Graphs and Network Flows
ISE 411

Lecture 1

Dr. Ted Ralphs

References for Today's Lecture

- Required reading
  - Sections 17.1, 19.1
- References
  - AMO Chapter 1 and Section 2.1 and 2.2

Introduction to Graphs and Network Flows

- What is a *graph*?
- What does *network flow* mean?
- The word network can be interpreted according to the dictionary definition.
  - electrical networks
  - communication networks
  - transportation networks (highways, railways, airline)
- The word flow comes from the movement of something from one point to another.
  - electricity
  - information
  - person or vehicle
  - inventory
  - money
  - Physical goods
- We will rigorously define the word graph shortly.

Graph Problems

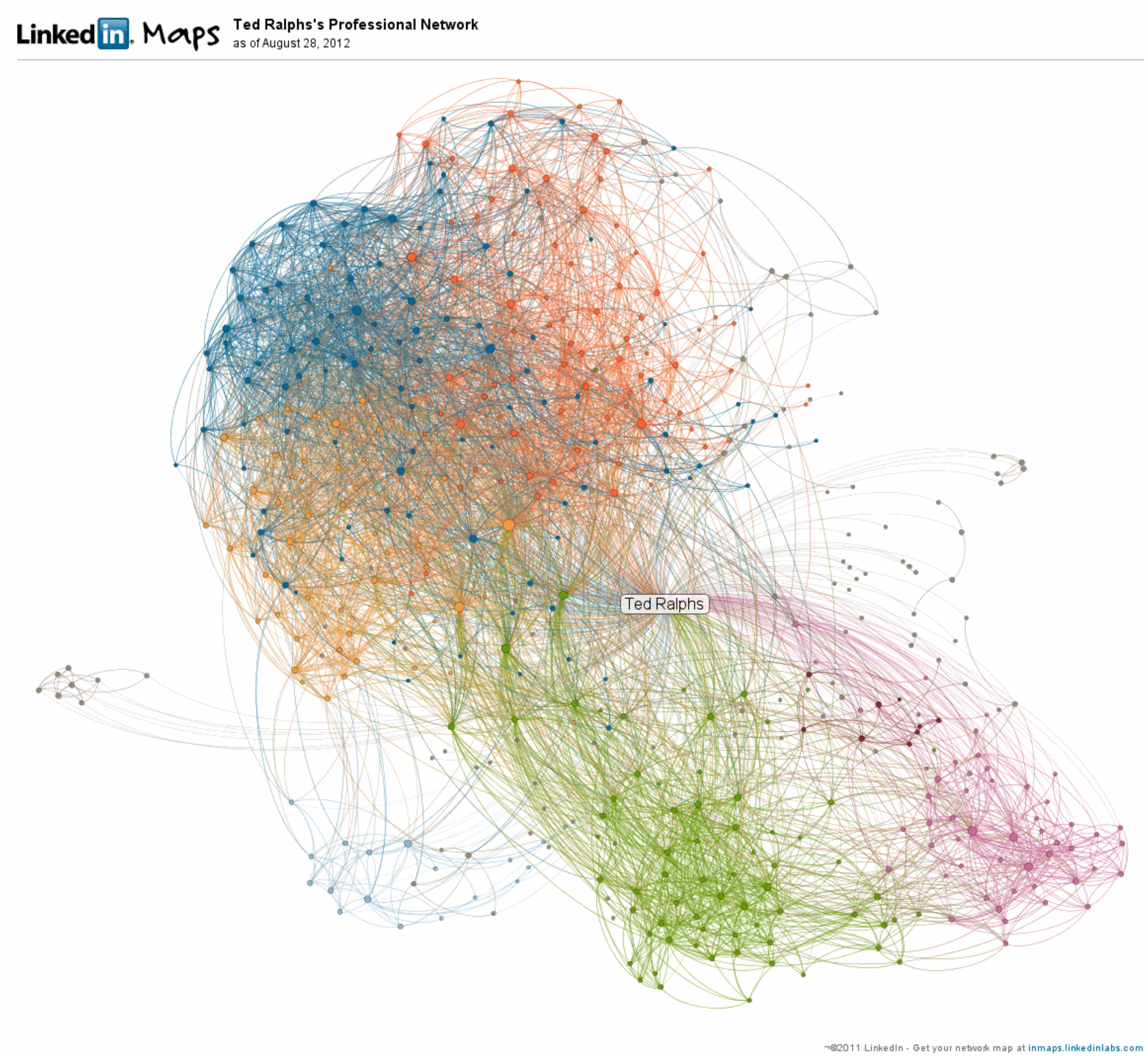
- Graphs model the *connectivity relationships* between items in a set.
- Specifically, we will specify that there is a direct link between certain pairs of items.
- In some cases, there will be a direction to the link.
- This will allow us to ask questions such as the following.
  - Is x connected "directly" to y?
  - Is x connected to y "indirectly," i.e., through a sequence of direct connections?
  - What is the set of of all items connected to x, directly or indirectly?
  - What is the shortest number of connections needed to get from x to y?

Applications of Graphs

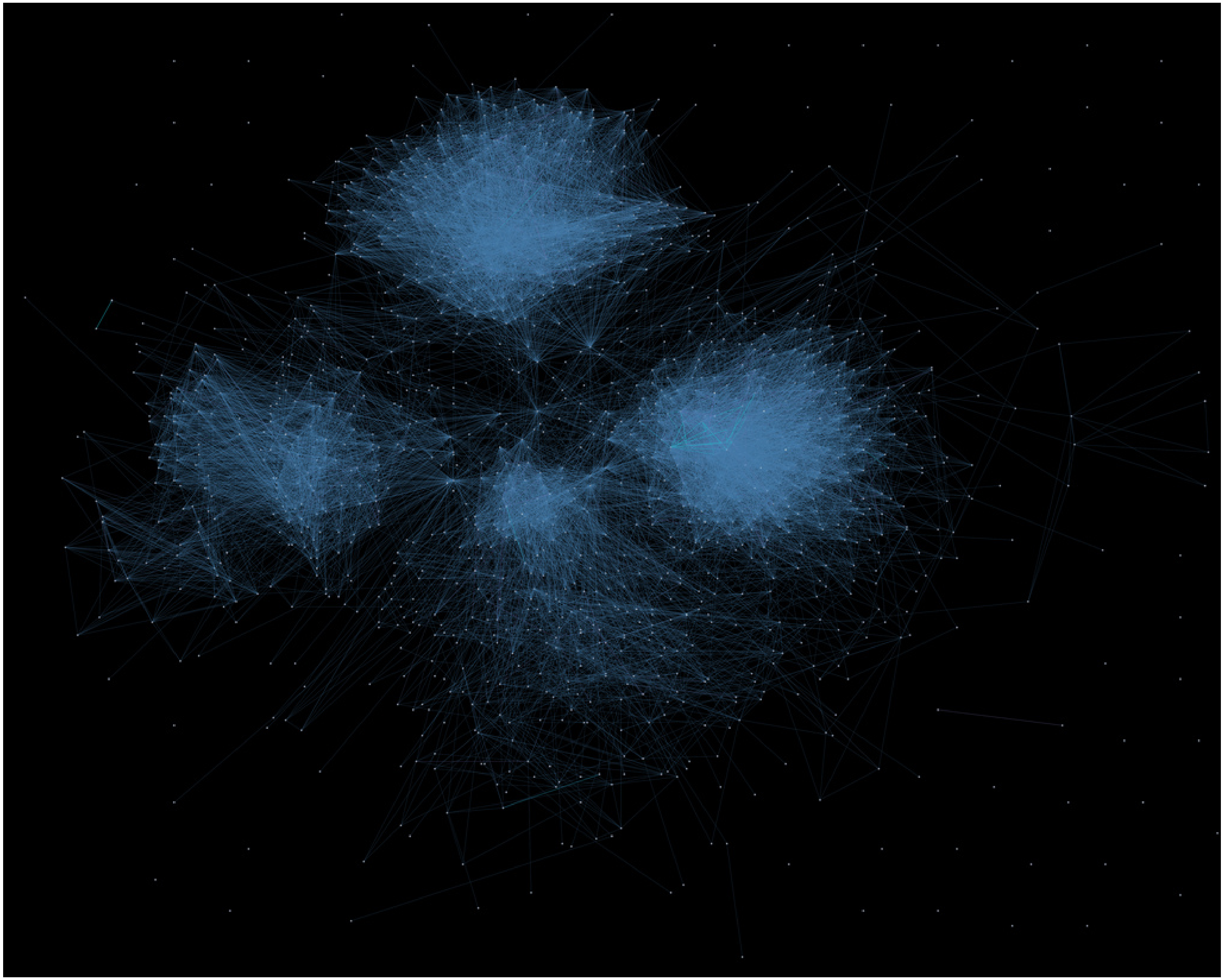
Applications of Graphs

- Maps
- Social Networks
- World Wide Web
- Circuits
- Scheduling
- Communication Networks
- Electricity Networks
- Biological Networks
- Matching and Assignment

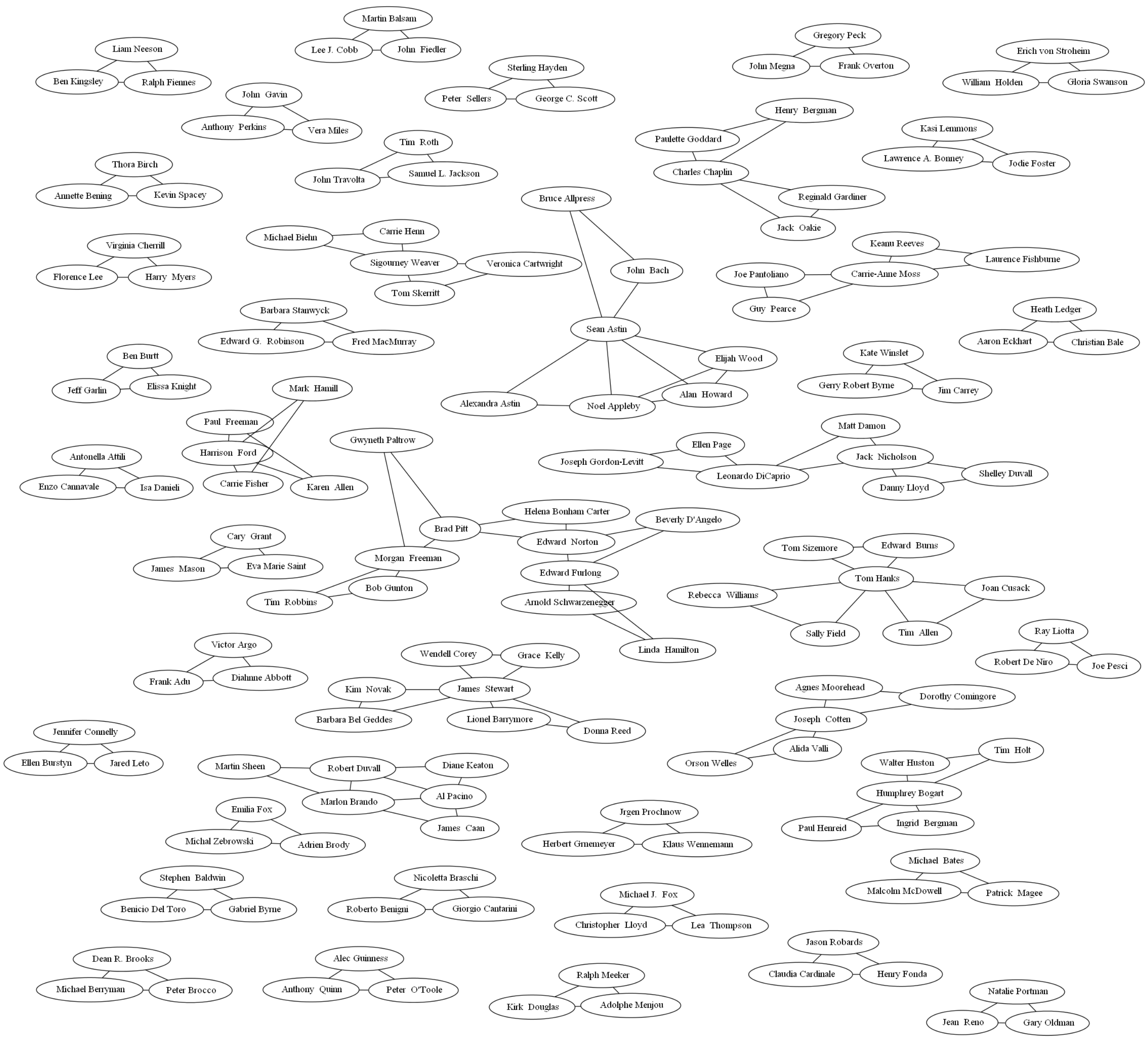
Graphs from Social Networks



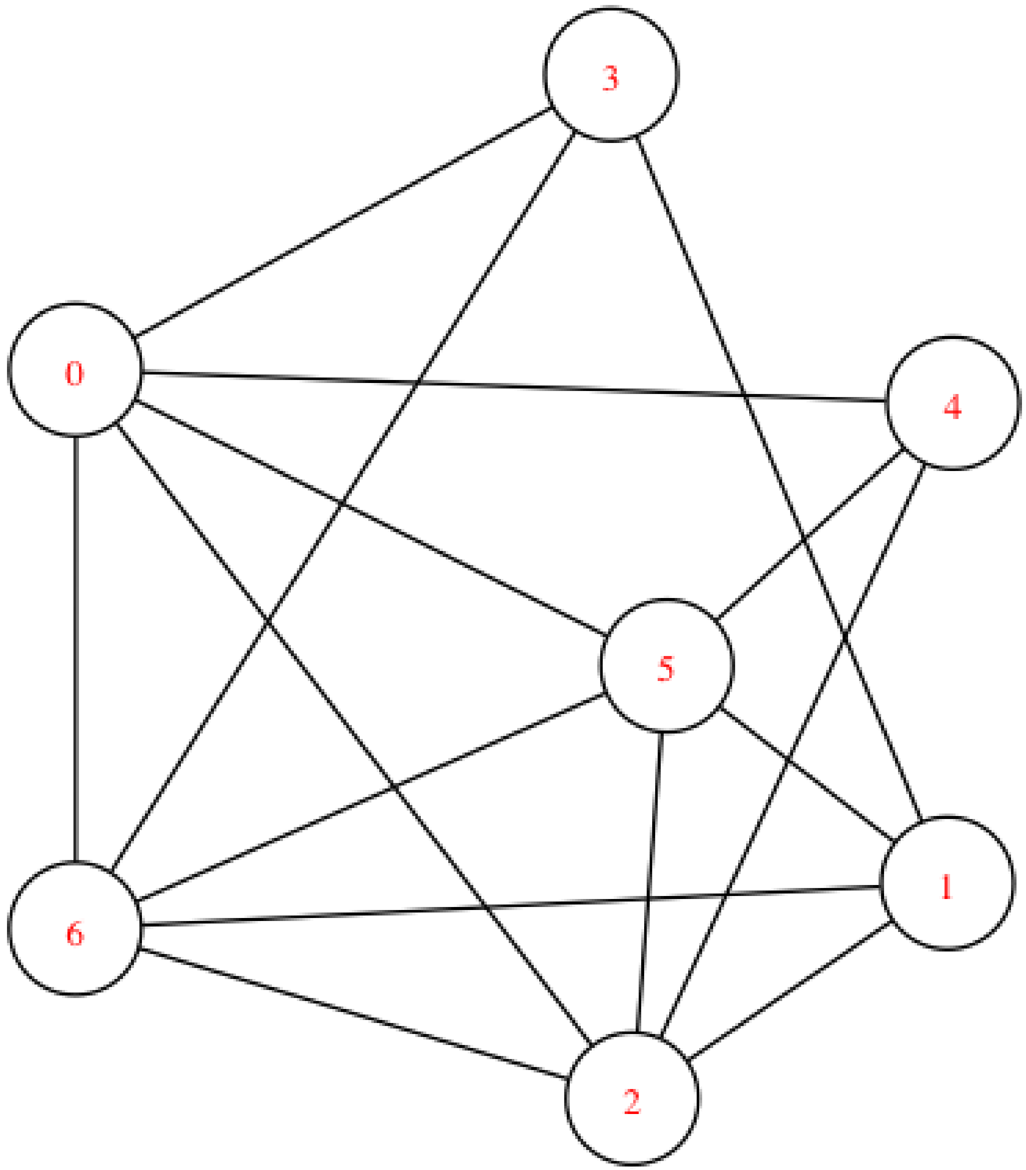
A Facebook Graph



Graphs for Fun



Example of an Abstract Graph



Network Flow Problems

- Network flow problems arise in many, many contexts
  - travel/transportation
  - logistics
  - manufacturing
  - telecommunications
  - chemistry/biology
  - finance
- Much of early work was *descriptive*.
- Our interest is in *prescriptive* models.

Basic Categories of Network Flow Problems

- Most of the problems we will encounter involve analyzing an existing network.
- Problem types
  - What is the shortest path from point to point?
  - What is the maximum throughput that can be achieved?
  - What is the minimum cost of moving a fixed quantity of goods from point to point?
- For the most part, the flow models we discuss will be *static*.
- We may discuss how to model dynamic flows near the end of the course.

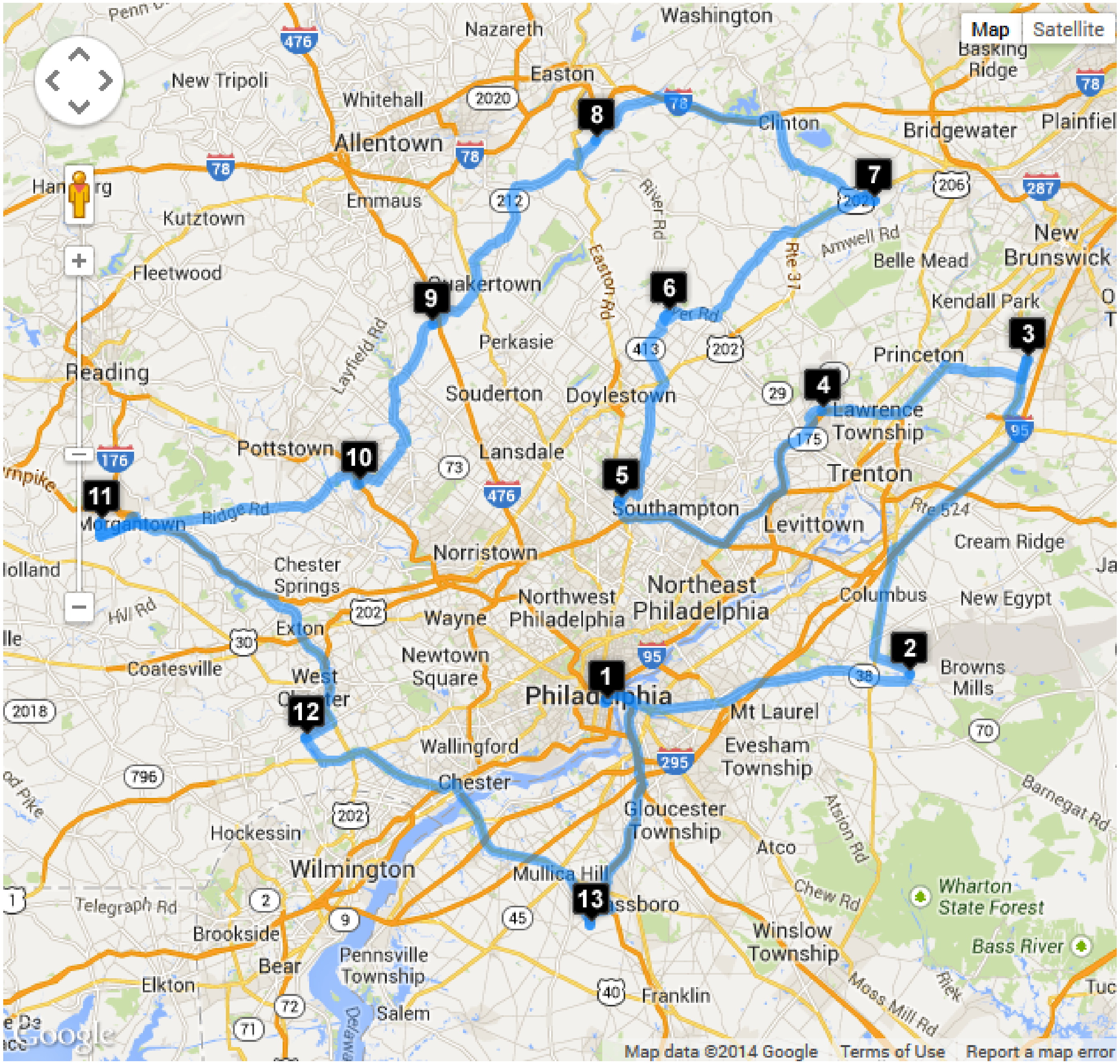
How Hard Are They?

- There is a well-defined sense in which flow problems are among the "easiest" optimization problems of practical importance.
- They are much easier to solve than general linear programs, for example.
- So why do we want to spend a whole course studying them?
- The applications of flow models require algorithms to be *extremely fast*.
  - GPS
  - Internet packet routing
- Algorithms for network flow problems are also embedded in a wide range of techniques for solving higher level models.
- For this reason, we will spend a substantial part of the course discussing data structures and other implementation issues.

Graphs

- A *graph* is an abstract object used to model connectivity relations.
- A graph consists of a list of items, along with a set of connections between the items.
- The study of such graphs and their properties, called *graph theory*, is hundreds of years old.
- Graphs can be visualized easily by creating a physical manifestation.
- There are several variations on this theme.
  - The connections in the graph may or may not have an *orientation* or a *direction*.
  - We may not allow more than one connection between a pair of items.
  - We may not allow an item to be connected to itself.
- For now, we consider graphs that are
  - *undirected*, i.e., the connections do not have an orientation, and
  - *simple*, i.e., we allow only one connection between each pair of items and no connections from an item to itself.

A Tour of a Graph Made from Map Data



Graph Terminology and Notation

- In an undirected graph, the "items" are usually called *vertices* (sometimes also called *nodes*).
- The set of vertices is denoted V and the vertices are indexed from 0 to $n-1$, where $n = |V|$.
- The connections between the vertices are unordered pairs called *edges*.
- The set of edges is denoted E and $m = |E| \leq n(n-1)/2$.
- An undirected graph $G = (V, E)$ is then composed of a set of vertices V and a set of edges $E \subseteq V \times V$.
- If $e = \{i, j\} \in E$, then
  - i and j are called the *endpoints* of e,
  - e is said to be *incident* to i and j, and
  - i and j are said to be *adjacent* vertices.

More Terminology

- Let $G = (V, E)$ be an undirected graph.
- A *subgraph* of G is a graph composed of an edge set $E' \subseteq E$ along with all incident vertices.
- A subset V' of V, along with all incident edges is called an *induced subgraph*.
- A *path* in G is a sequence of vertices such that each vertex is adjacent to the vertex preceding it in the sequence.
- A path is *simple* if no vertex occurs more than once in the sequence.
- A *cycle* is a path that is simple except that the first and last vertices are the same.
- A *tour* is a cycle that includes all the vertices.

Directed Graphs

- A directed graph $G = (N, A)$ consists of a set of nodes N and a set of arcs A whose elements are ordered pairs of nodes.
- In a directed graph, ordering is important. Consider an arc (i, j).
  - Node i is the *tail* and node j is the *head*.
  - We say (i, j) is incident to nodes i and j.
- Sometimes it will be convenient to refer to the number of nodes $(|N|)$ as n and the number of arcs $(|A|)$ as m.
- How do we change the definition to define an *undirected* graph?

Networks and Subgraphs

- A network is a (directed) graph whose nodes and/or arcs have associated numerical values.
  - costs
  - capacities
  - supplies or demands
- The graph $G' = (N', A')$ is a subgraph of $G = (N, A)$ if $N' \subseteq N$ and $A' \subseteq A$.
  - We say that $G' = (N', A')$ is the subgraph *induced* by N' if A' contains each arc of A with both endpoints in N', i.e., $A' = A \cap (N' \times N')$.
  - In the example graph, if $N' = \{1, 2, 3, 5\}$, what is the subgraph induced by N'?

Degree in Directed Graphs

- The degree of a node is the number of arcs to which it is incident.
  - The indegree of node i, $I(i)$, is the number of incoming arcs.
  - The outdegree of node i, $O(i)$, is the number of outgoing arcs.
  - $\sum_{i \in N} I(i) = \sum_{i \in N} O(i) = ?$

Node	Degree	InDegree	OutDegree
1	2	0	2
2	3	1	2
3	4	2	2
4	3	2	1
5	2	2	0

Directed Paths

- A path in a directed graph $G = (N, A)$ is a path in the *underlying undirected graph* (the graph obtained by ignoring the orientations of the arcs).
- A *directed* path requires the correct orientation of the arcs.

Directed Cycles

- As with paths, a cycle is cycle in the underlying undirected graph.
- In the example graph, $1 - 2 - 3 - 1$ is a cycle.
- A directed cycle is a cycle obeying the orientation of the arcs.
- The example graph has no directed cycle. What do we call such graphs?

Other Concepts

- What does it mean for a graph to be *connected*?
- What about *strongly connected*?
- What is a *cut*?
- What is an *s-t cut*?

Trees

- A *tree* is a connected, undirected graph that contains no cycle.
- Elementary properties of trees
  - A tree on n nodes has exactly $n - 1$ arcs.
  - A tree has at least 2 leaf nodes.
  - Every pair of nodes is connected by a unique path.
- What is a *subtree*?
- What is a *forest*?
- What are the directed analogues?

Application Example (Ahuja et al., 1.7)

- Have you ever thought about how word processing programs decide where to break lines?
- In TeX, an optimization procedure is used to decompose the words of a paragraph into lines.
- The objective of the optimization problem is to maximize the attractiveness of the paragraph.
- Suppose that a paragraph consists of n words and each word is assigned a sequence number from 1 to n. Let c_{ij} denote the attractiveness of a line if it begins with the word i and ends with the word $j-1$.

Shortest Path Problem

- Consider a directed network $G = (N, A)$.
  - Each arc $(i, j) \in A$ has an associated cost/length c_{ij}.
  - There is one distinguished node s called the *source node*.
- Define the *length* of a directed path as the sum of the lengths of the arcs in the path.
- The objective of the shortest path problem is to determine for every node $i \in N \setminus \{s\}$ a shortest length directed path from s to i.

Paragraph Problem as a Network Flow Problem

- Given the values of the c_{ij}'s, formulate a shortest path problem to decompose the paragraph into lines in such a way as to maximize the total attractiveness.

Reformulation and Modeling

- Given the values of the c_{ij}'s, construct an appropriate network and formulate an optimization problem on that network to decompose the paragraph into lines in such a way as to maximize the total attractiveness.
- General procedure for reformulation
  - Decide what nodes and arcs represent conceptually and what other network data is needed.
  - Decide what type of network problem the original problem represents.
  - Construct the network from the problem input data.
  - Solve the network problem.
  - Translate the solution to the network problem back into a solution to the original problem.
- What is the model and problem in this case?