Exercise: Exact Data Flow Analysis

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

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

The original program had a bug (`x[n-j]` instead of `x[j]` in statement `S3`). Sorry for that, but since no one sent me any mail, it probably means that people did not try hard enough.

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 $\mathcal{D}_2 = \mathcal{D}_4 = \{i \mid 1 \leq i \leq n-1\}$, and $\mathcal{D}_3 = \{i, j \mid 1 \leq i < j \leq n\}$. The folowing source analyses are needed: `s` and `x` in `S3`, and `s` in `S4`. The main point to note is thta the i loop *goes down* from $n-1$ to 1, and hence we have to be careful about our precedence order, $\prec$, and that we will not always look for the lexicographical *maximum*, Lmax (in the i dimension it will be minimum.

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]

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

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

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

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

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

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

[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.

segment is from i to n, and here the Last operation is equivalent to finding the lexicographic *minimum*. Hence $f_1(i,j) = i$.

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

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

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

$$f_2'(i,j) = \begin{cases} \{i,j \mid i = n-1; j = n\} & : \perp \\ \{i,j \mid 1 \le i < j \le n; i \le n-2\} & : (i+1,n) \end{cases} \qquad (8)$$

$$f_2''(i,j) = \begin{cases} \{i,j \mid 1 \le i = j-1 \le n-1\} & : \perp \\ \{i,j \mid 1 \le i < j-1 \le n-1\} & : (i,j-1) \end{cases} \qquad (9)$$

$$(10)$$

and hence

$$f_2(i,j) = \begin{cases} \{i,j \mid i = n-1; j = n\} & : \mathrm{Last}(\perp,\perp) \\ \{i,j \mid 1 \le i = j-1 \le n-2\} & : \mathrm{Last}((i+1,n),\perp) \\ \{i,j \mid 1 \le i < j-1 \le n-1; i \le n-2\} & : \mathrm{Last}((i+1,n),(i,j-1)) \end{cases} \qquad (11)$$

$$= \begin{cases} \{i,j \mid i = n-1; j = n\} & : \perp \\ \{i,j \mid 1 \le i = j-1 \le n-2\} & : (i+1,n) \\ \{i,j \mid 1 \le i < j-1 \le n-1; i \le n-2\} & : (i,j-1) \end{cases} \qquad (12)$$

Finally, remember that

$$
\begin{aligned}
\mathrm{Src}(\mathtt{s},\mathtt{S3}) &= \mathrm{Last}\left(\langle \mathtt{S2}, f_1(i,j)\rangle, \langle \mathtt{S3}, f_2(i,j)\rangle\right) \\
&= \mathrm{Last}\left(\langle \mathtt{S2}, i\rangle, \left\langle \mathtt{S3}, \left\{\begin{array}{ll} \{i,j \mid i=n-1; j=n\} & : \perp \\ \{i,j \mid 1 \le i = j-1 \le n-2\} & : (i+1,n) \\ \{i,j \mid 1 \le i < j-1 \le n-1; i \le n-2\} & : (i,j-1) \end{array}\right.\right\rangle\right) \\
&= \mathrm{Last}\left(\langle \mathtt{S2}, i\rangle, \left\{\begin{array}{ll} \{i,j \mid i=n-1; j=n\} & : \langle \mathtt{S3}, \perp\rangle \\ \{i,j \mid 1 \le i = j-1 \le n-2\} & : \langle \mathtt{S3}, (i+1,n)\rangle \\ \{i,j \mid 1 \le i < j-1 \le n-1; i \le n-2\} & : \langle \mathtt{S3}, (i,j-1)\rangle \end{array}\right.\right) \\
&= \left\{\begin{array}{ll} \{i,j \mid i=n-1; j=n\} & : \mathrm{Last}(\langle \mathtt{S2}, i\rangle, \langle \mathtt{S3}, \perp\rangle) \\ \{i,j \mid 1 \le i = j-1 \le n-2\} & : \mathrm{Last}(\langle \mathtt{S2}, i\rangle, \langle \mathtt{S3}, (i+1,n)\rangle) \\ \{i,j \mid 1 \le i < j-1 \le n-1; i \le n-2\} & : \mathrm{Last}(\langle \mathtt{S2}, i\rangle, \langle \mathtt{S3}, (i,j-1)\rangle) \end{array}\right. \\
&= \left\{\begin{array}{ll} \{i,j \mid i=n-1; j=n\} & : \langle \mathtt{S2}, i\rangle \\ \{i,j \mid 1 \le i = j-1 \le n-2\} & : \langle \mathtt{S2}, i\rangle \\ \{i,j \mid 1 \le i < j-1 \le n-1; i \le n-2\} & : \langle \mathtt{S3}, (i,j-1)\rangle \end{array}\right. \\
&= \left\{\begin{array}{ll} \{i,j \mid 1 \le i = j-1 \le n-1\} & : \langle \mathtt{S2}, i\rangle \\ \{i,j \mid 1 \le i < j-1 \le n-1; i \le n-2\} & : \langle \mathtt{S3}, (i,j-1)\rangle \end{array}\right. \qquad (13)
\end{aligned}
$$

We now have deermined 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 Src(`x`, `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 analysed, 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.

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

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

$$
\begin{aligned}
g_1(i,j) &= \mathrm{Last}\{(\mid i,j) \parallel 1 \le i \le n-1; j = n\} \\
& \left\{\begin{array}{ll} \{i,j \mid 1 \le i \le n-1; j=n\} & : \text{true} \\ \{i,j \mid 1 \le i \le n-1; i < j < n\} & : \perp \end{array}\right. \qquad (15)
\end{aligned}
$$

Similarly,

$$g_2(i,j) \quad = \quad \text{Last}\left\{ (i' \mid i,j) \left\| \begin{array}{l} i' \in \mathcal{D}_4 \\ (i,j) \in \mathcal{D}_3 \\ j = i' \\ \langle \mathtt{S4}, i' \rangle \prec \langle \mathtt{S3}, (i,j) \rangle \end{array} \right. \right\} \tag{16}$$

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

$$\begin{array}{rcl} g_2(i,j) & = & \text{Last}\{(i' \mid i,j) \parallel 1 \leq i < j = i' \leq n; i' < n\} \\ & = & \left\{ \begin{array}{lcl} \{i,j \mid 1 \leq i \leq n-1; j = n\} & : & \perp \\ \{i,j \mid 1 \leq i \leq n-1; i < j < n\} & : & j \end{array} \right. \end{array} \tag{17}$$

Hence

$$\begin{array}{rcl} \text{Src}(\mathtt{x}, \mathtt{S3}) & = & \text{Last}\left(\left\langle \mathtt{S1}, \left\{ \begin{array}{lcl} \{i,j \mid 1 \leq i \leq n-1; j = n\} & : & \text{true} \\ \{i,j \mid 1 \leq i \leq n-1; i < j < n\} & : & \perp \end{array} \right. \right\rangle, \right. \\ & & \left. \left\langle \mathtt{S4}, \left\{ \begin{array}{lcl} \{i,j \mid 1 \leq i \leq n-1; j = n\} & : & \perp \\ \{i,j \mid 1 \leq i \leq n-1; i < j < n\} & : & j \end{array} \right. \right\rangle \right) \\ & = & \left\{ \begin{array}{lcl} \{i,j \mid 1 \leq i \leq n-1; j = n\} & : & \text{Last}(\mathtt{S1}, \perp) \\ \{i,j \mid 1 \leq i \leq n-1; i < j < n\} & : & \text{Last}(\perp, \langle \mathtt{S4}, j \rangle) \end{array} \right. \\ & = & \left\{ \begin{array}{lcl} \{i,j \mid 1 \leq i \leq n-1; j = n\} & : & \mathtt{S1} \\ \{i,j \mid 1 \leq i \leq n-1; i < j < n\} & : & \langle \mathtt{S4}, j \rangle \end{array} \right. \end{array} \tag{18}$$

Similarly,

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

where

$$\begin{array}{rcl} h_1(i) & = & \text{Last}\left\{ (i' \mid i) \left\| \begin{array}{l} i' \in \mathcal{D}_2 \\ i \in \mathcal{D}_4 \\ \langle \mathtt{S2}, i' \rangle \prec \langle \mathtt{S4}, i \rangle \end{array} \right. \right\} \\ & = & \text{Last}\{(i' \mid i) \parallel 1 \leq i \leq i' \leq n-1\} \\ & = & i \end{array} \tag{19}$$

and

$$h_2(i) \quad = \quad \text{Last}\left\{ (i', j' \mid i) \left\| \begin{array}{l} (i', j') \in \mathcal{D}_3 \\ i \in \mathcal{D}_4 \\ \langle \mathtt{S3}, (i', j') \rangle \prec \langle \mathtt{S4}, i \rangle \end{array} \right. \right\}$$

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

$$
\begin{aligned}
&= \text{Last}\{(i' \mid i) \parallel 1 \le i \le i' \le n-1\} \\
&= \text{Last}\left\{ (i', j' \mid i) \,\middle\|\, \begin{array}{l} 1 \le i' < j' \le n \\ 1 \le i \le n-1 \\ i' \ge i \end{array} \right\} \\
&= \text{Last}\{(i', j' \mid i) \parallel 1 \le i \le i' \le n-1; i < j \le n\} \\
&= (i, n)
\end{aligned}
\tag{20}
$$

Hence

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