Exercise: Exact Data Flow Analysis

Exact data flow analysis for the following program that resolves an lower triangular system of linear equations using forward substitution:

```
S1: x[1] := b[1]/u[1,1];
    for i = 2 to n do
S2:  s := 0;
     for j = 1 to i-1 do
S3: s := s + x[j] * u[i, j];
     enddo
S4:  x[i] := (b[i] - s) / u[i,i]
   enddo
```

We have 4 statements, of which `S1` is special (it's domain is $\mathcal{D}_0 = \mathcal{Z}^0$ the empty polyhedron). The other relevant domains are $\mathcal{D}_2 = \mathcal{D}_4 = \{i | 2 \leq i \leq n\}$, and $\mathcal{D}_3 = \{i, j | 1 \leq j < i \leq n\}$. The following source analyses are needed: `s` and `x` in `S3`, and `s` in `S4`. Our precedence order, $\lhd$, always looks for the lexicographical *maximum*, Lmax since both i and j are ascending

Since two statements write into `s`, $\mathrm{Src}(\mathtt{s}, \mathtt{S3}) = \mathrm{Last}(\langle \mathtt{S2}, f_1(i,j)\rangle, \langle \mathtt{S3}, f_2(i,j)\rangle)$, where $f_1(i,j)$ and $f_2(i,j)$ are respectively[1]

$$f_1(i,j) = \mathrm{Last}\left\{ (i'|i,j) \left\| \begin{array}{l} i' \in \mathcal{D}_2 \\ (i,j) \in \mathcal{D}_3 \\ \langle \mathtt{S2}, i'\rangle \lhd \langle \mathtt{S3}, (i,j)\rangle \end{array} \right. \right\} \quad (1)$$

$$f_2(i,j) = \mathrm{Last}\left\{ (i',j'|i,j) \left\| \begin{array}{l} (i',j') \in \mathcal{D}_3 \\ (i,j) \in \mathcal{D}_3 \\ \langle \mathtt{S3}, (i',j')\rangle \lhd \langle \mathtt{S3}, (i,j)\rangle \end{array} \right. \right\} \quad (2)$$

Using the definition of the $\lhd$ relation[2], and our knowledge of the program text, we see that $\langle \mathtt{S2}, i'\rangle \lhd \langle \mathtt{S3}, (i,j)\rangle$ iff $i' \leq i$. Hence

$$f_1(i,j) = \mathrm{Last}\left\{ (i'|i,j) \left\| \begin{array}{l} i' \in \mathcal{D}_2 \\ (i,j) \in \mathcal{D}_3 \\ i' \leq i \end{array} \right. \right\} \quad (3)$$

$$= \mathrm{Last}\left\{ (i'|i,j) \| 2 \leq i' \leq i \leq n; 1 \leq j < i \leq n \right\} \quad (4)$$

This set of points is viewed as a family 1-D polyhedra (line segments indexed by i') *parameterized* by points in a 2-D parameter polyhedron, $\mathcal{D}_3$. At all points in $\mathcal{D}_3$, the line segment is from 2 to i, since theLast operation is equivalent to finding the lexicographic *maximum*. Hence $f_1(i,j) = i$.

Now, to find $f_2(i,j)$, we observe that $\langle \mathtt{S3}, (i',j')\rangle \lhd \langle \mathtt{S3}, (i,j)\rangle$ is equivalent to $i' < i \vee (i' = i \wedge j' < j)$, a disjunction of two sets of inequality constraints, and hence

[1] For notational simplicity, vectors are written as rows, and the bar separates the parameters, i.e., we seek the last i' as a function of i and j such that the stated constraints are satisfied.

[2] Observe the sense of the inequality.

$$
\begin{aligned}
f_2(i,j) &= \mathrm{Last}\left(\left\{(i',j'|i,j) \,\middle\|\, \begin{array}{l}(i',j')\in\mathcal{D}_3\\(i,j)\in\mathcal{D}_3\\i'<i\end{array}\right\}\right. \\
&\qquad\left.\bigcup\left\{(i',j'|i,j) \,\middle\|\, \begin{array}{l}(i',j')\in\mathcal{D}_3\\(i,j)\in\mathcal{D}_3\\i'=i; j'<j\end{array}\right\}\right) \qquad (5)\\
&= \mathrm{Last}\left(\mathrm{Last}\left\{(i',j'|i,j) \,\middle\|\, \begin{array}{l}(i',j')\in\mathcal{D}_3\\(i,j)\in\mathcal{D}_3\\i'<i\end{array}\right\},\right. \\
&\qquad\left.\mathrm{Last}\left\{(i',j'|i,j) \,\middle\|\, \begin{array}{l}(i',j')\in\mathcal{D}_3\\(i,j)\in\mathcal{D}_3\\i'=i; j'<j\end{array}\right\}\right) \qquad (6)\\
&= \mathrm{Last}\left(\mathrm{Last}\{(i',j'|i,j)\,\|1\le j'<i'<i\le n; 1\le j<i\le n\}\right. \\
&\qquad\left.\mathrm{Last}\{(i',j',|i,j)\,\|1\le j<i'=i\le n; j'<j\}\right) \qquad (7)
\end{aligned}
$$

This is $\mathrm{Last}(f_2'(i,j), f_2''(i,j))$ where

$$
\begin{aligned}
f_2'(i,j) &= \begin{cases}\{i,j|i=2; j=1\} & : \quad \bot\\ \{i,j|1\le j<i\le n; i\ge 3\} & : \quad (i-1,i-2)\end{cases} \qquad (8)\\
f_2''(i,j) &= \begin{cases}\{i,j|1=j<i\le n\} & : \quad \bot\\ \{i,j|1<j<i\le n\} & : \quad (i,j-1)\end{cases} \qquad (9)
\end{aligned}
$$

(10)

and hence

$$
\begin{aligned}
f_2(i,j) &= \begin{cases}\{i,j|i=2; j=1\} & : \quad \mathrm{Last}(\bot,\bot)\\ \{i,j|i\ge 3; j=1\} & : \quad \mathrm{Last}((i-1,i-2),\bot)\\ \{i,j|1<j<i\le n\} & : \quad \mathrm{Last}((i-1,i-2),(i,j-1))\end{cases} \qquad (11)\\
&= \begin{cases}\{i,j|i=2; j=1\} & : \quad \bot\\ \{i,j|j=1; i\ge 3\} & : \quad (i-1,i-2)\\ \{i,j|1<j<i\le n\} & : \quad (i,j-1)\end{cases} \qquad (12)
\end{aligned}
$$

Finally, remember that

$$
\begin{aligned}
&\mathrm{Src}(\mathtt{s},\mathtt{S3}) = \mathrm{Last}\left(\langle \mathtt{S2}, f_1(i,j)\rangle, \langle \mathtt{S3}, f_2(i,j)\rangle\right)\\
&= \mathrm{Last}\left(\langle \mathtt{S2}, i\rangle, \left\langle \mathtt{S3}, \left\{\begin{array}{lcl} \{i,j|i=2; j=1\} & : & \bot \\ \{i,j|j=1; i\geq 3\} & : & (i-1,i-2) \\ \{i,j|1<j<i\leq n\} & : & (i,j-1)\end{array}\right.\right\rangle\right)\\
&= \mathrm{Last}\left(\langle \mathtt{S2}, i\rangle, \left\{\begin{array}{lcl} \{i,j|i=2; j=1\} & : & \langle \mathtt{S3}, \bot\rangle \\ \{i,j|j=1; i\geq 3\} & : & \langle \mathtt{S3}, (i-1,i-2)\rangle \\ \{i,j|1<j<i\leq n\} & : & \langle \mathtt{S3}, (i,j-1)\rangle\end{array}\right.\right)\\
&= \left\{\begin{array}{lcl} \{i,j|i=2; j=1\} & : & \mathrm{Last}(\langle \mathtt{S2}, i\rangle, \langle \mathtt{S3}, \bot\rangle) \\ \{i,j|j=1; i\geq 3\} & : & \mathrm{Last}(\langle \mathtt{S2}, i\rangle, \langle \mathtt{S3}, (i-1,i-2)\rangle) \\ \{i,j|1<j<i\leq n\} & : & \mathrm{Last}(\langle \mathtt{S2}, i\rangle, \langle \mathtt{S3}, (i,j-1)\rangle)\end{array}\right.\\
&= \left\{\begin{array}{lcl} \{i,j|i=2; j=1\} & : & \langle \mathtt{S2}, i\rangle \\ \{i,j|j=1; i\geq 3\} & : & \langle \mathtt{S2}, i\rangle \\ \{i,j|1<j<i\leq n\} & : & \langle \mathtt{S3}, (i,j-1)\rangle\end{array}\right.\\
&= \left\{\begin{array}{lcl} \{i,j|1=j<i\leq n\} & : & \langle \mathtt{S2}, i\rangle \\ \{i,j|1<j<i\leq n\} & : & \langle \mathtt{S3}, (i,j-1)\rangle\end{array}\right. \qquad (13)
\end{aligned}
$$

We now have determined the precise source of `s` on the right hand side of statement `S3` as a function of its loop iteration indices (and the system parameters).

To solve for $\mathrm{Src}(\mathtt{x},\mathtt{S3})$ we use the same approach. There are again two potential statements. There is one additional complication (actually, it provides additional constraints which serve to simplify the domains manipulated). Since `x` is not a scalar, there is one more constraint to be added to the first equation below, to ensure that the *address* being read by $\langle \mathtt{S3}, (i,j)\rangle$ is the same as the one that the statement being analyzed, namely $\langle \mathtt{S4}, i'\rangle$ or `S1` wrote to. Hence, $\mathrm{Src}(\mathtt{x},\mathtt{S3}) = \mathrm{Last}(\langle \mathtt{S1}, g_1(i,j)\rangle, \langle \mathtt{S4}, g_2(i,j)\rangle)$, where we solve for $g_1(i,j)$ and $g_2(i,j)$ as follows.

$$
g_1(i,j) = \mathrm{Last}\left\{ (|i,j) \left\| \begin{array}{l} (i,j)\in\mathcal{D}_3 \\ 1=j \\ \mathtt{S1} \triangleleft \langle \mathtt{S3}, (i,j)\rangle \end{array}\right.\right\} \qquad (14)
$$

Observe that because of the textual order o the statements in the program the last condition is always true (the execution of `S1` in the original program precedes *all* instances of `S3`), and we can further simplify this to

$$
\begin{aligned}
g_1(i,j) &= \mathrm{Last}\{(|i,j)\|1<i\leq n; j=1\}\\
&\quad \left\{\begin{array}{lcl} \{i,j|1<i\leq n; j=1\} & : & \mathrm{true} \\ \{i,j|1<i\leq n; 1<j<i\} & : & \bot\end{array}\right. \qquad (15)
\end{aligned}
$$

Similarly,

$$g_2(i,j) = \text{Last}\left\{ (i'|i,j) \;\middle\|\; \begin{array}{l} i' \in \mathcal{D}_4 \\ (i,j) \in \mathcal{D}_3 \\ j = i' \\ \langle \texttt{S4}, i' \rangle \triangleleft \langle \texttt{S3}, (i,j) \rangle \end{array} \right\} \tag{16}$$

The last constraint reduces to[3] simply $i' < i$, and the above simplifies to

$$\begin{aligned} g_2(i,j) &= \text{Last}\{(i'|i,j) \| 1 \leq j = i' < i < n; i' > 1\} \\ &= \left\{ \begin{array}{lcl} \{i,j|2 \leq i \leq n; j = 1\} & : & \bot \\ \{i,j|2 \leq i \leq n; 1 < j < i\} & : & j \end{array} \right. \end{aligned} \tag{17}$$

Hence

$$\begin{aligned} \text{Src}(\texttt{x}, \texttt{S3}) &= \text{Last}\left(\left\langle \texttt{S1}, \left\{ \begin{array}{lcl} \{i,j|2 \leq i \leq n; j = 1\} & : & \text{true} \\ \{i,j|2 \leq i \leq n; 1 < j < i\} & : & \bot \end{array} \right. \right\rangle, \right. \\ &\qquad \left. \left\langle \texttt{S4}, \left\{ \begin{array}{lcl} \{i,j|2 \leq i \leq n; j = 1\} & : & \bot \\ \{i,j|2 \leq i \leq n; 1 < j < i\} & : & j \end{array} \right. \right\rangle \right) \\ &= \left\{ \begin{array}{lcl} \{i,j|2 \leq i \leq n; j = 1\} & : & \text{Last}(\texttt{S1}, \bot) \\ \{i,j|2 \leq i \leq n; 1 < j < i\} & : & \text{Last}(\bot, \langle \texttt{S4}, j \rangle) \end{array} \right. \\ &= \left\{ \begin{array}{lcl} \{i,j|2 \leq i \leq n; j = 1\} & : & \texttt{S1} \\ \{i,j|2 \leq i \leq n; 1 < j < i\} & : & \langle \texttt{S4}, j \rangle \end{array} \right. \end{aligned} \tag{18}$$

Similarly,

$$\text{Src}(\texttt{s}, \texttt{S4}) = \text{Last}(\langle \texttt{S2}, h_1(i) \rangle, \langle \texttt{S3}, h_2(i) \rangle)$$

where

$$\begin{aligned} h_1(i) &= \text{Last}\left\{ (i'|i) \;\middle\|\; \begin{array}{l} i' \in \mathcal{D}_2 \\ i \in \mathcal{D}_4 \\ \langle \texttt{S2}, i' \rangle \triangleleft \langle \texttt{S4}, i \rangle \end{array} \right\} \\ &= \text{Last}\{(i'|i) \| 2 \leq i' \leq i \leq n\} \\ &= i \end{aligned} \tag{19}$$

and

$$\begin{aligned} h_2(i) &= \text{Last}\left\{ (i',j'|i) \;\middle\|\; \begin{array}{l} (i',j') \in \mathcal{D}_3 \\ i \in \mathcal{D}_4 \\ \langle \texttt{S3}, (i',j') \rangle \triangleleft \langle \texttt{S4}, i \rangle \end{array} \right\} \\ &= \text{Last}\left\{ (i',j'|i) \;\middle\|\; \begin{array}{l} 1 \leq j' < i' \leq n \\ 1 < i \leq n \\ i' \leq i \end{array} \right\} \\ &= \text{Last}\{(i',j'|i) \| 1 \leq j' < i' \leq n; 2 \leq i' \leq i \leq n\} \\ &= (i, i-1) \end{aligned} \tag{20}$$

[3]Note that the inequality is now *strict* since S4 does not precede S3 in the *textual* order of the program

Hence

$$\begin{aligned} \mathrm{Src}(\mathtt{s}, \mathtt{S4}) &= \mathrm{Last}(\langle \mathtt{S2}, i\rangle, \langle \mathtt{S3}, (i, i-1)\rangle) \\ &= \langle \mathtt{S3}, (i, i-1)\rangle \end{aligned} \tag{21}$$