3 ;$
$\mathrm{a} 6=\mathrm{z} ;$
$\mathrm{c} 7=\mathrm{x}=\mathrm{y}+$
hash a list of possible values

## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

```
Current_val = \{
    \(" a ": 6, ~\)
\(" b ": ~ 4\)
\}
```


$\mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ;$
$\mathrm{b} 4=\mathrm{a} 3 ;$
$\mathrm{a} 6=\mathrm{z} ;$
$\mathrm{c} 7=\mathrm{x} 1+\mathrm{y} 2 ;$

## Local Value Numbering w/out adding registers

- Final heuristic: keep sets of possible values

$$
\begin{aligned}
& \mathrm{a} 3=\mathrm{x} 1+\mathrm{y} 2 ; \\
& \mathrm{b} 4=\mathrm{a} 3 ; \\
& \mathrm{a} 6=\mathrm{z} ; \\
& \mathrm{c} 7=\mathrm{b} 4 ;
\end{aligned}
$$

$$
\begin{aligned}
& \text { Current_val = \{ } \\
& \begin{array}{l}
" a{ }^{\prime \prime}=6, \\
" b ": 4
\end{array} \\
& \text { \} }
\end{aligned}
$$

## Local Value Numbering Pitfalls

- Consider a 3 address code that allows memory accesses

```
a[i] = x[j] + y[k];
b[i] = x[j] + y[k];
```

is this transformation allowed? No!
$a[i]=x[j]+y[k] ;$
$b[i]=a[i] ;$
only if the compiler can prove that a does not alias x and y

In the worst case, every time a memory location is updated, the compiler must update the value for all pointers.

## Local Value Numbering Pitfalls

- How to number:
- Number each pointer/index pair

$$
\begin{aligned}
& (\mathrm{a}[\mathrm{i}], 3)=(\mathrm{x}[\mathrm{j}], 1)+(\mathrm{y}[\mathrm{k}], 2) ; \\
& \mathrm{b}[\mathrm{i}]=\mathrm{x}[\mathrm{j}]+\mathrm{y}[\mathrm{k}] ;
\end{aligned}
$$

## Local Value Numbering Pitfalls

- How to number:
- Number each pointer/index pair
- Any pointer/index pair that are not proven not to alias must be incremented at each instruction

$$
\begin{aligned}
& (\mathrm{a}[\mathrm{i}], 3)=(x[j], 1)+(\mathrm{y}[\mathrm{k}], 2) ; \\
& (\mathrm{b}[\mathrm{i}], 6)=(\mathrm{x}[\mathrm{j}], 4)+(\mathrm{y}[\mathrm{k}], 5) ;
\end{aligned}
$$

## Local Value Numbering Pitfalls

- How to number:
- Number each pointer/index pair
- Any pointer/index pair that are not proven not to alias must be incremented at each instruction

$$
\begin{aligned}
& (\mathrm{a}[\mathrm{i}], 3)=(\mathrm{x}[\mathrm{j}], 1)+(\mathrm{y}[\mathrm{k}], 2) ; \\
& (\mathrm{b}[\mathrm{i}], 6)=(\mathrm{x}[\mathrm{j}], 4)+(\mathrm{y}[\mathrm{k}], 5) ;
\end{aligned}
$$

```
compiler analysis:
can we trace \(\mathrm{a}, \mathrm{x}, \mathrm{y}\) to
a = malloc(...);
\(\mathrm{x}=\) malloc(...);
\(\mathrm{y}=\) malloc(...);
// \(\mathrm{a}, \mathrm{x}, \mathrm{y}\) are never overwritten
```


## Local Value Numbering Pitfalls

- How to number:
- Number each pointer/index pair
- Any pointer/index pair that are not proven not to alias must be incremented at each instruction

$$
\begin{aligned}
& (\mathrm{a}[\mathrm{i}], 3)=(x[j], 1)+(y[k], 2) ; \\
& (\mathrm{b}[\mathrm{i}], 6)=(x[j], 4)+(y[k], 5) ;
\end{aligned}
$$

in this case we do not have to update the number
compiler analysis:
can we trace $\mathrm{a}, \mathrm{x}, \mathrm{y}$ to
a = malloc(...);
$\mathrm{x}=$ malloc(...);
y = malloc(...);
// $\mathrm{a}, \mathrm{x}, \mathrm{y}$ are never overwritten

## Local Value Numbering Pitfalls

- How to number:
- Number each pointer/index pair
- Any pointer/index pair that are not proven not to alias must be incremented at each instruction

$$
\begin{aligned}
& (\mathrm{a}[\mathrm{i}], 3)=(x[j], 1)+(\mathrm{y}[\mathrm{k}], 2) ; \\
& (\mathrm{b}[\mathrm{i}], 6)=(\mathrm{x}[\mathrm{j}], 4)+(\mathrm{y}[\mathrm{k}], 5) ;
\end{aligned}
$$

programmer annotations can also tell the compiler that no other pointer can access the memory pointed to by a

## Local Value Numbering Pitfalls

- How to number:
- Number each pointer/index pair
- Any pointer/index pair that are not proven not to alias must be incremented at each instruction

$$
\begin{aligned}
& (\mathrm{a}[\mathrm{i}], 3)=(x[j], 1)+(y[k], 2) ; \\
& (\mathrm{b}[\mathrm{i}], 6)=(x[j], 4)+(y[k], 5) ;
\end{aligned}
$$

in this case we do not have to update the number
restrict a
programmer annotations can also tell the compiler that no other pointer can access the memory pointed to by a

## Local Value Numbering Pitfalls

- How to number:
- Number each pointer/index pair
- Any pointer/index pair that are not proven not to alias must be incremented at each instruction

$$
\begin{aligned}
& (\mathrm{a}[\mathrm{i}], 3)=(x[j], 1)+(\mathrm{y}[\mathrm{k}], 2) ; \\
& (\mathrm{b}[\mathrm{i}], 6)=(\mathrm{a}[\mathrm{i}], 3) ;
\end{aligned}
$$

## Optimizing over wider regions

- Local value numbering operated over just one basic block.
- We want optimizations that operate over several basic blocks (a region), or across an entire procedure (global)
- For this, we need Control Flow Graphs and Flow Analysis


## Control Flow Graphs

A graph where:

- nodes are basic blocks
- edges mean that it is possible for one block to branch to another
end_if:
r4 = ...;


## Control Flow Graphs

A graph where:

```
start:
r0 = ...;
r1 = ...;
br r0, if, else;
```

- nodes are basic blocks

```
if:
r2 = ...;
br end_if;
```

- edges mean that it is possible for one block to branch to another

```
else:
r3 = ...;
br end_if;
```

```
end_if:
```

r4 = ...;

## Control Flow Graphs

A graph where:

```
start:
r0 = ...;
r1 = ...;
br r0, if, else;
```

- nodes are basic blocks
- edges mean that it is possible for one block to branch to another



## Control Flow Graphs

Simple analysis:

What property did we rely on for local value numbering?

```
start:
r0 = ...;
r1 = ...;
br r0, if, else;
```



## Control Flow Graphs

Simple analysis:

```
start:
r0 = ...;
r1 = ...;
br r0, if, else;
```

What property did we rely on for local value numbering?

we say that a node $b_{x}$ "dominates" another node $b_{y}$ iff:
are there any non-trivial domination relations

every path from the start to $b_{y}$ goes through $b_{x}$

## Control Flow Graphs

Simple analysis:

```
start:
r0 = ...;
r1 = ...;
br r0, if, else;
```

What property did we rely on for local value numbering?

we say that a node $b_{x}$ "dominates" another node $b_{y}$ iff:
are there
any non-trivial domination relations

every path from the start to $b_{y}$ goes through $b_{x}$

## Other examples

- The PyCFG tool draws CFGs for simple python code:
- Single statement basic blocks


## Dominance

- Given a CFG, determine for each node $b_{x}$, the set of nodes that dominate $b_{x}$


| Node | Dominators |
| :--- | :--- |
| B0 | B0 |
| B1 | B0, B1 |
| B2 | B0, B1, B2 |
| B3 |  |
| B4 |  |
| B5 |  |
| B6 | B0,B1,B5,B6 |
| B7 |  |
| B8 |  |




| Node | Dominators |
| :---: | :---: |
| B0 | B0 |
| B1 | B0, B1 |
| B2 | B0, B1, B2 |
| B3 | B0, B1, B3 |
| B4 | B0, B1, B3, B4 |
| B5 | B0, B1, B5 |
| B6 | B0, B1, B5, B6 |
| B7 | B0, B1, B5, B7 |
| B8 | B0, B1, B5, B8 |



## Computing Dominance

- Iterative fixed point algorithm
- Initial state, all nodes start with all other nodes are dominators:
- $\operatorname{Dom}(n)=N$
- Dom(start) $=\{$ start $\}$
iteratively compute:

$$
\operatorname{Dom}(n)=\{n\} \cup\left(\bigcap_{m \text { in } \operatorname{preds}(n)} \operatorname{Dom}(m)\right)
$$

## initial conditions

| Node | Initial |
| :--- | :--- |
| B0 | B0 |
| B1 | N |
| B2 | N |
| B3 | N |
| B4 | N |
| B5 | N |
| B6 | $N$ |
| B7 | $N$ |
| B8 | $N$ |



$$
\operatorname{Dom}(n)=\{n\} \cup\left(\cap_{m \text { in preds(n) }} \operatorname{Dom}(m)\right)
$$

| Node | Initial | $\mathbf{l} 1$ |
| :--- | :--- | :--- |
| B0 | B0 | $\ldots$ |
| B1 | $N$ | B0, B1 |
| B2 | $N$ | B0, B1, B2 |
| B3 | $N$ | B0, B1, B2, B3 |
| B4 | $N$ |  |
| B5 | $N$ |  |
| B6 | $N$ |  |
| B7 | $N$ |  |
| B8 | $N$ |  |


$\operatorname{Dom}(n)=\{n\} \cup\left(\cap_{m \text { in } \operatorname{preds}(n)} \operatorname{Dom}(m)\right)$

| Node | Initial | $\mathbf{I 1}$ |
| :--- | :--- | :--- |
| B0 | B0 | BO |
| B1 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1$ |
| B2 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 2$ |
| B3 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 2, \mathrm{~B} 3$ |
| B4 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 2, \mathrm{~B} 3, \mathrm{~B} 4$ |
| B5 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 5$ |
| B6 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 6$ |
| B7 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 6, \mathrm{~B} 7$ |
| B8 | $N$ | $\mathrm{BO}, \mathrm{B} 1, \mathrm{~B} 5, \mathrm{~B} 8$ |


$\operatorname{Dom}(n)=\{n\} \cup\left(\cap_{m \text { in } \operatorname{preds}(n)} \operatorname{Dom}(m)\right)$

| Node | Initial | $\mathbf{1 1}$ | $\mathbf{1 2}$ |
| :--- | :--- | :--- | :--- |
| B0 | B0 | B 0 | $\ldots$ |
| B1 | $N$ | $\mathrm{~B}, \mathrm{~B} 1$ | $\ldots$ |
| B2 | $N$ | $\mathrm{~B}, \mathrm{~B} 1, \mathrm{~B} 2$ |  |
| B3 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 2, \mathrm{~B} 3$ |  |
| B4 | $N$ | $\mathrm{~B}, \mathrm{~B} 1, \mathrm{~B} 2, \mathrm{~B} 3, \mathrm{~B} 4$ |  |
| B5 | $N$ | $\mathrm{~B}, \mathrm{~B} 1, \mathrm{~B} 5$ |  |
| B6 | $N$ | $\mathrm{~B}, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 6$ |  |
| B7 | $N$ | $\mathrm{~B}, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 6, \mathrm{~B} 7$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 5$ |
| B8 | $N$ | $\mathrm{~B}, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 8$ |  |


$\operatorname{Dom}(n)=\{n\} \cup\left(\bigcap_{m \text { in } \operatorname{preds}(n)} \operatorname{Dom}(m)\right)$

| Node | Initial | $\mathbf{l 1}$ | $\mathbf{1 2}$ |
| :--- | :--- | :--- | :--- |
| B0 | B0 | B 0 | $\ldots$ |
| B1 | $N$ | $\mathrm{B0}, \mathrm{~B} 1$ | $\ldots$ |
| B2 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 2$ | $\ldots$ |
| B3 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 2, \mathrm{~B} 3$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 3$ |
| B4 | $N$ | $\mathrm{B0}, \mathrm{~B} 1, \mathrm{~B} 2, \mathrm{~B} 3, \mathrm{~B} 4$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 3, \mathrm{~B} 4$ |
| B5 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 5$ | $\ldots$ |
| B6 | $N$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 6$ | $\ldots$ |
| B7 | $N$ | $\mathrm{B0}, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 6, \mathrm{~B} 7$ | $\mathrm{~B} 0, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 7$ |
| B8 | $N$ | $\mathrm{B0}, \mathrm{~B} 1, \mathrm{~B} 5, \mathrm{~B} 8$ | $\ldots$ |


$\operatorname{Dom}(n)=\{n\} \cup\left(\cap_{m \text { in } \operatorname{preds}(n)} \operatorname{Dom}(m)\right)$

| Node | Initial | 11 | 12 | 13 |
| :---: | :---: | :---: | :---: | :---: |
| B0 | B0 | B0 | ... |  |
| B1 | $N$ | B0, B1 | ... |  |
| B2 | $N$ | B0,B1,B2 | ... |  |
| B3 | $N$ | B0,B1, B2, 33 | B0,B1, B3 |  |
| B4 | $N$ | B0,B1, B2, B3, B4 | B0,B1, B3, $\mathrm{B}^{\text {4 }}$ |  |
| B5 | $N$ | B0,B1,B5 | .. |  |
| B6 | $N$ | B0,B1, B5, B6 | ... |  |
| B7 | $N$ | B0,B1,B5,B6,B7 | B0,B1, B5, 77 |  |
| B8 | $N$ | B0,B1,B5,B8 | ... |  |


$\operatorname{Dom}(n)=\{n\} \cup\left(\cap_{m \text { in } \operatorname{preds}(n)} \operatorname{Dom}(m)\right)$

| Node | Initial | 11 | 12 | 13 |
| :---: | :---: | :---: | :---: | :---: |
| B0 | B0 | B0 | $\ldots$ | ... |
| B1 | $N$ | B0,B1 | ... | ... |
| B2 | $N$ | B0,B1,B2 | ... | ... |
| B3 | $N$ | B0,B1, B2, B3 | B0, B1, B3 | ... |
| B4 | $N$ | B0,B1, B2, B3, B4 | B0,B1,B3, B4 | ... |
| B5 | $N$ | B0,B1,B5 | ... | ... |
| B6 | $N$ | B0,B1, B5, B6 | ... | $\ldots$ |
| B7 | $N$ | B0,B1,B5,B6,B7 | B0,B1,B5, B7 | ... |
| B8 | $N$ | B0,B1,B5,B8 | ... | ... |

No change so algorithm terminates!


## Next week

- Flow analysis examples continued:
- node traversal order for faster convergence
- Live variable analysis
- Generalized framework for flow analysis
- SSA form

