# Selected Topics in Algorithms
# Unique Shortest Paths

Kurt Mehlhorn

May 19, 2009

We investigate techniques for making shortest paths unique. We studied two methods in class. A third method was suggested to me by Thomasz Jurkiewicz.

**Perturbation by Infinitesimals** Order $E$ by $e_1 < e_2 < \ldots < e_m$. We set $w'(e_i) = w(e_i) + (\varepsilon - \varepsilon_i)$ where $\varepsilon \gg \varepsilon_1 \gg \varepsilon_2 \gg \ldots \gg \varepsilon_m > 0$ and the $\varepsilon_i$ and $\varepsilon$ are infinitesimals[1] With respect to $w'$, the weight of a path $p$ is

$$w'(p) = w(p) + |p|\varepsilon - \sum_e (e \in p)\varepsilon_i \ ,$$

where $(e \in p)$ is 1 if $e$ occurs in $p$ and is zero otherwise. Consider two distinct paths $p$ and $q$. We have $w'(p) < w'(q)$ if

- $w(p) < w(q)$ or
- $w(p) = w(q)$ and $|p| < |q|$ or
- $w(p) = w(q)$ and $|p| = |q|$ and $\min(p \setminus q) < \min(q \setminus p)$.

Advantage: mathematically clean
Disadvantage: costly to implement

**The Method by Hartvigsen and Mardon** A clever and efficient implementation of the above.
Advantage: efficient
Disadvantage: complex

**Random Perturbation** This was suggested to me by Thomasz. We set $w'(e_i) = w(e_i) + r_i\varepsilon$, where $\varepsilon$ is a positive infinitesimal and $r_i$ is a random integer in $[0..M-1]$ for a still to be determined $M$.

[1]Formally, distances are formal sums of the form $a + b\varepsilon + \sum_i c_i\varepsilon_i$, addition is as for polynomials with variables $\varepsilon$ and $\varepsilon_i$, and comparison is lexicographic.

**Lemma 1** *Consider a fixed source s. The probability that all shortest paths with source s are unique is at least* $e^{-2m/M}$.

**Proof:** Conceptually run Dijkstra's algorithm with weight function $w'$. Consider any fixed node $v$. The tentative distance $d(v)$ is initialized to $+\infty$. Whenever a neighbor $u$ of $v$ is deleted from the queue, $d(v)$ is updated to $\min(d(v), d(u) + w'(uv))$. Let $e = uv$. We call the choice of $r_e$ good if $d(v) \neq d(u) + w'(e)$, and bad otherwise. There is at most one bad choice for $r_e$ and hence at least $M - 1$ good choices.

We claim that if the choices for all edges are good, all shortest paths with source $s$ are unique. Assume otherwise. Then there must be a node $v$ with two shortest paths from $s$ to $v$ and such that these paths use different edges into $v$, say $xv$ and $yv$. Also assume that $x$ is removed from the queue before $y$. Since the choice for $e_{yv}$ is good, we have $d(y) + w'(yv) \neq d(x) + w'(xv)$.

The probability that all choices are good is at least

$$\left(\frac{M-1}{M}\right)^m = \left(1 - \frac{1}{M}\right)^{M(m/M)} \geq e^{-2m/M}$$

since[2] $(1 - 1/M)^M \geq e^{-2}$ for $M > 10$. ∎

**Theorem 1** *Let* $w'$ *be as above. With* $M = 8nm$, *the probability that shortest paths are not unique is at most* $1/2$.

**Proof:** For a fixed source $s$, shortest paths are unique with probability at least $e^{-2m/M}$. We cannot argue that therefore all shortest paths are unique with probability at least $(e^{-2m/M})^n$ as these probabilities are not independent[3]. The correct reasoning is as follows and leads almost to the same result.

The probability that for a fixed source, shortest paths are not unique is at most $1 - e^{-2m/M}$. Hence the probability that for some source, shortest paths are not unique is at most $n(1 - e^{-2m/M})$. For[4] $M = 8nm$, we have

$$n(1 - e^{-2m/M}) \leq n(4m/M) = 1/2\ .$$

∎

We run our favorite all-pairs algorithm for weight function $w'$. Let $d$ be the distance function computed. We perform the following check. For any pair $(u,v)$ with $v \neq u$, we check whether there are two neighbors $v'$ and $v''$ of $v$ with $d(u,v) = d(u,v') + w'(v'v) = d(u,v'') + w'(v''v)$. If this is the case for some pair $(u,v)$, we declare the perturbation a failure, choose new values $r_i$ and repeat. The check takes time $O(nm)$. We fail with probability at most $1/2$ and hence the expected number of trials is at most 2.

Advantage: Easy to implement and conceptually simple.

---

[2] Recall $\lim_{n\to\infty}(1 + x/n)^n = e^x$.

[3] Let us continue the incorrect reasoning. We have $(e^{-2m/M})^n = e^{-2nm/M} \geq 1 - 2nm/M$ since $e^x \geq 1 + x$. Thus the probability that shortest paths are not unique is at most $2nm/M$.

[4] Recall $e^x \approx 1 + x$ for $x$ small. More precisely, $1 + 2x \leq e^x \leq 1 + x$ for $|x| \leq 1/2$.