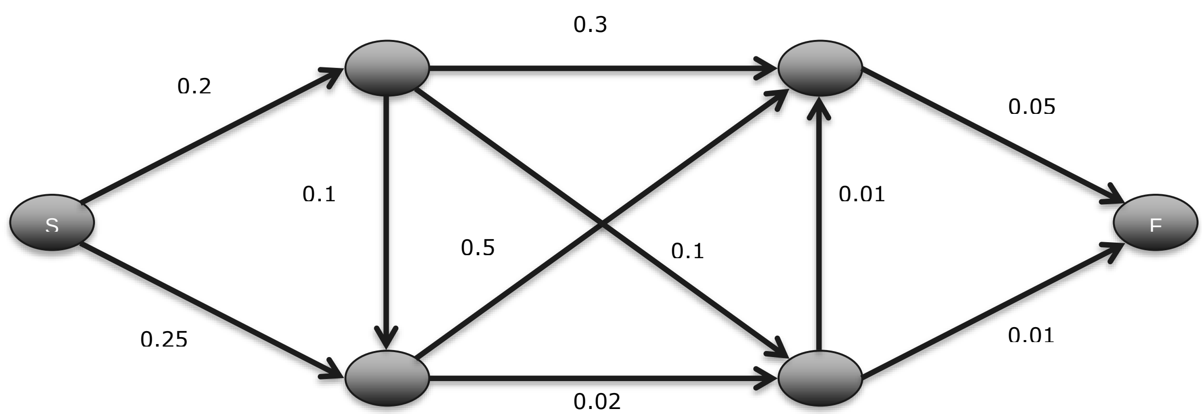
A Stochastic Dynamic Programming Model for Optimal Zombie Avoidance

Brian Denton

Industrial and Operations Engineering, University of Michigan

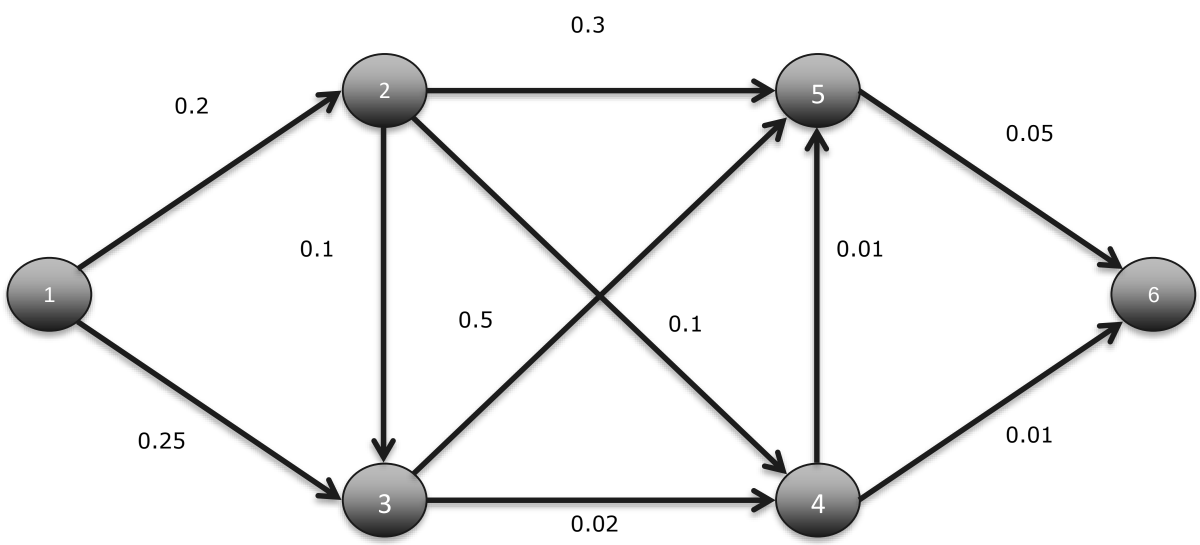
10/11/2016

Shortly after a zombie apocalypse you and your companions must traverse the following network in which edge weights are the probability of encountering a zombie along each edge. Your goal is to avoid encountering zombies as you move from the start vertex (S) to the finish vertex (F). At each vertex you select among the available edges with the following exception. If you encounter a zombie along a particular edge then due to the overwhelming anxiety associated with zombie encounters you hastily choose the next edge to traverse by randomly selecting among the options with equal probability. Future decisions are made deterministically provided you don't encounter another zombie.



This is a sequential decision making problem that can be modeled as a stochastic dynamic program. To develop a model formulation it is necessary to carefully define the decision maker's goal. In this case we may assume that the decision maker wants to maximize the probability of not encountering a zombie (or equivalently minimize the probability of an encounter). Thus, we can assume that every decision in the network (aside from those that are made randomly following a zombie encounter) has this goal in mind.

Next, we consider how to define the *state* of the system. This requires that we decide on the minimal amount of information needed to make a decision at any given point in the problem. In this problem the factors that affect the decision are: (a) your location in the network and (b) whether or not a zombie was encountered along the most recently traversed edge. To define the states we also need to define the sequence of stages for the problem. This can be done by labeling the vertices such that the order of vertices guarantees that the optimal value functions are computed in the order necessary to guarantee the optimality equations can be evaluated at each stage. A simple algorithm for this (discussed in class) is to arbitrarily select any vertex that has only outgoing edges and label this vertex 1; next, remove vertex 1 and all of its outgoing edges from the network. Arbitrarily select another vertex with only outgoing edges and label this vertex 2; next, remove vertex 2 and all its outgoing edges, and so on. Continue until all of the vertices have been labeled. For this problem the labeled network is as follows.



The next step is to define the *optimal value function.* For this problem a reasonable choice is to select the action in each state that will maximize the future probability of not encountering a zombie. To express the problem we let Z and $\bar{Z}$ represent encountering and not encountering a zombie, respectively. Thus the state at any stage of the problem is the combination of the vertex and either Z or $\bar{Z}$. The optimality equations for this problem can be written as follows:

$$v_t(\bar{Z}) = \max_{a \in A}\{\Pr(\bar{Z} \text{ along arc a})\, v_{t+1}(\bar{Z})\}$$

$$v_t(Z) = \sum_{a \in A} q(a) \Pr(\bar{Z} \text{ along arc a})\, v_{t+1}(\bar{Z})\},$$

for all t, and the boundary condition is $v_T(Z) = v_T(\bar{Z}) = 1$. In the above equations q(a) is the probability of selecting edge *a* which is $1/|A|$, i.e., one over the cardinality of the set of edges leaving the current vertex.

Applying this model we obtain the following:

Stage 6:

$v_6(\bar{Z}) = 1.0$, $v_6(Z) = 1.0$

Stage 5:

$v_5(\bar{Z}) = 0.95 v_6(\bar{Z}) = .95, \quad v_5(Z) = 0.95 v_6(\bar{Z}) = .95$

Stage 4:

$v_4(\bar{Z}) = \max_{\{5,6\}}\{0.99 v_5(\bar{Z}),\, 0.9 v_6(\bar{Z})\} = .94, \quad v_4(Z) = 0.5\big(0.99 v_5(\bar{Z})\big) + 0.5\big(0.99 v_6(\bar{Z})\big) = .96$

Stage 3: $v_3(\bar{Z}) = \max_{\{4,5\}}\{0.98 v_4(\bar{Z}),\, 0.5 v_5(\bar{Z})\} = .92, \quad v_3(Z) = 0.5\big(0.5 v_5(\bar{Z})\big) + 0.5\big(0.98 v_4(\bar{Z})\big) =$ .70

Stage 2:

$v_2(\bar{Z}) = \max_{\{3,4,5\}}\{0.9 v_3(\bar{Z}),\, 0.9 v_4(\bar{Z}),\, 0.7 v_5(\bar{Z})\} = .85$, $v_2(Z) = 0.33\big(0.9 v_3(\bar{Z})\big) + 0.33\big(0.9 v_4(\bar{Z})\big) + 0.33\big(0.7 v_5(\bar{Z})\big) = .74$

Stage 1: $v_1(\bar{Z}) = \max_{\{2,3\}}\{0.8 v_2(\bar{Z}),\, 0.75 v_3(\bar{Z})\} = .69$

Thus, there is a 0.69 probability you well never encounter a zombie if you follow the optimal policy. The complete optimal policy can be expressed as follows:

State	Optimal Action
1, $\bar{Z}$	Go to node 3
2, $\bar{Z}$	Go to node 4
2,Z	Randomly select an edge
3, $\bar{Z}$	Go to node 4
3,Z	Randomly select an edge
4, $\bar{Z}$	Go to node 5
4,Z	Randomly select an edge
5, $\bar{Z}$	Go to node 6
5, Z	Randomly select an edge

The optimal path if you never encounter a zombie: 1→ 3→4→5→ 6. If you do encounter a zombie along any edge then your posterior probability of not encountering a zombie becomes zero. This explains why the optimal value functions for the Z states associated with each of the nodes do not influence the optimal value function for the $\bar{Z}$ states. Upon encountering a zombie the decision maker is assumed to randomly select an arc and then make all future decisions with the goal of maximizing the probability of not encountering a zombie in the future.

Questions to consider:

1) Are the optimality equations above guaranteed to identify an optimal policy for this problem?
2) What other applications can you think of for this model?