Complex Socio-Technical Systems: Frames and Values
Symposium Program

Affiliated meeting at the 15th Congress of Logic, Methodology and Philosophy of Science (CLMPS 2015); University of Helsinki, Finland, 3—8 August 2015.

Organized by: the DLMPS commission on the Philosophy of Technology and Engineering Sciences. Organizer: Sjoerd Zwart, s.d.zwart@tudelft.nl

Aim and Focus

The notion of socio-technical (ST) system was introduced by Eric Trist, Ken Bamforth and Fred Emery while working at the Tavistock Institute in London during the postwar reconstruction of the British coal mining industry. Their version of the term referred primarily to the interplay between human and technological factors in working environments within organizations. Besides this intra-organizational meaning of the term, a second connotation of ST-systems has emerged, primarily in engineering and managerial contexts. It refers to large, often infrastructural systems embedded in society as a whole such as the various worldwide transport systems (railway, traffic, aviation, and waterway), large energy distribution and storage grids, water supply, sewage networks and digital information distribution grids such as the internet. This societal, extra-organizational version of the ST-system notion is rooted within systems engineering, and users of the notion focus on the design and management of large and complex ST-systems (CST). CST-systems are different from engineering systems because, besides being larger and more complex, they are deeply intertwined with social reality made up by human individuals (in various roles such as multi-purposed users, operators, service men, inspectors, etc.), groups of human beings (such as action groups, organizations, legislators, governments etc.), and other more general structures of social reality (such as norms and value systems, legislation, regulations, monetary systems, states, etc.).

CST-systems are philosophically interesting and relevant for at least epistemological and practical reasons. First, the ways in which the physical, social and the normative levels of reality are interwoven within CST-systems makes the explanation and prediction of their features extremely complicated and sometimes even impossible. CST-systems exhibit deep uncertainty, feature emergent properties, and develop on various levels on different time scales. On a more general level we should ask what the most appropriate ways are to get a cognitive grip on CST-systems. Anti-naturalists should decide whether they should be studied using the methods of the natural or the human sciences, and if both, how these methods should be combined. Naturalists will not come very far by only applying various versions of the empirical circle to get practical knowledge about how CST-systems behave. Unfortunately, sociology does not offer much help. Until today sociologists have been reluctant to study thoroughly the influence of technology on social structures.

Next, the unpredictability of CST-systems poses difficult problems for those who want to optimize, redesign and design them. Unfortunately, CST-systems do not let themselves to be designed and developed as refrigerators or cars. Simple models of instrumental design rationality fall hopelessly short for their development. To begin with, CST-systems are almost never designed from scratch. Many elements of the future CST- system often already existed before the CST-system has come into being and should be rearranged or adapted within the redesign process. Moreover the design requirements are not given at the start of the project but develop all along the way during the redesign. The goals are adapted to the means available and CST-systems have so many different kinds of roles and stakeholders that many of their goals may even be mutually exclusive.

Finally, many important societal problems are linked to CST-systems. Questions about information and transportation infrastructures and their security are often closely related to philosophical discussions about global justice; climate change, energy and mobility; sustainability and the environment. Moreover, these problems are here to stay—and are even likely to increase their impact in the future. In this symposium we focus on different frames on CST-systems and their related values. In the first four presentations we address the questions of how to model and conceptualize CST-systems and the various roles that they involve; after that we will turn to the practical and normative side of CST-systems.

In the first contribution *Nicola Guarino* asserts that an agent assumes an assigned role if she is supposed to fit a predefined behavioral pattern imposed to her (e.g., an employee); a non-assigned role is bestowed on an agent if a certain person just happens to fit a certain pattern of social expectations emerging from the interaction (e.g., a stakeholder). This difference in the direction of fit [Anscombe, Searle] has considerable implications for ST-system classifications. Next *Stefano Borgo* invites us to consider a context of CST-systems to be an ontologically mixed object that comprises two types of entities: (1) a description that lists the relevant entities and gives their relationships and roles (the conceptual framework) and (2) the physical, technical, social and information entities that actually instantiate, perhaps just in part, this description (the actual instantiation of this framework). He will discuss how the two-component interaction helps to model, analyze and compare CST-systems. After that, *Daniele Porello* and *Roberta Ferrario* will develop an ontological entanglement model of the normative, social and technical information from a sociotechnical stance. They will characterize legitimate sociotechnical ascriptions to several entity types within their ontology. Then, during the fourth and final contribution to the framing part, *Maarten Franssen* will explicate the complexity of CST-systems, using the multitude of the perspectives, actions, and goals within such a system and the question of which agent uses what means to achieve what end within the system. He will offer conceptual tools for dealing with this complexity, presumed to be available in engineering design. Franssen will go back to the barest outline of CST-systems, focusing on a few toy situations, such as the regulation of traffic on a crossroad through traffic lights and a taxi company offering taxi cabs for hire. This will end the part on the conceptualizing part of the meeting.

After the refreshments we will turn to normative issues relating to the design of, and autonomy within CST-systems. First Rafaela Hillebrand will discuss CST-system design for values and the role of procedural values in institutional design. She contends that CST-systems require an encompassing concept of value-sensitive design that has to look both at technology and the accompanying institutions. She illustrates her case by wind parks in the North Sea, where adequately designed institutions are necessary to ensure certain values, such as network stability, which are commonly perceived as "technological." What is perhaps even more important, she claims that certain institutions are needed to incorporate procedural values, e.g., by creating the institutional framework for stakeholder participation. Next, Sabine Thürmel focuses on participation in autonomous systems. She argues that in current CST-systems novel varieties of interplay between humans, robots and software agents are on the rise. Due to these developments we may speak of a participatory turn when assessing the current division of labor between humans and nonhumans. While the autonomy of the technical agents and their abilities increase over time human autonomy may be decreasing. Thürmel argues that to balance human and technical autonomy, we need a responsible innovation process that guides the modeling and employment of ST-systems.

Short abstracts of the planned talks

1. The Crucial Role of Stakeholders in CST- Systems
Nicola Guarino

A proper understanding of roles is of fundamental importance in ST-systems, where people, artifacts, organizations and norms interact one each other in different ways. Indeed, the way each of these components behaves in the system depends on the *role* it plays, i.e., on the specific relationship between the component and the system. So, a formal characterization of such different roles and their interplays can contribute to understand the *nature and structure* of ST-systems, i.e., their *ontology*.

In particular, most of the roles occurring in ST-systems are social roles, in the sense that bearing such roles presupposes some kind of intentionality external to the bearer: Being an *employee* or a *stakeholder* of an enterprise are typical prototypes of social roles.

There is however a crucial difference within social roles, which is very relevant to understand the nature of ST-systems. This is the difference between *assigned* and *non-assigned* roles. The *employee* role is assigned to a certain person, so that she is supposed to fit a pre-defined behavioral pattern *imposed* to her; the *stakeholder* role is not *assigned*, but just *recognized*, in the sense that a certain person *just happens* to fit a certain pattern of social expectations emerging from the interaction. So, the *direction of fit* [Anscombe, Searle] is different in the two cases. As a consequence, there is a striking difference between the two kinds of role, with very practical implications for ST-systems: only for assigned roles, and not for non-assigned ones, the bearer can be *replaced*: you can replace an employee, but you can't replace a stakeholder. Thanks to this formal difference between roles, we can give the following short answers to some of the questions raised in this Symposium:

1. An ST-system does necessarily include assigned roles (with both human and non-human bearers).

2. It is exactly the presence of non-assigned roles that distinguishes CST- from ST-systems (and makes it impossible to fully design them).

2. Context as a modelling device for CST systems
Stefano Borgo

Context, i.e. the set of facts or circumstances that surround a situation or event (WordNet) or the circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood (Oxford Dictionary), is one of those notions that are hard to avoid, even in scientific talk, when discussing complex entities. Unfortunately, the notion of context itself is unclear and, at the same time, very broad in meaning: it is hard to characterize what it means or the boundaries of its application. Multifarious terms like ‘circumstances’, ‘surrounding’ and ‘setting’ are widely used across dictionaries, research papers and other information resources when trying to ‘explain’ contexts.

Complex social-technical (CTS) systems cannot be fully analyzed in a linear way and people resort to *contextual talk* to introduce and discuss aspects of their structures, behaviors, multiplicity of roles, functions etc. To take advantage of this capacity, we aim to ontologically characterize contexts and to indicate how to formalize them in logical terms (so to make them part of a formal machinery) for this kind of use in CST-systems. Of course, even when addressing CST-systems one can easily build discourse environments[1] that reintroduce the variety of possible meanings of `context'. Nonetheless, by seeing a CST-system as a study subject and by just using contexts for its analysis, we can successfully anchor the notion.

We propose to see a context as an ontologically mixed object that comprises two types of entities (only the first of these is actually mandatory): a description or information entity that lists what entities are on focus and gives their relationships and roles (context as a conceptual framework to model expressions like ‘in the transportation context’, ‘in the context of an exam’), and the physical, technical, social, information entities that actually instantiate, perhaps just in part, the description (context as a situated framework to model, e.g., ‘in the context of the Italian transportation system’, ‘in the context of last week exam for class P101’). We will present how the two components interact from the (applied) ontological perspective, the potentiality of this definition and examples of how it helps to homogeneously model, analyze and compare general and specific cases of CST-systems.

3. The Socio-Technical Stance

Daniele Porello and Roberta Ferrario

In this work, we apply ontological analysis in order to answer the question: “what kind of entity can be categorized as a socio-technical system (STS)?” In particular, we will discuss the identification problems for STSs and try to understand whether and, in case, how their boundaries can be set out. The adjectives “socio-technical” or “socio-material” have been applied to a wide spectrum of very different things: organizations, facilities, institutions, even general social relations between individuals. We will argue that, rather than providing necessary and sufficient conditions for classifying something as a STS, it is more informative to investigate and model felicitous ascriptions of socio-materiality or socio-technicality to organizations, institutions, etc. For instance, an airport can be described as a STS but also as a complex technical artefact, a geographical area, a public or private company, an organization, a group of individuals. We claim that all these heterogeneous layers are not mutually exclusive, they are rather co-present in such a complex system as an airport and must be

[1] The terms ‘context’ and ‘environment’ are strongly related: informally speaking, the latter is a context which is maximal wrt some dimension or property.

taken into account in the analysis and in the modeling. None of these layers (not even their combination) is essentially a socio-technical system; the main claim of this work is that socio-technicality is something that can just be ascribed, when we analyze a complex system at a certain level of abstraction, by applying an analytical attitude that we would like to call "socio-technical stance". Obviously, socio-technicality cannot be applied to whatsoever, so the point is then to understand when we can legitimately ascribe it and what are the consequences of such an ascription. We focus in particular on an element that seems to be specific of STS: the entanglement of layers of heterogeneous information (for instance visual, technical, normative, social...). We will show that the entanglement is non-reducible in the relevant cases of STS. We will develop an ontological model of the entanglement of normative/social/technical information and we will define legitimate ascription of socio-technicality to a number of types of entities in our ontology.

4. Sociotechnical systems and their users

Maarten Franssen

Although there is no sharp, broadly accepted definition of a sociotechnical system, the term roughly indicates a complex entity consisting partly of technical devices and partly of people, in various roles, which has a particular function, i.e. through which some purpose is or some purposes are achieved. Typical real-life examples are large infrastructures (e.g. for energy transportation and distribution, transportation, or communication) or production companies (e.g. mining or manufacture). Such systems figure in an instrumental action context: there are users who realize goals through some form of use. In the case of sociotechnical systems, however, we are dealing with multiple users, multiple forms of use and multiple goals, and accordingly multiple perspectives on how well or poorly the system is functioning.

The concept of using something for some purpose is typically analysed from a single point of view: someone is, say, using a hammer to drive a nail through some pieces of wood. Now suppose this someone asks another person to hold a piece of wood stable during the hammering. Right away we have two perspectives, two actions and two sets of goals, and the question of who uses what to achieve what no longer has an obvious answer; it has to address a complexity completely absent from the initial situation of the lone hammerer.

This talk will explicate that complexity for sociotechnical systems and offer conceptual tools for dealing with it, in a descriptive sense presumed to be available in engineering design to account for the proper functioning of artifactual systems. It will do so by focusing on a few toy situations, such as the regulation of traffic on a crossroad through traffic lights or a taxi company offering taxi cabs for hire.

5. Design for Values and CST--Systems. The Role of Procedural Values and Institutional Design

Rafaela Hillebrand

Today ethics of engineering and technology does not contend itself with its original role as a retrospective technology assessment; rather it increasingly aims to incorporate ethical and societal values already in early design phases of new technologies. Here particularly the design-for-value approaches such as value-sensitive-design or design-for-sustainability have gained prominence. While these have been successfully applied to, for example, ICT technologies, we contend in this paper that a design-or-value approach for CST-systems requires a broader perspective than only on the design of the technological components. We argue for an encompassing concept of value-sensitive design that necessitates looking not only at the technology itself and its impacts, but also at the accompanying institutions.

In developing our argument we zoom in on energy systems with components as diverse as power plants, electrical grid(s), storage facilities as well as the users and other stakeholders. More specifically, we take the current development of offshore wind parks in the North Sea (including the accompanying changes in grid, storage etc.) as a study case. It is shown that in the case of CST-systems adequately designed institutions are necessary to ensure certain values that, such as network stability, are commonly perceived as "technological" values. Moreover, and maybe more importantly, certain institutions are needed to incorporate procedural values, e.g. by creating the institutional framework for stakeholder participation. We present a blueprint for such an encompassing design for values approach to a combined technological and institutional design of offshore wind parks in the North Sea, which aims to provide insights into a design-for-values approach to energy systems and CST-systems more generally.

6. Participation in Autonomous Systems

Sabine Thürmel

In current socio-technical systems novel varieties of interplay between humans, robots and software agents are on the rise: Software agents and robots may support humans, act on their behalf or even collaborate with them. Both regulation and control can be delegated to technical agents. Smart energy or Smart health systems are a case in point. In many distributed problem solving approaches humans and technical actors have become interaction partners. Emergency response systems based on multi-agent systems exemplify this development. Due to these developments we may speak of a participatory turn when assessing the current division of labor between humans and nonhumans. New capabilities in technical agents may emerge over time on the individual level due to machine learning algorithms. Even new cultural practices and novel policies may emerge: Big Data approaches may be employed for the optimization of individual behavior based on Big Personal Data or for the optimization of the behavior of a social system relying on Big Social Data. Autoadaption may occur on the individual and on the system level. Such social engineering is used in proactive health systems nudging the human users to (social) conformity with the predefined goals. While the autonomy of the technical agents and their abilities increase over time human autonomy seems to be decreasing. The governance embedded in these systems restricts the autonomy of the human participants and imposes an opaque guidance. In order to balance human and technical autonomy a responsible innovation process guiding the modelling and employment of such systems is essential.

Preliminary Program

1. **The Crucial Role of Stakeholders in CST- Systems**
   Nicola Guarino
   e-mail: nicolguar@gmail.com

2. **Context as a modelling device for CST systems**
   Stefano Borgo
   e-mail: stefano.borgo@cnr.it

3. **The Socio-Technical Stance**
   Daniele Porello and Roberta Ferrario
   e-mail: danieleporello@gmail.com]

4. **Sociotechnical systems and their users**
   Maarten. Franssen
   e-mail: M.P.M.Franssen@tudelft.nl

   -- Break --

5. **Design for Values and CST--Systems. The Role of Procedural Values and Institutional Design**
   Rafaela Hillebrand
   e-mail: rafaela.hillerbrand@gmail.com

6. **Participation in Autonomous Systems**
   Sabine Thürmel
   e-mail: Sabine.Thuermel@gmx.net