An Application of Maximum Flow:

The Baseball Elimination Problem

We are given the following tournament situation:

	Wins so far	Games still to play against these opponents				Games still to play
		Brown	Cornell	Harvard	Yale	
Brown	27		1	3	1	5
Cornell	28	1		0	6	7
Harvard	29	3	0		1	4
Yale	33	1	6	1		8

Note: No ties are allowed.

Question: Is Harvard eliminated or not?

(A team is eliminated if it can't be the first or tied for the first at the end of the tournament).

Preliminary Analysis

The maximum number of wins Harvard can end up with is 29 + 4 = 33 (by winning all its remaining games)

Suppose Harvard does win all its remaining games. Then it will not be eliminated if and only if

- Brown ends the season with no more than 33 wins. Since Brown currently has 27 wins, this means that Brown would have to win at most 6 of its remaining games.
- Cornell ends the season with no more than 33 wins. Since Cornell currently has 28 wins, this means that Cornell would have to win at most 5 of its remaining games.
- Yale ends the season with no more than 33 wins. Since Yale already has 33 wins, this means that Yale would have to lose all of its remaining games.

Consider the number of remaining games to be played among teams other than Harvard:

- 1 game between Brown and Cornell
- 1 game between Brown and Yale
- 6 games between Cornell and Yale

This gives a total of 8 games to be played among teams other than Harvard.

Visualizing the Baseball Elimination Problem as a Transshipment Problem

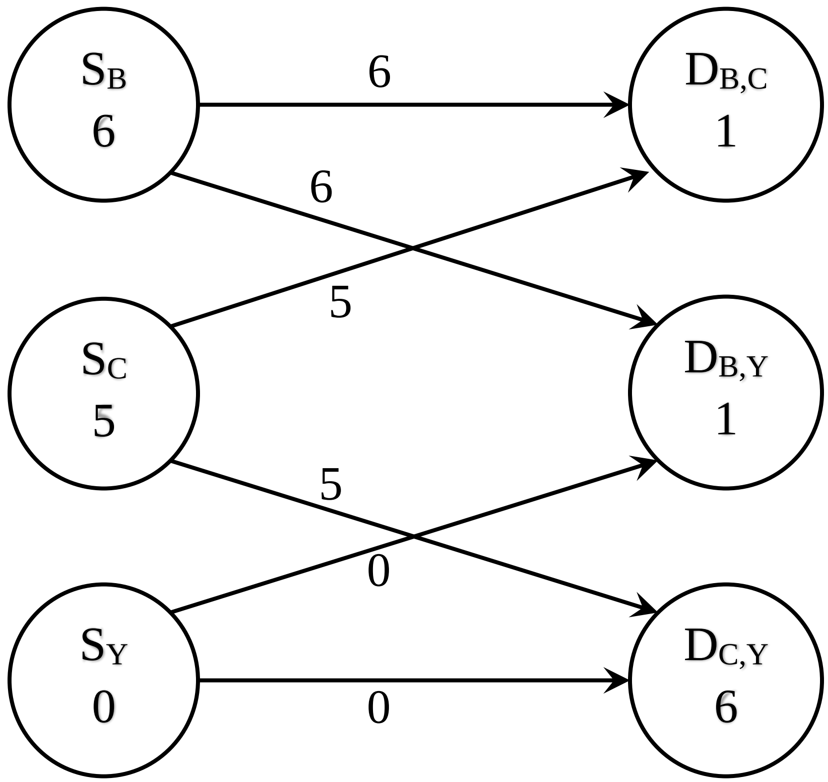
Imagine the remaining games between Brown, Cornell, and Yale as sink nodes that demand a certain number of wins.

- Brown and Cornell have 1 game to play. It will require 1 win.
- Brown and Yale have 1 game to play. It will require 1 win.
- Cornell and Yale have 6 games to play. Those six games will require a total of 6 wins.

And imagine each of the teams Brown, Cornell, and Yale as a supply node that can supply a certain number of wins.

- Brown can supply up to six wins (because if Brown wins more than six times, Harvard will be eliminated).
- Cornell can supply only five wins (because if Cornell wins more than five times, Harvard will be eliminated).
- Yale can supply no wins (because if Yale wins any games, Harvard will be eliminated).

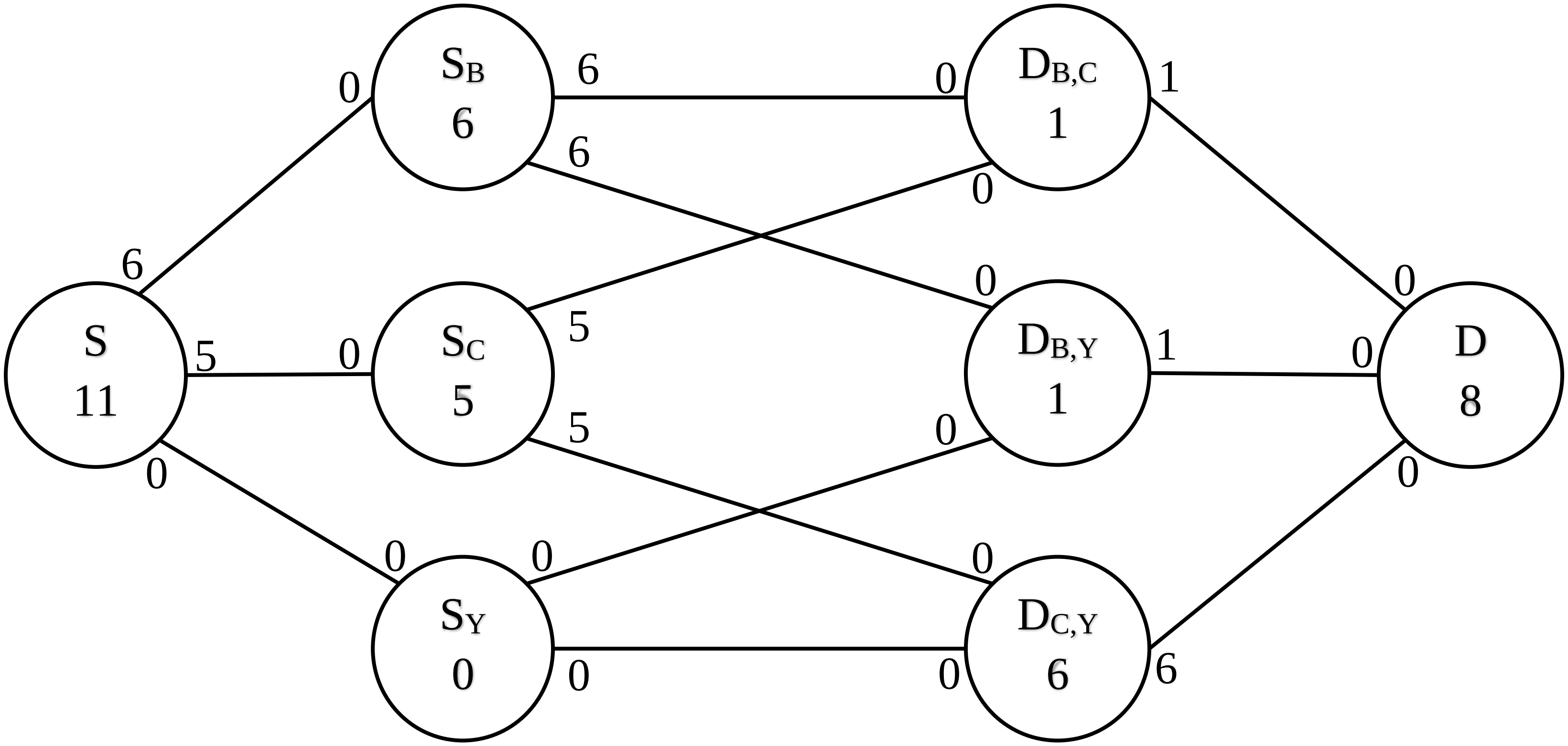
With this collection of supply and demand nodes, we get the following directed network representing a transshipment problem. (Notice that the flow capacity of each directed edge is set at the maximum flow possible from the source of that directed edge.)



At this point, we can see that the demand node $D_{C,Y}$ cannot be satisfied by the two supply nodes that feed it.

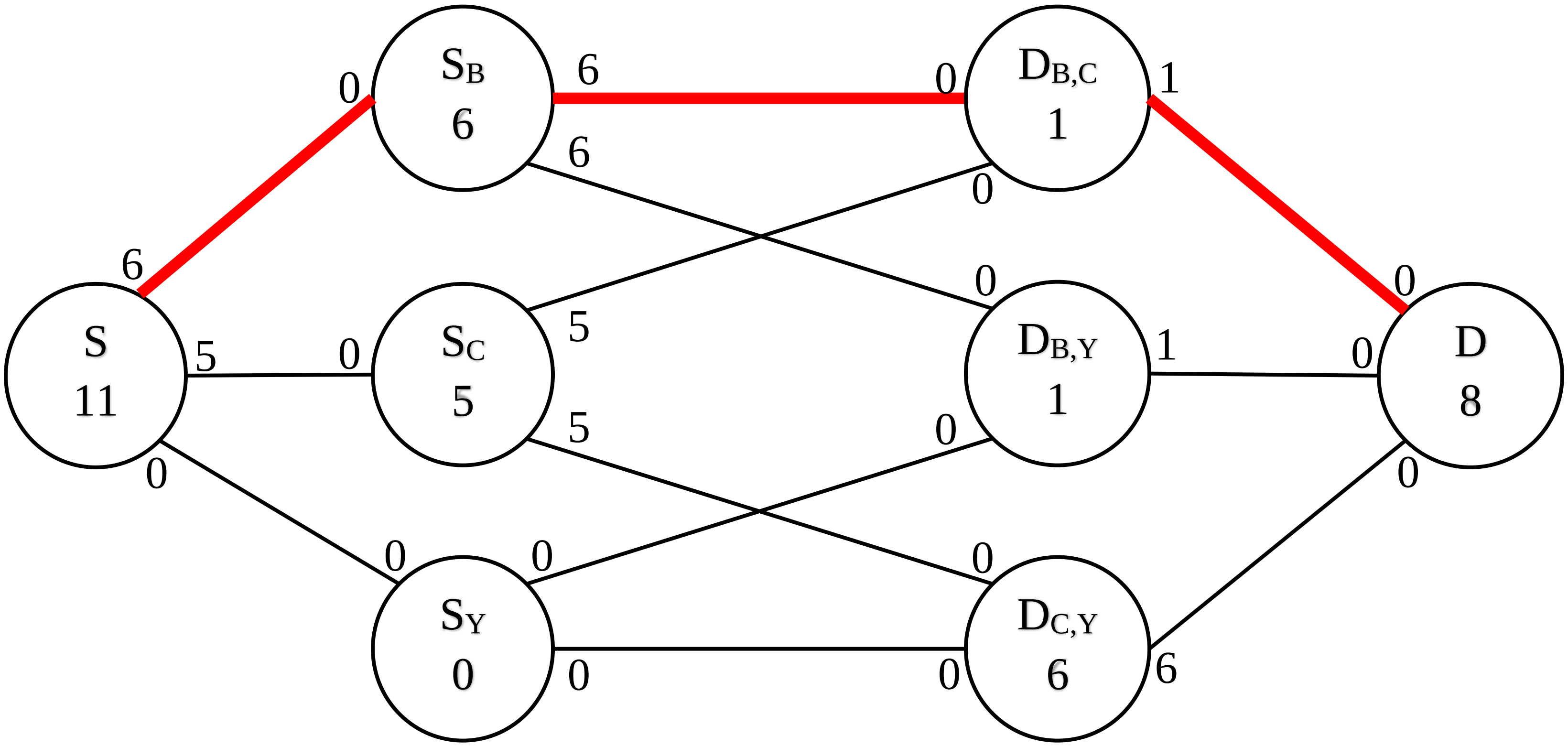
But it is still useful to review how the problem would be solved using Maximum Flow techniques.

Following the approach to the transshipment problem described in the previous lecture, we create a residual network of undirected edges, with a super source and a super sink.

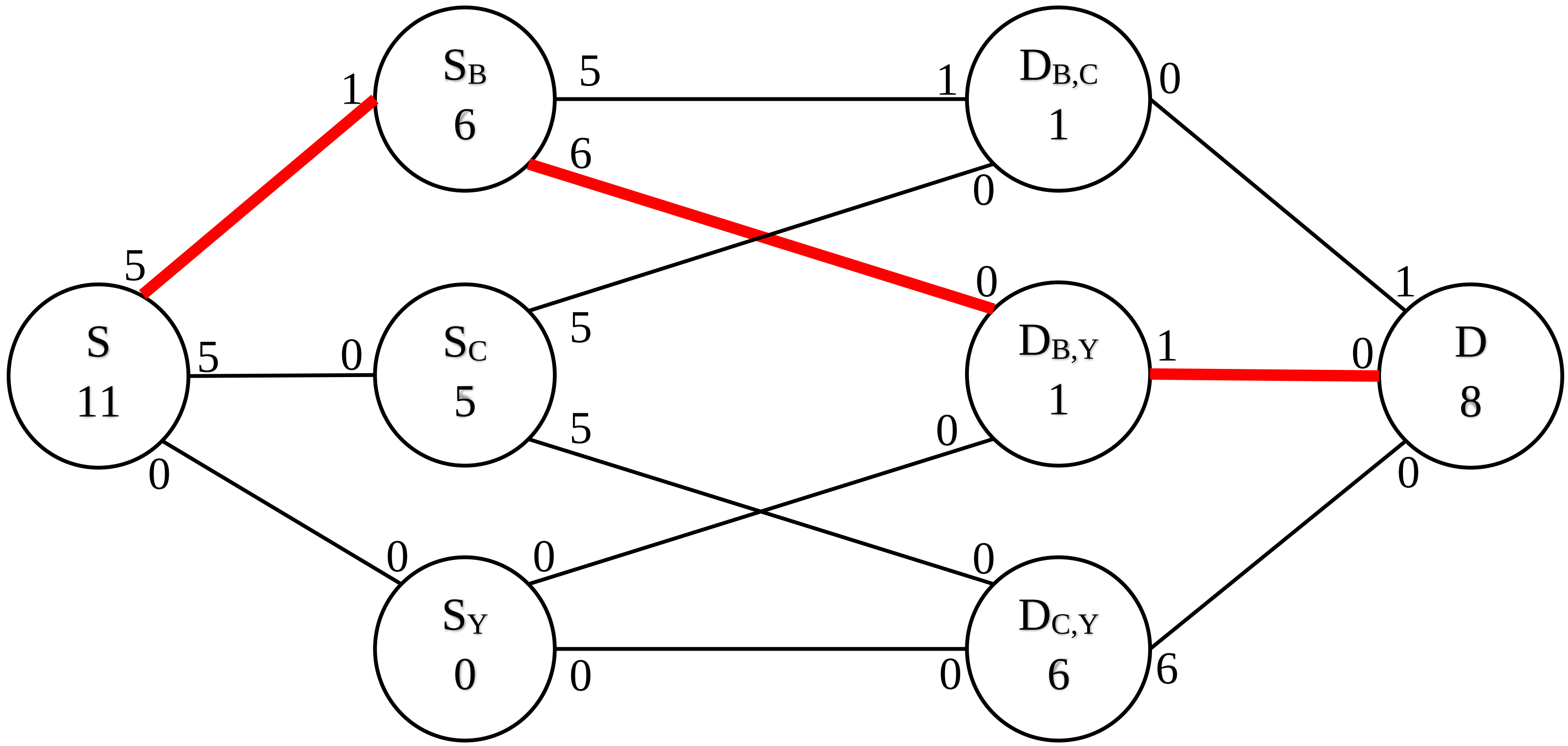


.

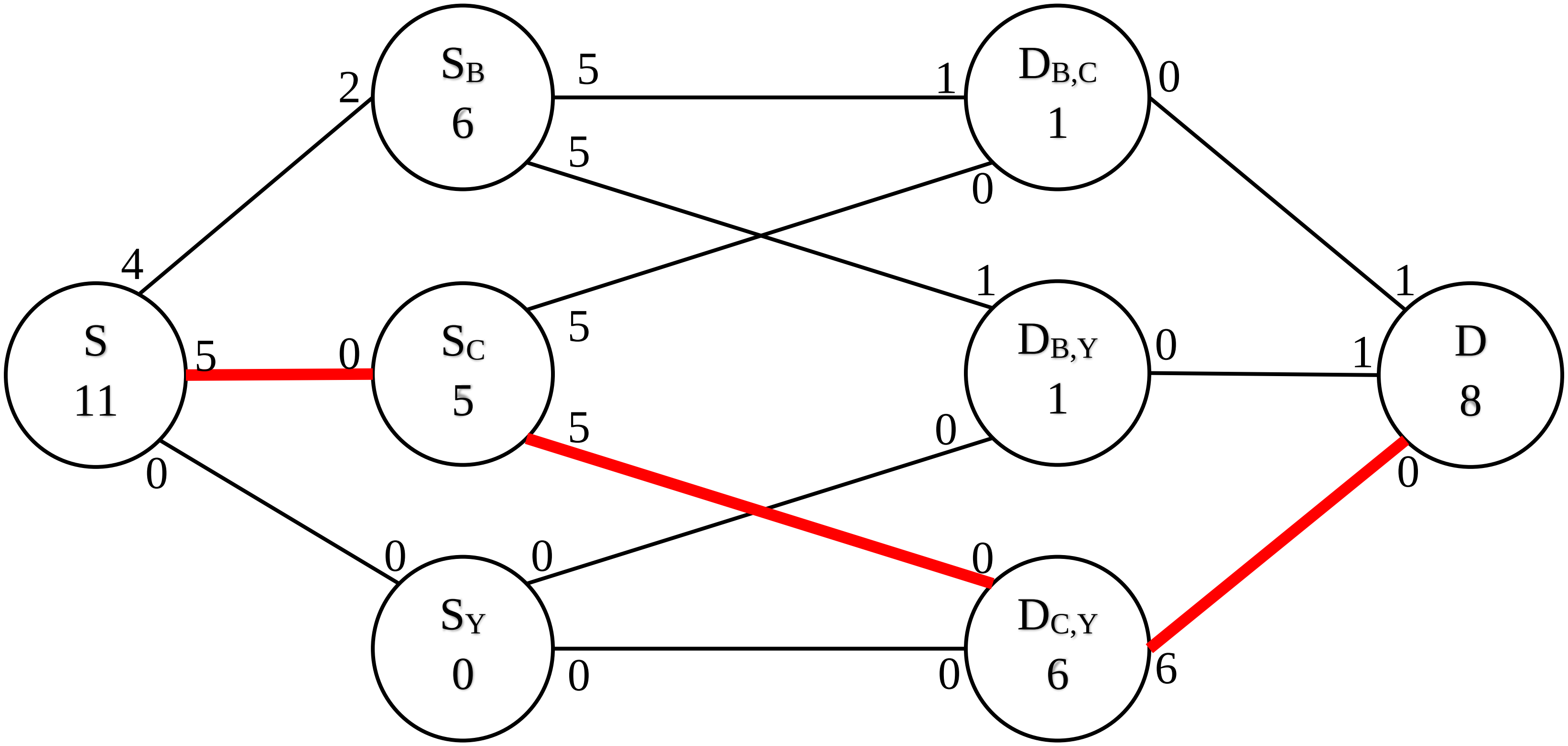
We identify an augmenting path capable of carrying 1 unit of flow.



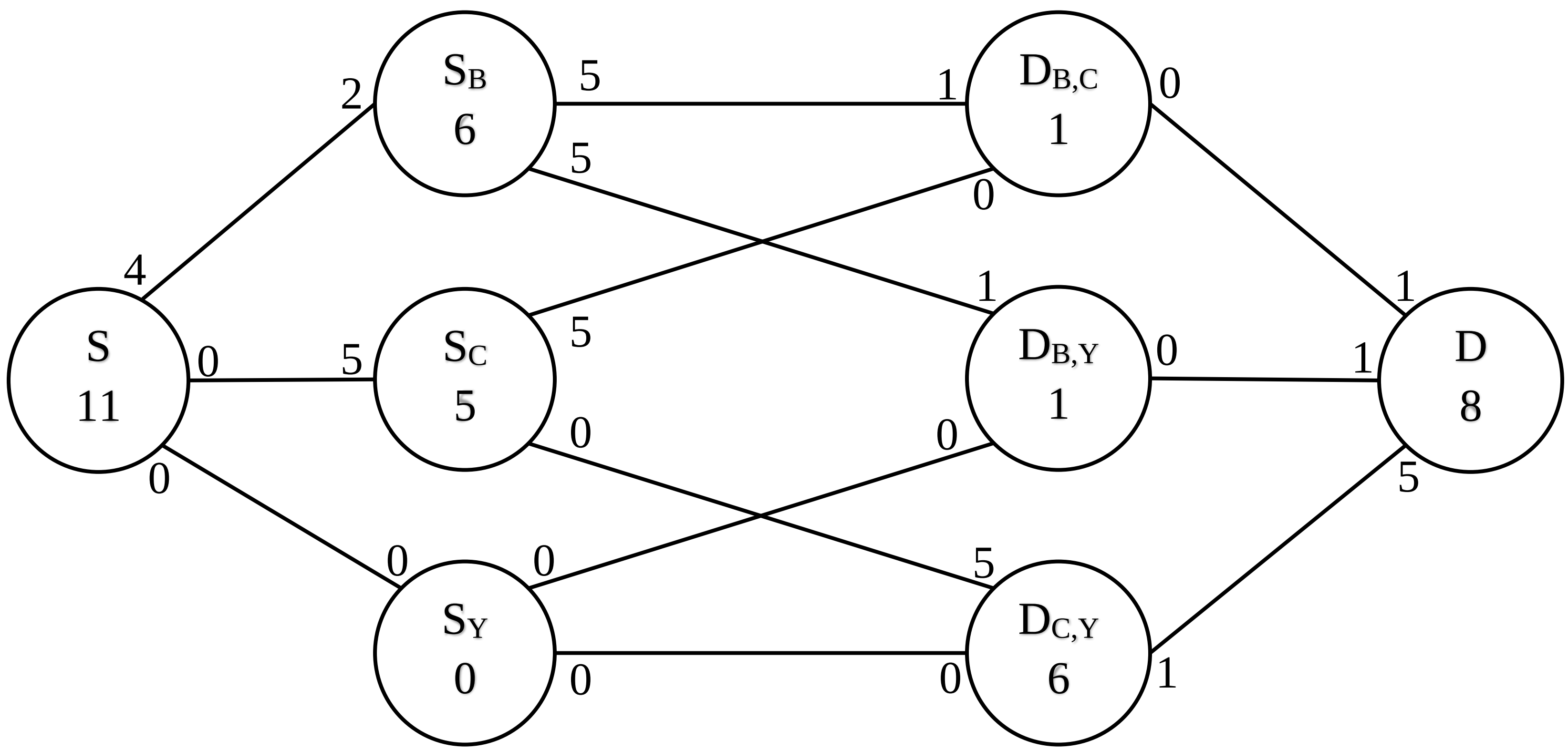
We turn on 1 unit of flow in the augmenting path identified above and update the residual network. In the new residual network, we identify an augmenting path capable of carrying 1 unit of flow.



We turn on 1 unit of flow in the augmenting path identified above and update the residual network. In the new residual network, we identify an augmenting path capable of carrying 5 units of flow.



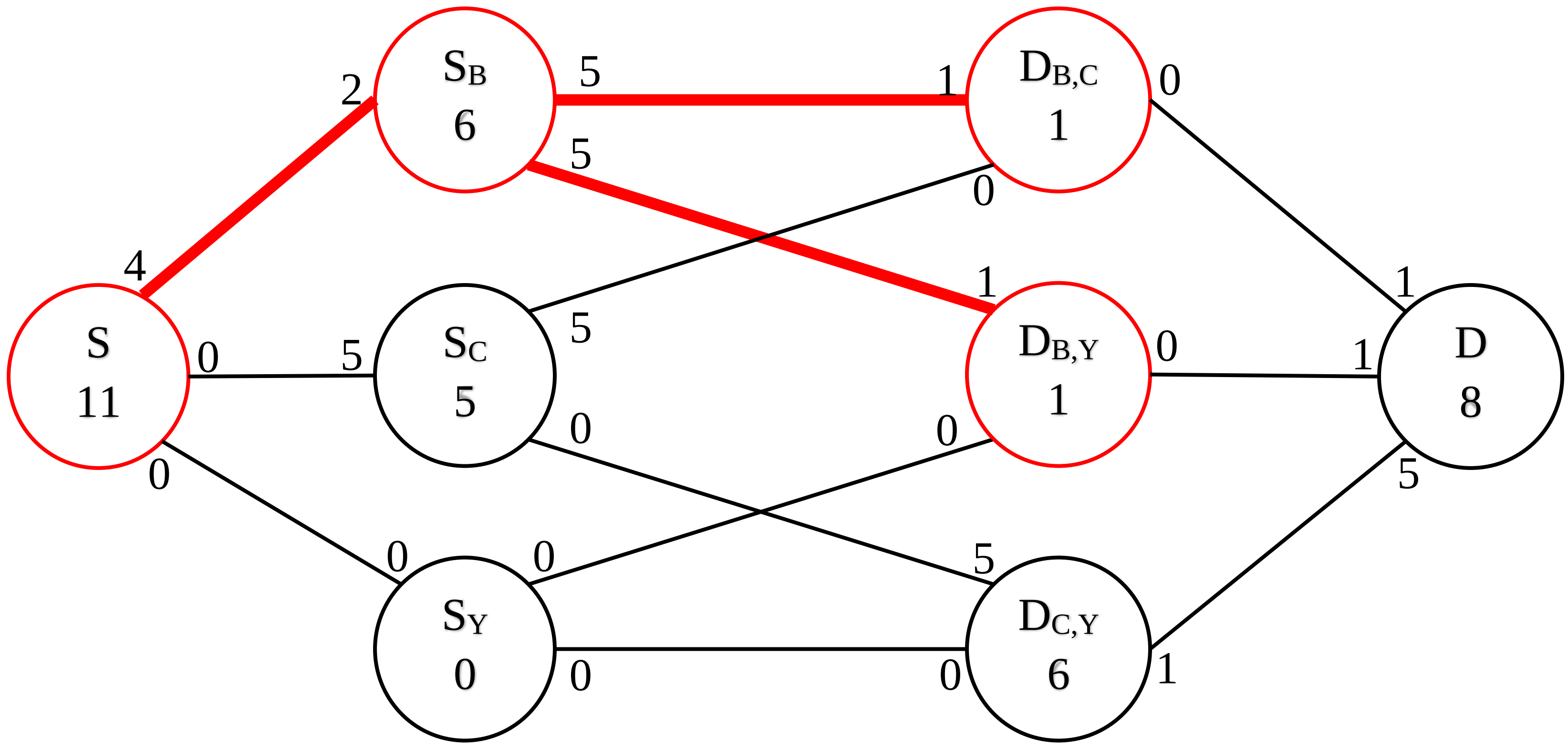
We turn on 5 units of flow in the augmenting path identified above and update the residual network.



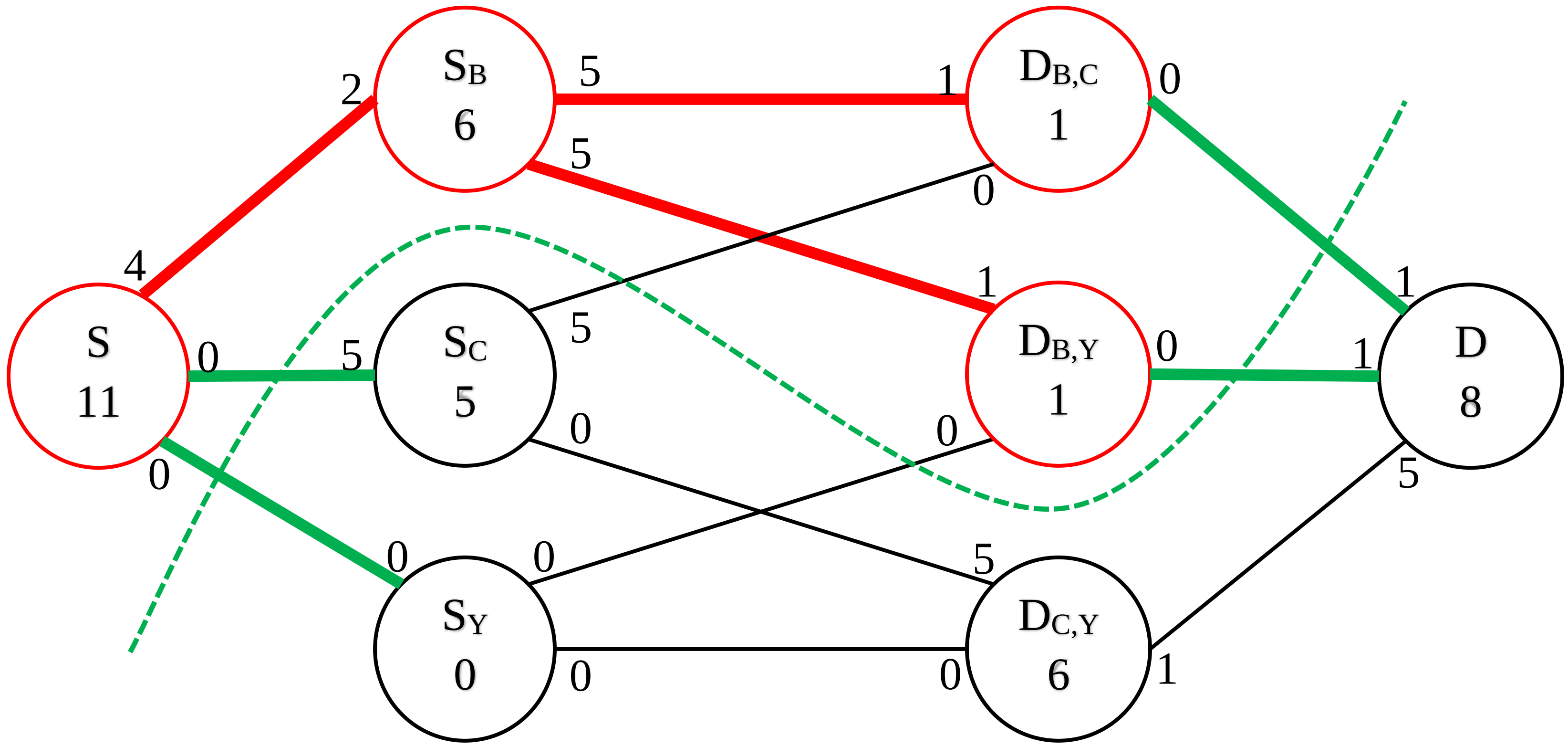
In the new residual network, there are no augmenting paths.

Since the total flow in the network is 7, which is less than the demand of 8, we conclude that the transshipment problem *does not* have a solution. That is, Harvard is eliminated from the tournament.

We have solved the problem (twice!), but for the purpose of review, it is instructive to consider a Minimum Cut for the residual network. In the set V of all nodes, we consider the subset $S \subset V$ of all nodes for which there exist augmenting paths from the super source node. (The set S is the collection of red nodes in the picture below. The set $V - S$ is the collection of black nodes.)



The set of all directed edges going from *S* to *V – S* is our *cut*. That is, the cut is the union of the four directed edges (S,S_Y), (S,S_C), $(D_{B,Y},D)$, $(D_{B,C},D)$, which are shown in green in the diagram below. Note that the directed edges $(S_C,D_{B,C})$ and $(S_Y,D_{B,Y})$ are not include in the cut, because they are going the wrong direction.



The capacity of the cut is the sum of the capacities of the arcs forming the cut. That is, the cut has capacity 0 + 5 + 1 + 1 = 7.

Notice that the capacity of this cut equals the max flow value found earlier.