Graph Analytics: Some Structural Tasks

- Degree Centrality
  - Degree centrality of *v*: *degree(v)* / |*E*|
    - "Important" nodes are those with high degree
- But: say we both have 5 friends
  - We have the same degree centrality
  - But what if your 5 friends are Barack Obama, Larry Page, Bill Gates, the Dalai Lama, and Oprah Winfrey?
  - Shouldn't you be considered more "important?"

Eigenvector Centrality (i.e., PageRank)

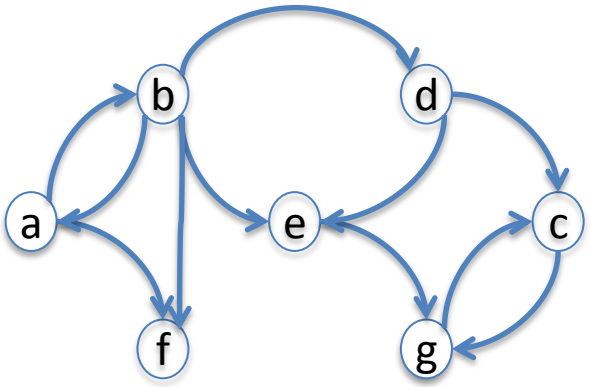
- Basic idea (but oversimplified)

```
while not converged:
  for each vertex v
    rank(v) = sum of ranks from incoming edges
```

Two problems:

1) If a page with millions of outgoing links links to me, that's less valuable than a page with only a few outgoing links.

2) If I'm 27 hops away from Barack Obama, he shouldn't really influence my rank. We need a *damping factor.*



PageRank: 2nd attempt

```
while not converged:
```

$$PR(A) = \frac{1-d}{N} + d\left(\frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)} + \cdots\right).$$

ensures ranks sum to 1

- Assume edges B→A, C→A, D→A, …
- *PR(X)* is the PageRank of vertex *X*
- *L(X)* is the number of outgoing links from *X*
- *d* is the damping factor
- *N* is the number of vertices