Complex Networks

Teachers: D. Garlaschelli, F. den Hollander, A. Plaat.
Written examination: Thursday, 29 January 2015, 10:00–13:00.

Open book exam: the lecture notes may be consulted, but no other material.

Answer each question on a separate sheet. Put your name, student number and the number of the question you are answering on each and every sheet. Provide full explanations with each of the answers.

Each question is weighted by a number of points, as indicated. The total number of points is 100. The final grade will be calculated as a weighted average: 40% for homework assignments and 60% for this exam.

Success!

1. [**5 points**] Explain what it means that many real-world networks are small worlds and have power-law degree distributions.
2. [**7 points**]
   - 2a. Give the definition of the Erdős-Rényi graph $\mathrm{ER}_n(p)$ with n vertices and parameter $p \in (0,1)$.
   - 2b. Does $\mathrm{ER}_n(\lambda/n)$, $\lambda \in (0,\infty)$, have a percolation transition for $n \to \infty$?
   - 2c. What is $\lim_{n\to\infty} f_{n,\lambda/n} = f_\lambda$ when $f_{n,\lambda/n}$ is the degree distribution of $\mathrm{ER}_n(\lambda/n)$?
3. [**8 points**]
   - 3a. Give the definition of the configuration model $\mathrm{CM}_n(\vec{k})$ with n vertices and degree sequence $\vec{k} = (k_1, \ldots, k_n)$ satisfying $k_1 + \cdots + k_n = \text{even}$.
   - 3b. Is the resulting random graph always simple or not?
   - 3c. Is there a simple graph with degree sequence $\vec{k} = (2,8,2,2)$, respectively, $\vec{k} = (2,2,6,2)$?
4. [**16 points**]
   - 4a. Draw a simple binary undirected graph with $N = 5$ nodes consistent with the degree sequence $\vec{k} = (1,2,1,2,4)$, and write down its adjacency matrix $\mathbf{A}$ (with entries $\{a_{ij}\}$).
   - 4b. On this graph, calculate the number of links (L), the average nearest-neighbour degree of each node i (k_i^{nn}), and the clustering coefficient of each node i (C_i). (For nodes with degree 1, conventionally set the

clustering coefficient to 0.)

4c. Assume that the network G you have drawn has been generated according to the Erdős-Rényi random graph model (with $N = 5$). Determine the value p^* of the connection probability that maximizes the likelihood of G given the model.

4d. Write down an approximate expression for $\langle k_i^{\mathrm{nn}} \rangle \equiv E[k_i^{\mathrm{nn}}]$ under the Erdős-Rényi model and, using p^*, calculate its value for each node i in G.

4e. Write down an approximate expression for $\langle C_i \rangle \equiv E[C_i]$ under the Erdős-Rényi model and, using p^*, calculate its value for each node i in G.

5. **[9 points]** Dijkstra's single-source shortest-path algorithm can be summarized as follows: *Pick the unvisited vertex with the lowest distance, calculate the distance through it to each unvisited neighbor, and update the neighbors distance if smaller. Mark visited (set to red) when done with the neighbors.* (The lecture notes provide the code for Dijkstra's algorithm on page 85.)

Of a graph the following is given:

	node	*distance*	*node*
1	a	7	b
2	b	10	c
3	a	9	c
4	a	14	d
5	c	2	d
6	d	9	e
7	e	6	f
8	b	15	f
9	c	11	f

5a. Draw the graph. Apply the steps of Dijkstra's algorithm to this graph. Write down the individual steps: What is the active node and what is the path length? Give the node sequence and the value of the shortest path between node a and node f in the graph.

5b. Assume that line 4 is changed to read as follows:

	node	*distance*	*node*
4	a	18	d

The graph is changed to reflect this new distance. What is the new shortest path (node sequence and path length), and how many steps of the algorithm would you need to arrive at this answer?

5c. Assume that line 4 is changed to read as follows:

	node	*distance*	*node*
4	a	-18	d

(Note the negative distance.) The graph is changed to reflect this new distance. What is the new shortest path (node sequence and path length), and how many steps of the algorithm would you need to arrive at this answer? Why?

6. **[3 points]**
Dijkstra computes the single-source shortest-path length, Girvan-Newman is a community detection algorithm, while Pagerank computes a centrality measure.

   6a. Can you use Dijkstra as part of Girvan-Newman? Why / Why not?

   6b. Can you use Girvan-Newman as part of Pagerank? Why / Why not?

7. **[10 points]** Consider the configuration model $\mathrm{CM}_n(\vec{K})$ with n vertices and degree sequence $\vec{K} = (K_1, \ldots, K_n)$ whose components are i.i.d. random variables with probability distribution function f satisfying $f(0) = 0$ (conditioned to satisfy $K_1 + \cdots + K_N =$ even). A "hacker" removes vertex i with probability $1 - \frac{1}{K_i}$ (together with all the edges incident to it). Under what condition on f does the "mutilated" random graph percolate for $n \to \infty$?

8. **[5 points]**

   8a. Give the definition of the contact process on $\mathbb{Z}$ with parameter $\lambda \in (0, \infty)$.

   8b. What is the meaning of the statement "$\lambda_1 \approx 1.6494$"?

   8c. What is the critical threshold for the contact process on $\Lambda_N = \mathbb{Z} \cap [0, N)$, $N \in \mathbb{N}$, with periodic boundary conditions?

9. **[16 points]**

   9a. Draw a simple undirected network consistent with the degree sequence $\vec{k} = (2, 2, 1, 3)$. Let G denote your graph. Assume that G has been generated according to the (simplified version of the) configuration model (CM) specified by the connection probabilities $p_{ij} = k_i k_j / 2L$ with $i \neq j$ and $p_{ii} = 0\ \forall i$. Write down the full 4×4 probability matrix $\mathbf{P}$ (with entries $\{p_{ij}\}$) explicitly.

   9b. Identify the realized edge(s) in G that have the maximum probability of occurrence under the CM model.

   9c. Identify the missing edge(s) in G that have the minimum probability of occurrence under the CM model.

   9d. Write down an approximate expression for $\langle k_i^{\mathrm{nn}} \rangle \equiv E[k_i^{\mathrm{nn}}]$ under the CM model and calculate its value for each node i in G.

   9e. Write down an approximate expression for $\langle C_i \rangle \equiv E[C_i]$ under the CM model and calculate its value for each node i in G.

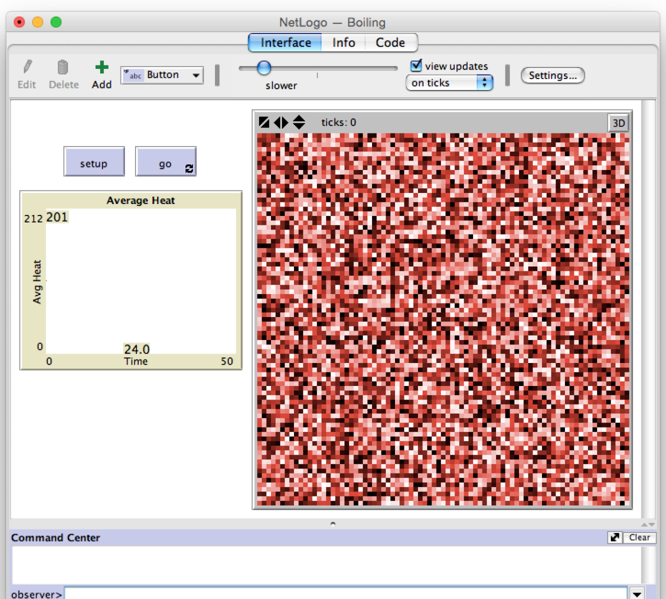
10. [**11 points**] The general public complains about inner city traffic levels. Pollution, noise, parking problems and accidents create a sense of unease among inhabitants, shopping public, shopkeepers and the general public. The city council wants to act. You are asked to draw a System Dynamics diagram for this situation (page 116 of the lecture notes contains a list of guidelines). Use the following entities:

    – local tax level
    – parking facilities
    – popularity of the council
    – city budget
    – attractiveness of city
    – traffic level
    – traffic complaints

    Provide a brief explanation for your diagram. (Points are awarded for correctness, clarity, and the identification of feedback loops.)

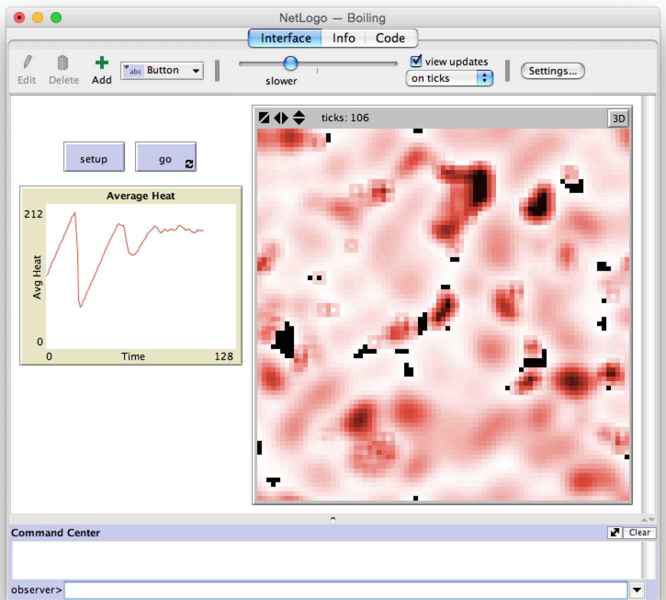
11. [**10 points**]

    11a. The Netlogo example application Boiling has the following initial state (after hitting Setup):

    Explain what we see: What are the light and dark pixels?

    11b. After hitting Go and waiting for a few steps, a sudden phase change occurs, and another, and another, until after some time the system stabilizes in a dynamic situation, and the following screenshot is taken:



Explain what we see: What are the large light and dark areas? What do the sharp lines mean?

11c. Describe the dynamic process. *Hint*: Explain how we can arrive at a state where high-level structure emerged while we started from a random state.