1

Introduction: Toward behavioral computational social science

The main assumption of this book is that individual behavior and social phenomena are somehow connected and that the investigation of that connection is central for all social sciences.

The work presented here can be classified as a methodological one since it deals with methods. With extreme synthesis, it presents the methods available for putting together the studying of individual behavior, as developed in behavioral sciences, with the many tools that today compose the approach called "computational social science" (from now on CSS).

The ideas and methods presented here have originated in different domains, and it is very difficult today to find an exhaustive and comprehensive description of them. The book thus aims not only at theoretically discussing a unified methodological approach but also at providing the readership with all the necessary information to experiment with the approach. Obviously, given its physical constraints, the preparation of this book has meant much selection and not all concepts and tools are explained from scratch and in details.

Behavioral Computational Social Science, First Edition. Riccardo Boero.

However, average readers with interests in the scientific explanation of social phenomena can surely comprehend what is discussed here and can use the many references provided to master all the topics and tools.

Before presenting how the book is organized and how selective reading can be conducted on it, this introduction focuses on the first question readers should ask: what is the use of this approach?

1.1 Research strategies in CSS

CSS has recently emerged because of many important technological and conceptual advancements.

Undoubtedly, it is the "Big Data" approach and the large availability of data associated with it that allow today studying large-scale social phenomena that were impossible even just a few years ago. At the same time, it is the availability of (cheap) computing power that allows storing, managing, and analyzing those datasets.

From the conceptual viewpoint, relatively new scientific tools such as those of social networks analysis, complexity, and other approaches finally find in such an abundance of social and behavioral data the chance to be applied and tested.

Along with the promises of the potentially fruitful integration of all those innovations, CSS seems today to give social sciences the possibility to overcome the well-known limits of more traditional approaches. Heterogeneity of individuals, nonlinearity of systems and behavior, and the lack of capability to effectively put in relationship social structures and social phenomena with individual behavior are just a few of the many examples of limits that potentially can be overcome.

Much of the research in CSS is today still aimed at exploring the potentials of the approach, but different research strategies have already emerged, and consequently, different methodologies have emerged too.

In particular, CSS can be considered today as a self-standing approach to social sciences because it provides tools and methods to pursue any kind of scientific research strategy.

Research strategies in science are only a few. Following an order that is not intended to imply any ranking of importance, it is firstly possible to explore the data in order to describe or classify it. It is an activity that is always needed in scientific investigations and that can provide first-hand and novel information about completely unknown phenomena and systems.

Second, it is possible to establish relationships in the data. Using statistical and other models, it is in fact possible to observe that some of the variables

appear to be connected, changing in similar or opposite ways. Variables that are somehow related are the first candidates to consider, select, and investigate deeper.

Most of the contemporary research works in CSS, for instance, the ones belonging to the Big Data and science of networks approaches, adopt one or both of these two research strategies.

CSS however allows also pursuing the third kind of scientific research strategy, which is the investigation of causality. Similarly to any other scientific domain, in social sciences, the investigation of causality requires the availability not only of data but also of tools for modeling. Modeling is the formalization procedure that ultimately allows developing, testing, and validating knowledge. In CSS, modeling is pursued by the approach of social simulation where adopted modeling tools are explicitly addressed at dealing with the peculiar complexities of social systems.

1.2 Why behavioral CSS

Behavioral CSS is aimed at investigating causality of social phenomena and it thus relies on modeling tools developed in social simulation.

Further, it refers to the need for behavioral information. The need for the integration of the two approaches is motivated by the unique form that causal relations have in social systems.

In fact, social phenomena can be conceived as different from phenomena in natural systems because of their complexity. The complexity of social phenomena, moreover, is surely characterized by the most complex object we know, which is the human brain, but not only by that. Social phenomena, in fact, are complex because their causes always involve both individual behavior and some of the many features of the social structure (e.g., institutions, social norms, ways of social interaction, etc.). These features of social complexity are the ones that the behavioral CSS approach explicitly acknowledges and that it tries to effectively deal with.

Most of the critiques to established research methods in social sciences are addressed at the same time. Causal explanations in traditional approaches are incomplete and often unreliable because the complexity of social systems is reduced, either from the perspective of individuals or from the one of the social structure. Established approaches are often obliged to apply reductionism because of the technical limits of the analytical tools that are adopted.

Today, CSS and behavioral sciences create the opportunity to overcome traditional limits and to finally deal with the individual, the social structure, and the relation between the two.

Behavioral CSS can thus be seen from three different perspectives. From the one of behavioral sciences, it can provide researchers with the tools to extend their investigations toward social contexts of interaction. Such an extension can improve knowledge about behavior because of the many feedbacks going from the social structure to individual behavior and *vice versa*.

From the perspective of researchers in social simulation, the integration of behavioral knowledge in their models can finally allow rigorous validation and external validity.

From the perspective of social sciences in general, behavioral CSS enriches the researcher's toolset with an approach that is explicitly addressed at the investigation of causality of social phenomena intended as a social causal mechanisms.

Because of what just said, the behavioral CSS approach presented in this book is aimed at readers in social sciences in general and particularly at those in behavioral sciences and social simulation.

The book discusses the approach and presents several tools in order to allow readers with different backgrounds experimenting with and perhaps extending it.

1.3 Organization of the book

Being aimed at presenting and discussing a methodological approach that puts together different tools and traditions, the book is organized in two parts. In the first one, there is the presentation of the main concepts and methods developed in the two approaches that are integrated (i.e., CSS and behavioral sciences). The first part also includes a short discussion of the advantages of the approach.

The second part takes a more applied and technical perspective and it presents methods for tool integration. In particular, because in order to investigate causality with behavioral CSS researchers have to rely on modeling, that part initially discusses how to integrate results of behavioral analyses in models. Secondly, there is a chapter that discusses how to model structures of social interaction.

In conclusion, the second part ends with the presentation of an example of application of the approach, mainly aimed at didactic purposes. In particular, it is presented the preparation and implementation of a model of a social phenomenon, starting from the collection of behavioral data, passing through its analysis, and arriving at the specification and formalization of behavior and interaction in the model. Some applications of the model are sketched too in order to give the reader the intuition of potential uses and results in terms of causal explanations.

Readers experienced in social simulation and agent-based modeling can probably read the book without spending much attention to Chapter 2, although the concepts related to causality that are presented there are crucial for the evaluation of the approach.

Readers coming from behavioral sciences can avoid Chapter 3 where common tools of that approach are shortly introduced.

Readers particularly doubtful about the analytical advantages provided by the approach should start reading Chapter 4 and then read the section presenting results in Chapter 8. If reading those chapters changes their minds, they can go backward to the rest of the book to better comprehend the reasons why the approach can be analytically effective and powerful.

Readers who intend to adopt the modeling approach presented here but who are not experienced in social simulation should spend particular attention to the second part of the book. The first three chapters of that part illustrate the algorithms that allow modeling behavior in agents and interaction between agents. The fourth chapter of that part provides examples of application of many of those algorithms. Finally, in the appendix, the technical implementation of what presented in Chapter 8 is reported and discussed.