Multiagent Socio-Technical Systems. An Ontological Approach*

Daniele Porello[1], Francesco Setti[1], Roberta Ferrario[1], and Marco Cristani[2]

[1] Institute of Cognitive Sciences and Technologies, CNR
[2] University of Verona
{daniele.porello,francesco.setti,roberta.ferrario}@loa.istc.cnr.it
marco.cristani@univr.it

Abstract. Socio-technical systems constitute a challenge for multiagent systems as they are complex scenarios in which human and artificial agents share information, interact and make decisions. For example, the design of an airport requires to interface information coming from automatic apparatuses as security cameras, conceptual information coming from agents, and normative information which agents' behavior must comply with. Thus, in order to design systems that are capable of assisting human agents in organizing and managing socio-technical systems, we need fine grained tools to handle several types of information. The aim of this paper is to discuss a general framework to describe socio-technical systems as cases of complex multiagent systems. In particular, we use a foundational ontology to address the problems of interoperability and conceptual analysis, we discuss how to interface conceptual information with low level information obtained by computer vision or perception, and we discuss how to integrate information coming from heterogeneous agents.

1 Introduction

This work concentrates on the mutual influence of vision, cognition and social interaction in socio-technical systems, i.e. technologically dense contexts, such as, for instance, airports, hospitals, schools, public offices [7]. The process of seeing a scene, forming a belief or an expectation and engaging in interaction with other agents are essential features of agents' (both human and artificial) behavior in such systems. The entanglement of several layers of information (e.g. individual vs collective, visual vs inferential, human vs artificial) poses a challenge to the modeling of such complex environments. The overall aim of the work initiated with this paper is to build a rich model of agents' interaction that is capable of providing assistance to real socio-technical systems. We believe that the multiagent systems paradigm is a valuable framework to set up the construction of such complex models, as what is at stake is not only how autonomous agents form beliefs and expectations, communicate and act within a norm-governed system, but also their interaction with decisions that must be taken at systemic level. Our claim is that all these layers that are required in order to describe agents' information in socio-technical systems can be represented and reasoned about by using

* Supported by the VisCoSo project grant, financed by the Autonomous Province of Trento through the "Team 2011" funding programme.

a rich ontological model, that is capable of specifying our conceptual hierarchy in a way that is general enough to describe a complex categorization including physical and social objects, events, roles and organisations. In order to be effectively connected with the visual systems of artificial agents, such ontological models must contain information about the external context, both in its physical and institutional aspects, and information about the agents that inhabit it, in their physical, perceptual, cognitive and social aspects. In this paper we try to address three research challenges:1) providing a rich and structured description of the domain in all its aspects that is usable and interoperable among agents, both human and artificial; 2) integrating visual information with knowledge representation and reasoning and 3) describing systemic information as information coming form heterogeneous agents.

The ontological model we propose to use is DOLCE [16], and such choice is motivated by a series of reasons. First, the ontological perspective allows us to specify the properties of the concepts that we deploy and the relations holding among them. Such properties and relations are obtained through a foundational analysis and are expressed as formal axioms [16]. Second, one of DOLCE's basic assumptions is its "cognitive bias", in the sense that it is meant to express the perspective of a cognitive agent on a given domain, rather than "how things really are in the world". Finally, given that it comprises a social [17, 3] and a cognitive module [9], it is capable of dealing with several layers of information.

Regarding the visual information, we rely on a probabilistic methodology based on graphical models such as Bayesian Networks (BN) [14]; essentially, they allow to process low-level information, such as video sequences, audio streams and multisensorial input through graph-based inferential mechanisms in a robust and formal way. In our framework, BNs capture the finer grained knowledge under the form of action detection (run, sit, drink) and social signals recognition (visual focus of attention, facial expressions) [8], all expressed under the form of posterior distributions; this way, the uncertainty associated to the noise of the sensors and the accuracy of the modeling can be delegated to the ontological engine. Even though information is processed with probabilistic means in BNs, we assume that these will give as output, via a mechanism based on thresholds, a discrete proposition, that will then be available to be reasoned upon with logical and ontological tools.

The relationship between individual and systemic information will be approached by means of techniques developed in multiagent systems, in particular in social choice theory, judgment aggregation, and belief merging [15, 4].

Ontologies for multiagent systems have been developed in particular in order to provide agents' communication languages, [12, 10]. Moreover, models of socio-technical systems based on goal models have been recently introduced in [5]. The original contribution of this paper is to provide a general multiagent model based on DOLCE to represent the entanglement of several layers of information in socio-technical systems.

The guiding examples of this paper are taken from the organization of an airport. Such socio-technical system involves several types of agents, surveillance cameras, security officers, customers, who interact and share information of different types, for example information coming from cameras, procedural information concerning rules, information that can only be indirectly inferred.

The remainder of this paper is organized as follows. In Section 2, we discuss the connection between low-level information coming from vision and high-level information contained in agents' knowledge bases. Section 3 presents our ontological modeling of socio-technical systems. In Section 4 we deal with the issue of how to integrate the information coming from different agents and we propose procedures that define systemic information.

2 Linking computer vision and propositional information

The connection between propositional information and visual information is a difficult task to approach. In computer vision, a label denoting a scene (e.g. "the plane has landed") is associated to a process that models the interaction of several features usually expressed by a graphical model [14]. More specifically, Bayesian Networks (BN) have been widely adopted in vision systems as they are applicable to all levels of processing, from the extraction of low-level actions (e.g., running, walking) to more complex high-level reasoning (e.g., Mark runs *and then* walks); essentially, they embed a mapping into a graph structure so that the nodes represent concepts or parameters of interest and dependencies are given by links. Their diffused use is due to fast numerical updating in singly connected trees and to the availability of techniques to decompose complex models into simpler ones, adopting heterogeneous learning techniques [19]. Bayesian networks can be learned and model dependencies for either static (BBN) or dynamic (DBN) domains. A simple DBN with single causal dependencies over time, the Hidden Markov Model, is often used for speech analysis [21] and has been extensively adapted for heterogeneous applications in the Computer Vision realm [18]. Bayesian Networks are the workhorses for the automated surveillance community: in most cases, different activities correspond to different BNs; they are trained in advance, and employed afterwards in classification tasks, aimed at discovering usual or abnormal activities. One of the main limitations of this philosophy, i.e., describing high-level activities employing BNs, is their limited expressivity: the usual architecture of a surveillance system builds upon a low-level layer in which simple actions are recognized (run, walk, sit); these outputs, expressed in the form of posterior probabilities, are fed into a mid-level layer of the network which connects them considering spatio-temporal relations: an individual can walk for a minute, and in the meantime he/she can talk with another guy, or he/she can make a phone call. In practice, this layer gives as output a set of BNs, one for each activity; in turn, each BN generates a posterior probability, depending on how well the structural knowledge embedded in the network fits the visual data. The high-level layer of the surveillance system performs the final classification, recognizing as ongoing activity the one which corresponds to the BN with the highest posterior probability. As a matter of fact, the kind of understanding provided by this architecture is limited: in one sense, it is fixed, enclosed in a graphical model which describes conditional dependencies among random variables, whose structure is decided a priori, drawn by hand from the researcher. In another sense, it is limited, since it amounts to a set of available activities that have to be recognized.

In order to interface visual and propositional information, we want to associate the set of BNs that describes the activity of a camera with a knowledge base consisting of a set of formulas that are defined by means of the predicates specified in our ontology. The knowledge base contains a set of low-level propositions that are directly connected with the BNs as well as higher level propositions that can be inferred by means of the ontological definitions. For example, "there is a queue at gate 8" is a perceptual proposition that is triggered by the BN, whereas "the boarding at gate 8 has started" is a proposition that can be inferred on the basis of visual information. One of the challenges of this work is to understand how the information coming from a BN, which is probabilistic, entails the assumption that a certain proposition holds in the knowledge base. In particular, a BN provides a degree of probability that the proposition that describes a scene is true or false (e.g. "an aircraft is landing on runway 4 with a probability of 75%"). Our approach is the following. As we are interested in providing human agents with a tool that assists their activity in monitoring the system, propositional information that becomes available to human agents should be as simple as possible. Thus, we associate probabilistic BNs with discrete information (true or false) concerning propositions, by defining a threshold of the degree of reliability provided by the BN that is sufficient to accept the proposition. The thresholding mechanism is also well founded under a Bayesian point of view, since it corresponds to the cost associated to a particular classification output [6]. Moreover, the reliability of the information coming from vision shall be discussed at systemic level as a problem of aggregating possibly divergent sources of information.

3 Ontological analysis: DOLCE

In order to describe agents' information in socio-technical systems, we need to integrate visual, conceptual, factual and procedural information. We propose to use the DOLCE ontology as integrating framework. The methodology employed in the construction of the DOLCE ontology is the following. Firstly, we define basic properties and relations, that are generic enough so to be common to all specific domains, like being an endurant (more simply, an object), being a perdurant (an event), being a quality or being an abstract (entity), one entity being part of another, an object participating to an event or having a certain quality... Then, we specify different modules, like the mental or the social module, that are composed by 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. Finally, we introduce domain specific concepts that specify more general concepts belonging to all these modules (like "an aircraft is a physical object").

We begin by presenting the general *ground* ontology; this 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[3]. We are here interested in presenting the descriptive features of DOLCE rather than in complexity or implementability issues. Notice that in

[3] Ideally, this is not the case, as in multiagent systems agents can be heterogeneous under many respects, including the adoption of different languages and also of different ontologies. The

[16] an appendix may be found with more implementable but less expressive versions of DOLCE, that are called DOLCE-Lite.

3.1 Foundational ontology

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 socio-technical system. For an introduction to DOLCE-CORE, we refer to [2], here we simply point at the relevant features.

The ontology partitions the objects of discourse, labelled *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 particular, we shall focus on the following categories.

Objects represent particulars that are mainly located in space, e.g. the aircraft 777, the gate 6, the queue at gate 6. On the other hand, events have properties that are mainly related to time, e.g. landing, the boarding of flight 717, the delay of flight 717. The relation that links objects and events is the *participation* relation: "an object x participates in event y at time t" $\mathrm{PC}(x, y, t)$.

Since individual qualities play an important role in modeling information coming from perception, or from different agents of the system, we shall take a closer look at them in the next paragraphs. An individual quality is an entity that we can perceive and measure that inheres to a particular (e.g. the length of runway A2 of Malpensa airport, the weight of Mark's luggage, the temperature inside waiting room 3…). The relationship between the individual quality and its (unique) bearer is the *inherence*: $\mathrm{I}(x, y)$ "the individual quality x inheres to the entity y". The category Q is partitioned into several *quality kinds* Q_i, for example, color, weight, temperature, the number of which may depend on the domain of application. Each individual quality is associated to (one or more) *quality spaces* $S_{i,j}$ that provides a measure for the given quality. We say that individual qualities are *located* at a certain point of a space S at time t: $\mathrm{L}(x, y, t)$: "x is the location of quality y at time t". Spaces allow for evaluating relationships between objects from the point of view of a given quality. For example, "the temperature inside room 3 (q) is higher than the temperature inside room 4 (q')" is represented in the ontology by assuming spaces of values with order relations and by saying that the location of the individual property q is lower than the location of q'. Spaces may be more structured objects and they may be specified along several dimensions[4]. The axioms that define the relationships between individual qualities, locations, and spaces state for example that every individual quality must be located in some of its associated spaces and that the location in a particular space must be unique, cf. [2]. E.g. the color of an object may be associated to color quality kinds with their relevant spaces such as hue, saturation, brightness.

---

strong requirement should be that their ontologies are well founded, so that their underlying assumptions are explicit enough as to enable communication and exchange of information via "connecting axioms". In the current paper, for the sake of simplicity, we will assume that all agents in the system share the same ontology, DOLCE.

[4] Quality spaces are related to the famous treatment of concepts in [11].

The category of regions R includes subcategories for spatial locations and a single region for time, denoted T: T(x) means "x is a time location" (e.g. October 10, 2012, 12:31 PM). The relation $PRE(x,t)$, where t is a time location, allows to specify that "x is present at time t". Note that in DOLCE-CORE we have that all entities exist in time: $\mathrm{PT}(x) \rightarrow \exists t(PRE(x,t))$.

The category of concepts shall be used in particular to model social objects. Concepts are reified properties that allow for viewing them as entities and to specify their attributes. In particular, concepts are used when the intensional aspects of a property are salient for the modeling purposes. The relationship between a concept and the object that instantiates it is called *classification* $CF(x,y,t)$ "x is classified by y at time t".

If we represent the DOLCE taxonomy as a tree, more specific categories, such as physical objects, mental objects and so on, can be plugged into the tree as children of the relevant categories. Summing up, there are three ways of understanding properties in DOLCE-CORE and therefore there are three ways to deal with different levels of information [2]. We can understand properties as extensional classes, as individual qualities, or as concepts. We shall apply this distinction to our modeling task: extensional predicates are used to model robust information (e.g. "waiting room 3 is located at gate 3"), individual qualities are used to model information coming from vision and perception, and concepts and roles are used to model information about norms, social objects and organizational properties.

3.2 Individual qualities and visual information

In order to integrate the information coming from computer vision or from perception of observers in systems with other types of information, we proceed as follows. We assume that agents, both human agents and artificial agents such as surveillance cameras, provide observation points on the system. For example, take a surveillance camera that is trained in the sense of Section 2 for a specific task. We represent the features that the surveillance camera is supposed to detect by means of individual qualities of the object/action/event that it is focusing on. As the information coming from visual detection may be revisable and depends on the perception of the observer, we represent it as specific subtype of "mental object" (MOB), namely, by means of a specific category VIS that represents visual objects[5]. Visual objects are representations of physical objects from the point of view of a given observer. The recognition of an object, the point of view of the specific observer, and the object itself are connected by means of the following relation: $V(i,x,v,t)$ that means "the camera i sees the physical object x as v at time t", where in particular v is in VIS and provides the visual representation of x[6]. For example $V(i,x,v_{airplane},t)$ means that a particular camera i sees x as $v_{airplane}$, namely as the image of an airplane. From this piece of information, we do not want to

[5] The cognitive module of DOLCE has been discussed in [9].

[6] This treatment presupposes the existence of the object that provides the focus of a given camera. Moreover, we are assuming that the individual qualities that trigger the recognition of an object as a visual representation are qualities of the *object itself* and not of the image (e.g. video sequence). This is motivated by the fact that the existence of the physical object is assumed to be the "same" focus of possibly divergent observers.

derive immediately the fact that there is an airplane at a specific time. The inference to factual information shall be done by means of "bridge rules" that link the recognition of an object as a certain entity and the endorsement in our knowledge base that the entity is actually located in a given place. The bridge rules are supposed to provide a thresholding for turning probabilistic information into factual propositional information and they may vary according to different scenarios or to the relevance of the particular piece of information.

The conditions that trigger propositions like $V(i, x, v, t)$ are represented by formulas that specify locations in a number of quality spaces. We assume that each observer i is associated with a set of individual qualities $q_{ix1}, \dots, q_{ixm}$ of an object x that represent the features that a specific observer is looking for, in order to detect that object. For example, such individual qualities represent pieces of information such as "the dimension of the object x from the point of view of camera i". By locating such qualities in specific regions of quality spaces, we can specify a set of preconditions that trigger the recognition of x as v from the point of view of i.

$$loc(q_{1x}, s_1, t_1) \wedge \cdots \wedge loc(q_{nx}, s_n, t_n) \rightarrow V(i, x, v, t)$$

For example, such conditions state that, according to camera i, if the dimension has a certain value, and the shape is of a certain type, and the color is such and such, then camera i recognizes x as an airplane. Of course the information required in order to model the locations of the individual qualities and the relevant spaces shall be provided by integrating the ontological analysis with the properties of the algorithm used in computer vision. Moreover, at least in principle, in case of human agents, we can represent the relevant cognitive aspects of vision by means of suitable qualities of objects and quality spaces.

The motivation for our treatment is that it allows for handling information coming from different observation points, or from a same observation point at different times, by spelling out the preconditions that trigger such information. An assumption we shall stick to is that the same camera cannot see an object in two ways at the same time:

$$V(i, x, v, t) \wedge V(i, x, v', t) \rightarrow v = v'$$

This amounts to assuming that the algorithm for visual detection is well-defined. However, at different times, the same camera can change the visual object that it provides, or it can even fail to recognize the object, thus we do not force more demanding constraints on V. Moreover, we do not presuppose that different observers of the same objects in the same location and at the same time have to agree on the same recognition. Thus, for example, a camera can see an object as a person whereas another fails to recognize it or classifies it as something else. We believe that this forms of mismatches of information have to be made explicit in our modeling and represented accordingly, as they are an important aspect of the interaction in socio-technical systems. We shall discuss how to handle possible mismatches of information in the next section.

3.3 Social objects and norms

One predicate that is particularly important for modeling socio-technical systems is the classification predicate: $CF(x, y, t)$ meaning that "x is classified as y at time t"[7]. By using CF, we can define a special type of social object, namely the notion of *role*. Roles are supposed to be contextual properties, that are characterised by *anti-rigidity* (AR) and *foundation* (FD): roles are concepts that classify entities at a certain point in time, but not necessarily classify them in each moment or each possible world in which they are present (AR) and that require a level of definitional dependence on another property (FD). For instance, someone who is a student at a certain point, not necessarily will be a student all throughout his/her life and there are possible worlds in which he/she is not a student; in order for someone to be classified as an employee, we need someone else who is classified as an employer.

$$AR(x) \equiv \forall y, t(CF(y, x, t) \rightarrow \exists t'(PRE(y, t') \wedge \neg CF(y, x, t'))$$

$$FD(x) \equiv \exists y, d DF(x, d) \wedge US(y, d) \wedge \forall z, t(CF(z, x, t) \rightarrow$$
$$\exists z'(CF(z', y, t) \wedge \neg P(z, z', t) \wedge \neg P(z', z, t)))$$

Given these characteristics, roles are essential to model organizations, as they allow to talk about properties that one acquires in virtue of the fact that one is member of an organization or has some rights/duties connected with the role he/she is playing in that very moment. Take for instance a security officer, who is allowed to carry a gun inside the airport terminal, but just when he/she is playing (or is classified by) the role of security officer; if the same person enters the airport while playing the role of passenger he/she is no more allowed to carry a gun. The same role can be played by many entities within the same domain (even entities of very different nature, like a human being and a software), the same entity can play more than one role, even simultaneously (like in airports with self-check-in stations, where the same person simultaneously plays the role of passenger and that of check-in operator).

Further developments of ontological analysis treat also norms and plans by means of DOLCE, cf. [1]. Here we just sketch some possible application. Our role-based analysis provides a way to connect low-level information (e.g. coming from cameras) with high-level information (e.g. coming from security protocols). For example, the concept of role allows for making explicit the conceptual dependence of a signal of alarm triggered by a particular scene that has been detected by cameras with the properties that are sufficient to trigger that signal. For example, a "suspect", according to our approach, can be modeled as a role. We can impose a constraint that specifies that only entities that are agentive physical objects can be classified as suspects in our scenario (according to the (FD) condition). Moreover, we can specify the description that defines "suspect" by spelling out a set of properties, like the participation to some kinds of events, e.g. "carrying a gun", "entering in an unauthorized area". Thus, if a person (an agentive physical object) is recognized (possibly by a computer vision system that locates a

[7] For an axiomatic definition of the predicates that we introduce, we refer to [17].

set of individual qualities in the relevant places) as possessing one or more of these properties, he/she is classified as a suspect and this triggers an alarm. In order to specify such a security protocol, the system should be capable of taking into account the various layers of information involved. Consider the following example:

$$\exists x(\exists i V(i, x, v_{\text{person}}, t) \wedge \exists j \exists y V(j, y, v_{\text{gun}}, t) \wedge \text{next}(x, y) \wedge \neg CF(x, \text{officer}, t)$$

$$\rightarrow \text{Suspect}(x, t))$$

The formula above means that if a camera detects a person and another detects a gun that is next to the person (i.e. $\text{next}(x, y)$), and that person has not the role of a security officer at that moment, then the person is a suspect.

Depending on the type of agent that is provided with this system, the reaction to the alarm could be of various kinds: either send a message to some other agent that has to follow the suspect, or activate another camera with a tracking system, for example. That depends on the security protocol that is implemented in the system. By extending the ontological treatment so to include norms, prescriptions and so on, the security protocols can be represented as formulas in our system. Obviously, a person can cease to be classified as suspect if further properties are discovered (for instance, a new video sequence may show that what at first appeared as a gun is in fact an umbrella, or it may show a police sign on the back of the person who eventually turns his/her back to the camera, and after that the same person who had been previously classified as "suspect" is subsequently classified as police officer). In particular, roles allow for linking a higher level property used by human agents involved in the system to low level properties that can be checked by means of perception (either direct, performed by human agents, or indirect, obtained by a camera).

4 A multiagent setting

We have described how to represent in an abstract way the pieces of information that are required in order to provide an analysis of socio-technical systems. In this section, we apply a multiagent perspective in order to deal with information at systemic level. We view agents as observation points in the system that are endowed with the reasoning capabilities provided by the ontology definitions in DOLCE. For example, cameras endowed with axioms that connect visual information with high-level organization concepts, or security officers that communicate pieces of information are viewed as agents in our system. The problem that we are going to tackle is how to integrate the information coming from possibly divergent agents into a collective information that is supposed to be made available at systemic level.

4.1 Modeling socio-technical systems

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 constants for particular individuals $\mathcal{C}_S$. For example, in the case of an airport, individual constants may refer to "the gate 10", "the flight 799", " the landing of flight 747", "the security officer at gate 10". According to our previous analysis of visual information, we need also constants for locations of individual qualities in their respective spaces. Moreover, the language contains a set of contextual predicates $\mathcal{P}_S$ that describe the pieces of information that agents may communicate in the intended situations. In case of an airport, we need for example to include predicates such as "being an aircraft", "being a queue", "being a gate", "being a delay", "being a landing", "being a security protocol" and so on. The set of predicates $\mathcal{P}_S$ and the set of individuals $\mathcal{C}_S$ are partitioned according to the basic categories, concepts, and individual qualities, etc. This is done by assuming a number of axioms that specify for each predicate the right category. We are taking here the suggestion to view DOLCE as a shared terminology, or Tbox, and let agents have possibly divergent knowledge bases, namely Aboxes.

4.2 Modeling agents' information

For the sake of simplicity, agents in our systems are modeled just as sets of (closed) formulas built by means of predicates that are either in DOLCE or in the specific language that we have sketched. This set may include information coming from vision, propositions concerning social objects, norms, plans, etc. We denote $\mathcal{L}_S$ the language of the agents in the system: it is defined as the set of atomic formulas or negations of atomic formulas defined on the alphabet given by $\mathcal{P}_S$, $\mathcal{C}_S$ and DOLCE. We denote an agent's set of propositions by $A_i \subset \mathcal{L}_S$. For example, in case A_i is a surveillance camera, it may contain a set of visual propositions $V(i, x, v, t)$ that are triggered by the detection of the relevant individual qualities[8]. The only general requirement that we put on the A_is is that each A_i is *consistent wrt the ontology*, namely they are consistent with the definitions provided in the ontological analysis. For example, A_i cannot contain a proposition such as "a security officer is a mental object", and so on[9]. Note that the amount of information that each agent shall submit at a given moment depends on the security protocol that is implemented in the system and on the situation at issue. For example, not all the visual information that is provided by all the cameras shall be continuously made available to the whole system. Moreover, we shall make the assumption that the propositions provided by the agents of our system can be synchronized by means of the temporal parameters that are attached to them. Thus, we assume that it is possible to talk about the state of the system at a particular moment or during a particular interval of time[10]. In this section, we shall abstract from those important issues and we simply assume that

[8] Note that we do not want the knowledge base to be closed under negative information, namely we have to endorse an *open world assumption on each A_i*. This is because the fact that a camera does not detect a man carrying a gun does not mean that we can claim that he is not carrying a gun.

[9] We are aware that the consistency assumption may be a highly demanding condition in case we model cognitive agents. We assume it here just for the sake of simplicity, in order to directly apply the model of the next section.

[10] We are aware that this is a demanding assumption. For example, synchronizing surveillance cameras and human agents' communications may require interfacing two different time seg-

at a given moment in time, we can take the sets of propositions provided by the agents of the system.

4.3 Modeling systemic information

We present now our modeling of *systemic information*. We want to be able to check the status of the system with respect to a number of parameters. As the sources of information are heterogeneous, namely each agent of the system provides his/her set of propositions, the problem of evaluating the state of the system as a whole can be viewed as a problem of integrating heterogeneous information. We shall model this issue by means of techniques developed in social choice theory [4], belief merging [13], and judgment aggregation [20, 15]. The reason is that, as we shall see, those techniques provide versatile tools to model aggregation of heterogenous types of information and they allow for spelling out the properties of each type of aggregation procedure.

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. Imagine three cameras that are pointing at the same scene, and such that two of them recognize an object as a gun, whereas the third does not. Furthermore, agents' knowledge bases can contain conflicting high-level information on the roles involved, and it is not clear where to place the source of disagreement.

The ontological analysis allows us to classify the types of information, thus the question is how to define suitable procedures to solve the different types of disagreement, e.g. normative, prescriptive, or visual. We briefly sketch our model. Suppose the system consists of n agents. Denote $A(\mathcal{L}_S)$ the set of all possible sets of atomic formulas in our language $\mathcal{L}_S$ that are consistent with the ontology. A profile of agents knowledge bases is given by a vector $(A_1, \ldots, A_n)$, we denote it $\mathbf{A}$. An aggregation procedure is a function $F : A(\mathcal{L}_S)^n \rightarrow \mathcal{P}(\mathcal{L}_S)$ that takes a profile of agents' knowledge bases and returns a single set of propositions. The set of propositions $F(\mathbf{A})$ represents then the systemic information according to the procedure F.

The ontological analysis allows us to partition the set of propositions in $\mathcal{L}_S$ into their respective types. For each predicate P in our language, we can easily check by means of DOLCE whether P is a social concept, a visual concept, a basic concept and so on. Since every proposition in the A_i is an atomic formula or a negation of an atomic formula, we can easily extend the classification of predicates in order to partition the agents' propositions into visual, conceptual and factual propositions.

Given a set of formulas A_i, we denote A_i^V, A_i^C, A_i^F , the visual, conceptual and factual propositions (respectively) that are contained in A_i. Accordingly, we partition profiles wrt their type of information; we denote them $\mathbf{A}^V$, $\mathbf{A}^C$, $\mathbf{A}^F$. We shall discuss aggregators that take profiles restricted to one of the types of propositions: $F^V : \mathbf{A}^V \mapsto A^V$, $F^C : \mathbf{A}^C \mapsto A^C$, and $F^V : \mathbf{A}^F \mapsto A^F$.

We introduce and discuss a number of properties of aggregators that have been widely studied in judgment aggregation and social choice theory. In particular, the application of social choice theory and judgment aggregation to ontology merging has

mentations of events. We abstract from this issue in order to present our analysis of systemic information.

been developed in [20]. In what follows, we present some arguments to evaluate to what extent those properties are relevant for our modeling tasks.

Unanimity. An aggregator F is called unanimous if $A_1 \cap \cdots \cap A_n \subseteq F(\mathbf{A})$ for every profile $\mathbf{A} \in A(\mathcal{L}_S)^n$.

Anonymity. An aggregator F is called anonymous if for any profile $\mathbf{A} \in A(\mathcal{L}_S)^n$ and any permutation $\sigma : N \to N$ of the agents, we have that $F(A_1, \ldots, A_n) = F(A_{\sigma(1)}, \ldots, A_{\sigma(n)})$.

Independence. An aggregator F is called independent if for any formula $\phi \in \mathcal{L}_S$ and any two profiles $\mathbf{A}, \mathbf{A}' \in A(\mathcal{L}_S)^n$, we have that $\phi \in A_i \Leftrightarrow \phi \in A'_i$ for all agents $i \in N$ implies $\phi \in F(\mathbf{A}) \Leftrightarrow \phi \in F(\mathbf{A}')$.

Neutrality. An aggregator F is called neutral if for any two formulas $\phi, \psi \in \mathcal{L}_S$ and any profile $\mathbf{A} \in A(\mathcal{L}_S)^n$, we have that $\phi \in A_i \Leftrightarrow \psi \in A_i$ for all agents $i \in N$ implies $\phi \in F(\mathbf{A}) \Leftrightarrow \psi \in F(\mathbf{A})$.

Monotonicity. An aggregator F is called monotonic if for any agent $i \in N$, formula $\phi \in \mathcal{L}_S$, and profiles $\mathbf{A}, \mathbf{A}' \in A(\mathcal{L}_S)^n$ such that $A_j = A'_j$ for all $j \neq i$, we have that $\phi \in A'_i \setminus A_i$ and $\phi \in F(\mathbf{A})$ imply $\phi \in F(\mathbf{A}')$.

Unanimity implies that if the agents of the system agree on a proposition ϕ, then ϕ is accepted at systemic level. We claim that unanimity is a desirable property of any aggregator, regardless 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* implies that all agents are treated equally, namely, that we have no reason to weight the information coming from an agent more than from another. 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 ground of the higher reliability of the first wrt the second. In case of visual information, anonymity may not be desirable. 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 coming from artificial agents and human agents are assumed to be more reliable than artificial ones. The condition of *independence* means that the acceptance of a formula at systemic level only depends on the pattern of acceptance in the individuals' sets (e.g. the number of agents who accept ϕ). That is, the reason for accepting ϕ should be the same in any profile. Independence is a 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, see [15]. For example, suppose that the security protocol of the airport prescribes to fire an officer if and only if conditions c_1 and c_2 hold. Suppose such conditions have to be checked by the relevant committee of agents. In that case, we do not want the outcome of the decision to depend on a particular scenario (i.e. profile), rather a form of impartiality should be respected. On the other hand, there are cases in which the number of agents supporting ϕ is not a good criterion for any profiles, we shall present an example below. *Neutrality* requires that all the propositions in the system have to be treated symmetrically. We believe that this is not desirable for our purpose in general, as we want to treat visual, factual and conceptual information according to different criteria. Moreover, there are reasons to weight certain propositions more than others even in case 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 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 consistency of the systemic information.

Consistency An aggregator F is *consistent* if for every profile $\mathbf{A} \in A(\mathcal{L}_S)^n$, the set $F(\mathbf{A})$ is consistent with the ontology.

It is well-known that not every aggregator that satisfies the properties that we have seen guarantees consistency. In particular, an aggregator that satisfies anonymity, independence, and neutrality may return inconsistent outcomes, cf. [15]. For example, merging information by means of the majority rule may lead to inconsistent sets of propositions[11]. For the sake of example, we introduce a class of aggregators to model systemic information that is adapted from [20]. We leave an exhaustive discussion of types of aggregators for future work.

Given a set of propositions $X \subseteq \mathcal{L}_S$, 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 $>_S$ ranks the propositions according to the number of agents supporting them: $\phi >_S \psi$ iff the number of agents supporting ϕ is greater than the number of agents supporting ψ (provided a tie-breaking rule for propositions with equal support). Moreover, we can define a priority order on propositions that depends on the reliability of the agents that support them. Given the set of agents N, we define the *expert* agents as a subset $E \subset N$. Thus, the reliability priority may be defined as $\phi >_R \psi$ iff the number of experts supporting ϕ is greater than the number of experts supporting ψ. We may also introduce more stringent conditions by imposing that ϕ has higher priority than ψ if the very experts that support ϕ also support ψ.

Definition 1 (Priority-based procedures). *Given a priority order $>_X$, the procedure based on $>_X$ is the aggregator mapping any profile $\mathbf{A}$ to $F(\mathbf{A}) := S$ for the unique set $S \subseteq \mathcal{L}_S$ for which (i) $\phi \in S$, where ϕ is the top proposition according to $>_X$; (ii) if $S \cup \{\psi_2\}$, $\psi_1 \in S$, $\psi_1 >_X \psi_2$ and is consistent, then $\psi_2 \in S$.*

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

[11] These results depend on the structure of the language that the agents use. It is enough to include some minimal logical connection to generate inconsistent outcomes, cf. [15]. Even if the propositions in the agents' sets are atomic, we are evaluating consistency wrt the ontology, that contains complex propositions.

violated (because ϕ may cease to be accepted if a formula it is contradicting receives additional support). Moreover, such procedures are consistent by construction. The priority order is supposed to represent the importance of the property for the system. For example, the proposition that states the recognition of an object as a gun should receive high priority in our ordering. Moreover, priority based procedures allow for weighting the information according to the reliability of different sources. For example, we can weight the information coming from security officers, that are viewed as experts, more than information coming from surveillance cameras. Thus, priority based procedures may be used to define aggregators that provide collective information on visual propositions: $F^V : \mathbf{A}^V \mapsto A^V$. Moreover, a priority order based on the reliability of agents can be used to merge factual information $F^V : \mathbf{A}^F \mapsto A^F$, provided we single out the right class of experts in our system. 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. Moreover, it is interesting to point at 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 analyse incompatibility of norms or concepts defined in the system with the collective information gathered by the agents.

5 Conclusion

We have depicted and discussed a number of important elements in order to model a complex scenario such as a socio-technical system. We have seen that in order to provide a faithful representation of knowledge, we need to model agents endowed with visual, cognitive and social capabilities as well as systemic procedures that handle complex interactions. We have argued that the ontological analysis allows for specifying the types of information involved in the system and we have proposed the application of techniques from social choice theory and belief merging in order to define and analyze several concepts of systemic information. Future work shall focus on two directions. Firstly, we will try to extend the ontological analysis to model agents that are endowed with a set of actions that depend on the information state. For example, agents can send an alarm signal in case they can infer that a person is a suspect, they can communicate pieces of information to other agents, they can ask questions to other agents, they can ask other agents to perform tasks, they can prescribe actions to be taken (e.g. "close the gate 12"), etc. For instance, an observation point i can see that a person is getting close to a security area and it sends this information to agent j who can check if such information holds also on the basis of his/her visual input. Moreover, i can ask the other agents to track the path of the person who has been recognized as a suspect. Secondly, we plan to extend our treatment of systemic information by discussing more general classes of functions that aggregate agents' information and by introducing procedural aspects of agents' interaction, e.g. negotiation, dialogues, deliberation.

Bibliography

[1] G. Boella, L. Lesmo, and R. Damiano. On the ontological status of plans and norms. *Artif. Intell. Law*, 12(4):317–357, 2004.

[2] S. Borgo and C. Masolo. Foundational choices in dolce. In S. Staab and R. Studer, editors, *Handbook on Ontologies*. Springer, second edition, 2009.

[3] E. Bottazzi and R. Ferrario. 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, 2009.

[4] F. Brandt, V. Conitzer, and U. Endriss. Computational social choice. In G. Weiss, editor, *Multiagent Systems*. MIT Press, Forthcoming.

[5] F. Dalpiaz, P. Giorgini, and J. Mylopoulos. Adaptive socio-technical systems: a requirements-based approach. *Requir. Eng.*, pages 1–24, 2013.

[6] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley, New York, 2nd edition, 2001.

[7] F. E. Emery and E. Trist. Socio-technical systems. In *Management Sciences: Models and Techniques*, volume 2, pages 83–97. Pergamon, 1960.

[8] M. Farenzena, L. Bazzani, V. Murino, and M. Cristani. Towards a subject-centered analysis for automated video surveillance. In *International Conference on Image Analysis and Processing*, pages 481–489, 2009.

[9] R. Ferrario and A. Oltramari. Towards a computational ontology of mind. In A. C. Varzi and L. Vieu, editors, *Formal Ontology in Information Systems, Proceedigs of the Intl. Conf. FOIS 2004*, pages 287–297. IOS Press, 2004.

[10] R. Ferrario and L. Prévot. Special issue: Formal ontologies for communicating agents. *Applied Ontology*, 2(3-4), 2007.

[11] P. Gärdenfors. *Conceptual spaces - the geometry of thought*. MIT Press, 2000.

[12] M. Hadzic, P. Wongthongtham, T. S. Dillon, and E. Chang. *Ontology-Based Multi-Agent Systems*, volume 219 of *Studies in Computational Intelligence*. Springer, 2009.

[13] S. Konieczny and R. Pino Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.

[14] S. L. Lauritzen. *Graphical Models*. Oxford University Press, 1996.

[15] C. List and C. Puppe. Judgment aggregation: A survey. In *Handbook of Rational and Social Choice*. Oxford University Press, 2009.

[16] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. Wonderweb deliverable d18. Technical report, CNR, 2003.

[17] C. Masolo, L. Vieu, E. Bottazzi, C. Catenacci, R. Ferrario, A. Gangemi, and N. Guarino. Social roles and their descriptions. In *Proc. of the 6th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR-2004)*, pages 267–277, 2004.

[18] H. Q. Minh, M. Cristani, A. Perina, and V. Murino. A regularized spectral algorithm for hidden markov models with applications in computer vision. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 2384–2391, 2012.

[19] A. Perina, M. Cristani, U. Castellani, V. Murino, and N. Jojic. Free energy score spaces: Using generative information in discriminative classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(7):1249–1262, 2012.

[20] D. Porello and U. Endriss. Ontology merging as social choice. In *Proceedings of the 12th International Workshop on Computational Logic in Multiagent Systems (CLIMA-2011)*, volume 6814 of *LNAI*, pages 157–170. Springer-Verlag, July 2011.

[21] L. Rabiner. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proc. of IEEE*, 77(2):257–286, 1989.