Chapter 5

Cognitive Systems Engineering

Gordon Baxter, University of St Andrews

Summary

Cognitive systems engineering is an approach to socio-technical systems design that is primarily concerned with the system's behaviour – what it does and why? This contrasts with other perspectives on socio-technical systems whose emphasis is on structure and relationships between the system agents. It has been primarily used in domains where systems are safety-critical. In use, it focuses on an analysis of how people cope with complexity, understanding how artefacts are used and understanding how people and artefacts work together to create Joint Cognitive Systems.

Background

Cognitive Systems Engineering (CSE) was proposed in the 1970s but was only fully formulated in the early 1980s by people from the field of human factors and ergonomics, and the cognitive sciences. CSE applies a functional approach to the study and development of human-machine systems, focusing on what the system does, rather than how it does it. When CSE practitioners talk about machines, they mean any artefact that has been designed for a particular purpose. Systems are analysed in terms of their joint cognitive systems in order to emphasise the central idea of the co-agency of the human and the machine, rather than simply focusing on the interaction between the two.

CSE largely grew out of work in industrial control systems, particularly nuclear power and now mostly concerns itself with complex application domains (aviation, health care, nuclear power and so on). Its focus is on how systems cope with complexity, how artefacts are used, and the joint cognitive systems that are involved. The goal of CSE is to identify the things (people, resources, artefacts and so on) that are necessary to support the observed ways in which people and systems work, and to use these insights to inform system development.

CSE was developed to deal with three particular problems that were becoming apparent as computer-based systems came into widespread use in the 1970s:

- The increasing complexity of socio-technical systems, which was largely due to large-scale computerisation.
- The ways that the new technologies were being inadequately deployed, leading to a rise in problems and failures (such as the nuclear accident at Three Mile Island in 1979).
- The limitations in existing models for designing and describing systems. These models were mostly linear, and limited in their applicability, such as the paradigm that viewed people as relatively simple information processors.

One of the aims of CSE was to try to bring about a paradigm shift in the way that people thought about developing interactive systems. The early work on interactive systems had focused on the more physical aspects of the work that needed to be done, such as the studies of socio-technical systems carried out by the Tavistock Institute on long wall coal mining. The increased use of automation and computers, however, had started to shift the emphasis from physical work to cognitive work, which was something that CSE directly addressed. CSE also explicitly recognised the concept of emergent behaviour, noting that systems were more than just the sum of their parts, and how system behaviour is an ongoing process that follows a continuous cycle, rather than made up of single (apparently) unconnected interactions.

CSE considers all systems to be socio-technical systems, so can be considered as another perspective on socio-technical systems design. Rather than focus on structural aspects and the relationships between the agents in the system, CSE focuses on the system's behaviour: what it does, and why. CSE therefore focuses on what it describes as cognitive systems, and how these are comprised of joint cognitive systems.

CSE uses a range of methods for collecting data. These include observation (which may be recorded), interviews (including Knowledge Elicitation techniques, such as the Critical Decision Method), verbal reports (think aloud protocols) and instruments to collect ratings (surveys and questionnaires) and categorisation data (card sorting). There is some overlap with the methods that ethnographers use, but CSE practitioners use these methods to look at specific issues, often over limited time periods, and analyse the data to produce results that can be used to directly inform system design.

The most recent (and stable) incarnation of CSE is described in detail in the two books by its originators, Erik Hollnagel and David Woods (Joint Cognitive Systems: Foundations, and Joint Cognitive Systems: Patterns in Cognitive Systems Engineering), and developments continue to take place in this area, where there is an active global community. There are strong links between CSE and Resilience Engineering, and these are reflected in the overlapping communities of practitioners and academics.

Cognitive Systems

In formal terms, a cognitive system is any system that can modify its behaviour on the basis of experience so as to achieve specific anti-entropic needs. In other words, a cognitive system can control what it does. Under this definition most living organisms and some kinds of machines are cognitive systems. Machines are a subset of the more general class of artefacts.

Organisations can be considered to be cognitive systems, because they can control what they do. They can also be considered as artefacts too, albeit artefacts that have been designed for a particular purpose, even though organisations are of a social, rather than a physical nature.

The technological aspects of cognitive systems are mainly of interest because of how they are used. These systems are invariably embedded in a socio-technical context: people (and organisations, more generally) are involved in designing, building, testing and using cognitive systems. From the CSE practitioner's viewpoint all systems are considered to be socio-technical systems. A distinction is drawn, however, between the technological system, in which the technology plays a central role in determining what happens, and the organisational system, in which people play that particular role. CSE practitioners are mostly concerned with applications in complex dynamic domains, such as aviation, industrial process control, healthcare and so on.

Studies of these domains are based around three identifiable, interleaved threads:

- Understanding how people cope with complexity, particularly the complexity that has arisen through advances in technology and socio-technical changes.

- Understanding how artefacts are used, particularly how they have become an inherent part of people's activities (both work and leisure).

- Understanding how people and artefacts can be described as Joint Cognitive Systems, and hence how they can work together.

Joint Cognitive Systems

The joint cognitive system (JCS) is the basic unit of analysis in CSE, and is used to emphasise the central idea of co-agency. In other words the human and machine have to be considered together, rather than as separate entities linked by human-machine interaction. CSE uses a functional approach to the analysis of systems, focusing on what a JCS does and why, rather than how it does it.

It is important with JCSs (as it is with any system) to define the boundaries clearly, and make them explicit. The boundary of a JCS is determined by the purpose of the JCS, and the focus of the analysis of that JCS.The distinction between a cognitive system and a JCS is clearest at the level of the single individual: a person is a cognitive system, but cannot be a JCS. A person using an artefact, however, is a JCS, and a group of two or more people is a JCS too. In practice, CSE focuses on JCSs which exhibit at least one of the following characteristics:

- **The functioning of the JCS is non-trivial**. In other words, obtaining a response from the JCS requires more than a simple action. In those cases where more complex artefacts are involved, it may be necessary to engage in some form of planning in order to produce a response.

- **The functioning of the artefact is at least partly unpredictable**. This may be attributable to ambiguities in the design, for example, which make it hard to identify what a particular widget is supposed to do, or could be due to a lack of knowledge or training on the part of the users.

- **The artefact itself entails a dynamic process**. In such cases the pace of events and the way that they develop is not driven by the user. The corollary of this is that time is regarded as a scarce resource.

CSE starts with trying to understand the patterns of work, using observation of JCSs. It then uses this understanding to guide the search to identify the things (people, resources and so on) that would be useful and necessary to support the types of work that have been observed. The results can be used as a basis for (innovative) design, in participation with others, to support work and the processes of change.

Retrospective

There are inherent links between the fields (and the communities) of CSE and resilience engineering, at least in part because Erik Hollnagel and David Woods have been involved in both from the earliest days. CSE grew out of the cognitive sciences and human factors, however, whilst resilience engineering grew directly out of safety engineering. So one of CSE's aims was to promote a more cognitive-based view of systems, whilst Resilience Engineering was aiming to change the view of safety from the avoidance of failure avoidance to one that emphasises success and learns from that. CSE ideas have been applied in areas such as aviation, and industrial process control, and more recently in healthcare. The research into CSE continues alongside research in Resilience Engineering, and in some cases is being used to provide a foundation for engineering resilience into new systems, particularly in healthcare (and particularly in the USA). This work is likely to continue for the foreseeable future.