Chapter 8
Trends in Social Science: The Impact of Computational and Simulative Models

LABSS-ISTC-CNR

Rosaria Conte, Mario Paolucci, Stefano Picascia and Federico Cecconi

8.1 The Survey: Method and Caveats

The simple method that we used is based on a survey conducted with the help of Google Scholar (Beel 2009; Chen 2010) comparing the growth (or decrease) of publications' citations for different disciplines in the social and computational sciences in the last decade with the relative growth of agent-based simulation and social simulation (Squazzoni 2008, 2010) within each discipline. This is quite a simple approach, indeed, which tells us nothing about progress in research quality or achieved breakthroughs but may help us compare the expansion pace of different research fields.

We performed search-engine queries using one "computational/simulation" tag and one discipline label. The queries were expressed in the form *tag* + *discipline label* + *year*. For each query, the raw citation number for every year ranging 2006 to 2011 was recorded. Subsequently, the variation between the years was calculated and the variation in the number of citations in each discipline/tag records compared.

We used eight tags:

- Agent-based simulation
- Monte carlo simulation
- Network analysis
- Neural network
- Numerical simulation

R. Conte (✉) · M. Paolucci
Institute of Cognitive Sciences and Technologies, CNR LABSS. Laboratory of Agent-Based Social Simulation, Via Palestro 32, 00185 Rome, Italy
e-mail: rosaria.conte@istc.cnr.it

M. Paolucci
e-mail: mario.paolucci@istc.cnr.it

F. Cecconi
Institute of Cognitive Science and Technology – CNR, Rome, Italy

F. Cecconi (ed.), *New Frontiers in the Study of Social Phenomena*,
DOI 10.1007/978-3-319-23938-5_8

- Reinforcement learning
- Game theory
- System dynamic

The first six tags are computational tags. *Agent-based simulation* identifies the class of simulation using artificial agents interacting with one another. *Monte carlo simulation* refers to the generic class of stochastic simulation models. *Network analysis* indicates the set of models that refers to complex network environment. *Neural network*, *numerical simulation,* and *reinforcement learning* refer to computational frameworks typically applied to the study of learning and adaptation in dynamic environments. *Game theory* and *system dynamics* are not *strictly* computational tags that we inserted in our dataset for the sake of comparison.

To effectuate our search, we used to Google Scholar (GS), mainly for convenience and ease of access. Since its introduction in 2005, GS has elicited mixed feelings in the scientific community. The first research papers reporting on its coverage found GS wanting (Neuhaus and McCulloch 2006). However, as was to be expected, Google improved and enlarged its coverage, and the current literature (Chen 2010) reports that sources that in 2005 had low coverage (ranging from 30 to 88 %) have now reached between 98 and 100 % coverage. In addition, GS is known to index sources, such as conference proceedings, working papers, and technical reports usually not included in other metrics (like ISI Web of Science).

It should be noted that our investigation suffers from some limitations, the most serious of which concerns the way we performed our queries, which can only detect literal matches but has no semantics. The text is searched as such in the articles' title and body, to ensure, for example, that a query for "Engineering social simulation" will not return articles *in* the engineering field, but articles generically referring *to* engineering. Nonetheless, GS's rough number of citations can be considered not only as a good proxy for the real number of occurrences of a particular keyword in scientific papers, but can be replaced with a more specific search in a yet-to-be-completed search engine.

Another caveat stems from the design of Google Scholar, which was not made with the purpose of retrieving large amounts of data, but to find specific papers—thus, we are stretching the tool's usage in a way that might cause some retrieval artifacts. To partially compensate, we ran the queries twice at two months' distance, finding substantial agreement.

8.2 Results

We analyzed seven disciplines—economics, history, philosophy, physics, psychology, sociology, and statistics. The criterion was to start a comparison between the traditional disciplines of nature (physics) and those of society (philosophy, economics, sociology, and psychology). Statistics and history were included to facilitate the understanding of the effect of some tags (see below). For each subject, we used

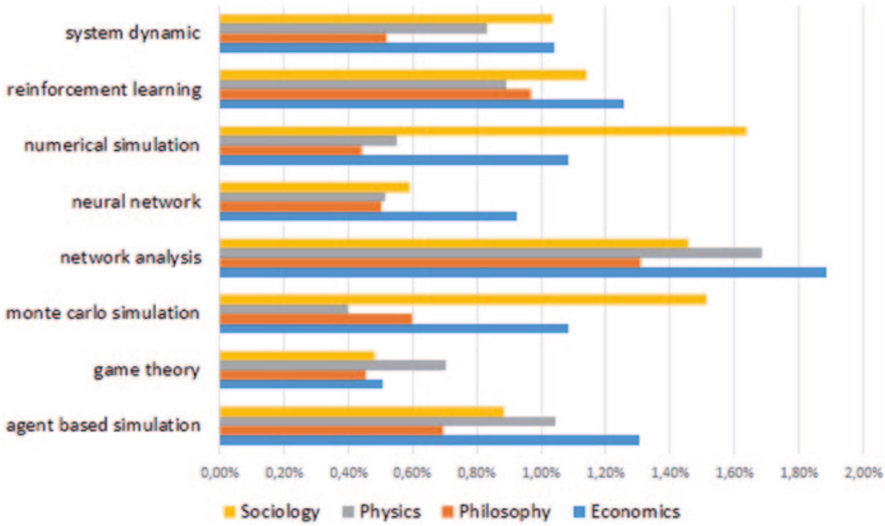

Fig. 8.1 The increase of citations during a 2-month period

eight different tags. We collected data at two different times separated by about two months. Figure 8.1 shows how the number of citations increased during the sampling period.

The increase shown in Fig. 8.1 was about 0.75 % on average, and this seems reasonable considering both the amount of scientific production and the indexing of further findings in these fields. Figure 8.2 shows the average (in percentage) of the differences of occurrences between the year 2005 and the listed year, up to 2011, for the various tags independent of disciplines.

Figures 8.3–8.6 show the trend of citations for economics, philosophy, psychology, and sociology (average on citations, Fig. 8.3), physics (Fig. 8.4), statistics (Fig. 8.5), and history (Fig. 8.6). All tags, in different measures, grow percentually in the number of citations. We can recognize a few overall trends: the first is the sharp increase in *network analysis* that dominates all the other ones through disciplines.

The second and third positions are occupied by the tags *agent-based simulation* and *reinforcement learning*, with *agent-based simulation* coming first in Figs. 8.4–8.6. The performance of the remaining tags depends on the discipline: In Fig. 8.3, numerical and Monte Carlo simulations come but game theory, which comes last also in history, but interestingly stays at the top of this last five tags group for physics. Monte Carlo simulation performs well in the social sciences (Fig. 8.3) but jumps between last and next-to-last position in all other ones. In statistics (Fig. 8.5), we found huge growth in *network analysis*. (note that we rescaled the y-axis of this graph, which has a maximum at 250 %, while the previous graphs had a maximum at 1808.4. %)

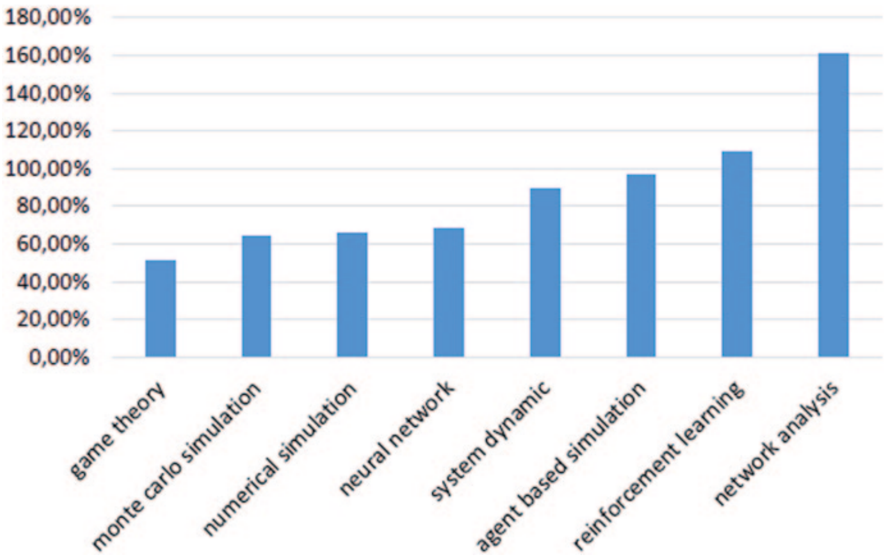

Fig. 8.2 The difference in citation for different tags

8.3 Discussion and Some Conclusions

Results show that some tags are more associated than others with the growth of the various disciplines (see Fig. 8.2). The difficult thing is to try to answer the question why. What does the clear affirmation of network analysis depend on? And why does

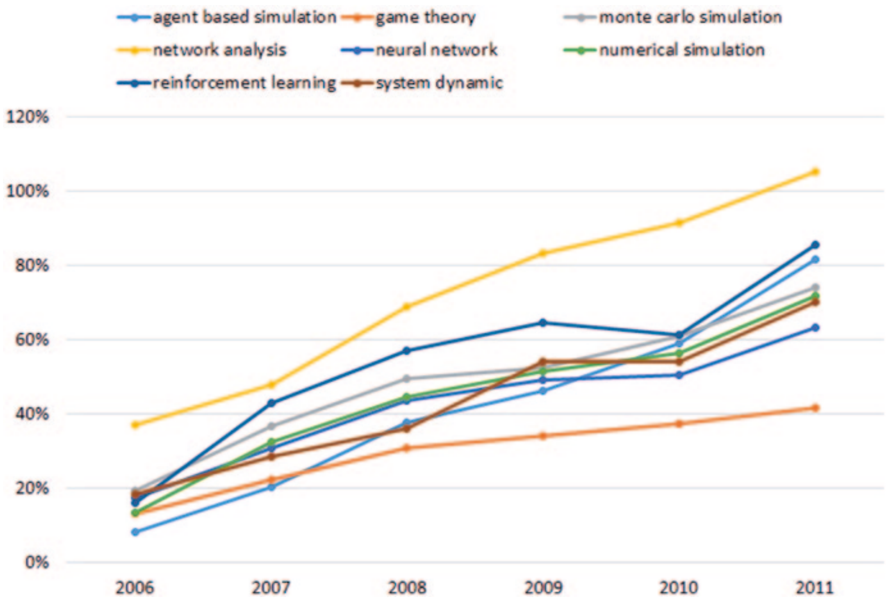

Fig. 8.3 Trends in economics, philosophy, psychology, and sociology

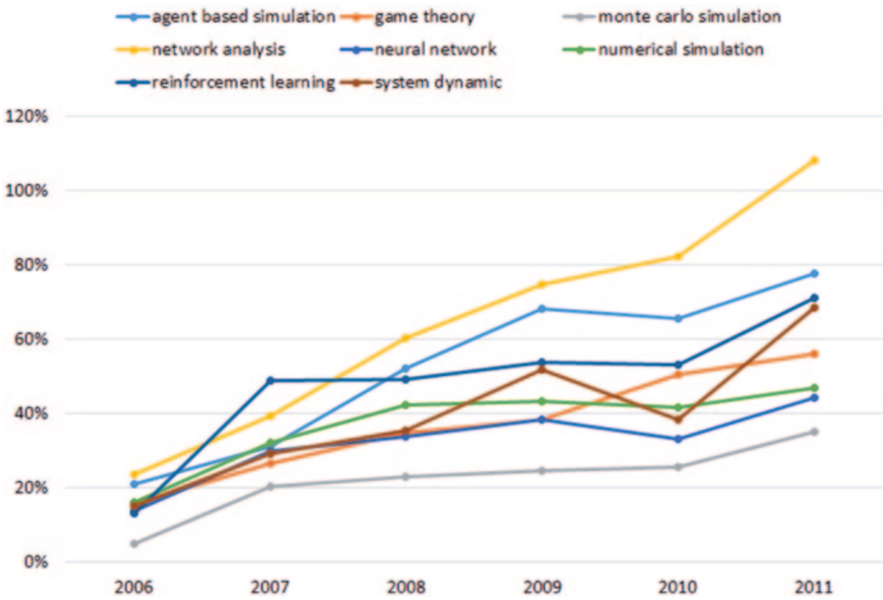

Fig. 8.4 Trend in physics

agent-based simulation, despite being relatively young compared, for example, to game theory, predominate in the disciplines of behavior that focus on these matters?

The larger increment is with the *network analysis* tag. With *network analysis* there is a convergence of interests on the part of disciplines, e.g., information and

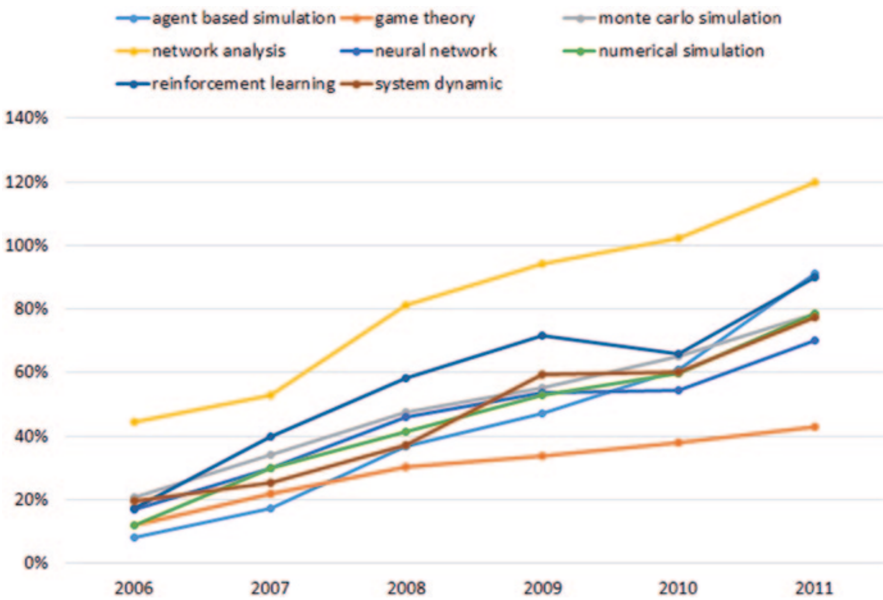

Fig. 8.5 Trend in statistics

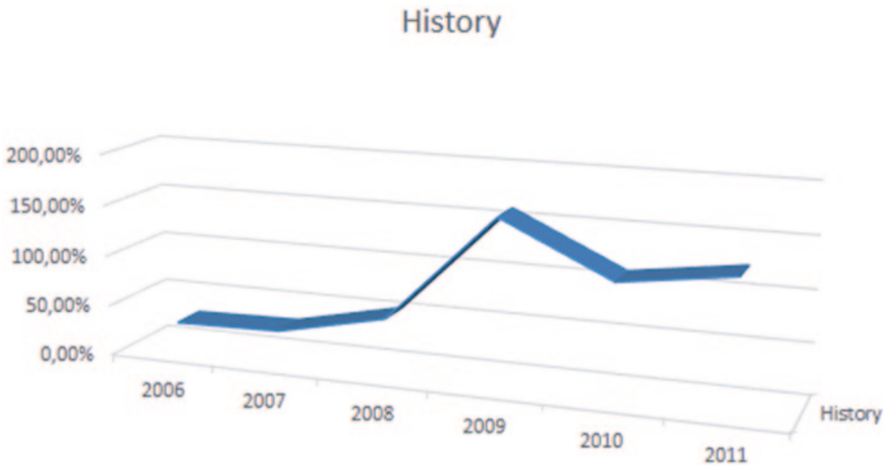

Fig. 8.6 History, using "agent based simulation" tag

knowledge, which study such matters,. Thanks to the study of both artificial and natural networks (the network of chemical reaction in a protein, the *web*, the network of citations among scientists), it was possible to highlight a new set of issues, in particular the concept of networks with topology different from the random topology, linking the idea of universality related to topological parameters (Albert and Barabási 2002; Barabasi et al. 1999, 2000; Strogatz 2001).

The situation is different for *agent-based simulation*, in which the increase does not depend on the universal character of the tag, but rather on a characteristic of the method. In fact, in our view, the use of models based on agents is increasing not because it yields new "universal" models, but rather because they these models to be *the only models that can work in a certain field of application* (Bankes 2002; Helbing 2012). Evidence in support of this argument lies in the value that the tag *agent-based simulation* adds to physics (see Fig. 8.4). Certainly, the interest of physics in the study of socio/economic phenomena (the so-called *econophysics* and *sociophysics* frameworks) has increased considerably in the last decade. Hence the necessity to find quantitative methods that can handle phenomena with highly heterogeneous agents; that are able to account for social influence; and that exhibit the ability to imitate, evolve, and evaluate different alternatives. ABM represents at least an attempt to find a solution to this problem (Bonabeau 2002).

The ascent of *reinforcement learning* and *Monte Carlo simulation*, on the other hand, is not clear: why do such *dated* tags persist over time? To clarify this apparent paradox, we should investigate possible correlations between them and a successful tag, like *agent-based simulation* and *reinforcement learning*. In fact, a main feature of agent-based modeling is the use of learning/adaptive agents (Andrighetto 2010; Campennì et al. 2009; Cecconi 2010; Epstein 2011).

History shows, for agent-based simulation, a typical pattern in these types of empirical studies: its sudden growth around 2009 could be explained as an artifact, a side effect of papers discussing the "*history of agent-based simulation*." In that period, indeed, we found a great deal of scientific production concerned with the "founding" problems of ABM. But it could be genuinely due to a historian's finding a new methodology (Dean 2000). The question is open for further investigation. One way to proceed might be to develop some algorithm to evaluate the text of the abstracts for the "top-ranked" papers, and to assign different weights to different semantic structures (for example, *history of agent-based simulation).*

In this paper we have shown how some *computational/simulation* tags, including agent-based simulation, occupy a larger space than other tags in the scientific literature, even in well- structured fields. This could indicate transversal trends in the growth of scientific paradigms.

The work is only beginning: it will be necessary to discover connections between scientific fields. For example, is it true that econophysics and sociophysics studies tend to abandon traditional methods of investigation—one-for-all statistical mechanics—in favor of simulations-based artificial agents? To give an answer it is probably necessary to try to understand the current trends in the scientific production and those that the scientific community will adopt in the near future both within the science of nature and that of society.

References

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47.

Andrighetto, G., Campenni, M., Cecconi, F., & Conte, R. (2010). *The complex loop of norm emergence: A simulation model 'Simulating interacting agents and social phenomena'* (pp. 19—35). Berlin: Springer.

Bankes, S. C. (2002). Agent-based modeling: A revolution?. *Proceedings of the National Academy of Sciences of the United States of America, 99*(Suppl 3), 7199–7200.

Barabási, A.-L., Albert, R., & Jeong, H. (1999). Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and Its Applications, 272*(1), 173–187.

Barabási, A.-L., Albert, R., & Jeong, H. (2000). Scale-free characteristics of random networks: The topology of the world-wide web. *Physica A: Statistical Mechanics and Its Applications, 281*(1), 69–77.

Beel, J. (2009). Google Scholar's ranking algorithm: The impact of citation counts (An empirical study) 'Research Challenges in Information Science, 2009. RCIS 2009. International Conference on Information, 160–164, Technology: New Generations.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems, *Proceedings of The National Academy of Sciences of the United States of America, 99*(Suppl 3), 7280–7287.

Campenni, M., Andrighetto, G., Cecconi, F., & Conte, R. (2009). Normal = Normative? The role of intelligent agents in norm innovation. *Mind & Society, 8*(2), 153–172.

Cecconi, F., Campenni, M., Andrighetto, G., & Conte, R. (2010). What do agent-based and equation-based modelling tell us about social conventions: The clash between ABM and EBM in a congestion game framework. *Journal of Artificial Societies and Social Simulation, 13*(1), 6.

Chen, X. (2010). Google scholar's dramatic coverage improvement five years after debut. *Serials Review*, *36*(4), 221–226.

Conte, R. (2000). The necessity of intelligent agents in social simulation, *Advances in Complex Systems, 3*(01n04), 19–38.

Conte, R., Edmonds, B., Moss, S., & Sawyer, R. K. (2001). ‘Sociology and social theory in agent based social simulation: A symposium’. *Computational & Mathematical Organization Theory, 7*(3), 183–205.

Dean, J. S., Gumerman, G. J., Epstein, J. M., Axtell, R. L., Swedlund, A. C., Parker, M. T., & McCarroll, S. (2000). Understanding Anasazi culture change through agent-based modeling, *Dynamics in human and primate societies: Agent-based* modeling of *social and spatial processes*, 179–205

Epstein, J. M. (2011). *Generative social science*: *Studies in agent-based computational modeling*. Princeton: Princeton University Press.

Helbing, D. (2012). *Agent-based modeling social self-organization.* Berlin: Springer, pp. 25–70.

Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual review of sociology,* 143–166.

Neuhaus, J. M., & McCulloch, C. E. (2006). Separating between- and within-Cluster covariate effects by using conditional and partitioning methods. *Journal of the Royal Statistical Society, Series B (Statistical Methodology), 68*(5), 859–872.

Squazzoni, F. (2008). The micro-macro link in social simulation. *Sociologica, 2*(1), 0–0.

Squazzoni, F. (2010). The impact of agent-based models in the social sciences after 15 years of incursions. *Sociological Methodology, 2*(2)*,* 1–23.

Strogatz, S. H. (2001). Exploring complex networks. *Nature, 410*(6825), 268–276.