Migration and Social Networks - An Explanatory Multi-evolutionary Agent-Based Model

Hugo S. Barbosa Filho
Polytechnic School of Pernambuco
University of Pernambuco
Email: hsbf@ecomp.poli.br

Fernando B. de Lima Neto
Polytechnic School of Pernambuco
University of Pernambuco
Email: fbln@ecomp.poli.br

Wilson Fusco
Economics and Population Research Center
Joaquim Nabuco Foundation
Email: wilson.fusco@fundaj.gov.br

Abstract—Sociodemographic studies on human migration phenomena are mostly based on surveys and censuses, which significantly increases the research costs. This scenario becomes even worse when the study involves migration and social networks, which often lacks on representative data and consensually accepted concepts by demographers and sociologists. In this paper we propose a new multi-evolutionary agent model dedicated to social simulations, mainly for those problems where higher order dynamic behaviors (e.g. secondary emergent phenomena) are important to the investigated phenomenon. Its usefulness lies on its multilevel evolutionary adaptability which enables it to capture multiple parallel phenomena. To verify our hypothesis we applied the model to Brazilian internal human migration phenomenon and to the influence of social networks on migration flows. This is followed by a comparative analysis against simulations carried out based on a non-evolutionary cognitive agent model. Results show that the proposed model was able to rise secondary migration-related phenomena such as countermigration flows . Experiments with the cognitive agent also produced the emergence of migration flows but no secondary phenomenon was observed which was the case with our approach. Furthermore, results also pointed out a significant influence exerted by information exchange inside social networks on migration flows.

Index Terms—Agent-based modeling, social simulations, artificial societies, migration, social networks

I. INTRODUCTION

Quantitative demographic studies can easily become very costly and difficult to be performed, mostly those which involves migration [1]. To reduce such difficulties, mathematical models have been used as a tool that can be combined to the statistical methods commonly used. Even though such approach may obtain satisfactory results, these outcomes are commonly implausible as a representation of a social system [2] .

As an alternative to analytical models, the application of agent-based models have shown significant growth as a tool for the construction of artificial societies. However, several phenomena have social behaviors of highest orders, emerging from the synergistic interactions among existing social phenomena. Actually, most of the social phenomena with scientific interest are of complex nature and have emergent behaviors of some higher orders. For example, urban violence is a phenomenon that can emerge from socio-economic inequalities (which in turn may arise from ineffective educational systems) or religious fanaticism.

In most models of agents commonly found in the literature, the representation of phenomena is based on rules (e.g. social, environmental and cultural) in a rather low level manner, resulting in the emergence of a particular macro phenomenon [3]. This may or may not produce emergence as one would obtain in real social interaction.

In this paper, we argue that a more elaborate manner to incorporate realistic behavior, based on social interaction (i.e. synergy) of large population of agents, would favor plausibility. Moreover, we argue that highest other behaviors should be indirectly caused by such interaction of agents.

II. BACKGROUND

In this section, we first provide a brief background on demography and human migration-related phenomena. Second, we present the most recent advances in migration modeling and how this might also be quite relevant to this area.

A. Migration

In demography, the scientific studies are mostly related to human population and its dynamics. It encompasses features such as structures, sizes, distributions and behaviors or phenomena which can change those aspects over space or time.

Actually, population in a given region can change quantitatively as consequence of the (1) births, (2) death and (3) migrations. As the first and second are more stable over the time, making it easier to to perform predictions, the latter, conversely, is quite unpredictable phenomenon hence the subject of many investigations in demography [1].

Births and death are natural and mainly ruled by biological aspects. However, migration is more affected by socio-economic factors and human behavioral subjectiveness. Although births and death rates are also influenced by socio-economic factors (such as education, poverty or diseases) and migration can be biologically motivated (like in natural disasters), in general, migration flows are more unstable and dynamic over time.

Ravenstein, in 1885, published the first work which proposes a well empirically grounded description about the general aspects of the human migration [4] phenomenon. In this work, Ravenstein stated that international migration might be described according to 11 laws, which were latter named as "Ravenstein's Laws of Migration"). As in this work we are not

focusing only in international migration but in general aspects of the phenomena, four of the most relevant are listed below:

- every migration flow generates a return or countermigration;
- the majority of migrants move a short distance;
- migrants who move longer distances tend to choose big-city destinations;
- families are less likely to make international moves than young adults.

After Revenstein, several quantitative models of migration flows and the variables that affect those flows were proposed. More on classical models of migration can be seen in [5].

In last decades other social factors are also being investigated as related to migration flows such as social networks [6], [7]. However, scientific researches carried out in order to establish the role played by social networks on migration flows are mostly based on surveys, census and official immigration data, which has problems and limitations [1], [8]. Therefore, the proposed method aims at producing more accurate social simulations as our approach is able to incorporate details that could not be easily included in the traditional social simulations.

B. Analytical vs. Agent-Based Migration Models

In order to reduce the uncertainties and unpredictability on migration-related studies, several migration models were proposed and can be found in literature. One of the most known was proposed in 1969 by Todaro [9]. Since that seminal work, migration was investigated from a wide range of perspectives and from different approaches. From classical physical approaches - where migration is mostly related to distances between origin and destination- to neoclassical economics models where migration emerges from individuals search for more satisfactory economic conditions like higher wages or better job opportunities seems to be more realistic - refer to figure 1
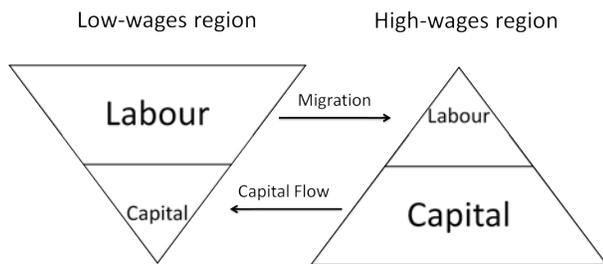

Fig. 1. Neo-classical economic model of migration.

Several mathematical and computational tools have been used in order to model and simulate social phenomena. However, Agent Based Social Simulation (ABSS) are regarded as the state-of-the-art in social simulations today [10]. Works such as Game Theory [11] afford important insights on the understanding of human collective behavior, but like most mathematical models, it assumes a perfect rationality and homogeneous population.

Simulations of social phenomena are referred as agent-based when individuals are represented as artificial agents in a computer architecture. The major motivation to use agent-based models is the possibility of modeling and controlling the model in different granularity levels, from environmental spatial characteristics to behavioral and cognitive individual aspects [12]. This approach enables the social scientist to produce highly heterogeneous and sophisticated models of artificial societies. Axtel and Epstein have pointed out compelling arguments supporting the use of ABSS instead of the analytical models [13].

In ABSS agents may be defined in function of their social abilities, states and rules. All of these elements are fundamental to the simulation social dynamics since they also influence the behavior of individual agents [10]. Thus, recognizable phenomena or patterns of behavior may emerge. Observing the emergence of such phenomena is an important objective of the social scientist when performing such simulations.

Agent-based models were already applied in human migration experiments but in other contexts or approaches such as rural-urban migration [14] or housing search motivated migration [15].

III. The Agent Model Architecture

The cognitive agent model proposed in this paper is characterized by the existence of multiple evolutionary levels. Each of these levels represents an evolutionary process - with its own dynamics and mechanisms - in which the decision-making processes occur. In general, the evolutionary levels can be described as follows:

- Genetic - represent the innate characteristics of the agent whose evolutionary process will occur over the generations.
- Social - individual perception and interpretation of collective values which evolution occurs throughout the life cycle of the individual.
- Phenotypic - expression of genetic traits and social values of an individual whose development will occur either within the same generation and across generations.

As a matter of fact, we propose that the decision mechanism of the agent is to be based on the response of the phenotypic layer as a result of the combined influences from social and genetic layer. A diagram representing the hierarchical structure of evolutionary layers - with examples of possible internal elements - is in Figure 2.
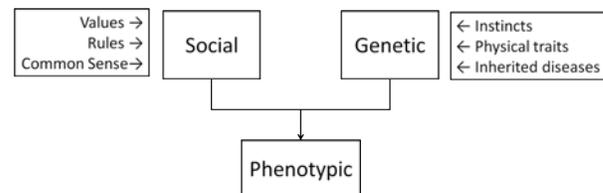

Fig. 2. Multi-layer evolutionary model of an agent

Each layer has its own evolutionary processes that is guided by the behavior of the agent according to the actual problem to

be studied. The phenotypic and social evolutionary processes will be influenced by the outcome of the agent's actions throughout its existence as genetic evolution will occur from the crossings between the players over the generations.

The cognitive module determines the actions to be performed by the agent, under the influence of phenotypic module (responsible for determining the physical and behavioral expression of the agent) and their perceptions about the environment and of itself.

From their views on the environment, the evaluation of its performance given an objective function (depending on the problem to be tackled) and the history of his previous actions, the agent updates its phenotypic elements. This update will be responsible for phenotypic evolution of the module and will be responsible for guiding agents in their search for striking the appropriate balance between social and individual motivations for decisions.

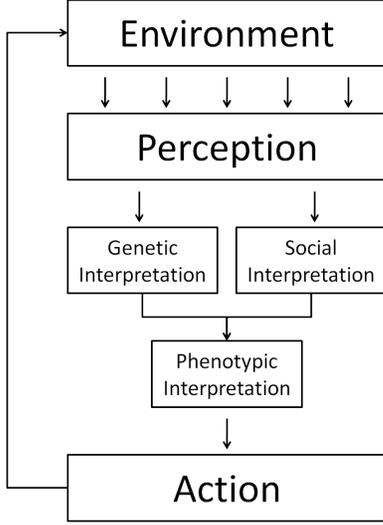A diagrammatic representation of the agent decision making process is in Figure 3.

Fig. 3. High-level representation of the decision-making mechanism performed by a multi-evolutionary agent.

A. Implementation

In the proposed model, each layer λ_i is computationally represented as a vector of size n where $i = 3$ and $\lambda_i = \left(\lambda_{i_1}, \ldots, \lambda_{i_j}, \ldots, \lambda_{i_n}\right)$ with $\lambda_{i_j} \in \mathbb{R}$. The parameter n is dependent of the problem since each element λ_{i_n} represents the most basic evolutionary unit. As in other evolutionary computation techniques, each layer here performs at least the crossing over operation; nonetheless, other operators (e.g. mutation or elitism) may be implemented. Actually, only the layers λ_1 e λ_2 (i.e. Genetic and Social) performs the crossing-over operation. In the outermost layer (i.e. Phenotypic) evolutionary processes occur along with environmental responses to all performed action. More details on each part of the model are described below.

1) Perception: At a given instant, each agent has a perception about the environment. This perception is a combination of several elements present in the neighborhood of the agent. Each of these elements can have characteristics that are treated by one of the initial layers of perception (i.e. genetic and social). Perception may be seen as multiplexed signal S comprising two distinct signals given by (1), where S_G is the genetic component and S_S is the social component.

$$S = (S_G, S_S). \tag{1}$$

To illustrate, suppose an environment with limited food where an agent is faced with a food that is owned by another agent. That perception will be decomposed into its genetic and social components. The survival instinct (genetic) of the agent determines that it consumes resources although the rules impose a penalty if he does (social). What action will be taken is determined by phenotypic module.

2) Interpretation: The phenotypic layer can be seen as an array size of N_P where $N_P = N_G + N_S$ (where N_G and N_S are the number of *genes* in genetic and social layers). The output of phenotypic layer is given by

$$f_P(t) = [f_G(t), f_S(t)] \tag{2}$$

where

$$f_G(t) = [P_1(t) \cdot G_1, \ldots, P_i(t) \cdot G_i, \ldots, Pn(t) \cdot G_N], \tag{3}$$

$$f_S(t) = [P_{n+1}(t) \cdot S_1, \ldots, P_j(t) \cdot S_j, \ldots, Pm(t) \cdot S_M]. \tag{4}$$

$$P_i(t) = [x_{P_1}(t), \ldots, x_{P_i}(t), \ldots, x_{P_n}(t)] \tag{5}$$

$$P_j(t) = [x_{P_{n+1}}(t), \ldots, x_{P_j}(t), \ldots, x_{P_m}(t)] \tag{6}$$

$$G_i = [x_{G_1}, \ldots, x_{G_i}, \ldots, x_{G_n}] \tag{7}$$

$$S_j = [x_{S_{n+1}}, \ldots, x_{S_j}, \ldots, x_{S_m}] \tag{8}$$

In (6), x_{P_j} is in the range $[-1, 1]$ and in (7) and (8) , $x_{G_i} \in \mathbb{R}$ and $x_{S_j} \in \mathbb{R}$. Positive values of x_{P_i} means a positive response (gain) to x_{G_i} while negative values means a penalty to x_{G_i}. Similarly, positive values of x_{P_j} means a positive response to x_{S_j} and the negatives, a penalty.

3) Action: Each action A in the model have an array $[A_g, A_s]$ where A_g represents the gain (or penalty) to the agent most basic needs while A_s represents the social response to the action A.

After defining the vector of gains, the layer (genetic or social) with the biggest gain will have a higher probability of being chosen.

Returning to the illustration made above, suppose that the phenotypic answer to a perception was $[y_g, y_s]$ with $y_g > y_s$. So the probability of the agent to ignore the social values (e.g. laws and moral rules) and perform the action which offers the highest gain to their most inner primitive needs (i.e. surviving, feeding) will be higher.

The description in this section is just the general architecture of the proposed model. Other mechanisms (e.g. social dynamics and knowledge representation) are more problem-related and will be described in following sections.

IV. Experiment

The experiments carried out in this paper were implemented in NetLogo 4.1 [16] and are mainly centered in the comparison between the proposed multi-evolutionary agent model against a cognitive agent model. The objective is to verify whether the proposed agent is more suitable to represent not only the phenomenon itself but also if higher orders or macro level phenomena can be observed.

The social phenomena investigated here is the emergence of internal migration flows between two very similar regions. To make it more clear, let's suppose that the simulation is about the migrations flows between two small urban areas in the same country and state, with identical culture, rules and similar economic aspects.

As pointed out in section II-A, migration is one of the most imponderable (and sometimes counterintuitive) social phenomenon investigated in demography. Actually, we are not mainly interested in those aspects which are widely accepted theoretically as relevant to migration flows (e.g. differences of labor market and expected incomes). Instead, we are interested in how synergistic phenomena can arise as a consequence of migrations and social networks phenomena. As previously commented, social networks plays a key role in migration-related phenomena such as return migration and some inter-national migration flows - even against economic pressures.

The artificial society created for this experiment may be described as follows:

- two regions in a country are distant from each other enough to not allow daily flows;
- each region has its own citizens, workplaces, houses and social places;
- citizens may work, interact and establish new social ties;
- according to socio-economic attributes, citizens' happiness will may affected.
- low happiness levels may trigger different actions;

More details on each model entity and its inner dynamics is given in following sections.

A. Environment

In this experiment, the *environment* is the top-level structure, working as a container for other elements. Figure 4 depicts a hierarchical representation of the environment and its inner structures. For simulation purposes, the environment has no micro-level dynamic and it is just the space within the other entities act. Indeed, macro-level behaviors emerges from the dynamics of the entities within each region.

The environment is formed by two distinct squared regions, namely *R1* and *R2* . Both regions have their own socio-economic attributes. However, as we are simulating two very similar regions, some of those parameters are identical or very similar.

Actually, both regions were calibrated against real data of Pernambuco state in Brazil obtained by Brazilian Demographic Census carried out in 2000 and the statistical analysis were
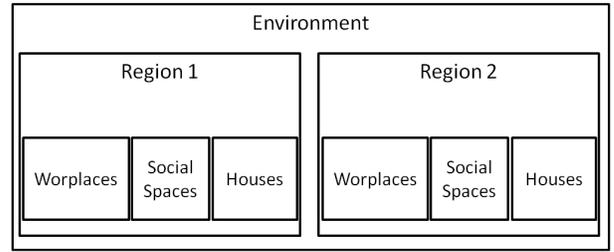

Fig. 4. Hierarchical representation of the environment

TABLE I
REGIONAL SIMULATION PARAMETERS

Parameter	Scenario dependent	R1	R2
Average wages	Yes	$(10, 15)$	$(10, 15)$
Monthly economic growth (%)	Yes	$(0, 0.1, 0.25)$	0
Initial unemployment rate	No	0.1	0.1

performed using SPSS Statistics 17[1]. Moreover, most likely probability distributions used to generate the random values (as well as their parameteres) for calibration purposes were estimated from the real data with EasyFit Professional[2].

In both regions, Citizens ages were generated by a gamma distribution

$$a_i \sim \Gamma(\alpha, \lambda), \tag{9}$$

where $\alpha = \frac{\mu^2}{\sigma^2}$, $\lambda = \frac{\sigma^2}{\mu}$ and μ and σ^2 stands for mean and variance respectively. Here, these values were $\alpha \approx 1.83$ and $\lambda \approx 0.068$

Moreover, education level, represented by the (e_i) was generated by the exponential distribution

$$s_i \sim Exp(\lambda) \tag{10}$$

with $\lambda \approx 3.82$.

Other parameters were defined according to the simulated scenario. For instance, all regional parameters are:

B. Citizens

Citizens are individuals that live in one region, may have a family, a job and a social life. Each iteration citizens will follow the routine depicted in the algorithm 1. Citizens have socio-economic attributes. Some of them generated based on real data from Pernambuco state provided by Brazilian Demographic Census - 2000 [17].

Citizens attributes are:

- **Age** defined according to (9);
- **Education** year of study defined according to (10);
- **Wages** determined by the Workplace proportionally to Education;
- **Knowledge** information about wages paid in the other region.
- **Relationships** list of citizens which is connected to.

Citizens may interact with other citizens and create new social ties (i.e. relationship). Through these ties, in certain

simulation scenarios, citizens may transmit information about its current wages to their acquaintances in other region.

Algorithm 1: Citizens general cycle

---

while *Stop condition(s) not true* **do**
  foreach *Citizens* **do**
    Go home;
    Interact;
    if $R_s \leq P_S$ **then**
      if *Citizen is employed* **then**
        Go to workplace;
      else
        Do nothing;
    else
      Go to one of social spaces in same region chosen randomly;
    Communicate with other known agents;
    Update happiness;
    Think;

- **Interact** - in this routine, the citizens may exchange their social information or reproduce according to the *crossover* rate.
- **Communicate** - the citizen will inform to friends the average wage paid in its job;

Algorithm 2: Citizen cognitive procedure

Set Avg as the average happiness of all friends;
if *Age is between* 14 *and* 70 **then**
  if $Happiness < Avg$ **then**
    if $Wages \geq Migration\ cost$ **then**
      Migrate;
      Look for a job;
    else
      Look for a new job;

One of the most important procedures in this experiment is the *happiness update procedure*. This procedure is based mainly on the neoclassical economics approach to migration decision mechanisms. This approach views the migration phenomenon as a simple sum of cost-benefit decisions undertaken to maximize expected income [1]. As seen before, this cost-benefit is not only about higher wages or more job opportunities but it is also take in to account the social relationships (quantity and quality). It means that someone with good relationship with family and friends is most unlike to migrate than another who does not have.

In 3, W_i is the actual wages, C_r is the cost of life in current region, E_{Wj} is the expected wages paid in the other region. S_G is the *Genetic* component and S_S is the social component. F_{S_i} and F_{S_l} is the sum of n highest relationship strengths

---

Algorithm 3: Happiness-update procedure

---

if *Age between* 14 *and* 70 **then**
  Set $F_e = [(W_i - C_r) - (E_{W_j} - C_r)] \times S_G$;
  Set $F_s = [(f_{S_l} - f_{s_e})] \times S_S$;
  Set $Happiness = F_e - F_s$
else
  Set $Happiness = F_s$;

---

with the citizen's friends in same region. Meanwhile F_{S_e} is the sum of n highest strengths with friends in the other region.

In other words, to calculate the *social force* agent will take into account its closest acquaintances in each region. How much this force will affect the decision is the result of the product between the social force and its social component. Similarly, the economic force is the difference between wages and expected wages in other region.

The evolutionary layers are updated as in other evolutionary algorithms. The genetic layer will remain the same during the whole agent life. As we are not interested in population growth, the number of citizens in the environment is fixed. The genetic evolution will occur when a family produces a new agent which will combine genetic elements from both parents. To keep the number of agents constant, always when a new citizen is generated, one of the oldest agents in the environment will die. One may think that this routine can affect the wages labor market and wages distribution but both citizens (the newborn and the old) are not in the age range of 14 and 70 and so they actually are not working.

The social component is updated several times during the social interactions following the cross-over rate. The social information is always transmitted from the happier agent to the other. In the following algorithm, $P_c \sim U(0, 1)$ and c_r is the crossover rate.

Algorithm 4: Social crossover

Given two citizens C_1 and C_2;
if $P_c \geq c_r$ **then**
  if C_1 *is happier than* C_2 **then**
    C_2 will copy part of C_1 social component proportional to c_r;
  else
    C_1 will copy part of C_2 social component proportional to c_r;

Each one has a initial population composed by 300 citizens. They also have social places, namely: (i) Houses (100), (ii) Social Public Spaces (20) and (iii) Workplaces (40), all with its own attributes and processes. Numbers in parentheses correspond to how many instances of the structure is in each region.

V. RESULTS

In this section we compare the results obtained in each simulation scenario focusing on to extract from the output data differences in macro-level behaviors in the models.

In figures 5 and 6, we can follow the progress of migrants population of each region over the time. Figure 6 depicts the migration progress with the multi-evolutionary agent model while 5 was produced by a regular cognitive model. Both scenarios are identical except by the agent model. R1 and R2 average wages are 15 and 10 respectively.
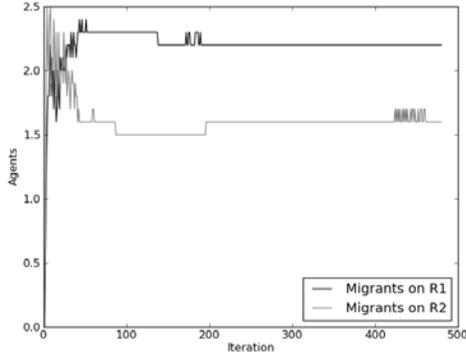
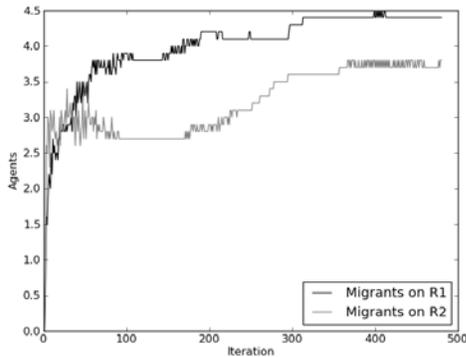Fig. 5. Migrants on each region over time with *cognitive agent* model.

Fig. 6. Migrants on each region over time with *multi-evolutionary agent* model and 1% crossover rate.

In both images, we can see that the models have shown the expected convergent behavior. However, the cognitive model, after the first iterations converged, while the multi-evolutionary model shown secondary migration flows (around iteration 150). Although the proposed model shown some sort of resonant behavior, none of them were able to clearly show migration flow without communication.

When the communication inside the social network is enabled, a new pattern of convergence is seen. Figure 7 and 8 depict two scenarios with communication inside the social networks. When compared figures 7 and 8 suggest that increasing communication rate reduces the time of convergence.
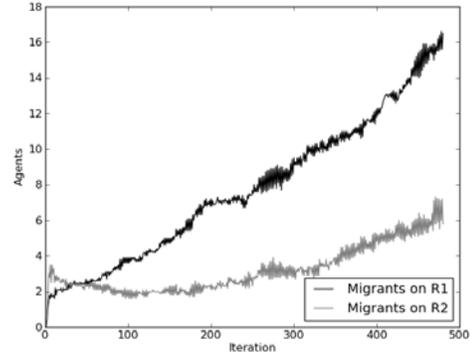
Fig. 7. Migrants with multi-evolutionary model, social network information flow enabled and 1% of communication probability.
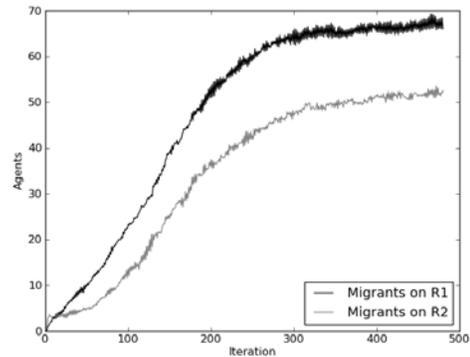
Fig. 8. Migrants with multi-evolutionary model, social network information flow enabled and 2.5% of communication probability

Both figures 7 and 8, when compared to figure 5 suggest that the communication inside social networks, combined to social transmission gave rise to a migration flow which was kept active during the simulation.
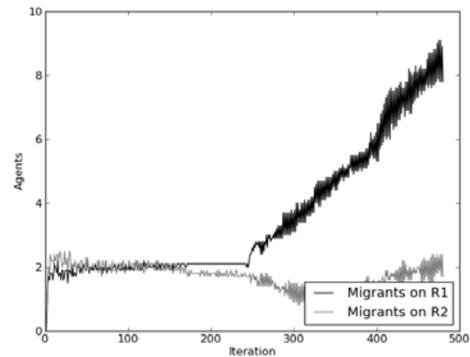
Fig. 9. Migrants with multi-evolutionary model and social network communication from iteration 240 and 1% of communication rate.
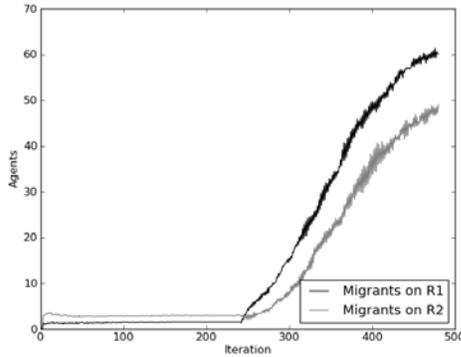
Fig. 10. Migrants with cognitive agent model and social network communication enabled from iteration 240 and 1% of communication rate.

VI. Conclusions

In this work, we presented a novel multi-evolutionary agent model for explanatory social simulations. Armed with the model, we investigated the effects of social networks on the migration flows. The proposed model was compared against a cognitive model in order to analyze the emergence of macroscopic behaviors with each model. The hypothesis is that the multi-evolutionary agent is more suited to simulate social phenomena with higher order macroscopic emergences (not only migration).

A. Discussion

Results have shown that the proposed model was able to represent migration dynamics with same expected behaviors pointed by other classic models in literature. However, the model was also able to capture second order migration-related macroscopic behaviors. Actually the model shown two distinct second order phenomena: (1) migration flows originated and maintained mainly by the inner social networks communications and (2) return migration.

Nowadays, the influence of social networks on migration flows has being investigated, mostly with empirically grounded theories. However, researches in demography lacks of a more objective tool in order to reduce the subjectiveness inherent to the human-related phenomena.

B. Future Works

Investigations such as cultural influences between regions or the role of common sense on other aspects related to migration flows may be carried out . The model may also be applied on other contexts such as urban violence, its origins and consequences.

References

[1] D. S. Massey, J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor, "An Evaluation of International Migration Theory: The North American Case," *Population and Development Review*, vol. 20, no. 4, p. 699, Dec. 1994.

[2] N. Gilbert, "Emergence in social simulation," pp. 144–156, 1995.

[3] R. Sun, "Individual action and collective function: From sociology to multi-agent learning," *Journal of Cognitive Systems Research*, vol. 2, pp. 1–3, 2001.

[4] E. G. Ravenstein, "The Laws of Migration," *Journal of the Statistical Society of London*, vol. 48, no. 2, pp. 167 – 235, 1885.

[5] J. Stillwell and P. Congdon, *Migration Models*. John Wiley & Sons Inc, 1993. [Online]. Available: http://books.google.com/books?id=J0mwPQAACAAJ&pgis=1

[6] W. Fusco, "Redes sociais na migração internacional: o caso de governador Valadares," 2002.

[7] D. Fazito, "Análise de redes sociais e migração: dois aspectos fundamentais do" retorno," *Rev. bras. Ci. Soc*, vol. 25, no. 72, pp. 89–176, Feb. 2010.

[8] P. Redfern, "An Alternative View of the 2001 Census and Future Census Taking," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 167, no. 2, pp. 209 – 248, 2004. [Online]. Available: http://www.jstor.org/stable/3559859

[9] M. Todaro, "A model of labor migration and urban unemployment in less developed countries," *The American Economic Review*, vol. 59, no. 1, pp. 138–148, 1969.

[10] N. Gilbert and K. G. Troitzsch, *Simulation for the Social Scientist*, 2nd ed. Open University Press, 2005.

[11] J. V. Neumann and O. Morgenstern, *Theory of games and economic behavior*. Wiley, 1964.

[12] N. Gilbert and R. Conte, *Artificial Societies: the computer simulation of social life*, 1995.

[13] J. Epstein and R. Axtell, "Growing artificial societies," p. 208, 1996.

[14] A. L. Espíndola, J. J. Silveira, and T. J. P. Penna, "A Harris-Todaro agent-based model to rural-urban migration," *Brazilian Journal of Physics*, vol. 36, no. 3a, Sep. 2006.

[15] S. Sun and S. M. Manson, "An Agent-based Model of Housing Search and Intraurban Migration in the Twin Cities of Minnesota," *Transport Economics*, 2010.

[16] U. Wilensky, "NetLogo," *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL*, 1999. [Online]. Available: http://ccl.northwestern.edu/netlogo/

[17] IBGE, "Censo Demográfico 2000: Características da População e dos Domicílios," *IBGE, Rio de Janeiro*, 2001.