Progetto dei Corsi di

Complex Systems and Network Science

Corso di Laurea Magistrale in Informatica

Analyzing Complexity

Corso di Laurea Magistrale in Relazioni Internazionali

A.A. 2017-2018

Dipartimento di Informatica — Scienza e Ingegneria
Università di Bologna

Ozalp Babaoglu   Alessio Netti

April 26, 2018

1 General Information about the Project

As part of your course requirement, you are to complete the project described below, which must be carried out individually. Submission of the project for evaluation must be done via email to the address: `alessio.netti@unibo.it`.

The deadline for submission is **23:59:59 hours on July 1, 2018**. The email must be sent from your University address (`name.surname@studio.unibo.it`) and the subject field of the email must be *CSNS Project 2018*.
You will receive a confirmation message within a few days of your submission. The email should contain an archive (in .zip or .tar.gz format) containing the following:

1. The source code that was developed (either in NetLogo or PeerSim);
2. A short paper, in PDF format, describing the model that was implemented, the experiments that were carried out using it, and a discussion explaining the results that were obtained.

Your full name, email address and student ID number (matricola) must be included in all of the source files, in the paper, and in the submission email that you send. The source code should be well documented and formatted, following good programming practices. The paper can be written in Italian or in English, and should be structured like a technical paper, thus containing a title, abstract and bibliography. It is strongly suggested that you limit the length to 16 pages and that you follow the Springer format for *Lecture Notes in Computer Science* (LNCS). Templates are available for both Word[1] and LaTeX[2]. You can use any text processing system you prefer (even though LaTeX is suggested) to write the paper as long as you submit the result as a PDF file.

The project must be done *individually*: no sharing of papers or source code is permitted. You are, of course, encouraged to discuss issues and solutions with fellow students or with the instructors.

2 Grading Policy

For your project to be satisfactory, it must satisfy the following requirements:

- The project must implement the specifications that follow. You are allowed (and encouraged) to apply modifications and extensions to the project, but they must be proposed to the instructors beforehand and approved by them.
- The model's implementation, and all of the related simulations and experiments, must be carried out using either the NetLogo or PeerSim software systems. If PeerSim is chosen, the cycle-driven simulation engine should

[1] Link to .doc template
[2] Link to .tex template

be used, and the simulator must be configurable by means of the standard PeerSim configuration file.

- Your paper must thoroughly describe the model that was implemented and justify all significant design decisions and extensions that were applied to it. You should also discuss the expected behavior of the model, by making previsions. Most importantly, you have to explain the experiments that you performed in terms of methodology and the results that you obtained. Significant implementation details can be inserted, if important in the context of the model, but should otherwise be kept as comments in the code itself.

You are encouraged to focus on a simple model and to apply extensions to it only if you completely understand the behavior of the base model. This can be achieved by working in modular fashion, thus incrementally (and carefully) adding new features, enriching your model. Ending up with a complex, unpredictable and difficult to understand model is very easy. On the contrary, you should prove through your experiments that you fully understand the behavior of your model and that you can interpret the results you obtained, and are able to relate them with real-world phenomena. Finally, you should try to find tipping points or interesting equilibrium states in your model.

If you are interested in these topics (e.g. you want to build better models or study other systems of this kind), do not hesitate to contact us when looking for a thesis topic.

3 Description of the Project

The purpose of this year's project is to create a model where *culture dissemination* and *self-segregation* within a given population can be observed. The basic hypothesis is that individuals usually prefer to interact and be close to other individuals that are culturally similar to them (*homophily*). The emerging social behaviors from this simple individual tendency are complex, and largely reflect the structure of real-life societies.

The tools for our framework are constituted by two famous models: Axelrod's model for culture dissemination [?], and Schelling's self-segregation model [?]. Without going into formal details, Axelrod's model describes the diffusion of cultural traits in a given population, represented as a *network*, in which each individual is characterized by a *cultural code*. Neighboring individuals that share similar cultural codes tend to imitate each other, becoming more and more similar as time passes. Depending on the parameters of the model and the features of the network, the population can either converge to a few, or even a single cultural code (cultural *homogeneity*) or retain its diversity by forming different cultural clusters (cultural *mixity*).

Schelling's model tries instead to model how different cultural groups tend to isolate from each other, for example in large cities, thus forming large independent clusters and resulting in self-segregation. From a certain point of view, this model is similar to Axelrod's model, but it includes the notion of *movement*. In fact, Schelling's model uses a checkerboard grid in which each cell can be either empty or populated by an individual. Individuals require a fraction of their neighbors to be similar to themselves: if this fraction falls below a given *tolerance* threshold, the individual will move to a new random location in an attempt to find greater happiness. Depending on the given tolerance threshold, more or less segregated populations will appear on the checkerboard. Even though Schelling's model was not originally designed for networks, it can be easily adapted for them.

3.1 Formal Specification

As mentioned earlier, our model is a combination of Axelrod's and Schelling's models [**?**]. We model a small community through a *network* of N nodes: nodes represent *sites* (houses, blocks of flats etc.), while edges represent *neighbor* relations. Each site can either be *empty* with probability e, or it can be *inhabited* by a single individual with probability $(1-e)$. Therefore, the average number of empty sites is expected to be Ne, while that of individuals is $N(1-e)$.

Each individual i that is associated with one of the non-empty sites in the network is characterized by a *cultural code* $\sigma(i)$, which is a vector of length F where each element corresponds to a different *cultural trait*. Cultural traits are represented as integer values in the interval $[0, q-1]$. When a new individual is generated, each of the F cultural traits in their cultural code $\sigma(i)$ is picked at random uniformly within this interval. The F parameter defines the *complexity* of the cultural code, while the q parameter defines how many different cultures are present in the population: as such, higher values of q will result in more mixed, heterogeneous starting populations, while lower values will result in culturally more uniform ones.

Once the network and the population have been generated, the model can be executed. At every simulation step, for each individual i a random individual j occupying one of the neighboring sites of i is chosen. At this point, we compute the *cultural overlap* $\omega_{i,j}$ between i and j, which is defined as:

$$\omega_{i,j} = \frac{1}{F}\sum_{k=1}^{F}\delta(\sigma_k(i), \sigma_k(j)) \tag{1}$$

where $\sigma_k(i)$ is the kth cultural trait of i and δ represents the Kronecker's delta function that is 1 if its two arguments are equal, and 0 otherwise. In other words, the cultural overlap $\omega_{i,j}$ corresponds to the proportion of cultural traits that are the same for i and j, and it is a measure of how *culturally similar* they are. As the next step, i randomly chooses a trait k out of the F cultural traits of j, and copies it with probability $\omega_{i,j}$. Mathematically, the copying operation is accomplished as:

$$\sigma_k(i) = \sigma_k(j) \tag{2}$$

In practical terms, Equation **??** says that individual i imitates a culture trait of individual j, and is more likely to do so, the more culturally similar they are. This is a reasonable assumption since individuals are usually more open and well-disposed towards other individuals they feel similar to (once again, homophily). If, on the other hand, imitation does not occur (with probability $(1 - \omega_{i,j})$), individual i will compute the *average cultural overlap* $\overline{\omega_i}$ with its entire neighborhood, which is defined as the average of the $\omega_{i,j}$ values for each individual j that is a neighbor of i. This second index defines the cultural similarity of an individual compared to their neighborhood as a whole and once again, empty sites in the neighborhood of i are not considered. At this point, if $\overline{\omega_i}$ is lower than a fixed threshold T, individual i will decide to move to a random new site that is empty, leaving its previous site empty. The parameter T represents a *tolerance* threshold of individuals towards other cultures, and is used to model stricter or more adaptable societies.

The execution stops when, at a given step, no interactions (cultural imitations or moves on the network) are performed by individuals. Also note that if an individual is surrounded by only empty sites at a given step, and thus cannot perform any of the actions described above, it will move directly to a new site.

3.2 Discussion

You are required to implement the model described above, and to perform experiments with it. Each experiment should employ a particular configuration of the model's parameters: motivate the choices of parameters, and try to relate them with real-world societies. Also, be sure to repeat the same experiment multiple times, to ensure that your outcomes are not due statistical anomalies. You should make previsions about the expected behavior for a given configuration, and then explain why the results you obtained confirm or contradict your previsions. Highlight emergent social behaviors and make assertions that can be associated with real-world scenarios and phenomena. The following are examples of assertions that you could test as being true or false:

- Small, scarcely-connected neighborhoods tend to be highly segregated and uniform from a cultural point of view;
- An hyper-connected society leads to cultural homogeneity and to the disappearance of cultural minorities.

In order to achieve the above, you should experiment creatively with the model parameters including F and q that determine the cultural diversity of the starting population, parameter e that sets the probability for empty sites in the network, and T, the tolerance threshold that triggers agents to move to a new site. Most importantly, you should experiment with network *topologies* generated with different methods and characterized by different features: remember that this is a network-based model in which the spatial placement of sites has primary importance for its behavior.

References

[1] Axelrod, Robert. "The dissemination of culture: A model with local convergence and global polarization." Journal of conflict resolution 41.2 (1997): 203-226.

[2] Gracia-Lázaro, C., et al. "Residential segregation and cultural dissemination: An Axelrod-Schelling model." Physical Review E 80.4 (2009): 046123.

[3] Schelling, Thomas C. "Dynamic models of segregation." Journal of mathematical sociology 1.2 (1971): 143-186.