Chapter 1

Motivations, aims and structure

1.1 Introduction

As we have already hinted in the introductory sections of this thesis we start from the assumption of having a society of individuals. Such a society has formed in some way that we do not investigate and may be either **flat** or possess some **inner structure**. In both cases we assume the presence of static structures that outlive the individuals and allow us to speak of generations and of a time counter with periods and epochs.
The fact of disregarding the evolutionary history of a society is in accordance with the approach proposed by Axelrod ([12] and [13]). It is easy however to imagine an extension that could be able to account, at least in part, for such evolutionary history. Such an extension could take two forms:

- a unidimensional form;

- a multidimensional form.

Under the **unidimensional form** we can associate to each individual a single numeric value either on an ordinal scale or on a cardinal scale ([96]).
In the former case that number defines the start order of the individuals so that some of them (those with the lower numbers) have an advantage on the others (those with the higher numbers) and ties are possible.
In the latter case we can see the numerical value as an endowment of money (or of any other numerary good) for each player so that the higher is the value associated with an individual the more he is rich and so the more he is assumed to be better off. In this way we measure the richness of an individual (and not his utility that may even vary inversely with his richness) according

to a common cardinal scale among the individuals.
If we would like to consider **utility** instead we would need a function u_ι mapping the endowment of money of the individual ι on a numerical quantity called utility. In order to do so we could follow the procedure suggested by Myerson ([85]). In this case we should solve the following problems or:

- if every individual ι has his function u_ι or do we have only one function for all the individuals;
- if such functions represent a common knowledge among all the individuals or within groups of individuals or are a private knowledge of every individual;
- if the values that such functions produce can be compared among themselves or can be summed so to produce a global utility or a social welfare.

In the present thesis we use rarely the concept of utility and only in combination of Pareto concepts such as efficiency, optimality and dominance ([39]). From this premise we derive that we allow every individual ι to have his function u_ι that produces values that can be compared in the Pareto way. For what concerns the nature of the knowledge we will be more specific wherever this will be needed and the same is true also for the issue of the welfare of a set of individuals up to the welfare of the whole society.
Under the **multidimensional form** we associate to each individual a vector of properties, each measured according to its own scale. In this way we associate to every individual a point in a multidimensional space in order to prevent any form of compensation among different dimensions.
If, for instance, we consider five properties we may have that the properties are the following:

- wealth or richness;
- health;
- mood;
- working status;
- education level.

In this case only the first property is a commodity according to the economic theory and is measured according to a common cardinal scale whereas the other four are measured over ordinal scales that we can assume as common

and as represented through ordinal numerical values ([96]). For instance we can measure health on a ordinal scale with values, in increasing order, "bad", "poor", "sufficient", "average" and "good" to wich we associate, respectively, the values -3, -1, 1, 3 and 5 .
We can behave in similar ways for properties such as "mood", "working status" and "education level" by using either ordered string labels or numeric labels on an ordinal scale. Also in this case we can say of two labels or values which is better but not how much it is better, both in general terms and for every single individual.
So doing we associate to each individual a point in the space $\mathbb{R}^5$ and, at the same time, we prevent any compensation between properties. With this we mean that we do not allow any compensation, for instance, between health and wealth or between mood and working status and so on ([23]).
Once the positions of the various individuals have been defined we can identify:

- the dominant individuals and so those that are no worse off on all the dimensions and are better off on some of them;

- the dominated individuals and so those that are no better off on all the dimensions and are worse off on some of them;

- the equivalent individuals and so those that are better off on some dimensions, worse off on some other dimensions and tied on the remaining dimensions.

In this thesis we usually do not use such tools but in few cases where we use the multidimensional form.

1.2 Some introductory concepts

For what concerns **the structure** of our societies we have that in the **flat case** the society can be seen as a set of peers according to an approach that is used in Game Theory (GT, [89], [85], [18], [19]).
In this **flat case** we have a set[1] $\mathbb{I}$ of individuals. Every individual $\iota \in \mathbb{I}$ is characterized by a type θ_ι that takes values from a set of possible types Θ_ι (or **type set**) so that $\theta_\iota \in \Theta_\iota$, according to a classical approach of GT ([68], [89], [85]).

[1]With symbols like $\mathbb{I}$ we denote set of individuals and of their specifications so that $\mathbb{D}$ denotes a set of deciders, $\mathbb{S}$ a set of stakeholders and so on, with obvious variations that are, anyway, properly defined in the due places.

We assume that every type in a type set defines both the actions available to an individual of that type and the preferences on the alternatives that he has at his disposal.
At this point we may consider the following cases ([60]):

- private values case;

- common values case.

In the **private values case** every set Θ_ι is a private information of each individual so that whenever he interacts for the very first time with another individual the interaction occur under conditions of **ignorance**[2] ([48], [49]). Whenever the interactions with the same individual occurs repeatedly an individual can start guessing the type of the other from the other's actions and preferences (how they are revealed by his actions and choices) so that the interaction becomes an interaction under conditions of **uncertainty** though it hardly ever turns into either **risk** or **certainty**.
In the **common values case** we must differentiate at least two possible cases.
In the former or **loose** case we have that the sets Θ_ι are publicly revealed but an individual has no feedback that his revelation has been noticed and recorded in some way so that it is not forgotten as time passes by unless it is repeated with a certain frequency.
In the latter or **binding** case ([85], [60]) we have that the sets Θ_ι are publicly revealed and each individual has a feedback of this acknowledgment that he acknowledges and so on ad infinitum. This infinite recursion of mutual acknowledgments involves all the individuals so that their population is homogeneous, at least from this point of view. In practical cases we may have subpopulations where such common knowledge is segregated so that individuals belonging to distinct subpopulations act as in the private values case with regard to the individuals of other subpopulations.
In the case where there is **some structure**, on the other hand, we may refer to Figure 1.1 where we show the simplest case where we have only two levels. In this case we have[3]:

- the set $\mathbb{S}$ of the **Stakeholders**;

- the set $\mathbb{SP}$ of the **Social Planners**.

[2]For the definitions of the concepts of ignorance, uncertainty, risk and certainty see Appendix B.

[3]In what follows we are going to use uppercase initials to denote the sets and lowercase initials when we refer to the corresponding concept.

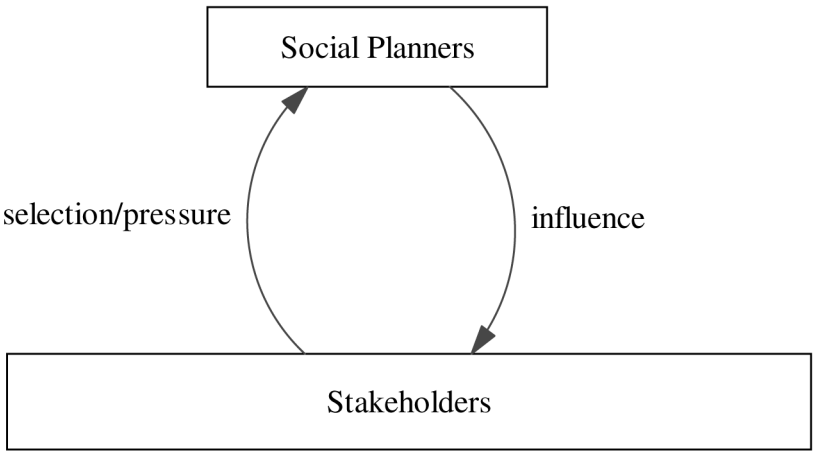

Figure 1.1: *Interactions between stakeholders and social planners*

As it is shown in Figure 1.1 the stakeholders can both **select** the social planners (as it happens in all the democracies through elective procedures) and exert **pressure**s on them in both direct and indirect ways. As to the direct ways we mention strikes, riots, opinion or pressure movements and referendum or laws from people initiative. As to the indirect ways we mention press campaigns and electoral competitions.
On the other hand the social planners, with their decisions and policies, **influence** the stakeholders by affecting their welfare and their both social and natural environment ([34]).
We note how we have $\mathbb{S} \subseteq \mathbb{I}$ and $\mathbb{SP} \subseteq \mathbb{I}$. Moreover we can have the following two cases:

(1) $\mathbb{S} \cap \mathbb{SP} = \emptyset$,

(2) $\mathbb{S} \cap \mathbb{SP} \neq \emptyset$.

In the case (1) we have that the social planners are not directly influenced (as stakeholders) from their decisions and their policies so they are more immune from conflicts of interests or from the fact that they could favor themselves through publicly intended decisions or policies.
In the case (2) we have that the social planners are also stakeholders so that they are directly influenced as stakeholders and we are in presence of a potential conflict of interests.
Both stakeholders and social planners (see Figure 1.2) are inserted in a stable structure that we call, for simplicity, **Society**.
Society is to be seen at the macrolevel ([107]) and is composed of all the

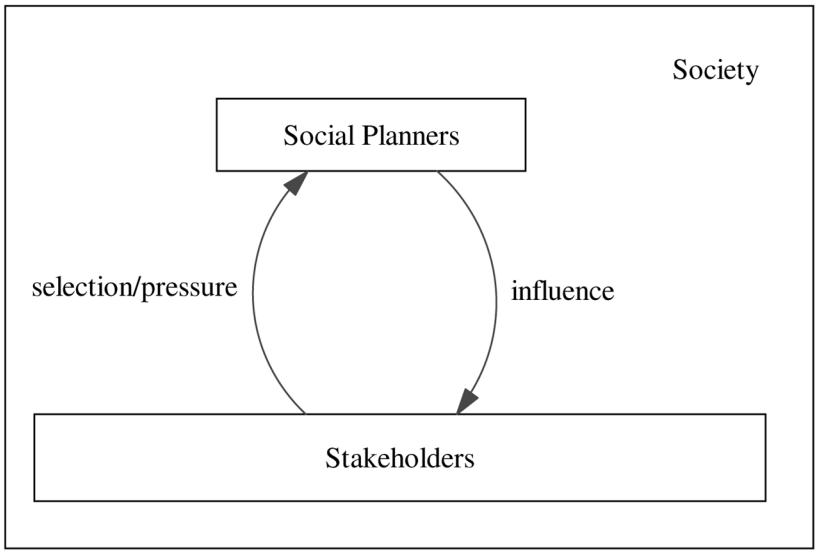

Figure 1.2: *Society as the general background and high level framework*

structures that outlive both groups and coalitions and individuals. Such structures have both static and dynamic features since they represent the framework within which the individuals act but are also modified by the action of the individuals (see Chapter 3 for further details).
From our point of view we consider the society as composed by norms and rules as well as habits and customs that tend to exert both normative and prescriptive influences at both the mesolevel and the microlevel (see Figure 1.3).
On the other hand both individuals at the microlevel and coalitions and groups at the mesolevel behave on the **expectations** (see the dashed arrows in Figure 1.3) that they have on how society is going to change or can be modified through their actions (see Figure 1.3). In Figure 1.3 we represent only the relations between the lower levels and the Society but similar relations exist also between the mesolevel and the microlevel though they are less strong and more variable essentially owing to the lower stability of the coalitions and the groups that can be found at the mesolevel.
As we have already seen in the flat case, also in this case every individual $\iota \in \mathbb{I}$ has associated a set Θ_ι of his types. In addition to this set every individual has associated a **role** ρ_ι or a variable that may assume a value, for instance, in the set:

$$\{s, sp\} \tag{1.1}$$

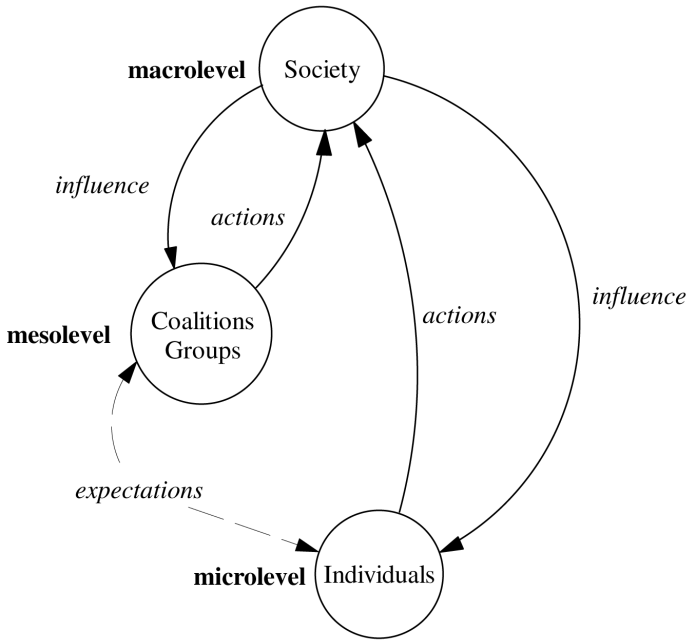

Figure 1.3: *Top-down and bottom-up interactions*

so that $\rho_\iota = s$ identifies ι as a stakeholder whereas $\rho_\iota = sp$ identifies ι as a social planner. We note how, to account for the case (2) above, the role is a variable or dynamic entity that assumes different values depending on the context and on time[4] (see section 4.2.2).
This feature imposes us to consider the following conditions (see Table 1.1):

- a condition of **perfect recall** (PR) that is concerned with the past;
- a condition of **perfect prevision** (PP) that is concerned with the immediate future.

Both conditions refer to the present time so, if we measure the time according to a succession of discrete steps and denote the present with[5] t_h, the immediate future is t_{h+1} whereas the past is represented by the values t_k with $k < h$. An individual $\iota \in \mathbb{I}$ has a **perfect recall** (see the entry labeled y in the first column of Table 1.1) if he can remember the roles that he has played in the past up to the present time otherwise he has no such capability (see the entry labeled n in the first column of Table 1.1).

[4]If we consider a discrete time t_k with $k \in \mathbb{N}$ in order to make explicit this dependence we can write $\Theta_\iota(k)$ or $\Theta_\iota(t_k)$ and $\rho_\iota(k)$ or $\rho_\iota(t_k)$.

[5]It should be obvious how the variables h and k assume positive integer values.

PR - PP	y	n
y	(A)	(B)
n	(C)	(D)

Table 1.1: *Perfect Recall and Perfect Prevision*

On the other hand, with similar conventions, an individual $\iota \in \mathbb{I}$ has a **perfect prevision** if he can anticipate for sure the role he is going to play the next time otherwise he has not such capability.
On the ground of these definitions we have the following four cases (see Table 1.1):

- (A) where an individual ι has both PR and PP so that he can incur in a potential conflict of interests;
- (B) where an individual ι has PR but no PP and so a typical situation that occurs in the case of repeated interactions with other individuals ([12], [13]);
- (C) where an individual ι has no PR but has PP so that he has no memory of the past but can foresee the immediate future so that he can incur in a potential conflict of interests;
- (D) where an individual ι has no PR no PP so that he has a reactive behavior ([123]) that prevents him from incurring in conflicts of interests.

In what follows we are going to see applications of the four cases with the possible exception of the case (C).
For what concerns the actions available to each individual ι they form a set $\mathscr{A}_\iota$ that define, in a certain sense, the possible strategies of ι as a player. The set $\mathscr{A}_\iota$ depends, in general, from the current type θ_ι so we should write $\mathscr{A}_\iota(\theta_\iota)$ (if it is time independent) or $\mathscr{A}_\iota(\theta_\iota, t_k)$ (if it is time dependent).
Such a set is usually seen as **exogenously** defined and fixed, as it occurs in GT where the players cannot modify their actions set to adapt it to changing situations neither in the case of repeated games ([89], [85], [60]). It is, moreover, assumed to be common knowledge, in a classical sense, among all the players ([89], [85], [60]).
Another possibility is that this set is **endogenously** defined and modified by every player ι according to his needs and to either the current situation or the history of the past situations as $\mathscr{A}_\iota(\theta_\iota, t_k)$ with $k < j$ if t_j represents the present time.

In this case we must state if this set is either a private or a common knowledge and if there is a delay and of which entity with which its knowledge spreads among the other involved individuals.

1.3 The starting points and the possible extensions

In section 1.2 we have, therefore, defined our starting points and so:

- a flat structure society;

- a two layers society.

In both cases society acts as the macrolevel.
In the former case we have neither any hierarchy nor any structure (but for those of the society as a whole) so that the individuals are peers that act as singletons. In this case we have only the microlevel and the macrolevel since the individual act anyway in a sort of normative context, at least in one of those that are established by[6] $NCGT$ and not in a normative emptiness.
In the latter case (see Figures 1.1 and 1.2) we have a simple hierarchy and the individuals play a role that defines some of them as capable to issue orders/commands to the others that are free to either obey or disobey.
They (or at least some of them) obey if they find those orders as attuned with their will, interests and values. They (or at least some of them) disobey in every other case.
As possible extensions to this simple scheme we can consider the following:

(1) the presence of coalitions among the stakeholders, the so called **parties**;

(2) the presence of groups among the stakeholders;

(3) the presence of groups of social planners;

(4) the presence of a hierarchy among the social planners;

(5) the presence of roles for the social planners, that may be termed as government and opposition.

[6]With the acronym $NCGT$ we denote the Non Cooperative Game Theory whereas with the acronym CGT we rather obviously denote the Cooperative Game Theory.

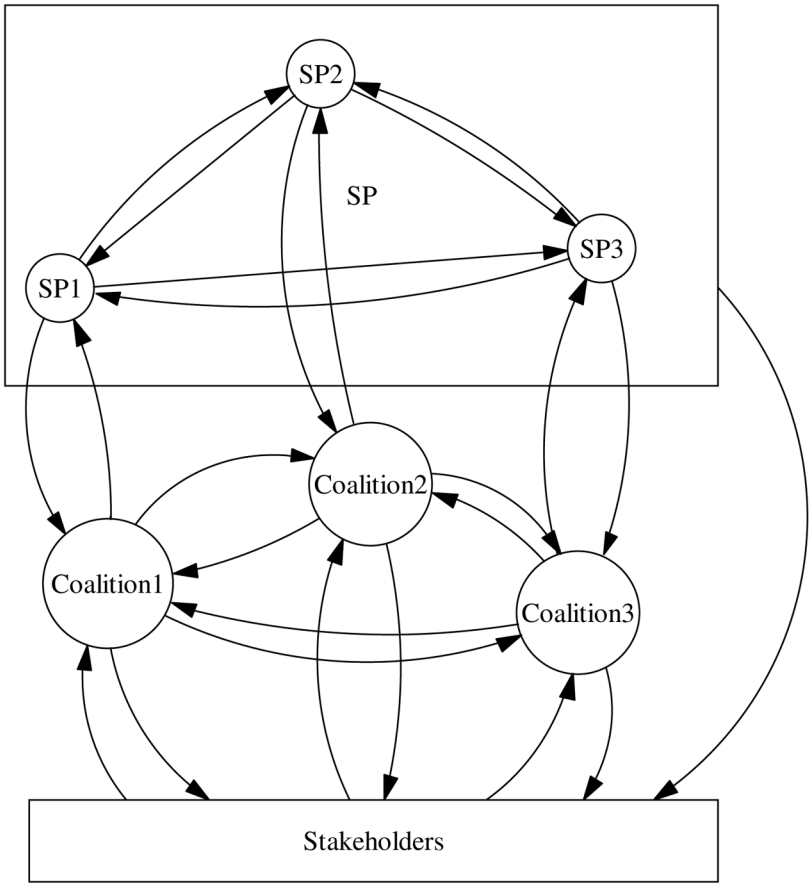

Figure 1.4: *Interactions and hierarchies*

In Figure[7] 1.4 we represent an example that refers to the points (1) and (3). In such a figure we have the stakeholders of the set $\mathbb{S}$ that form three coalitions that, in their turn, define three groups of social planners. The three coalitions are not necessarily a partition[8] of the set $\mathbb{S}$ so we can have stakeholders that do not belong to any coalition but that, in the current framework, risk of not being represented by any of the social planners.
In this case we have that, for instance, $Coalition1$ defines the group $SP1$

[7]In Figure 1.4 we use a directed arrow to represent a relation between the elements at its ends. If we have $A \longrightarrow B$ this may mean that we can have, depending on the context, an influence of A over B or a migration from A to B. If we have $A \longrightarrow B$ and $B \longrightarrow A$ this may give rise to a feedback loop but also to successive migrations from A to B and from B to A and so on, depending on the context.

[8]Whenever we speak of a partitioned set we mean a subdivision of that set in a certain number of disjoint subsets whose union is equal to the given set.

and is, in some way, influenced by its behavior. The same holds also for the other two coalitions.
We note how (see Figure 1.4):

- the coalitions influence each other;
- the coalitions get their members from the set $\mathbb{S}$ but these members can pass from one coalition to another and also exit from any coalition;
- the groups $\mathbb{SP}_i$ of social planners interact among themselves and influence, as a single set $\mathbb{SP}$, the whole set of the stakeholders;
- the sets $\mathbb{SP}_i$ can be formed even by only one social planner;
- the stakeholders can become social planners only passing through the forming of a coalition.

For what concerns the point (3) we can implement it through the definition of the following concepts:

- the concept of level;
- the concept of fitness;
- the concept of complexity.

The concept of **level** allows us to partition the set $\mathbb{SP}$ in a certain number of subsets $\mathbb{SP}_i$. If we have n levels we have n subsets among which we can define the following ordering:

$$\mathbb{SP}_0 \downarrow \mathbb{SP}_1 \downarrow \ldots \downarrow \mathbb{SP}_n \tag{1.2}$$

where $A \downarrow B$ means that the set A is subordinate to the set B or that the members of A are at a level below those of the set B. We say that a set is subordinated to another if the members of the latter have the power to reverse (bottom-up) or the capability to frame (top-down) the decisions of the members of the former set. A system with levels may be with or without balancing. In the former case there is at least one neutral level to which a subordinate level can appeal to oppose to decisions and recommendations from subordinating levels. In the latter case there is no possibility to make such form of opposition.
The concepts of **fitness** and **complexity** work together as follows. Every subset $\mathbb{SP}_i$ has the capability to deal with issues of certain complexities so that if an incoming issue has a complexity that falls in the range of a certain level the members of that level can issue a fitness message and ask for being

authorized to deal with that issue.
If no other subset issues a fitness message the only issuing subset is authorized to deal with the current issue otherwise the authorization is given to the issuing subset that is at the highest level among the applicants.
If no subset issues a fitness message we assume that such an issue is dealt with by the members of $\mathbb{SP}_n$ according to a sort of closure principle or an appeal to the highest level authority.
This procedure, in order to be implemented, requires that the issues have a declared complexity as a numerical value. Such a value can be defined as directly proportional to the number of the effectively affected stakeholders so that the higher this number is the higher the complexity of the associated issue is.
For what concerns the point (4) we note how the set $\mathbb{SP}$ (or every subset $\mathbb{SP}_i$ if we have a hierarchy) can be seen as partitioned in two subsets:

- the government subset $\mathbb{G}$;

- the opposition subset $\mathbb{O}$.

If we adopt a simplified setting we have that the members of $\mathbb{G}$ make the proposals that must be approved also by the members of $\mathbb{O}$ through the use of some to be specified electoral mechanism in order to become effective.
For the moment we do not give any further detail since we deserve to deal with these issues again in other chapters of this dissertation.

1.4 The motivations and the aims of the thesis

This thesis is grounded on a simple seminal idea: that is possible to induce selfish and self interested individuals to act in ways that produce collectively desired behaviors without using or, rather, with a very limited use of coercive or punitive tools. We underline how such behaviors must be collectively desirable and not desirable by a single individual (a dictator) or by a small set of individuals (an oligarchy).
Moreover the level of desirability must be high enough but not absolute so that we consider a small number of free-riders[9] as a physiological calamity and their full elimination either too costly or too time consuming or both in

[9]The concept of free-rider will be defined in Chapter 3. For the moment we define a free-rider as an individual that enjoys the benefits without paying any of the costs associated to the production of such benefit.

order to be accomplished.
From this seminal idea that is the basic motivation, we derived other motivations (to be pursued possibly in indirect ways) such as:

- to prove that some one-shot bargaining solutions are based on unrealistic assumptions and cannot work in real cases ([91], [89], [85], [93]);

- to propose iterative procedures for decision and negotiation ([39]), [41], [47]) among competing individuals;

- to devise distributed protocols and strategies of cooperation and coordination without the intervention of any central authority;

- to prove that coercive and punitive tools have, in many cases, little deterrence and have costs that are disproportionate with respect to what they allow to obtain;

- to prove that the individuals in many cases self organize as complex systems in ways that are unforeseen and unavoidable through a top-down planning ([65], [113], [27]);

- to show how control mechanisms work the better the more they are distributed both as localization and as responsibility.

On the other hand, the aims of the thesis are the following:

- to prove that our motivations were sound so that their pursuing had a sense;

- to show that our motivations can produce usable tools for their accomplishments (a sort of self-realizing prophecy so dear to the economists).

In addition to these aims another aim of the thesis was to prove that the possible conflicts between individual behaviors and collective desired outcomes are not in the fabric of the world (and so are unavoidable) but depend on contingencies and can be solved or prevented through the proper use of incentives and rewards.

1.5 The structure of the thesis

The thesis has the following coarse grain structure.
In Chapter 2 we describe the approaches that we are going to adopt in the thesis as well as the tools that we are going to use and the levels at which we

are going to move. This chapter develops some of the concepts that we have already sketched in the opening sections of the thesis as well as in the present chapter and refers heavily to the concepts that we present in the appendices.
In Chapter 3 we introduce a sort of (concise) sociological framework for this thesis under the form of the main features of the public policies as normative tools that the social planners can devise in order to shape the behavior of the stakeholders. Moreover in this chapter we discuss the problem of control as a dynamic relation between controllers and controllees.
In Chapter 4 we present the top-down approach where a set of **social planners** devise some mechanisms in order to have the stakeholders behave according to certain desired patterns of behavior or collective behaviors. In this chapter we use concepts that we recall in the Appendices B and C but we also present some properties that we wish the proposed mechanisms satisfy and show the possible inconsistencies among them under the form of classical[10] impossibility theorems.
After the presentation we have done in Chapter 4, in Chapter 5 we discuss the use of simulation techniques at the macrolevel (or at the level of the whole society) and we introduce the concepts we are going to use in Chapter 9.
In Chapter 6 we present the bottom-up approach and discuss how a set of individuals can interact whenever they face a common problem to which they have to find a solution. We present and analyze the various ways of interaction and the properties that the individuals wish are satisfied by the possible solutions. We also examine if and how the individuals can keep under control the presence of free-riders through the use of distributed mechanisms.
In Chapter 7 we discuss the interactions between the two approaches we presented in Chapter 4 and in Chapter 6. In this chapter we also discuss the presence of the mesolevel under the form of coalitions and groups and of both a hierarchy among the social planners and a differentiation of the social planners according to types, as we have seen in section 1.3.
After the presentation we have done in Chapter 7, in Chapter 8 we discuss the use of simulation techniques at the mesolevel (or in presence of either groups or coalitions) and at the microlevel (or at the level of the individuals acting as singletons) and we introduce the concepts we are going to use in the proper sections of Chapter 9.
In its turn, chapter 9 is devoted the description of some simulation models that we use to verify the correctness of the assertions we have made in the foregoing chapters. In this chapter we present some models where we define

[10]Such theorems are termed as classical since they are well known form the literature of both decision and voting theory.

agents endowed with simple capabilities and let them interact under more or less binding constraints in order to see if they can show emerging behaviors to be compared with some intended (by one or more social planners) behaviors. We plan also to introduce in the models, at least in part, the features we have outlined in Chapters 4 and 6.
The models we present in this chapter have been written using programs such as NetLogo ([86]) and Vensim ([117]).
Last but not least Chapter 10 is devoted to the drawing of some conclusions about the lessons we have learned during the writing of the thesis and to the listing of [some of] the open problems.
In addition to these chapters, that represent the core part of the dissertation, we have three appendices where we have inserted some background materials.
In Appendix A we recall the basics of *Multiagent Systems* together with some concepts related to *Simulation* and *Artificial Life Systems*.
In Appendix B we revise some very basic concepts of *Game Theory* and of *Decision Theory*. In this appendix we discuss also some concepts of both *Mechanism Design* and *Implementation Theory* and present a set of properties that are our wish list in such cases.
In the closing Appendix C we present a brief discussion of some principles of ethics and collective behaviors with the aim of showing how such principles can play a useful role within our framework. Our intention is indeed to use ethics and distributed social control ([109]) as a way to modify the behaviors of the individuals in order to produce desired collective behaviors without causing excessive losses to the involved individuals.
Such control may either derive from the interactions among the individuals ([109]) or from a sort of self discipline ([55]) that the individuals impose upon themselves: in the former case the control arise from the interactions as a structuring force of a society whereas in the latter it arises through a limitation of the alternatives that an individual sees as feasible for himself under certain circumstances and conditions.

Chapter 2

Approaches, levels and tools

2.1 Introduction

In this chapter we present what we can define as the meta-tools as well as some of the tools that we are going to use in this thesis.
The **meta-tools** have already been sketched in the opening parts of this thesis. They represent descriptive tools that we use essentially to describe the interactions among the individuals and the rules and strategies that they can follow within the society in which they live as well as within the groups and the coalitions they form.
As meta-tools we consider:

- the **approaches**,
- the **levels**,
- the **simulation tools**.

The meta-tools are therefore in our hands as tools that help us in describing the various types of individuals and their interactions within various frameworks.
The **tools**, on the other hand, consist of, essentially, **Implementation Theory** (**IT**) , **Mechanism Design** (**MD**) and **Game Theory** (**GT**) though we can use, occasionally, also auction models, bargaining models and negotiation models.
Such tools are in the hands of the social planners (as it occurs in the top-down case, see Chapter 4) or of the individuals (as it occurs in the bottom-up case, see Chapter 6) or of both (as it occurs in presence of coalitions, groups, hierarchies, layers and types, see Chapter 7).
The presentation in this chapter is based on the materials of Appendices B

and C and is targeted to the full understanding of the different approaches that we are going to use in the following chapters, from chapter 3 to chapter 7.

2.2 The approaches

The **approaches** can be said to represent the types of lenses that we can put on to observe the various mechanisms we are dealing with. They also shape the types of questions that we can raise about the issues we are going to investigate ([94], [23]).
In this thesis we are going to use the following approaches:

(1) descriptive;

(2) normative;

(3) prescriptive.

Such approaches are not fully independent one from the others ([85]) and can be ranked according to the constraints they pose on the involved entities (that can be either individuals or coalitions and groups or even the whole society) as:

$$descriptive \prec normative \prec prescriptive \tag{2.1}$$

where the binary relation $\prec$ is to be read as "poses less constrains than" and is endowed with classical and intuitive properties.

2.2.1 The descriptive approach

If we use a **descriptive approach** ([94]) we move within descriptive decision theory so that we ask how the individuals do behave in given choice or decision situations. In this case ([94]) we have to raise questions about what we need to know in order to be able to predict their behavior in such situations. The motto in this case is "what is" and an example of a descriptive approach is the use of the lexicon of a language for every days conversations. If we detail this approach we can state that descriptive decision theory ([94]) is typically **inductive** in nature. In this case ([94]) we start with observations of how individuals behave in given concrete classes of decision situations then we try to describe them with general rules and lastly we verify which general properties such rules satisfy. The properties we are interested in may concern **rationality** ([98], [21]), **fairness** and **equity** ([24], [25], [124], [125], [38]).
From these premises we can state that in this case we use a sort of **bottom-up** approach where we try to discover and evaluate the rules that govern the

decisions and this approach ([94]) requires that we classify both the actors[1] and the various situations in which they are involved.
According to this approach we need to derive the utilities (or at least the gains) that every individual assigns to the alternatives from the decisions they take whenever they face the problem of choosing among such alternatives. In this way we act according to a revealed preferences approach ([85]). If we are able to derive such utilities we can use them (assuming they do not change with time[2]) in order to predict the future choices of the actors.
In this case the quality of a decision ([94]) is referred to the systems of values of the deciders that generally are not known and cannot be postulated but can only be inferred in accordance with the inductive nature of the approach.

2.2.2 The normative approach

If we decide to adopt a **normative approach** ([94]) we move within normative decision theory so that we ask how the individuals ought to behave in given choice or decision situations.
The motto in this case is "what ought to be" and an application of this approach is in the specification of the use of the lexicon in certain professional or otherwise typical situations.
Within this approach, in order to understand the behavior of the individuals we must introduce the concepts of both preferences and values as well as the concepts of rationality and rational decisions in all their variations. In this case one of our "commandments" may be stated as follows ([94]):

> *if an individual is rational then he will behave in a way that is foreseen by our theories of rational behavior and in order to maximize his expected benefit or utility.*

Within this approach we can consider the following cases:

- one decider;

- two or more deciders.

In the former case we are within classical decision theory whereas in the latter case we can use the tools of GT though, if the deciders are more than two, we must consider the possible presence of coalition among them so that we must use both $NCGT$ and CGT.

[1] In general we use the terms actor, player and individual as synonyms though we prefer the former whenever we need a more neutral term and we use the term player whenever we want to stress a GT context.

[2] This is a rather strong assumption indeed.

If we detail this approach we can state that normative decision theory ([94]) represents a sort of **top-down** approach since it is deductive by its very nature. It indeed ([94]) postulates criteria of optimality or of rationality or of equity and ([94]) derives strategies or methods of allocation or of aggregation of preferences that are supposed to satisfy these criteria.
It is obvious that the verification of such postulates may occur only ex-post and, at the same time, it is well known the presence of impossibility theorems ([101], [102], [87], [32]) that prevent that certain set of properties (or wish lists) are all satisfied at the same time.
Within this approach we use a normative model where the guiding criterion is, in most of the classical approaches, that of optimality over known values and preferences (as expressed through utility values).
To deal with a normative approach we can use axiomatic approaches ([85]) as ways to shape the reality though in this way we may face some paradoxes ([85]) since there may be discrepancies between the top-down predictions made by the theory and the bottom-up realizations from real actors in real situations. The alleged paradoxes are also inherent in the theory of the competitive approach ([87], [101], [102]).

2.2.3 The prescriptive approach

If we decide to adopt a **prescriptive approach** we pose strong constraints on the normative approach so that the motto in this case becomes "what must be".
According to this approach we consider how individuals must behave according to prescribed norms and rules. The presence of prescriptions usually causes conflicts with the preferences and the values of the individuals so to trigger violations that require the devising of control tools and penalties if such violations are detected.
Examples of prescriptive approaches involve:

- the rule of the driving side in a given country;
- the rule of the writing direction for a given language;
- the rules to be adopted at crossroads with or without traffic-lights.

This approach can be characterized according to a layered framework. Within this layered framework we may identify:

- a general or top level framework;
- some distinct domains within that top level framework;

- the proper rules, with penalties and sanctions and dedicated authorities, for each domain.

An example of this structuring can be found in a given religion with its areas of prescriptions and the commandments for each area. Another far less binding example is a political creed or ideology with its areas of concern and the dogmas or the mottoes of the "founding fathers" for each area.

2.3 The levels

The term **level** needs to be disambiguated in the context of this thesis since it assumes at least two meanings.
It can indeed be used to describe the **granularity** of the interactions but it can also be used to describe a **position** in a hierarchy though in this thesis it can assume even other meanings in other contexts, both to be specified in due time.
If we use the term level to denote the **granularity of the interactions** we can identify the following levels ([105]):

(1) microlevel;

(2) mesolevel;

(3) macrolevel.

At the **microlevel** we see the individuals acting as singletons, at the **macrolevel** we see the society of the individuals or, in the jargon of GT, the possible grand coalition whereas at the **mesolevel** we allow the formation of coalitions and groups with their dynamics ([40]). In this section 2.3 we comment a little also the approaches suggested in [62].
We recall how with the term **society** we denote both the whole set of the individuals that, under certain circumstances, may form the so called grand coalition (dynamic aspect) and (see Figure 1.2) all the structures that outlive both groups and coalitions and individuals (static aspect).
If we consider the static aspect we usually capitalize the word whereas in other cases, mainly whenever we consider the dynamic aspect, we refer to the context to clarify which is the intended meaning. In any case the grand coalition may prove either stable or unstable.
If it is **stable** we say that the society has a strong cohesion and its members can safely share among themselves, according to a TU approach, the values of the games they are involved in.
If it is **unstable** there is no such cohesion so that the grand coalition tends to

split in smaller coalitions. In this case such coalitions behave has singletons in $NCGT$ games so to obtain, according to a TU approach, a value from every game in which they are involved and share such value among their members in a subsequent CGT step ([40]).
If we use the term level to denote the **position in a hierarchy** we use it:

- as a power index,

- to denote a position within a decision stream.

In the role of a **power index** a level defines the relative importance of every level with regard to both its superior and inferior levels.
In the role of a **position denoter** a level simply denotes the position of its belonging members within a two way decision stream from the lower levels to the upper ones (up to the uppermost level) and vice versa. For further details we refer to Chapter 4.
In both cases ([70]) the levels are not part of the fabric of the reality but are useful tools for both description and analysis (see Figure 2.1).
If we use the term level to denote the **granularity of the interactions** we have that the levels also are related to the types of observations we can make and each of them is more or less suitable for our approaches (see section 2.6 for further details).

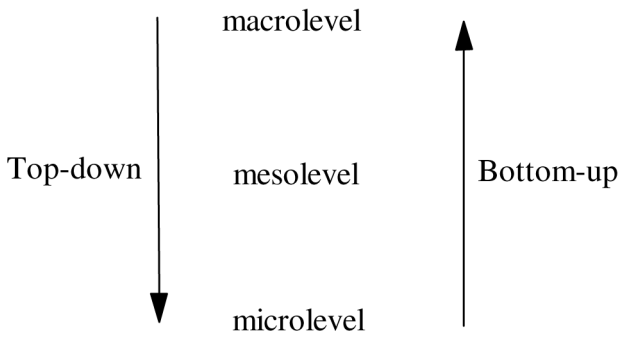

Figure 2.1: *Top-down and Bottom-up*

We recall how the mesolevel and the macrolevel provide the contexts for individual behaviors or actions under the form of both advantages and constraints and allow us to define links between individual behaviors and groups/collective behaviors in order to explain the emergence of collective phenomena at such levels.
In Figure 2.1 we give a sketch of both the Top-down and the Bottom-up flows

or approaches.
According to the Top-down approach (that we deal with in Chapter 4) we want to investigate how upper levels can influence the lower levels (see the thin arrows in Figure 2.2) and so how[3]:

- the macrolevel can influence both the mesolevel and the microlevel;
- the mesolevel can influence the microlevel.

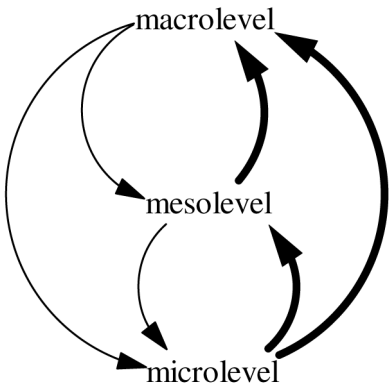

Figure 2.2: *Influence and emergence*

According to the Bottom-up approach (that we deal with in chapter 6) we want to understand in which ways and under which forms the behaviors at lower levels can **emerge** at higher levels (see the thick arrows in Figure 2.2) and so how:

- certain individual behaviors at the microlevel can give rise to certain aggregate behaviors at the mesolevel and at the macrolevel;
- certain group or coalition level behaviors at the mesolevel can give rise to certain aggregate behaviors at the macrolevel.

As a side effect of such **emergence** we have possible conflicts between norms and rules and realized behaviors. We call such conflicts the micro-macro

[3]We underline that in what follows we use some shortcuts so when we say "how the macrolevel can influence the mesolevel" it must be read as "how social norms and rules can shape the behaviors of groups or coalitions of individuals". Similar considerations hold also in the other cases and also when we speak of emergence since the levels are analysis tools and are not elements of the fabric of a reality.

conflict ([105], [106]) and we introduce it more formally in section 2.7.
The concept of **emergence** is very complex and is central to the sociological investigation. In this thesis we are going to use it only to denote the fact that certain phenomena cannot be explained at a level but need to be explained at a higher level under the hypothesis that there is no upper level to a society so defining a sort of closure or self-reference. With this we do not deny the existence of feedback links between the levels (that we represent in Figure 2.2 and that we aim at describing) but we simply deny both meaning and existence to any meta-society.

2.3.1 The microlevel

At the microlevel we see the individuals and their individual behaviors ([70]) within a history of repeated interactions that can be either with memory or memoryless ([12], [13], [50]).
In the expected case each individual has a more or less perfect recall of his past interactions. He is therefore able to remember, with a certain degree of precision, the past attitudes of all the individuals with whom he has interacted in the past. In this case an individual is able to attune his behavior in the current interaction with the expected behavior of the other individual with whom he is going to interact.
In the implemented case each individual has no memory of his past interactions so that he can choose its current attitude or behavior only according to:

- the outcome of the last interaction if he has a limited form of memory;
- the value of a cumulative performance index if he has a sort of integrative memory;
- the expected behavior of the individual with whom he is going to interact.

The last tactics ([89], [85], [18]) resembles fully the strategies adopted by the players in classical GT where each of them fully knows the games he is going to play and the pay-offs for him and for the other players in all the possible combinations of strategies.
At the microlevel, therefore, the individuals interact among themselves according to more or less complex strategies ([12], [13], [50], [105]) that depend on the expected or implemented strategies adopted by the others.
In the former case the individual is guided by his immediate future whereas in the later case he is guided by his more or less immediate past.
During their interactions each individual:

- can play one or more roles;

- can assign to each role a different weight or importance;

- can act strategically.

At this level we note that acting strategically means that an individual prefers to select a less preferred alternative if he fears that a even worse alternative might result as the final outcome or as the winning alternative within a decision process that involves more than one player ([101], [102]).
The individual behaviors can, moreover, depend on the presence of the groups and the coalitions since:

- groups and coalitions make riskier decisions than individuals;

- conformity increases with group and coalition size;

- inhibition decreases with group and coalition size.

From this we derive the descriptive but also normative (if not prescriptive) presence of the mesolevel.
On the other hand ([70]) the individual behavior depends on:

- preferences,

- decisions,

- actions

as the **fundamental units** on which higher levels patterns of behavior depend.
Such fundamental units can be used within a $NCGT$ approach in repeated games that start at an arbitrary initial time t_0 ([12], [13]) when the players start choosing their strategies for the interactions with other players[4].
At this level we can use the classical tools of $NCGT$ ([89], [85], [60]) in order to describe such interactions and which are the best (or equilibrium) strategies that the players can adopt. We note that the players are to be seen as inserted in an environment ([12], [13]) where they repeatedly interact with other players in games that involve two or more players but where they cannot form any coalition (that we assume appear at the mesolevel, section 2.3.2).

[4]This is one of the differences that we have with real cases where no initial conditions can be either fixed or known.

2.3.2 The mesolevel

The mesolevel is a connection layer between the microlevel (where the individuals interact as singletons) and the macrolevel where we see the behavior of the individuals at an aggregate level of detail.
The **mesolevel** is therefore an analytical tool for the description of the grouping tendencies of the individuals. To describe such tendencies we firstly introduce the concept of **group** and then we specialize it with the concept of **coalition** as it is used in Game Theory ([89], [85], [37]).
In the present thesis we consider the following alternative though related descriptions:

- one based on the use of groups,

- one based on the use of coalitions.

A group $\mathscr{G}_i$ is a subset of the set $\mathbb{I}$ of the individuals(see section 4.2.2). The groups do not form a partition of the set $\mathbb{I}$ but rather a covering. Every group is characterized by a **role** (see section 1.2) whereas each individual $\iota \in \mathbb{I}$ can hold more than one role at the same time, each with its relative importance or weight[5], so that an individual may belong to more than one group at the same time.
We note that:

- an individual must hold at least one role;

- if an individual dismiss a role he leaves the corresponding group;

- if an individual acquires a role he joins the corresponding group;

- if an individual plays a very particular role he may be the only member of the corresponding group.

We can impose that every individual $\iota \in \mathbb{I}$ holds only one role at a time so that the groups are disjoint and are termed as coalitions. In this case an individual:

- can change his role so to switch from one coalition to the one corresponding to his new role;

[5]We underline that only the roles with a non negative weight have a meaning for a given individual and that the higher a weight is the more important is the corresponding role. Since real individuals have limited capabilities we assume that the weights sum to one so that no individual can hold too many roles nor can give too importance to many of them.

- can assume a very particular role so to implement a singleton coalition.

In this case the coalitions form a partition of the set $\mathbb{I}$ of the individuals (see section 4.2.2).
The main difference, on which we are going to give more details shortly, between a group and a coalition is the following feature that can be used when we simulate the behaviors at the mesolevel:

- a group can never reduced to a single representative player,
- a coalition, if it is stable, can be represented as a single representative player.

According to the groups approach, the mesolevel we consider group behaviors as:

- the interactions within every group or **intragroup**;
- the interactions among the groups or **intergroups**.

The interactions of the latter type can be either **bilateral** (or between two groups at a time) or **multilateral** (or between more than two groups at a time).
In both cases such interactions are constrained by the fact that the members of a group belong to other groups as well moreover, within one or more interactions, a group may either shrink or grow.
It may shrink since some of its members dismiss his defining role since it causes a conflict with the interests associated to some other more important (and so with a higher weight) role. It may grow since an individual acquires one more role in addition to the ones he already holds and so enters in the corresponding group.
Let us see some examples. Take an individual $\iota \in \mathbb{I}$ that holds two roles, the one of citizen and the one of wholesaler. During an interaction she may say: "as a citizen I agree with this proposal but as a wholesaler I must consider my business and so I am of the opposite opinion". In this case it is as if he dismessed for a while the role of citizen so to keep only the role of wholesaler.
Another example is the one of an individual $\iota \in \mathbb{I}$ that holds three roles, the one of citizen, the one of mother and the one of worker. During an interaction he may say: "as a mother and as a citizen I think that that polluting factory must be closed though as a worker I am in solid for the people working there and that, owing this closing, will lose their jobs". As a last example we may consider the one of an individual $\iota \in \mathbb{I}$ who holds the role of student and then assumes the one of part-time worker so to be able to earn money and

pay his studies.
If the groups are **disjoint** we speak of coalitions (see section 4.2.2). In this case, at this level ([40]) we can use CGT to describe the interactions within the coalitions and $NCGT$ to describe the interactions among the coalitions where each coalition acts as an individual player (assuming it has a sufficient stability).
In the simplest synchronous cases ([64], [40]) we may have a cycle with three phases:

(1) one in which the coalitions interact among themselves in order to obtain a payoff (and so a gain if it is positive or a cost if it is negative),

(2) one in which each coalition shares among its members what it got at the previous step;

(2) a rearrangement phase where each coalition can either attract new members from other (possibly singleton) coalitions or lose members that pass to other coalitions or form singleton coalitions.

In phase (1) we can use the classical tools of $NCGT$ ([89], [85], [60]) in order to describe such interactions and which are the best (or equilibrium) strategies that the players can adopt.
In phases (2) and (3) we can use classical tools of CGT ([53], [89], [85]) to understand how the coalitions behave as to the relations among their members and how they can either attract new members or lose members toward other coalitions.
The cycle can be assumed to last forever[6] or to have a termination that is either fixed or depends on endogenous conditions and so on conditions that depend on the interactions themselves.
If we would admit asynchronous interactions we could not devise a single cycle for all the coalitions but we could have concurrent cycles involving subsets of coalitions that could migrate at their will from one interaction setting to another. In this thesis we do not deal with this complex framework any further so that we are going to move in the simplified and rather unrealistic setting of the synchronous three phases cycle.
We now go back to the issue of the representative player and make some more comments on it. As we have already stated a group cannot be replaced with a representative player so that we must resort to ad hoc [simulation]

[6]Whatever this may mean in a real finite world. Apart from any irony, with this we mean that the involved players have no idea about how long they will be going to interact among themselves in the future neither they have any memory about when they started their interactions.

techniques if we want to describe the interactions among groups. This impossibility derives from the very nature of the groups as a covering of the set of the individuals so that an individual can belong to different groups depending on his many roles.
What about coalitions? Under which conditions can a coalition be replaced by a representative or equivalent player? The key condition is a condition of stability.
A coalition is termed **stable** if it is of constant size or if it neither acquires nor loses members otherwise it is termed **unstable**. We assume therefore that an unstable coalition firstly must attain its stability and then can interact with the other coalitions in order to get a share of the whole game in which they involved.
Any stable coalition ([40]) can be replaced with a representative player so that, according to a Transferable Utility approach:

- during every interaction such representative player gets a payoff that is the worth for the coalition that he represents;

- when such a worth has been obtained it is shared among the members of the coalition according to some equity or fairness criterion ([38], [24], [25], [125]).

For the sharing of the worth for a coalition ([89], [85]) the members of a coalition can resort to mechanisms such as the **core** (or similar mechanisms) or the **Shapley value** (or similar mechanisms).
In both cases if the core is not empty every allocation of the worth among the members of a coalition is stable so that no sub-coalition has an incentive to leave the coalition to which it belongs in order to get a better share. This is obviously true if we act under the assumption of stable coalitions so that the Shapley value is only one of the possible ways to distribute the worth of a coalition among its members that belongs to the core.
In every case where we replace the coalitions with their representative players we can devise some parameters to describe their interactions in the same setting as the one we can use at the microlevel.
Such parameters include:

- the **strength** as the number of the players within a coalition;

- the **gain** or the total worth gained by a coalition during its past interactions;

- the **threshold** or the minimum level of gain from future interactions that assures the stability of the coalition.

2.3.3 The macrolevel

At the macrolevel we consider goal oriented collective behaviors so that we can state that:

- we have a **success** if the desired and planned collective behavior is implemented by the individuals (within acceptable intervals of tolerance),
- we have a **failure** and a micro-macro conflict (see section 2.7).

At this level we set both norms and rules and we consider the possible collective patterns of behavior. In this case we can have the following cases:

(1) we have one dominant behavior or strategy[7];

(2) we have two or more co-existing behaviors.

In the case (1) the dominant behavior is destined to prevail over the other (dominated) behaviors independently from their initial distributions among the individuals of a society. This means that in the long run all the individuals tend to adopt the same behavior and this adoption occurs earlier the stronger is the dominance of a behavior over the others. This fact may either depend on an intentional planning from the social planners or represent an emergence and so a behavior that cannot be explained but a result of cumulative interactions.
In the former case we usually have no micro-macro conflict (see section 2.7) whereas in the latter case such conflict occurs almost always since the emerging behavior conflicts with the behaviors that the social planner wished they were implemented by the stakeholders.
In the case (2) we have two or more behaviors that co-exist at either constant levels (possibly after a transitory period) or at oscillatory levels (as it occurs, in general, in the case of prey-predator systems represented with Lotka-Volterra type differential equations, [59]). Such behaviors depend heavily on the distributions of the types of the individuals within a population and on the frequencies of the encounters between individuals of either different or identical types.
At this level we can use the tools of *CGT* ([53], [89], [85]) with the aim of analyzing the stability of the **grand coalition**.
If the grand coalition is stable we obtain one or more co-existing behaviors from the stakeholders that can be either attuned to or in contrast with the

[7]In this dissertation we consider the terms behavior and strategy as synonyms since a strategy is expressed as a behavior and a behavior is the representation of a strategy though a given behavior can correspond to distinct strategies.

behaviors desired and planned by the social planners. In the former case we have no conflict whereas in the latter we have a micro-macro conflict (see section 2.7).
At the **macrolevel** the simulation is carried out at the population level as we are going to discuss in Chapter 5. In this case we can introduce the following work hypothesis:

- an open world hypothesis (see section 5.4) where we analyze the interactions among types under the assumption that new individuals can enter the interaction whereas others can disappear from it;

- a closed world hypothesis (see section 5.6) where we analyze only the migrations among types in order to understand if they can coexist or if one is destined to dominate over the others that, on the long run, die out.

2.4 The simulation tools

The simulation tools play a central role within our frameworks ([64]). They come in many forms and may be suitable at different levels of granularity ([2], [4], [9], [10]).
Though we are going to focus our attention on these tools mainly in chapters 5, 8 and 9 and present some background materials in section A.2 at this level we may state that:

- if they allow the description of individual behaviors they can be used at the microlevel and, under the hypothesis of stable coalitions, at the mesolevel;

- if they allow the introduction of more or less strong links between the individuals in order to describe more or less binding agreements and the rising of unstable coalitions and groups they can be used at the mesolevel;

- if they describe the set $\mathbb{I}$ at an aggregate level possibly in presence of co-existing behaviors they can be used at the macrolevel.

We recall how a behavior is a strategy used by the individuals of a certain type and that, depending on the distribution of the types over $\mathbb{I}$ we can have:

- two or more co-existing behaviors;

- one dominant behavior.

A way to have co-existing behaviors is through **segregation** ([105]). Segregation, or the grouping of the individuals based on some common knowledge features such as race or religious creed[8], may be either spontaneous (as a consequence of the autonomous and independent choices made by the individuals of different types) or forced by prescriptive social rules or norms (as it occurred with the segregation of the Jews and the "Negroes" in the ghettos and of the lepers in the leper hospitals).
If we use System Dynamics tools such as Vensim ([116], [95]) we can move only at the macrolevel and possibly at the mesolevel if we have to describe a small number of groups or coalitions. Examples of the use of such tools are represented by:

- the description of the interactions between two populations, one of so called preys and the other of so called predators where a type is fixed and characterizes the individuals of a given population ([64]);

- the description of the interactions between two (such as doves and hawks) or three types of strategies (such as doves, hawks and law-abiders) within a constant size population where each strategy can be adopted by the individuals depending on which types of individuals they interact with ([64]).

We recall how the former case is an example of what we have called the open world hypothesis (5.4) whereas the latter case is an example of what we have called the closed world hypothesis 5.6.
If, on the other hand, we use a tool like NetLogo ([86]) we can describe the behaviors of the individuals through the definition of simple rules of interaction either with other individuals or with the environment.
In this case the individuals are assumed to move (in many cases at random) in a finite world that they can modify in some way and in which they encounter other individuals with whom they interact. Such encounter cause variations in the status of every individual and these variations account for modifications of his future behavior (see chapter 8 for more details).
With tools like this one we can move at the microlevel and, under the hypothesis of stable coalitions, at the mesolevel. In order to do so we represent the various types of the individuals:

- internally through procedures that describe the interaction capabilities and local values of global variables that define the potentialities of each individual within each type;

[8]Or some other rather easily ascertainable features such as language, political membership or social class.

- externally through different shapes and colors that allow the differentiation among the various types.

If we have to describe the interactions either of unstable coalitions or of groups we can introduce more or less strong links among some of the individuals represented as agents.

2.5 The tools

2.5.1 Some preliminary remarks

As we have already hinted in section 2.1 both Mechanism Design (MD) and Implementation Theory (IT) are in the hands of the social planners whereas Game Theory (GT) defines the games that the stakeholders have to play in the various situations they can face since the preferences and utilities of the stakeholders usually are either only imperfectly known or are not known at all by the social planners. This is one of the reasons for the occurrence of the micro-macro conflict (see section 2.7). In sections 2.5.2 and 2.5.3 we very concisely present such tools and describe at which level they can be used and for what purposes.

2.5.2 Mechanism Design and Implementation Theory

MD ([54], [85], [89]) and IT ([82], [81]) are related with each other[9] since they are both concerned with the design of game forms such that it is a dominant strategy for each of the involved players to reveal truthfully his preferences. All this is in accordance with the so called **revelation principle** that focuses the attention on equilibria characterized by an **incentive compatibility** or by the fact the all the participants to a mechanism are better off if they truthfully reveal any private information they possess and that it is necessary for the working of the mechanism. Moreover in [89] they are considered as synonyms according to an approach that we are going to follow also in the present dissertation.
Both such tools can be used at the macrolevel since they represent the main tools that the social planners can use to shape the **game forms** of the games that they desire will be played by the groups and the coalitions (at the mesolevel) and by the players (at the microlevel). Roughly speaking a **game form** ([91]) is defined when the involved players are known as well as the strategies at their disposal. In order to have a **game** we must specify the

[9] In the present section we use also information loosely extracted from Wikipedia.

level of rationality of the players and their preferences among the available strategies. The players' preferences are expressed usually as numerical utilities or payoffs whereas their rationality is assumed to be full.
In some more recent versions of the theory it is introduced the possibility that the players are not fully rational and may make involuntary mistakes when they either choose one of their available strategies or perceive the strategy chosen by one of their opponents.
In many cases, however, the knowledge that the social planners have of the individuals and of the groups and the coalitions dynamics is rather fuzzy and incomplete so the game forms they prepare when turned into games usually give rise to outcomes that, at the best, only partially coincide with the expected or desired ones (see section 2.7).
The full treatment of MD and IT is out of the scopes of this dissertation so in this section we only outline their distinguishing features from an unitary perspective. With this we mean that we follow [89] (and so consider MD and IT as synonyms) but use one term or the other depending on the source me are referring to.
MD is based on a **normative approach** and ([54]) tries to understand what can happen if a given mechanism (or game form defined as a set of rules) is designed with the aim of achieving some desirable objective.
We underline how in classical MD the designer must take as a given the set of the player whereas he can choose at his will (for every player $\iota \in \mathbb{I}$):

- the **choices** available to the players or their sets of **strategies** S_ι,
- the consequences of the players' actions or the **outcomes** of the game form O_ι for any profile of strategies:

$$S = \Pi_{\iota \in \mathbb{I}} S_\iota \tag{2.2}$$

We underline how, in classical MD theory, both S_ι and O_ι can be quite arbitrarily chosen by the designer at the moment when he decides to design a given mechanism for his purposes. This may not be feasible within our context where both sets may be constrained from outside the design process from pre-existing and unavoidable constraints.
Moreover the designer acts under the following constraints:

- incentive compatibility;
- individual rationality.

With **incentive compatibility** we mean that each player chooses the strategies designed for his own type preferring them to those designed for other

types of players.
With **individual rationality** we mean that there is no coercion and that the participation in a mechanism is voluntary. More precisely the **individual rationality** constraint means that no player would willingly participate to a mechanism in which his best utility is lower than the utility he would get by not participating or, in other terms, means that a player is expected to have a non negative expected utility if he decides to participate to a mechanism.
In the designing of a given mechanism the classical theory of MD provides the designer with the **revelation principle** that says that there is always a direct revelation mechanism that corresponds to any generic mechanism so that the two mechanisms produces the same outcomes. This principle allows the designer to consider only the Nash Equilibria that are characterized by incentive compatibility and individual rationality.
In this way if the designer wants to implement some outcome or some property he can consider only the so called **direct revelation mechanisms** or mechanisms where:

- the players have incentives to get involved;
- the players have incentives to reveal truthfully their preferences on the possible outcomes of a game.

In this case ([54]) we speak of direct and truthful mechanisms. If at least one of them exists then the designer will be able to implement a given outcome or property otherwise such implementation is destined to fail.
When a mechanism is turned in a game the players, in full autonomy from the designer, introduce within such abstract context:

- their types τ_ι;
- their utilities or payoffs $\pi_\iota(s, o_\iota, \tau_\iota)$ for every $s \in S$.

In this way the players define, through their preferences that have been assumed in the design but that are not perfectly known by the designer, their own Nash Equilibria among which they select one equilibrium, possibly according to principles of Pareto dominance.
The general problem with MD ([24]) is that the players' revelation of their truthful valuations may not be a Nash Equilibrium (so that one or more players have incentives of deviating from adopting it). A possible solution may be to motivate the use of truthful revelation through dominant strategies. Notwithstanding this, in many cases ([24]) honesty induced according to the revelation principle by using direct mechanisms does not provide efficient solutions so that there are solutions where at least one of the players is

better off without none of the others being worse off.
Other problems that the designer may face, especially within our context, are the following:

- when transforming a game form in a game the players devise unforeseen strategies and, correspondingly, new Nash Equilibria;

- the strategies, by their very definition, involve individual players whereas, in our context, we are in presence of players that can devise group or coalition strategies that boycott the strategies assumed by the designer.

2.5.3 Game Theory

GT by its very nature ([89], [85], [37]) deals with the interactions among rational players possibly with the intervention ([68]) of an external neutral player, the so called nature, that distributes at random the types to the players before the start of a game.
GT comes in two "flavors"[10]: *NCGT* and *CGT* ([89], [85]).
In *NCGT* we are interested in the strategic interactions among the players and focus on the attainment of a Nash Equilibrium or of one of its variants. *NCGT* can be used:

- at the microlevel to describe the interactions among the individuals as competing players;

- at the mesolevel to describe the interactions among stable competing coalitions.

As we have already seen (see section 2.3.2), at the mesolevel each stable coalition can be replaced by a representative player so that the interactions among m coalitions can be seen as the interactions among m players under the constraint that they cannot form any coalition.
The main features of *NCGT* we are interested in are:

- the repeated interactions among players through repeated games;

- the behaviors that no player or no coalition has an interest to abandon unilaterally or the equilibria of the games;

[10]We recall how tha acronym *NCGT* denotes Non Cooperative Game Theory whereas the acronym *CGT* denotes Cooperative Game Theory.

- the ways to identify the outcomes of the games through the possible definition of the strategies or the behaviors that the players or the coalitions will adopt.

In CGT we are interested to the formation of coalitions among players, to their stability conditions and to how each coalition shares among its members the gains it gets from the interactions with other coalitions ([40]).
CGT, in its turn, comes in two "sub-flavors"([89], [85]): with transferable utilities (TU) or with non-transferable utilities (NTU) (see Appendix B for further concise details).
In both its "sub-flavors" it can be used:

- at the mesolevel for the single coalitions;

- at the macrolevel for the grand coalition.

The main features of CGT we are interested in are:

- how the individuals join to form coalitions;

- under which conditions such coalitions are stable or can grow by attracting new members or can shrink by losing members ([40]);

- how each of them can share among its members the gain or worth or whatever it gets through the interactions with other coalitions.

The main difference between the TU and the NTU versions is represented by the ways in which the aforementioned worth is shared.
In the TU case the worth is associated to the grand coalition or to smaller coalitions and is shared among their members according either to set based solutions (such as the core or the stable set) or to point-wise solutions (such as the Shapley value).
In the NTU case every player claims his own worth so that the coalition to which he belongs must attain an outcome that guarantees him of obtaining such value.
In this dissertation we are going to adopt only the TU approach so to focus our attention on equity or fairness issues ([38]).

2.6 The relations

In this section we establish the possible relations firstly between the three approaches we have presented in section 2.2 then between the three levels we have introduced in section 2.3 and lastly between the approaches and the

levels.
For what concerns the relations between the three approaches we refer to Figure 2.3 where the directed arrows must be read as "assumes" so that $A \longrightarrow B$ means A assumes B or that B is a prerequisite for A.

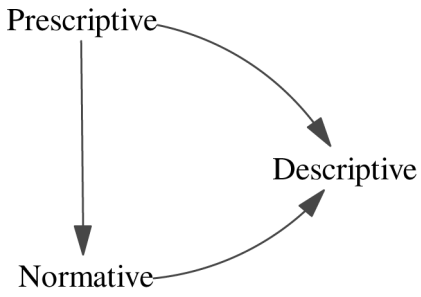

Figure 2.3: *Relations between the approaches*

If we recall that a **descriptive** approach can be used to make predictions about the behaviors of the individuals as well as of the coalitions considered as individual players it is easy to see how:

- a normative (N) approach assumes a descriptive (D) approach;
- a prescriptive (P) approach assumes a descriptive approach;
- a prescriptive approach assumes a normative approach.

Paraphrasing, in order to understand how things must or should go[11] we must firstly understand how they go and how they will probably go in the near future. In a similar way in order to understand how things must go we must firstly understand how they should go since the prescriptive approach is a constrained version of the normative approach.
From such considerations and from Figure 2.3 it is easy to see how the binary relation $\longrightarrow$ satisfies transitivity. Moreover from Figure 2.3 we derive how the descriptive approach can be seen as the basic step also in order to reduce the incidence of the micro-macro conflicts (see section 2.7).
In any case such approaches can be used one independently from the others. The compliance with the foregoing relations assures a higher robustness of each approach since, for instance, if we consider the relation:

$$N \longrightarrow D \tag{2.3}$$

[11] And so in order to evaluate the possible outcomes of either normative or prescriptive policies.

we (as social planners) may be able to devise norms that are more coherent with or that conflict less with the preferred/adopted behaviors of the stakeholders or that contrast less with their interests or that contain the necessary incentives.
As an example we can consider a situation where the social planners want to impose a door-to-door garbage collection system to the stakeholders in substitution of a bins based collection system without performing any descriptive analysis of the situation so to understand the objective problems of the stakeholders that should keep the garbage at their homes to have it collected at scheduled times. If such a system is adopted as a rule without such descriptive step it will almost surely fail since the stakeholders will tend to dispose of their garbages in the wrong places and at the wrong times.
If the pre-requisites are neglected the devised norms, rules and prescriptions are contrasted and violated so that, in order to have them respected, it is necessary to introduce controls and sanctions with additional costs that weigh on the set of the stakeholders as a whole.
If we consider the prescriptive (P) approach as a constrained version of the normative approach we can make a concise comparison between the normative approach (N) and the descriptive approach (D) as follows:

- central problem:

  N: how the deciders ought to act in a given situation;

  D: how the deciders actually act in a given situation;

- systems of values:

  N: are assumed so that we can define good and bad decisions;

  D: are inferred as an assignment of utility values to the outcomes of a decision process;

- utilities:

  N: are given a-priori;

  D: are derived a-posteriori from the outcomes of the decisions;

- rationality:

  N: is assumed for every decider[12];

[12]The definition of the concept of rationality is one of the central task of the theory ([94]). The rationality involves the decisions and is represented through the preferences over a set of alternatives A. In common cases we define a preference relation P that in order to

D: is derived from the outcomes of a decision process.

For what concerns the relations between the three levels we underline how we put the analysis at the more suitable level:

- for our purposes;

- for the conditions we face.

If, for instance, the grand coalition is stable we can move at the macrolevel and use the mesolevel to gain insight in the dynamics of the subcoalitions (if we can associate each subcoalition to one of the dominating behaviors) and at the microlevel, if we want to understand the motivations of the individuals so to maintain stability of certain behaviors until this is necessary.
If the grand coalition is unstable so that we are in presence of smaller coalitions (and so of conflicting behaviors) it is better for us to move at the mesolevel and at the microlevel.
If no coalitions form we have only the microlevel at our disposal.
Last but not least for the last purpose (and so a description of the relations among the approaches and the levels) we use Table 2.1 where we have that:

s means strong usefulness;

a means average usefulness;

w means weak usefulness.

	descriptive	normative	prescriptive
microlevel	s	s	w
mesolevel	s/a	s/a	a
macrolevel	w	w	a

Table 2.1: *Approaches and levels*

From an analysis of Table 2.1 if we proceed by rows (so to see, given a level, which approach is more appropriate to that level) we see how:

be said rational must satisfy the following criteria of consistency (where $a, b, c \in A$):

- asymmetry so if aPb then $\neg\, bPa$;
- transitivity so if aPb and bPc we have aPc;
- completeness so for any pair a, b we have either aPb or bPa.

- at the microlevel we can expect that both the descriptive and the normative approaches prove strongly useful (as it is shown for instance by the uses of GT in such contexts) whereas the prescriptive approach is much less useful;

- at the mesolevel both the descriptive and the normative approaches lose some usefulness since the interplay of the coalitions and groups hide the independent behavior of the individuals whereas the prescriptive approach gains some usefulness since it is easier to put constraints on groups rather than on individuals;

- at the macrolevel we can describe the whole society so both the descriptive and the normative approaches have the lowest utility whereas the prescriptive approach keeps its usefulness unchanged.

If, on the other hand, we proceed by columns (so to see, given an approach, at which level it is more appropriate) we see how:

- the descriptive and the normative approaches lose their usefulness from the microlevel to the macrolevel;

- the prescriptive approach gains some usefulness from the microlevel to the mesolevel.

2.7 The micro-macro conflict

At the core of this dissertation we have the analysis of the micro-macro conflicts ([105], [106]). The clearest expression of a micro-macro conflict is a more or less strong mismatch between:

- the desired behaviors B_d from the social planners

  and

- the effectively implemented (by the stakeholder) behaviors B_i.

With this we do not mean that the elements of the set B_d are right and those of the set B_i are wrong or vice versa, whatever it may be the meanings of right and wrong. We only say that the two sets do not coincide or that:

- some desired behaviors are not implemented,

- some undesired behaviors are implemented.

A measure of this mismatch may be represented as:

$$\rho = \frac{|B_d \cap B_i|}{|B_d \cup B_i|} \in [0, 1] \tag{2.4}$$

In this way we have that:

- if $\rho = 0$ the mismatch is the highest since the sets B_d and B_i have no element in common;
- if $\rho = 1$ the mismatch is the lowest since the sets B_d and B_i have only common elements;
- if $\rho \in (0, 1)$ the mismatch is variable from the highest to the lowest since the sets B_d and B_i have only some elements in common.

If any mismatch is present it is revealed by the social planners only in an ex-post phase when the implemented behaviors are detected, essentially through their revealed effects. We note, moreover, how there may be present transitory phases that make harder the revelation of such mismatches.
With this we mean that a corrective action may require time to be devised, implemented and exert its effects so too early detections may give rise to:

- over-corrections,
- under-corrections (less frequently)

that, at the end, may result in other unwanted behaviors.
In addition to all this, such revelation requires the availability of control mechanisms and structures and suffers from inevitable delays and costs though the presence of such structures may or should act as a deterrent[13] against undesired behaviors.
Moreover we must consider that not all the effects can be revealed and associated to given behaviors though in this thesis we keep the analysis at a simple level so to discard such refinements.
The presence of the mesolevel introduces a further complexity in this scheme since (as it will be examined in the chapters from 6 inclusive to 9 inclusive) the groups and the coalitions act as "mediation" elements that:

[13]We note how the effect of deterrence increases with the fear of being discovered and punished in effective ways but decreases whenever past violations, though detected, have not been punished but rather remitted in some way. In these cases ethics can play relevant roles.

- may facilitate the adoption of norms and prescriptions by the stakeholders that belong to certain groups or coalitions (as it occurs, for instance, in case of professional, political, cultural or environmental associations);

- my hinder such adoption through corporative actions of any sort.

In the former case the groups and the coalitions may act as second level controllers so to enforce the adoption of the planned and desired behaviors. In the latter case, on the other hand, they may act as structures that enforce deviant behaviors by giving them a sort of legitimacy and some "common behavior" justifications.

2.8 The nature of the micro-macro conflict

In this closing section of the chapter we present the main reasons ([16]) for the arising of the micro-macro conflict as we have presented it in section 2.7 and whose analysis and treatment is the core part of the present dissertation. This section, of course, can give only some hints of the complexity of the interactions between the norms[14] that are devised by the social planners and those that arise from the interactions among the stakeholders (or social norms, see further on) and that can conflict, in various ways ad to various degrees, with the norm of the former type.
As we have seen in section 2.7 the micro-macro conflict can be seen as a measurable mismatch between:

(1) some desired behaviors

and

(2) some effectively implemented behaviors.

The behaviors of the (1) type are fixed by the social planners within either a normative or a prescriptive perspective. They depend on formal norms ([16]) that are written and codified and that are designed as a part of the social planners' activity (as we are going to describe it in Chapter 4) and then are imposed and enforced through both incentives and sanctions.
The **formal norms** have both prescriptive (or positive) and proscriptive (or negative) features that characterize both the normative and the prescriptive

[14]Under the term **norms** we include formal elements such as laws, rules and regulations as well as informal elements such as the social norms as they are defined in [16]. The context makes clear what we refer to together with its being either formal or informal.

approach.
Their **prescriptive features** define positive aspects or how the stakeholders are either expected to behave or must behave and so what they may or must do.
Their **proscriptive features** define negative aspects and so what the stakeholders cannot do so that these features are characterizing of a prescriptive perspective.
We note here how the absence of some proscriptive features may be assumed by the stakeholders as a justification of unwanted (from the social planners) individual behaviors but this conflict, in absence also of the corresponding prescriptive features, cannot be seen as an example of a micro-macro conflict. In other words we have that such a conflict can arise only in presence of explicitly stated either positive or negative norms that are contradicted by the effective behaviors of the stakeholders.
On the other hand ([16]), the behaviors of the (2) type are implicit in the operations of a society and arise from the interactions among the stakeholders.
Essentially such behaviors depend on self fulfilling expectations in the sense that ([16], page *ix*) if some stakeholders believe that a sufficiently large number of other stakeholders uphold a given norm then, under the right circumstances, they will conform to it.
These behaviors, therefore, depend heavily on **informal norms** that emerge either within or from the interactions among the agents and are not designed and imposed by the social planners. These informal norms are termed in [16] **social norms** and develop spontaneously from the interactions of the agents without either an explicit planning or a deliberate design.
The **social norms** are therefore social constructs that arise in certain circumstances and have a life of their own since a sufficient number of agents believe they exist and so act accordingly.
In order to understand how this can happen in the simulation oriented chapters that follow we are going to present some models where a small set of individuals of a certain type is "inoculated" in a society with the aim of showing if their behavior spreads or vanishes in the overall population.
It is obvious how in many cases a social norm conform to a corresponding formal norm. This may happen for the following reasons:

- **normative expectations** so that the preference for conformity is conditional on the expectations to conform from the others;

- **prescriptive expectations** so that the preference for conformity is conditional on the presence of punishments of and sanctions for the transgressions;

- **legitimacy expectations** so that the preference for conformity is conditional on the recognition of the legitimacy of the others' expectations that an agent will follow a norm.

In all these cases we have different types of agents with conditional preferences for conformity. What is common to such agents is the belief that conforming to a norm is worthwhile if enough other agents are already obeying the norm.
The foregoing expectations are grounded on:

- personal direct knowledge from past interactions,

- indirect knowledge from observed behaviors.

Such reasons do not exclude the existence of deviant behaviors but, in the case of conformity, such behaviors are confined within a small set of[15] free-riders or defectors and do not spread within the society to larger and larger subsets of stakeholders.
If such a conformity is absent this may depend on the lack of the foregoing expectations in any possible combination. With this we mean that some agents behave differently from a formal norm since:

- they do not assign any legitimacy to the conformity of the others;

- they do not believe that transgressions can be detected but, even if they are detected, they will be remitted in some way;

- they do not assign any legitimacy on the expectations of conformity.

In addition to social norms we could also consider also ([16]):

- descriptive norms,

- conventions.

Descriptive norms (such as fashions and fads) involve both coordination and imitation among the agents whereas conventions induce coordination among the agents. In the present dissertation we focus our attention mainly on social norms where we usually have no coordination among the agents but there is rather a tension between individual and collective utilities.
The main reason for this choice is twofold: firstly we had neither the time nor the competences to engage ourselves in such a wide spectrum task and secondly we had no chance to make better than Cristina Bicchieri in her book ([16]).

[15] A free-rider is an individual who defects from a policy in order to get a known ex-ante gain whereas a defector is an individual who behaves against a policy in order to get an ex-post potentially uncertain gain.