# Optimal Discounted Cost in Weighted Graphs

Ashutosh Trivedi

September 22, 2017

**Definition 1.** *A weighted finite graph $\mathcal{G}$ is a tuple $(V, E, A, w)$ where:*

- *$V$ be a finite set of vertices,*
- *$A$ be a finite set of actions,*
- *$E : V \times A \to V$ be the set of edges, and*
- *$w : V \times A \to \mathbb{N}$ be the weight function.*

A path of $\mathcal{G}$ is a sequence $\langle v_1, a_1, v_2, a_2, \ldots, v_{k+1}\rangle$ of vertices and actions such that for all $1 \leq i \leq k-1$ we have that $E(v_i, a_i) = v_{i+1}$. Similarly an infinite path of $\mathcal{G}$ is an infinite sequence $\langle v_1, a_1, v_2, a_2, \ldots\rangle$ such that for all $i \geq 1$ we have that $E(v_i, a_i) = v_{i+1}$. We say that a path starts from a vertex $v$ if the first vertex in the path sequence is $v$. Let $\text{FPath}(v)$ and $\text{Path}(v)$ be the set of all finite and infinite paths, resp., starting from vertex $v \in V$. We write $\text{FPath}(v, N)$ for the set of all paths starting from $v$ of length $N$.

## 1 Finite-Horizon Total Reward Problems

Given a path $\pi = \langle v_1, a_1, \ldots, v_{k+1}\rangle$ we define its cost $\text{Cost}(\pi)$ as

$$\text{Cost}(\pi) = \sum_{i=1}^{k} w(v_i, a_i).$$

For a vertex $v$ and step $N$, we write $\text{Cost}_N(v)$ as minimum cost of all paths starting from $v$ of length $N$, and is defined as

$$\text{Cost}_N(v) = \min_{\pi \in \text{FPath}(v,N)} \text{Cost}(\pi).$$

Consider the following set of *optimality equations*:

$$\mathcal{C}(v, N) = \begin{cases} 0 & \text{if } N = 0 \\ \min_{a \in A} \{w(v, a) + \mathcal{C}(v', N-1) \;:\; E(v, a) = v'\} & \text{otherwise.} \end{cases}$$

**Theorem 1.** *If a function $C : V \times [\![N]\!] \to \mathbb{N}$ is a solution of optimality equations 1, then $C(v, N) = \textsc{Cost}_N(v)$.*

*Proof.* The proof is in two parts. First we show that for every path $\pi \in \textsc{FPath}(v, N)$ of length $N$ we have that $C(v, N) \leq \textsc{Cost}(\pi)$, and then we show that there is a path $\pi_* \in \textsc{FPath}(v, N)$ such that $C(v, N) = \textsc{Cost}(\pi_*)$.

— Consider an arbitrary path $\pi = \langle v_1, a_1, v_2, \ldots, v_{N+1} \rangle \in \textsc{FPath}(v, N)$ starting from $v$, i.e. $v = v_1$. We have that:

$$\begin{aligned}
C(v_1, N) &\leq w(v_1, a_1) + C(v_2, N-1) \\
&\leq w(v_1, a_1) + w(v_2, a_2) + C(v_3, N-2) \\
&\leq \cdots \\
&\leq \sum_{i=1}^{N} w(v_i, a_i) + C(v_{N+1}, 0) \\
&= \sum_{i=1}^{N} w(v_i, a_i) = \textsc{Cost}(\pi).
\end{aligned}$$

— Now consider the path $\pi_* = \langle v_1, a_1, v_2, \ldots, v_{N+1} \rangle \in \textsc{FPath}(v, N)$ starting from $v$, i.e. $v = v_1$ such that each $a_i$ is carefully chosen using the solution of optimality equations in the following fashion:

$$a_i = \operatorname*{arg\,min}_{a \in A} \left\{ w(v_i, a) + C(v', N-i) \ : \ (v_i, a, v') \in E \right\}.$$

Now it is clear to see that for this path we have:

$$\begin{aligned}
C(v_1, N) &= w(v_1, a_1) + C(v_2, N-1) \\
&= w(v_1, a_1) + w(v_2, a_2) + C(v_3, N-2) \\
&= \cdots \\
&= \sum_{i=1}^{N} w(v_i, a_i) + C(v_{N+1}, 0) \\
&= \sum_{i=1}^{N} w(v_i, a_i) = \textsc{Cost}(\pi).
\end{aligned}$$

From these two parts it follows that $C(v, N) = \textsc{Cost}_N(v)$. The proof is now complete. $\square$

**Theorem 2.** *The finite horizon optimal cost problem can be solved in $O(N \cdot |V| \cdot |A|)$.*

## 2 Finite-Horizon Discounted Reward Problems

For this section let us fix a discount factor $0 \leq \lambda < 1$ Given a path $\pi = \langle v_1, a_1, \ldots, v_{k+1} \rangle$ we define its discounted cost $\textsc{DCost}(\pi)$ as

$$\textsc{DCost}(\pi) = \sum_{i=1}^{k} \lambda^{i-1} \cdot w(v_i, a_i).$$

For a vertex $v$ and step $N$, we write $\textsc{DCost}_N(v)$ as minimum discounted cost of all paths starting from $v$ of length $N$, and is defined as

$$\textsc{DCost}_N(v) = \min_{\pi \in \textsc{FPath}(v,N)} \textsc{DCost}(\pi).$$

Consider the following set of *optimality equations*:

$$\mathcal{D}(v, N) = \begin{cases} 0 & \text{if } N = 0 \\ \min_{a \in A} \{w(v, a) + \lambda \cdot \mathcal{D}(v', N-1) \ : \ E(v, a) = v'\} & \text{otherwise.} \end{cases}$$

**Theorem 3.** *If a function $D : V \times [\![N]\!] \to \mathbb{R}$ is a solution of optimality equations 2, then $D(v, N) = \textsc{DCost}_N(v)$.*

## 3 Infinite-Horizon Discounted Reward Problems

For this section let us fix a discount factor $0 \leq \lambda < 1$ Given an infinite path $\pi = \langle v_1, a_1, \ldots \rangle$ we define its discounted cost $\textsc{DCost}(\pi)$ as

$$\textsc{DCost}(\pi) = \sum_{i=1}^{\infty} \lambda^{i-1} \cdot w(v_i, a_i).$$

For a vertex $v$, we write $\textsc{DCost}(v)$ as minimum discounted cost of all infinite paths starting from $v$, and is defined as

$$\textsc{DCost}(v) = \min_{\pi \in \textsc{Path}(v)} \textsc{DCost}(\pi).$$

Consider the following set of *optimality equations*:

$$\mathcal{D}(v) = \min_{a \in A} \left\{ w(v, a) + \lambda \cdot \mathcal{D}(v') \ : \ E(v, a) = v' \right\}$$

**Theorem 4.** *If a function $D : V \to \mathbb{R}$ is a solution of optimality equations 3, then $D(v) = \textsc{DCost}(v)$.*

*Proof.* The proof is in two parts. First we show that for every path $\pi \in \textsc{Path}(v)$ we have that $D(v) \leq \textsc{DCost}(\pi)$, and then we show that there is a path $\pi_* \in \textsc{Path}(v)$ such that $D(v) = \textsc{DCost}(\pi_*)$.

— Consider an arbitrary path $\pi = \langle v_1, a_1, v_2, \ldots \rangle \in \textsc{Path}(v)$ starting from $v$, i.e. $v = v_1$. We have that:

$$
\begin{array}{rcl}
D(v_1) & \leq & w(v_1, a_1) + \lambda D(v_2) \\
& \leq & w(v_1, a_1) + \lambda w(v_2, a_2) + \lambda^2 D(v_3) \\
& \leq & \cdots \\
& \leq & \displaystyle\sum_{i=1}^{\infty} \lambda^{i-1} w(v_i, a_i) = \mathrm{D}\textsc{Cost}(\pi).
\end{array}
$$

— Now consider the path $\pi_* = \langle v_1, a_1, v_2, \ldots \rangle \in \textsc{Path}(v)$ starting from $v$, i.e. $v = v_1$ such that for every vertex $v_i$ the action $a_i$ is carefully chosen using the solution of optimality equations in the following fashion:

$$
a_i = \arg\min_{a \in A} \left\{ w(v_i, a) + \lambda D(v') \;:\; (v_i, a, v') \in E \right\}.
$$

Notice that such path would be eventually periodic. In particular, every time a vertex is visited, same action can be chosen. Now it is clear to see that for this path we have:

$$
\begin{array}{rcl}
D(v_1) & = & w(v_1, a_1) + \lambda D(v_2) \\
& = & w(v_1, a_1) + w(v_2, a_2) + \lambda^2 D(v_3) \\
& = & \cdots \\
& = & \displaystyle\sum_{i=1}^{\infty} \lambda^{i-1} \cdot w(v_i, a_i) = \textsc{Cost}(\pi).
\end{array}
$$

From these two parts it follows that $C(v, N) = \textsc{Cost}_N(v)$. The proof is now complete. $\square$