Problem Set 1 FAQ

- **Perl tutorials**:
  - http://archive.ncsa.uiuc.edu/General/Training/PerlIntro/
  - http://proquest.safaribooksonline.com/main.asp?cat=9

- **Traceroute with AS numbers**. Several of you asked specifically about where to get a version of traceroute that supports the "-A" option.

  Note: You can do this assignment without having "traceroute -A" (although this does make the assignment easier. Use `whois -h whois.ra.net` and look at the `origin` field in the results.

- **Problems 7 and 8**. You can get away without unzipping route-views.bgp.20020903.gz. If you want to process the file (necessary for PS1 #8), you can pipe it to your program/script like so "zcat route-views.bgp.20020903.gz | ", where is your program that expects standard input.

  There are also is a more elegant way to do this in perl (and probably other languages, too), called a "pipe open". An example of this is given in the tutorial notes. The official documentation for pipe opens is at: http://www.perldoc.com/perl5.6/pod/perlopentut.html#Pipe-Opens. Alternatively, you can "man perlopentut".

  The other less-clever approach is to simply use temporary files that don't reside in your athena locker to unzip everything into temporarily while you're working on them.
  Possible places to use are:
  /tmp
  /var/tmp
  /mit/bitbucket
  /mit/bitbucket2
  You can't expect anything you put in any of these locations to persist for any length of time, but should be fine for a single session.

- **Problem 8.1** You may notice some notation such as {A,B} in a few of the AS paths in the Routeviews Table dump.

  It's a (somewhat obscure) aggregation trick called an "AS-Set". What happened there was that AS 10796 heard routes that it saw that it could aggregate coming from both 11060 and 12262.

  Essentially what it means is that the path could go through *either* 11060 or 12262. The reason they are both kept as a set is so that information is not lost in aggregation. Otherwise, routing loops might show up.

  The right thing to do here is to count this as 10796-11060 and 10796-12262. It doesn't happen often enough in the table to really affect the answers, (plus it's not something you should be expected to know) so if you've already handled these another way, don't sweat it (just tell us how you counted them).

  More information is available here:
  http://www.cisco.com/warp/public/459/aggregation.html

- **Problem 9.3**. You will (hopefully) notice that some of the IP addresses in the file that contains byte counts (i.e., 20001206.byte.summary.gz) do not correspond to any prefix in the routing table we provide (one of your tasks in part 4 is to explain why this is the case). In other words, you may find some IP addresses for which there are no longest prefix matches in the routing table.

  Thus, it is difficult for you to assign this portion traffic to either the Genuity or the Internet2 link. For this portion of traffic, list the destination as "uncertain". So, your answer for part three should look like something of the form: x% Genuity, y% Internet2, z% uncertain -- with, of course, correct values for x, y, and z. :)