Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institut für Technische Informatik und Kommunikationsnetze

Prof. L. Thiele

Hardware/Software Codesign - HS 15

Solution for Exercise Sheet 4

Issue Date: 14 October 2015
Discussion Date: 21 October 2015

4.1 Closeness Function

The nodes of the data flow graph shown in Fig. 1 represent arithmetic operations. The edges are labeled with the bit widths of the required data types. The following closeness function is given:

$$Closeness(o_i, o_j) = \left(\frac{Conn(o_i, o_j)}{TotalConn(o_i, o_j)}\right) + + \left(\frac{FU_{cost}(o_i) + FU_{cost}(o_j) - FUGROUP_{cost}(o_i, o_j)}{FUGROUP_{cost}(o_i, o_j)}\right)$$

The different terms in this function have the following meanings:

- $Conn(o_i, o_j)$: the number of common wires between objects o_i and o_j.
- $TotalConn(o_i, o_j)$: the sum of all wires that connect to objects o_i and o_j. Common wires are counted twice - once for each object.
- $FU_{cost}(o_i)$: the cost of the functional unit that implements o_i.
- $FUGROUP_{cost}(o_i, o_j)$: the cost of the minimal number of functional units that are required to execute objects o_i and o_j. (For example, should both o_i and o_j be additions, they would be assigned to one functional unit of type adder.)

4.1.a) Physical interpretation

Find a physical interpretation of the closeness function.

4.1.b) Hierarchical clustering

- Compute the closeness values for all pairs of objects in Fig. 1 according to the given closeness function. The cost of an adder and the cost of a subtractor are 1.
- Define a suitable closeness function for clustering hierarchical objects based on the function given above. Cluster the graph until you receive a bi-partitioning.

Solution:

The physical interpretation of $Closeness(o_1, o_2)$ is that it denotes the relative benefit of clustering o_1 and o_2. There are two considerations for the clustering: wires between clusters and the functional unit costs. For either consideration, we have an expression that denotes the gain due to clustering divided by a normalizing term.

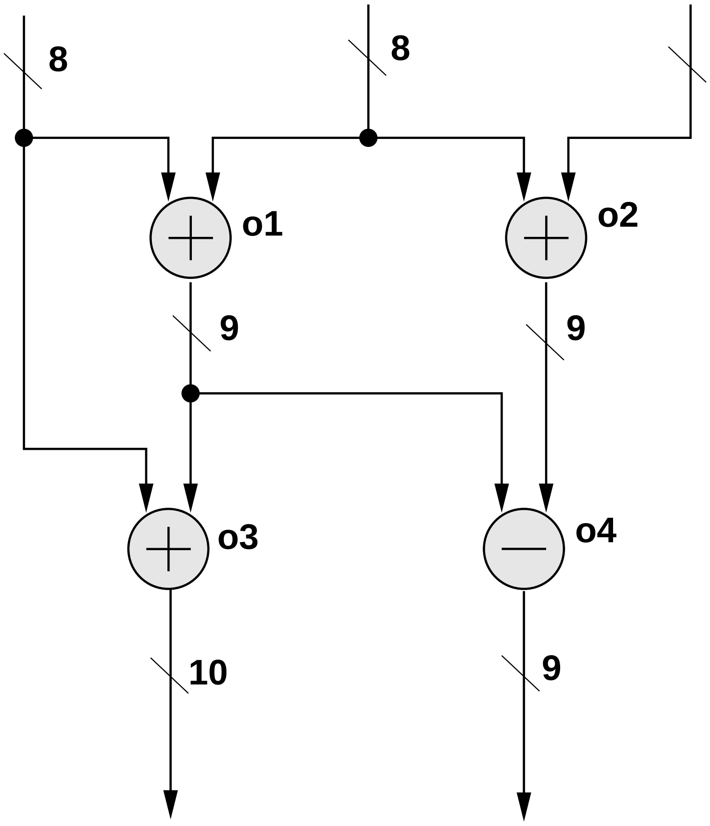

Figure 1: Data flow graph

$$
\begin{aligned}
c(o_1, o_2) &= \frac{8}{50} + \frac{2-1}{1} = 1\frac{4}{25} = 1.16 \\
c(o_1, o_3) &= \frac{17}{52} + \frac{2-1}{1} = 1\frac{17}{52} = 1.327 \\
c(o_1, o_4) &= \frac{9}{52} + \frac{2-2}{2} = \frac{9}{52} = 0.173 \\
c(o_2, o_3) &= 0 + \frac{2-1}{1} = 1 \\
c(o_2, o_4) &= \frac{9}{52} + \frac{2-2}{2} = \frac{9}{52} = 0.173 \\
c(o_3, o_4) &= \frac{9}{54} + \frac{2-2}{2} = \frac{1}{6} = 0.167
\end{aligned}
$$

One possible solution for finding a hierarchical clustering is to cluster operations o_1 and o_3 into a "supernode" o_{13} and treat this node like any initial node. This involves re-computing all closeness values and repeating the clustering procedure.

$$
\begin{aligned}
c(o_{13}, o_2) &= \frac{8}{60} + \frac{2+1-1}{1} = \frac{2}{15} + 2 = 2.133 \\
c(o_{13}, o_4) &= \frac{9}{62} + \frac{2+1-2}{2} = \frac{9}{62} + \frac{1}{2} = 0.645 \\
c(o_2, o_4) &= \frac{9}{52} + \frac{1+1-2}{2} = \frac{9}{52} = 0.173
\end{aligned}
$$

The updated closeness values suggest to put o_{13} and o_2 into one new partition o_{123}. We have now a bi-partition o_{123} and o_4.

```
P = {{}, O}; /* all in HW */

PROCEDURE PARTITIONING
        REPEAT
                P_old=P;
                FORALL o_i ∈ HW
                        AttemptMove(P, o_i);
                ENDFOR
        UNTIL (P == P_old)
END PROCEDURE

PROCEDURE AttemptMove(P, o_x)
        IF SatifiesPerformance(Move(P, o_x)) AND
           (f(Move(P, o_x)) < f(P))
                        P = Move(P, o_x);
                        FORALL (o_y ∈ Sucessors(o_x))
                                AttemptMove(P, o_y) ;
                        ENDFOR
        ENDIF
END PROCEDURE
```

Figure 2: Pseudo code for a greedy HW/SW partitioner

4.2 HW/SW partitioning

Fig. 2 shows the pseudo code of a greedy algorithm for HW/SW partitioning. The algorithm starts with a partition where all objects are realized in hardware. Then, objects are migrated to software as long as the performance requirement is satisfied (function SatisfiesPerformance) and the cost of the new partitioning is lower (function f). If an object is migrated, the algorithm also tries to migrate all successor nodes (function Successors).

- Apply the algorithm to the sequence graph shown in Fig. 3. The function SatisfiesPerformance(P) should return `TRUE` if P satisfies the latency bound $\bar{L} \leq 7$. To determine the latency of a partitioning, you have to construct a valid schedule. The execution times of start- and end nodes of the sequence graph are 0, all other node execution times are given in Fig. 3, split into HW (d_{HW}) and SW (d_{SW}). For a communication between HW and SW, a delay of 0.5 per edge has to be accounted for. For HW nodes there are no resource constraints, i.e., all ready nodes can be executed in parallel. The SW nodes have to share one processor. The function f determines the cost. For a SW node the cost is 0, and for a HW node the cost is 1.

Solution:

We assume in this solution that the main procedure processes hardware elements in the increasing order of their indexes. This problem intrinsically has many solutions because the main routine, at a certain step, allows the free choice of a node.

Communication delay between hardware and software always adds to the latency and cannot be hidden by other operations because software tasks are running strictly sequential.

Latency	Cost	Tasks in SW	Comment
L0 = 4	C0 = 11	-	move op1 to SW
L1 = 5.5	C1 = 10	op1	move op3 to SW
L2 = 6.5	C2 = 9	op1,op3	move op4 to SW
L3 = 6.5	C3 = 8	op1,op3,op4	move op5 to SW
L4 = 6	C4 = 7	op1,op3,op4,op5	try to move op2 to SW
L5 = 8	C5 = 6	op1,op3.op4,op5,op2	Terminate, no other node can be moved to SW without violating Latency constraint

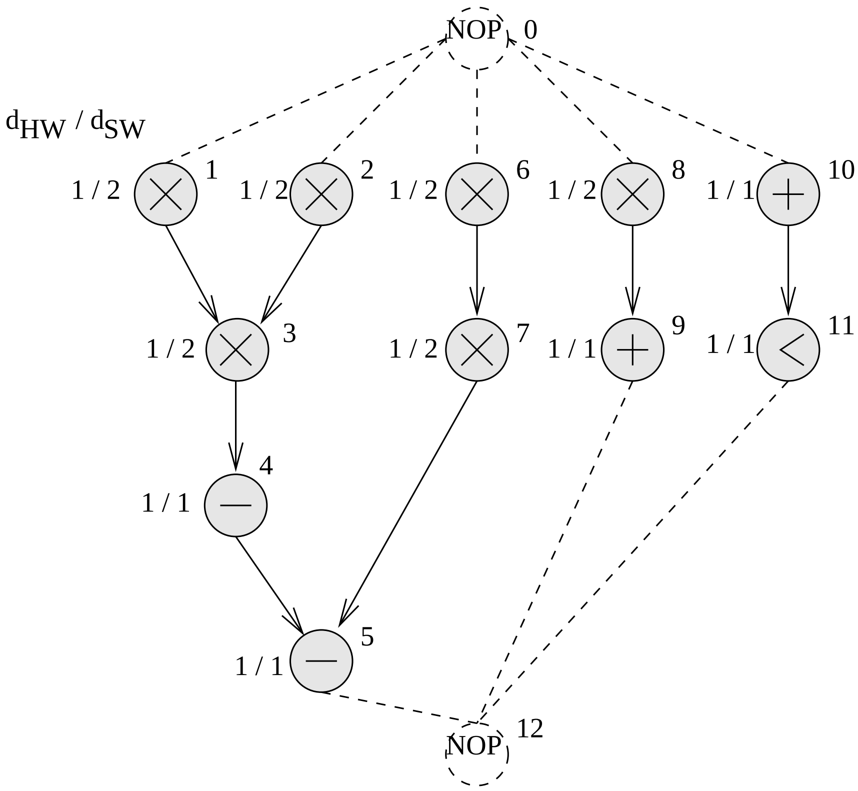

Figure 3: Data flow graph

Result: Nodes 1,3,4,5 are implemented in software and nodes 2,6,7,8,9,10,11 are implemented in hardware.

Note again that the solution is given when the main procedure processes elements in increasing order of their indexes. However, if we do not stick to this assumption, other nodes may also be added to the software partition. For example, node 10 can also be moved to software implementation.