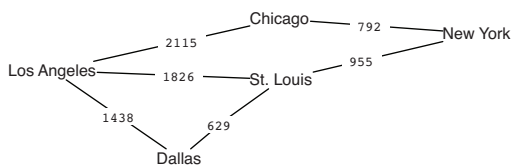
Introduction

In 1973, researchers at the University of South Florida began performing word association tests. A word association test is a quiz in which the subject is given a word and asked to say the first thing that comes to mind. For example, if I say *Halloween*, you might reply *pumpkin*. For the following two decades, researchers conducted around 750,000 of these tests on about 6,000 individuals.

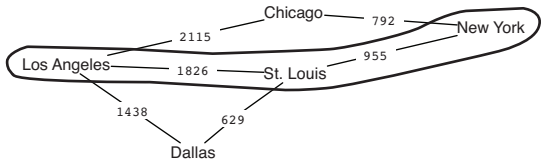
In 2008, while researching word association, I came across the study's results on the University of South Florida's website. The data was collected in a series of text files available online. They were almost indecipherable in their raw form, however, so I started writing simple programs that could interpret the data.

I arranged the words as a network. That is, *Halloween* and *pumpkin* are nodes, and because many people associated *pumpkin* with *Halloween,* the two are connected. *Pumpkin,* in turn, is connected to *pie*, *pie* to *crust*, and so on. The result is a structure with about 10,000 words and around 75,000 connections between them.

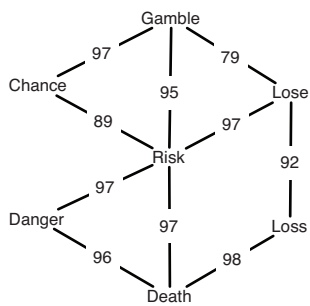
I then implemented an algorithm called Dijkstra's Shortest Path. In computer-science textbooks, the algorithm is often introduced like this:



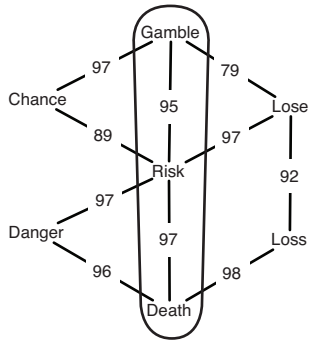
"What's the shortest distance between New York and Los Angeles?" The algorithm will find the most direct path:



The network I created with the word associations is very similar to the cities example. In the above example, the distances between New York and LA represent physical measurements, but in the word-association network, distances represent the strength of the association. In the illustration below, the smaller the number between the words, the stronger the association.



Thus, if you take the word *gamble,* the shortest path to the word *death* is through the word *risk*, based on the study's data.



This book contains the shortest path from every word in the USF study to *death*. The words are ordered by their associative distance from *death*.