Chapter 9
On the Quality of Collective Decisions in Sociotechnical Systems: Transparency, Fairness, and Efficiency

Daniele Porello

9.1 Introduction

Decision-making in organization is a wide area that usually relies on formal methodology such as decision theory and game theory and on empirical investigations of actual decision-making in organizations. The aim of this paper is to propose a rather different question and to introduce a methodology to approach it: How can we conceptualize the quality of collective decisions made within the context of a complex sociotechnical system? Sociotechnical systems (STS) are complex organizational scenarios in which human agents interact in a normative constrained environment with themselves and with artificial agents (Emery and Trist 1960). For example, an understanding the organizational structure of an airport requires understanding the interaction between agents operating with metal detectors, sensors, and security cameras, as well as interacting with customers in a normatively specified way.

Defining STS is a complex task. Here we have decided to highlight the features of STS that are significant for understanding decision-making in this case. We view the complexity of STS as due to the entanglement of several layers of information—e.g., normative, perceptual, factual, conceptual—as well as of information sources, e.g., human, artificial, normative.

The quality of collective decisions in STS is evaluated by using the following three fundamental concepts: *transparency*, *fairness,* and *efficiency.* The key role of transparency in sociotechnical design was first stressed in (Guarino et al. 2012) and it has been argued that transparency is very important to enhance the adaptivity and resiliency of systems.

We conceptualize the *transparency* of a collective decision in terms of the entitlement of the agents involved in the systems to a justification of the decision made by that system. That is, the agents involved in the system (e.g., employee,

D. Porello (✉)
Laboratory for Applied Ontology, Institute of Cognitive Sciences and Technologies (ISTC-CNR), Via alla Cascata 56 C, 38123 Trento, Italy
e-mail: danieleporello@gmail.com

F. Cecconi (ed.), *New Frontiers in the Study of Social Phenomena,*
DOI 10.1007/978-3-319-23938-5_9

customers, users) are entitled to know the procedure that has been used to make the decision. Moreover, the choice of such a procedure has to be justified to them. Thus, a transparent decision has to be justified to those who are affected by the decision.

We conceptualize *justifications* of decisions in terms of *fairness* and *efficiency*. Intuitively, fairness is understood as non-arbitrary discrimination between the sources that are involved in the collective decision. For instance, a fair decision among stakeholders does not arbitrarily weight one's vote more than another. *Efficiency* is related to the rationality of the outcome. In decision theory or game theory, it is related to a maximization of an expected desirable value that is attached to the collective decision (Neumann and Morgenstern 1944).

We shall model fairness and efficiency conditions by means of techniques developed in welfare economics that have been recently used also in Multiagent Systems and Artificial Intelligence (Boella et al. 2011; Brandt et al. 2013; Woolridge 2008). In particular, we propose approaching the problem by using the methodology of *social-choice theory* (SCT) (Arrow 1963; Taylor 2005). SCT is a branch of welfare economics that studies the procedure for aggregating a number of possibly different individual preferences or choices into a collective preference or choice. An example of application of social-choice theory is voting theory, that is, the study of the property of voting procedures such as the majority rule. The reason that social-choice theory is a good methodology for investigating collective decisions is that it allows for specifying in a formal and clear way a number of properties that capture qualitative aspects of decisions. Those properties express, for instance, whether a procedure discriminates between individuals, whether the criterion of the choice has to be valid regardless the context of the decision, whether any issue to be decided has the same weight, and so on.

Moreover, social-choice theory provides an abstract treatment of collective decision-making that can be instantiated in a number of scenarios and allows us to check whether a certain procedure satisfies a number of qualitative desiderata. In particular, we shall use social-choice theory and judgment aggregation. The reason is that, as we shall see, those techniques provide versatile tools to model the aggregation of heterogenous types of information, and they allow for spelling out the properties of each type of aggregation procedure. The properties of aggregation procedure, or of decision procedures, then provide tools to model the concepts of justification of decisions that we look for.

Collective decisions are defined here not only as decisions made by a group or a team of individuals, such as committees, but also decisions that are made by the chief of a sector within the organization that is supposed to decide after gathering information coming from heterogeneous sources.

The application of social-choice theory to model collective decisions in sociotechnical systems requires a careful examination of the matter of possible decisions.

As we have recalled, a fundamental aspect of sociotechnical systems is the entanglement of heterogeneous layers of information. Therefore, we need to describe in an abstract and general way the types of information that are involved in complex sociotechnical systems.

In order to address and conceptualize this type of information, we shall use a foundational ontology. In particular, we shall exemplify our treatment by using DOLCE (Masolo et al. 2003, 2004) because it is capable of addressing the interconnection between different modules that gather different types of information, e.g., social, perceptual-mental, physical, organizational (cf. Boella et al. 2004; Bottazzi and Ferrario 2009; Porello et al 2014; Porello et al 2013).

The remainder of this paper is organized as follows. In Sect. 9.2, we informally discuss the background of social-choice theory and judgment aggregation. In Sect. 9.3, we present a model of judgment aggregation and we discuss the properties that formalize conceptions of fairness and efficiency. Section 9.4 presents our treatment of heterogeneous information in sociotechnical systems by means of DOLCE ontology. Section 9.5 approaches the problem of assessing the quality of decisions in sociotechnical systems by instantiating the methodology of judgment aggregation to possible scenarios of rich information entanglement.

9.2 Background on Social-Choice Theory and Judgment Aggregation

Social-choice theory originated through the seminal work of Kenneth Arrow (Arrow 1963), who provided a general framework for preference aggregation, namely, the problem of aggregating a number of individual conflicting preferences into a social or collective preference.

Take the following example: Suppose that a committee of three individuals (label them 1, 2, and 3) has to decide which security protocols to implement among three possible alternatives say: *a*, *b,* and *c*. In many settings of social-choice theory, preferences are assumed to be linear orders, that is, individual preferences are supposed to be *transitive* (an agent prefers *x* to *y* and *y* to *z*, then she/he should prefer *x* to *z*), *irreflexive* (an agent does not prefer *x* over *x*), or *complete* (for any pair of alternatives, agents know how to rank them, *x* is preferred to *y* or *y* is preferred to *x*).[1]

Suppose agents' possibly conflicting preferences can be faithfully represented by the following rankings of the options. Preference profiles are lists of the divergent points of view of the three individuals, as in the following example:

1. a > b > c
2. b > a > c
3. a > c > b

In the scenario above, the agents have conflicting preferences and there is no agreement on which is the best policy to be implemented. Since the policies are alterna-

[1] These conditions are to be taken in a normative way. They are not, of course, descriptively adequate, as several results in behavioral game theory show. However, the point of this approach is to show that even when individuals are fully rational—i.e., they conform to the rationality criteria that we have just introduced—the aggregation of their preferences is problematic.

tive, 1 and 3 would pursue *a*, whereas 2 would pursue *b*. In order to decide a collective option, we need a procedure that can settle the possible disagreement.

Suppose now that the individuals agree on a procedure to settle their differences; for example, they agree on voting by *majority* on pairs of options. Thus, agents elect the collective option by pairwise comparisons of alternatives. In our example, *a* over *b* gets two votes (by 1 and 3), *b* over *c* gets two votes (by 1 and 2), and *a* over *c* gets three votes. The majority rule defines then a social preference $a > b > c$ that can be ascribed to the group as the group preference.

The famous Condorcet paradox shows that it is not always the case that individual preferences can be aggregated into a collective preference. Take the following example:

1. a > b > c
2. b > c > a
3. c > a > b

Suppose agents again vote by majority on pairwise comparisons. In this case, *a* is preferred to *b* because of 1 and 3, *b* is preferred to *c* because of 1 and 2; thus, by transitivity, *a* has to be preferred to *c*. However, by majority also *c* is preferred to *a*. Thus, the social preference is not "rational," according to our definition of rationality, as it violates transitivity.

Kenneth Arrow's famous impossibility theorem states that Condorcet's paradoxes are not an unfortunate case of majority aggregation; rather they may occur for any aggregation procedure that respects some intuitive fairness constraint (Arrow 1963). In the next section, we shall discuss in more detail the formal treatment of the intuitions concerning fairness and we shall define a number of properties that provide normative desiderata for the aggregation procedure.

A recent branch of SCT, Judgment Aggregation (JA) (List and Pettit 2002; List and Puppe 2009) studies the aggregation of logically connected propositions provided by heterogeneous agents into collective information. The difference with preference aggregation is that in this case anti-type propositional attitudes can in principle be taken into account.

For example, take three sensors whose behavior can be described by the following propositions *C* "the alarm triggers" whenever *A* "metal is detected" or *B* "liquid is detected." In propositional logic this amounts to assuming that each sensor satisfies the constraint: $A \vee B \rightarrow C$

Suppose the three sensors 1, 2, and 3 provide different responses, each compatible with the above constraint.

	A	A ∨ B	B	A ∨ B → C	*C*
1	Yes	Yes	Yes	Yes	Yes
2	No	No	No	Yes	No
3	No	Yes	Yes	Yes	Yes

In this case, a conflict may emerge from the fact that the three sensors may have divergent sensitivities on detecting *A* or *B*. One can study the aggregation procedure in order to define a notion of collective information provided by the aggregated behavior of the detectors.

In order to do that, one can choose a number of policies to aggregate sensors' information in order to define a sort of collective sensor. If we select *unanimity* in the example above, no proposition, besides the constraint, is elected as collective information, thus the collective sensor does not trigger any alarm. If the *majority rule* is used, then the collective information is given by all the propositions at issue; therefore the alarm triggers.

Analogously to the case of Condorcet's paradox in preference aggregation, situations of inconsistent aggregations of judgments have been individuated. These paradoxical situations have been labeled in the literature *doctrinal paradoxes* or *discursive dilemmas*. It is important to notice that such paradoxical situations actually occurred in the deliberative practice of the U.S. Supreme Court (Kornhauser and Sager 1993). This problem has been perceived as a serious threat to the legitimacy of group deliberation and it has been considered a seminal result in the recent debate on the rationality of democratic decisions (Kornhauser and Sager 1993; Pettit 2001).

We show an example of such a paradox by slightly modifying the previous example. Suppose 3 rejects B because she/he rejects the premise A.

	A	$A \vee B$	B	$A \vee B \rightarrow C$	C
1	No	No	No	Yes	No
2	Yes	Yes	No	Yes	Yes
3	No	Yes	Yes	Yes	Yes
Majority	*No*	*Yes*	*No*	*Yes*	*?*

By majority, A and B fail, so they are collectively false; however, the $A \vee B$ pass, which is inconsistent in classical logic. That would mean that the alarm triggers even in the case that none of A and B is collectively satisfied.

Such paradoxes does not exclusively concern the majority rule; they also apply to any aggregation procedure that respects some basic fairness desiderata. This is the meaning of the theorem proven by Christian List and Philip Pettit (List and Pettit 2002).

Therefore the notion of collective decision and collective information requires a careful examination of the aggregation procedures that provide viable solutions. In the next sections, we shall sketch a model for defining collective decisions, and we shall place it within sociotechnical systems.

9.3 A Model of Judgment Aggregation

We present the main elements of the formal approach of judgment aggregation (JA). The reason we focus on JA is twofold: on the one hand, it considered to be more general than preference aggregation (List and Pettit 2002); on the other hand, it has been claimed that JA can provide a general theory of aggregation of propositional attitudes (Dietrich and List 2009). Propositional attitudes, such as beliefs, desires, preference, and judgments, model the relationship between an agent and a sharable content.

Propositional attitudes have been extensively discussed in analytic philosophy, and formal languages for modeling propositional attitudes have been proposed by several contributions in philosophical logic (e.g., van Benthem 2011). Therefore, JA provides the proper level of abstraction for placing our model of decisions based on heterogeneous types of information.

Throughout this section, we shall refer to the individual sources of information in the system as individuals, who may represent actual human agent of the systems as well as sensors.

The content of this section is based on List and Pettit (2002) and Endriss et al. (2012) and builds on them. Let P be a set of propositional variables that represent the contents of the matter under discussion by a number of agents. The language L is the set of propositional formulas built from P by using the usual logical connectives (e.g. $\neg, \wedge, \vee, \rightarrow$).

Definition 1 An agenda X is a finite nonempty subset of L that is closed under (non-double) negations.

An agenda is the set of propositions that are evaluated by the agent in a given situation. In the examples of the previous section, the agenda is given by $A, B, A \vee B, A \vee B \rightarrow C, C,$ plus their negations that allow us to model rejection of a certain statement: The rejection of a matter A is then modeled by an agent accepting $\neg A$. We define individual judgment sets as follows.

Definition 2 A judgment set J on an agenda X is a subset of the agenda J. We call a judgment set J **complete**, if for every formula in the agenda X, either A is in J or $\neg A$ is in J. We call J **consistent** if there exists an assignment that makes all formulas in J true.

We assume the notion of consistency that is familiar from logic. These constraints model a notion of rationality of individuals; i.e., individuals express judgment sets that are rational in the sense that they respect the rules of (classical) logic.

Denote with $J(X)$ the set of all complete consistent subsets of the agenda, namely, $J(X)$ denotes the set of all possible (rational) judgment sets on the agenda.

Given a set $N = \{1, \ldots, n\}$ of individuals, denote with $\boldsymbol{J} = (J_1, \ldots, J_n)$ a profile of judgment sets, one for each individual. A profile lists all the judgments of the agents who are involved in the collective decision at issue.

We can now introduce the concept of aggregation procedure. The domain of the aggregation procedure is given by $J(X)^n$, namely, the set of all possible profiles of individual judgments. The value of the aggregation function is assumed to be a set of judgment, i.e., an element of the power set $P(X)$.

Definition 3 An aggregation procedure for agenda X and a set of N individuals is a function $F: J(X)^n \rightarrow P(X)$.

An aggregation procedure maps any profile of individual judgment sets to a single collective judgment set. Given the definition of the domain of the aggregation procedure, the framework presupposes individual rationality: all individual judgment sets are complete and consistent. Note that we did not yet put any constraint on the collective judgment set, i.e., the result of aggregation, so that at this point the procedure may return an inconsistent set of judgments.

This is motivated by our intention to study both consistent and inconsistent collective outcomes. For example, in the doctrinal paradox of the previous section, the majority rule maps the profile of individual judgments into an inconsistent set. The consistency of the output of the aggregation is defined by the following properties.

Definition 4 An aggregation procedure *F*, defined on an agenda *X*, is said to be *collectively rational* if *F* is

- *complete* if $F(\boldsymbol{J})$ is complete for every profile $\boldsymbol{J}$ in $J(X)^n$;
- *consistent* if $F(\boldsymbol{J})$ is consistent for every profile $\boldsymbol{J}$ in $J(X)^n$.

That is, collective rationality forces the outcome of the procedure to be rational in the same sense of the individual rationality. Of course, the case of doctrinal paradox violates collective rationality.

We now introduce a number of properties— usually called axioms in social-choice theory—that provide a mathematical counterpart of our intuition on what a fair aggregation procedure is. The following are important axioms for JA discussed in the literature (Kornhauser and Sager 1993; List and Pettit 2002):

Unanimity (**U**): If for all agents *i,* a formula *A* is in J_i, then *A* is in $F(\boldsymbol{J})$.

Anonymity (**A**): For any profile $\boldsymbol{J}$ and any permutation of the individuals $\sigma: N \rightarrow N$, we have that $F(J_1, \ldots, J_n) = F(J_{\sigma(1)}, \ldots, J_{\sigma(n)})$.

Neutrality (**N**): For any formula *A* and *B* in the agenda and profile $\boldsymbol{J}$, if for all *i* we have that *A* is in J_i iff *B* is in J_i, then *A* is in $F(\boldsymbol{J})$ iff *B* is in $F(\boldsymbol{J})$.

Independence (**I**): For any formula *A* in the agenda and profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$, if for all *i*, *A* is in J_i *iff* *A* is in J'_i, then *A* is in $F(\mathbf{J})$ iff *A* is in $F(\boldsymbol{J})$.

Monotonicity (M): If for any agent *i*, formula *A* in the agenda, and profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ such that coincide on every judgment set except for J_i, we have that if *A* is not in J_i and *A* is in J'_i then if *A* is in $F(\boldsymbol{J})$, then $F(\boldsymbol{J}')$.

Such properties capture and formalize a number of intuitions concerning the fairness of the aggregation procedure. Unanimity entails that if all individuals accept a given judgment, then so should the collective. Anonymity states all individuals should be treated equally by the aggregation procedure. Neutrality is a symmetry requirement for propositions that prescribe that all the issues in the agenda have an equal weight. Independence says that if a proposition is accepted by the same subgroup under two distinct profiles, then that proposition should be accepted either under both profiles or under neither profile. Monotonicity entails that by adding support for a proposition, its acceptance does not change.

This fairness condition may be used to model the arguments that justify the collective decision to the individuals. For instance, it is well known by May's theorem (Taylor 2005) that the majority rule can be characterized in terms of those axioms:

the majority rule is the aggregation function that satisfies (A), (M), (N), plus a minimal rationality requirement (Endriss et al. 2012).

Therefore the justification of a decision made by majority may appeal to axioms such as (A), by saying that majority does not discriminate between individuals' opinions.

Of course there are situations in which the majority rule is not appropriate. For instance, when we know that the individuals providing information are not equally reliable, one may appeal to other axioms in order to justify the decision. A case for refraining from deciding by majority is when there are inconsistent outcomes. The methodology of judgment aggregation and social-choice theory allows us to know in advance what are the possible situations and the possible aggregation procedures that may lead to inconsistent outcomes. The impossibility theorem of List and Pettit (List and Pettit 2002) is as follows:

Theorem 1 (List and Pettit 2002) There are agendas such that there is no aggregation procedure that satisfies (A), (N), (I) and collective rationality.

In particular, for any aggregation procedure that satisfies (A) and (S), there is a profile of judgment sets that returns an inconsistent outcome. The majority rule that we have seen in the examples satisfies (A) and (N) and (I); accordingly, the discursive dilemma shows a case of inconsistent aggregation. Very simple agendas may trigger inconsistent outcomes, one example being the agenda of the doctrinal paradox that we have presented. Technically, any agenda that contains a minimal inconsistent set of cardinality greater than 2 may trigger a paradox.

A solution that would guarantee a rational outcome would be to use a dictatorship, i.e., a procedure such that a single individual in any possible scenario decides the outcomes. Such procedures are not desirable because, besides violating important intuitions concerning fairness, they amount to discharging all the relevant information of a given scenario.

The methodology of JA can be extended to treat many voting procedures and characterize whether they may return inconsistent outcomes. Moreover, since the notion of aggregation procedure is very abstract, one can in principle model more complex procedures or norms, such as those that define decision-making in organizations.

9.4 Ontological Analysis of Information in STS

A crucial aspect of decision-making in sociotechnical system is that decisions may concern and may be based on heterogeneous types of information. For instance, suppose a personnel director has to decide whether to fire an employee on the grounds that the employee is accused of theft. Further suppose that surveillance cameras seem to support the accusation, whereas human witnesses are against the accusation of theft. Moreover, such an accusation has a number of normative and procedural constraints that have to be satisfied in order to be effective. In such a case, a personnel director is faced with a decision that has to weight information coming from security cameras, human agents, and normative constraints, and then decide what to do.

In order to describe the complex layers of information that are possibly involved in sociotechnical systems, we need to integrate the perceptual, conceptual, factual, and procedural information into a harmonious system. We propose to use the DOLCE ontology as integrating framework (Masolo et al. 2003). After defining basic properties and relations that are generic enough to be common to all specific domains—like being an *object*, being an *event*, being a *quality,* or being an abstract (entity)—DOLCE specifies different modules, like the mental or the social module, that are composed of entities that share some characterizing features. For example, mental entities are characterized by being ascribable to intentional agents, and social entities are characterized by the dependence on collectives of agents. These conceptual relations specify the definitions of the basic entities in our ontology; e.g., roles are properties of a certain kind that are ascribable to objects (e.g., being employed by an organization).

In order to apply the ontology to a specific domain, we introduce domain-specific concepts that specify more general concepts belonging to all these modules (e.g., “an aircraft is a physical object”).

The general ground ontology is meant to be not-context-sensitive and to provide a shared language to talk about some fundamental properties of concepts and entities. In this sense, the ontology provides a general language to exchange heterogeneous information and may be used as vocabulary to define communication languages for agents and to make explicit the matters of decisions.

We present some features of DOLCE-CORE, the ground ontology, in order to show that they allow for keeping track of the rich structure of information in a sociotechnical system.

The ontology partitions the objects of discourse, labeled particulars (PT) into the following six basic categories: objects, O; events, E; individual qualities, Q; regions, R; concepts, C; and arbitrary sums, AS. The six categories are to be considered rigid—i.e., a particular cannot change category through time. For example, an object cannot become an event.

In order to describe a concrete scenario for applying our ontological analysis, we enrich the language of DOLCE by introducing a specific language to talk about the scenario at issue. The language contains a set of *individual constants* for particular individuals. For example, in case we want to talk about an airport, individual constants may refer to “the gate 10,” “the flight 799,” “the landing of flight 747,” or “the security officer at gate 10.” Moreover, the language contains a set of *contextual predicates* that describe the pieces of information that agents may communicate in the intended situations (e.g., being a passenger, being a sensor, being a preference of an agent).

The language consisting of simple propositions can be partitioned according to the module they belong. For instance, we know that the predicates such as *passenger*, *customer*, *officer,* and *employee* can be accurately conceptualized as *roles.* Roles are social concepts that are characterized by the fact that they are anti-rigid (e.g., a passenger may cease to be a passenger) and dependent on other concepts (e.g. the concept of passenger requires the concept of person) (Masolo et al. 2004).

That is, in our specific ontology, we assume the axiom: RL (*employee*), that states that employee is a role. When we apply the predicate employee to an individual in our domain, e.g., *Employee* (Beatrix), we are building an atomic proposition that states some simple fact. This type of information can be retrieved by means of the ontological classification of the predicate. In this case, since employee is a role, it is a piece of social information belonging to the social module.

In a similar manner, we can list artificial sensors in our domain, e.g., *Sensor*(s1); categorize them as artificial agents, e.g., *ArtificialAgent*(s1); and model the output of a sensor as perceptual information coming from artificial agents.

We can easily extend the classification of predicates in order to partition all the (atomic) propositions into the relevant classes. For the sake of example, we can split here the possible types of propositions into *perceptual*, *social,* and *factual* propositions.

In Fig. 9.1, we depict a number of categories for an ontology developed in DOLCE for classifying information.

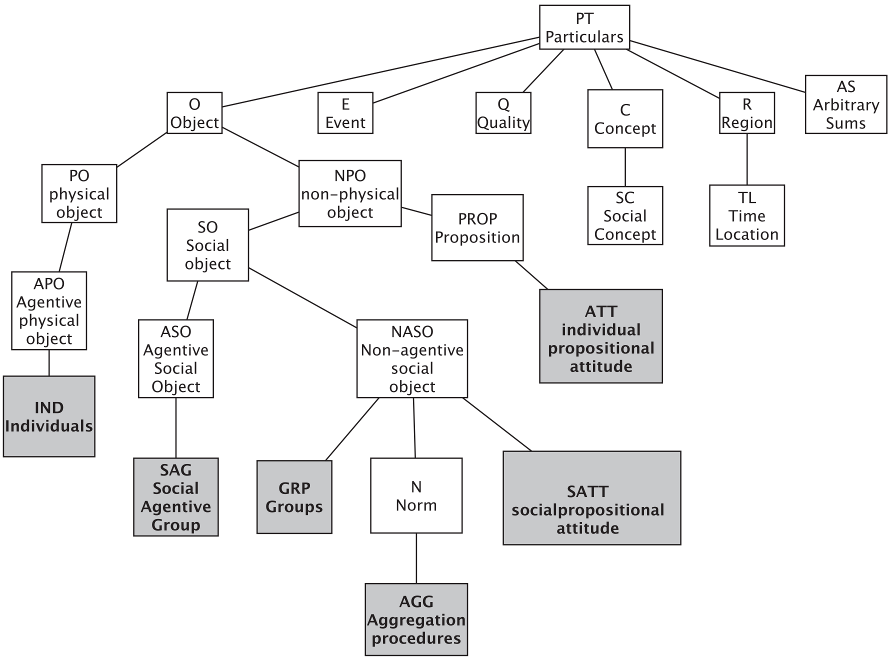

Fig. 9.1 An excerpt from DOLCE ontology

9.5 Assessing the Quality of Collective Decisions in Sociotechnical Systems

We have discussed how to represent in an abstract way the pieces of information that are required in order to provide an analysis of decisions and collective information in sociotechnical systems. We view agents as observation points in the system that are endowed with the reasoning capabilities provided by the ontology definitions in DOLCE and by logical reasoning. In this section, we present how to apply the methodology of JA to describe complex decisions in sociotechnical systems.

The properties of aggregation procedures that we have discussed in Sect. 9.3 provide a qualitative evaluation of the collective information or decision made within the system in a given moment. In a complex system like the one we are depicting, there may be several sources of disagreement between agents. For example, a possible disagreement may be at the level of perceptual information, as in the example of the sensors discussed in Sect. 9.2.

The ontological analysis allows us to classify the types of information; thus the question is how to evaluate the procedures that actually lead to collective decisions.

We briefly sketch our model. Suppose that we are able to list the agent—the information points—that are relevant for a certain decision. Call such a set of agents N of n agents. Denote as $A(L)$ the set of all possible sets of atomic formulas in our language L that are consistent with the ontology. We are presupposing that all the agents of the system agree on the definition provided by the ontological level. They may, however, disagree on matters of fact.

A profile of agents' propositional attitudes is given by a vector of sets of sentences, denoted $\boldsymbol{A}$. An aggregation procedure is a function F that takes a profile of agents' attitudes and returns a single set of propositions. The set of propositions $F(\mathbf{A})$ represents then the outcome of a collective decision of the system according to the procedure F.

For example, consider the case of the personnel director. Suppose there are three different security cameras and two human witnesses. Suppose proposition C means that "the accusation of theft is valid".

Agents	C
Camera 1	No
Camera 2	Yes
Camera 3	Yes
Human witness1	*Yes*
Human witness 2	*No*
Collective decision	*C in F(A)?*

Understanding what the procedure has been used to make the decision concerning C is crucial for the transparency of the system. We are not going to argue about which procedure is the best in this particular scenario. We claim only that social-choice theory and judgment aggregation, as well as the ontological analysis of information, allow for understanding and formalizing qualitative aspects of collective decisions in STS.

We now discuss a number of important concepts in evaluating collective decisions. In particular, we focus on the concept of *transparency*, the concept of *fairness,* and the concept of *efficiency* of decisions.

Firstly, a decision is *transparent* whenever the procedure F by means of which the decision has been taken is accessible to the agents involved in the decision.

In the example of the personnel director, the procedure is in fact *dictatorial*, because it is the director who has to take such a decision. However, what requires an explication, or even better, a justification, is the reason why the decision has been taken. That is, a dictatorial decision, such as the one taken by a single decision-maker, can nonetheless be a transparent decision, once it has been explained and justified to the relevant agents. One way of justifying such a decision is to mention how different information and different inputs affected the decision, which is equivalent to deciding which aggregation procedure a single decision-maker has followed with respect to different inputs. That is, in the example, the personnel director should make explicit whether the information coming from artificial agents outweighs the information coming from human agents.

The concept of *fairness* is quite debatable. However, the literature on social-choice theory is exactly about formalizing conceptions of fairness of an aggregation procedure. Therefore, the evaluation of fairness can be understood as the investigation of the properties of the decision procedures, for instance, whether the decision has been unanimous or anonymous with respect to the sources of information.

Unanimity implies that the agents of the system agree on a proposition. We claim that unanimity is a desirable property of any collective decision, regardless of the specific type of propositions. As agents are the observation points of the system, and our knowledge of the system is provided by means of agents' information, a violation of unanimity would amount to discharging information for no apparent reason (i.e., no agent against).

Anonymity, as we saw, implies that all agents are treated equally—we have no reason to weight the contribution coming from one agent more than the contribution coming from another one. This requirement is desirable when we cannot (or we do not want to) distinguish the reliability of agents. For example, we may not want to distinguish the information provided by two security officers that are communicating on the grounds of the higher reliability of the first compared to the reliability of the second. There are cases in which anonymity may not be a desirable property. For example, we want to weight the information coming from a trained security officer more than the information coming from a surveillance camera. Whenever appropriate, this is intended to model the fact that human agents may double-check outcomes from artificial agents, and human agents are assumed to be more reliable than artificial ones, at least at a number of tasks.

The condition of *independence* means that the acceptance of a formula at the systemic level only depends on the pattern of acceptance in the individuals' sets (e.g., the number of agents who accept). That is, the reason for accepting should be the same in any profile. Independence is a much more demanding axiom than the previous two; whether or not it should be imposed is debatable. A domain of application for which it is desirable is to merge normative information, where one expects impartiality across decisions.

Neutrality requires that all the propositions in the system have to be treated symmetrically. We believe that this is not desirable in the general case of heterogeneous information such as a STS. The reason is that we want in principle to treat visual, factual and conceptual information according to different criteria. Moreover, there are reasons to weight certain propositions more than others even when they belong to the same class. For example, the proposition that states that an object has been seen as a gun by a surveillance camera should be considered as highly sensible, and therefore it should be taken into account at systemic level. *Monotonicity* implies that agents' additional support for a proposition that is accepted at systemic level will never lead to it's being rejected. This property is desirable in most of the cases, provided the relevant agents are involved.

A further requirement that is usually viewed as a desirable property is the rationality of the collective decision. In particular, we focus on consistency: An aggregator F is consistent if for every profile, the set $F(\boldsymbol{A})$ is consistent with the ontology. As we saw, not every aggregator that satisfies the properties that we have seen guarantees consistency. For example, merging information by means of the majority rule or by a quota rule may lead to inconsistent sets of propositions.

The concept of consistency models a very weak notion of *efficiency* and more demanding views on efficient decisions can be modeled by adding further constraints.

We conclude by presenting a class of procedures that can be tailored for aggregating information in the scenario of STS. Those procedures are discussed in detail in (Porello and Endriss 2014) and (Taylor 2005).

Given a set of propositions X, we define a priority order on formulas in X as a strict linear order on X. Several priority orders can be defined on X, for example, a *support* order ranks the propositions according to the number of agents supporting them. Moreover, a *relevance* order ranks types of propositions (e.g., factual, perceptual, normative) according to their importance for the decision at issue. Moreover, we can define a priority order on propositions that depends on the *reliability* of the agents that support them. Thus, the reliability priority may be defined as a proposition A is more reliable than B if the number of experts supporting A is greater than the number of experts supporting B.

Thus, a priority-based procedure tries to provide a consistent outcome by checking the relevant information according to the priority. That is, the procedure tries to discharge conflicting information with a lower priority. For priority-based procedures, neutrality or anonymity may be violated by the priority order. Independence is also violated (because it may cease to be accepted if a formula it is contradicting receives additional support). Moreover, such procedures ensure consistency by construction.

Priority-based procedures allow for weighting the information according to the reliability or the relevance of different sources. For example, we can weight the information coming from security officers, who are viewed as experts, more than information coming from surveillance cameras. Moreover, we can weight the reports of cameras that are closer to the location at issue more than the information coming from other cameras. Note that it may be hard to compute the systemic information, given the required consistency check. The complexity depends of course on the language that we use to implement our ontology (a study of the complexity of computing problems related to judgment aggregation was presented in Endriss et al. 2012).

It is interesting to point out an application of non-consistent aggregators, namely, aggregators that return inconsistent sets of propositions. By using the analysis of aggregators provided by judgment aggregation, it is possible to pinpoint the places where the inconsistencies in the system are generated. In particular, aggregators that may return inconsistent information are useful to pinpoint causes of normative or conceptual disagreement, namely, to analyze incompatibility of norms or concepts defined in the system with the collective information gathered by the agents.

9.6 Conclusions

We have presented some basic elements for developing a model for assessing the quality of collective decisions in sociotechnical systems. We argued that we need a precise ontological understanding of the pieces of information involved in decisions and that welfare economics, social-choice theory, and judgment aggregation provide important tools for understanding fairness and efficiency of decisions. Therefore, foundational ontology plus the study of aggregation procedures provide important elements for developing a theory of justification of collective decision.

As a conclusion, we can view transparency as a necessary condition in order to make an assessment of the quality of decisions possible. Transparency amounts to making the procedure and the motivation of a collective decision accessible. That is, the first thing we need to demand in a system is transparency. We conceptualized transparency as a form of entitlement of the agents involved in the system to a justification of the decision made by the system. Future work has to investigate this concept in detail. For instance, one further condition on justifications is that they have to be addressed to real agents; that is, they have to be accessible to them—for instance, they have to be cognitively adequate to their addressees. Moreover, justifications have to be acknowledgeable by real agents; they should appeal to reasons that are shared among agents.

Acknowledgments D. Porello is supported by the VisCoSo project, financed by the Autonomous Province of Trento, "Team 2011" funding program.

References

Arrow, K. (1963). Social choice and individual values. Cowles foundation for research in economics at Yale University, Monograph 12. Yale: Yale University Press.

Boella, G., Lesmo, L., & Damiano, R. (2004). On the ontological status of plans and norms. *Artificial Intelligence and Law, 12*(4), 317–357.

Boella, G., Pigozzi, G., Slavkovik, M., & van der Torre, L. (2011). Group intention is social choice with commitment. In Proceedings of the 6th international conference on coordination, organizations, institutions, and norms in agent systems, COIN@AAMAS'10, pp. 152–171, Berlin, Heidelberg, Springer-Verlag.

Bottazzi, E., & Ferrario, R. (2009). Preliminaries to a DOLCE ontology of organizations. *International Journal of Business Process Integration and Management, Special Issue on Vocabularies, Ontologies and Business Rules for Enterprise Modeling, 4*(4), 225–238.

Brandt, F., Conitzer, V., & Endriss, U. (2013). Computational social choice. In G. Weiss (Ed.), *Multiagent systems*. Cambridge: MIT Press.

Dietrich, F., & List, C. (2009). The aggregation of propositional attitudes: Towards a general theory. Technical report.

Emery, F. E., & Trist, E. L. (1960). Socio-technical Systems. In C. W. Churchman & M. Verhulst (Eds.), *Management science, models and techniques* (Vol. 2, pp. 83–97). Pergamon.

Endriss, U., Grandi, U., & Porello, D. (2012). Complexity of judgment aggregation. *Journal of Artificial Intelligence Research, 45,* 481–514.

Guarino, N., Ferrario, R., & Sartor, G. (2012). Open ontology-driven sociotechnical systems: Transparency as a key for business resiliency. In M. De Marco, D. Te'eni, V. Albano, & S. Za (Eds.), *Information systems: Crossroads for organization, management, accounting and engineering*. Berlin: Springer.

Kornhauser, L. A., & Sager, L. G. (1993). The one and the many: Adjudication in collegial courts. *California Law Review, 81*(1), 1–59.

List, C., & Pettit, P. (2002). Aggregating sets of judgments: An impossibility result. *Economics and Philosophy, 18*, 89–110.

List, C., & Puppe, C. (2009). Judgment aggregation: A survey. In P. Anand, C. Puppe, & P. Pattanaik (Eds.), *Handbook of rational and social choice*. Oxford: Oxford University Press.

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., & Oltramari, A. (2003). Wonderweb deliverable d18. Technical report, CNR.

Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., & Guarino, N. (2004). Social roles and their descriptions. In Proc. of the 6th Int. Conf. on the principles of knowledge representation and reasoning (KR-2004), pp. 267–277.

Neumann, J. V., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Pettit, P. (2001). Deliberative democracy and the discursive dilemma. *Philosophical Issues, 11*(1), 268–299.

Porello, D., & Endriss, U. (2014). Ontology merging as social choice: Judgment aggregation under the open world assumption. *Journal of Logic and Computation, 24*(6), 1229–1249.

Porello, D., Setti, F., Ferrario, R., Cristani, M. (2013). Multiagent socio-technical systems: An ontological approach. In Proceedings of COIN@AAMAS/PRIMA 2013, pp. 42–62

Taylor, A. D. (2005). *Social choice and the mathematics of manipulation*. New York: Cambridge University Press.

van Benthem, J. (2011). *Logical dynamics of information and interaction*. Cambridge: Cambridge University Press.

Woolridge, M. (2008). *Introduction to multiagent systems*. New York: Wiley.