Human-centred assembly: a case study for an anthropocentric cyber-physical system

Constantin-Bala Zamfirescu\textsuperscript{a,b}, Bogdan-Constantin Pirvu\textsuperscript{a,*}, Dominic Gorecky\textsuperscript{a}, Harish Chakravarthi\textsuperscript{a}

\textsuperscript{a}DFKI - Innovative Factory Systems, Trippstadter Strasse 122, Kaiserslautern, Germany
\textsuperscript{b}Lucian Blaga University of Sibiu, Department of Computer Science and Automatic Control, Emil Cioran 17, Sibiu, Romania

Abstract

To engineer the factory of the future the paper argues for an anthropocentric cyber-physical reference model that assimilate in an integrated, dynamic, structural and functional way all the required components (i.e. physical, computational and human) of a synthetic hybrid system. This is due to the real need to design large-scale complex systems that accommodate the latest achievements in factory automation where the human is not merely playing a simple and clear role inside the control-loop, but is becoming a composite factor in a highly automated system (“man-in-the-mesh”). The concept is demonstrated by instantiating our anthropocentric cyber-physical reference model in a concrete case study, dealing with the cognition augmentation of the human operator in a manual assembly workstation.

Keywords: Cyber-Physical Systems; Human-Maschine Interaction; Augmented Reality; Virtual Reality

1. Introduction

In the last decade the advances in factory automation became aware of the fact that any significant improvement may be achieved only by considering the tight integration of computational, physical and social elements [1]. Due to the relevant structural interactions among these elements, the integration presumes a clear depart from the traditional

\* Corresponding author. Tel.: +49-631-20575-4849; fax: +49-631-20575-3402.
\textit{E-mail address:} Bogdan-Constantin.Pirvu@dfki.de
reductionist approach in science and engineering with the aim to create synthetic hybrid systems that can achieve goals beyond the inherent capabilities of its composite elements (i.e. human, physical and computational). Today, this comprehensive outlook can be found in very dissimilar research areas (e.g. IBM’s Smarter Planet), including the smart factory concept [2].

The integration of physical and computational elements is well-reflected in the standard view of cyber-physical systems (CPSs). A CPS poses some exclusive features that differentiate it from the conventional systems (i.e. embedded systems, sensor networks, etc.) [3][4][5]: integrality (the CPS’s functionalities are relying on the unified composability of its elements with self-organization capabilities, such as learning, adaptation, auto-assembly, etc.), sociability (the ability to interact with other CPSs via different communication technologies, not only device-centred but human-centred as well in an open mixed network environment), locality (the cyber and physical capabilities of a CPS are bounded by the spatial properties of the environment), irreversibility (self-referential timescale, sensed as dynamics, not discrete, nor spatial), adaptive (with self-organization and evolving capabilities), autonomous (control loop must close over the life-cycle of a CPS, including the assimilation of human factor who is constantly closing the loop of any engineered artefact, despite its automation degree), and highly automated (as a key driving-force of eroding the boundaries between its composite elements and favouring their structural interactions). Even if it is not explicitly stated, the human factor plays a crucial role in a CPS to display the above mentioned features. Some recent studies are trying to give a more comprehensive view over CPS that, besides the classical computational and physical dimensions, includes the social one as an integral part of a CPS. This may be observed in the new emerging concepts (e.g. cyber-physical-social systems [6], human system integration [7], smart environments [8], etc.) that are paving the way towards the old vision of symbiotic man-machine systems [9]. For that reason, in [5] we defined the anthropocentric cyber-physical system (ACPS) as a reference model for factory automation that integrates the physical component (PC), the computational/cyber component (CC) and the human component (HC). The key characteristic of an ACPS reference model is its unified integrality which cannot be further decomposed into smaller engineering artefacts without losing its functionality.

From the engineering stance, the ACPS concept emphasizes the adaptive and dynamic division of labour among the ACPS components as a result of their continuous interactions. Consequently, the function allocation cannot be fixed at the design time as in the Fitts’s list [10], but oscillate between different levels of automation (from fully manual to fully automated). In other words, a priori assumptions as regards the function allocations among the ACPS components are useless when the precise future of the ACPS is known only a posterior. As in the latest Seridan’s taxonomy for automation [11], the dynamic allocation mechanism is a consequence of the innate complexity and bounded rationality conditions an ACPS operates. In line with the cybernetics’ Law of Requisite Variety [12], it requires real-time evaluations for the physical/cognitive status of the human operator (in terms of risk, reliability, and costs of automation) to grasp all possible inputs that may affect the ACPS behaviour. Moreover, the adaptive and active allocation of tasks between the ACPS components to provide an optimal workload balance is not restricted to the operational requirements - as considered in the research topic of adaptive automation [13], but includes also the symbolic integration of man and machines in a closed-loop - as considered for example in the augmented cognition research field [14]. Consequently, ACPS is “adaptive automation” + “augmented cognition”.

The paper presents some insides regarding the interaction-based architectural design of a representative workstation from the SmartFactoryKL demonstrator that requires adaptive and dynamic division of labour among the ACPS components, namely manual assembly module. Consequently, the next section will summarize the ACPS reference model of our anthropocentric cyber-physical reference architecture for smart factories (ACPA4SF). A brief description of the SmartFactoryKL manual assembly workstation for customizable products (i.e. so-called keyfinders) is presented in the third section. In the fourth section the workstation will be detailed as an instantiation of an ACPS type from ACPA4SF, i.e. ACPS production system. The paper is summarized in the last section.

2. ACPA4SF reference model

In [5] we defined the ACPA4SF as a composition of four ACPS types that are self-sufficient to describe and engineer any manufacturing control system, from fully centralized to fully decentralized approaches. It includes the (Fig. 1a):
- **ACPS production system** (the production resources available in the factory, i.e. machines, transportation, and storage systems);
- **ACPS product design** (all the necessary product design knowledge and engineering tools to manufacture a product, i.e. manufacturing operations workflow for a product type);
- **ACPS planning and control** (i.e. the orders from the customers in terms of product instances); and
- **ACPS infrastructure** (the contextual data and control elements required by the previous ACPS types to operate in the real factory environment, i.e. buildings, rooms, technological infrastructure, etc.).

To get a product instance manufactured, there is a continuous interactions flow for exchanging relevant knowledge among these ACPS types. For example the **ACPS planning and control** that reflects a product instance (i.e. intelligent product) has to manage its itinerary through the factory by negotiating with other ACPS types to get produced. Consequently, it needs to know: from the **ACPS product design** how to manufacture the product instance, from the **ACPS production system** where and when to execute the processing operations, and from the **ACPS infrastructure**, if the identified manufacturing resources are available at reasonable costs. Similarly, the **ACPS product design** needs to know: from the **ACPS production system** which are the possible manufacturing operations available in the plant and from the **ACPS infrastructure** in what context their availability is valid. These basic ACPS types are widely accepted in the decentralized approaches for factory automation (i.e. multi-agent, holonic, service-oriented, etc.) where the decomposition reflects as accurate as possible the physical reality and not the desired functionalities that may emerge from many interacting components. Their coherent description is beyond the scope of this paper, and is presented/instantiated in [5]/[15].

All these ACPS types that compose the ACPA4SF inherit the ACPS reference model which defines in an abstract way the core relationships among its composite entities (Fig. 1b). The interactions between these components are usually made via adaptors (optional in many cases) that translate the signals into the specific format of the interacting component. For example, between the PC and HC there are special displays or meters to measure the working parameters of a machine; between the CC and HC there are the classical human-computer interaction (HCI) devices, such as screens, mouse, keyboard, etc.; while between the PC and CC there are special transducers or analog-to-digital converters. These components are connected outside the ACPS to their specific dimension: physical (e.g. via mechanical slots), computational (e.g. via computer-specific communication standards), and social (e.g. via natural language). It does not exclude the mediated interactions between the ACPS components (HC to CC.
via PC, and so on) in a so called “smart environment” [8], when the components will play an active role in shaping these hybrid interactions. The ACPS components participate on a role-basis in an ACPS and their inclusion in a concrete instance depends on the engineering compromises that should be accommodated in its real implementation. A role implements one or more interaction protocols (or methods) that accommodate multiple adaptation loops between:

- **PC-HC** - in today factories humans must supervise an automated system or closely co-work with them in a “man-in-the-mesh” manner. All the studies from the broad topic of human-machine interaction provide clear evidence for the continuous mutual adaptation loop between HC and PC in the manufacturing process. Moreover, distributed cognition argues that the acquisition, propagation and processing of information is a distributed process that always happen across a network of humans and artefacts [16]. Consequently, human cognition is situated and embodied, with the power to detect fine-grained patterns and correlations between millions of multi-dimensional signals that are impossible or impractical to transfer into computational systems;

- **HC-CC** - in information science the most widely used research framework for longitudinal studies is the adaptive structuration theory [17]. Basically, the theory criticizes the techno-centric view of using a technology and stresses the anthropocentric aspect, by observing that humans who are using a technology for their work create individual perception about its role. This perception is very dissimilar across groups and has a great influence on how the technology is further used;

- **PC-CC** - the integral relationship between the PC and CC has been studied extensively in robotics, rejecting the traditional assumption of dualism that sees matter and mind/control as independent constituents of reality. For instance in [18] are many examples demonstrating that the behaviour of any system is not merely the outcome of an internal control structure, but it is also affected by the environment in which the system is physically embedded, by its morphology, and by the material properties of the composite elements.

### 3. Manual assembly workstation from the SmartFactoryKL demonstrator

This section describes the manual assembly workstation from the production system for assembling customizable products, so-called **key-finders** that can connect to smartphones to track down missing keys. The key-finder product includes casting-covers, printed circuit board equipped with LED, loudspeaker and a Bluetooth module. The production line is part of the SmartFactoryKL demonstration facility [2] and was designed to test advanced paradigms for manufacturing control (Fig. 2a). It comprises four workstations that implements automation solutions with different degrees of autonomy, from almost complete automation to fully manual operation. In this paper we are focusing on the last workstation from the SmartFactoryKL demonstration facility (labelled “4” in Fig. 2a, where human operator completes the key-finders assembling (e.g. custom-builds). At this station the quality control is performed and/or acts as a backup solution when the automated modules (i.e. the first workstations, labelled “1”, “2” and “3” in Fig. 2a) are down or not properly working. The description of the first modules is beyond the scope of this paper, and its instantiation in terms of ACPA4SF is presented in [15].

The tasks from the workstation 4 consist of manually assembling the key-finders by using the available sub-components which are stored in special part containers located within the workstation. With increasing complexity and variability of products, the main difficulty for the operator is to follow the right assembly procedure to correctly manufacture the desired product. To support this task, there are many possible alternatives to set up a smart working environment (i.e. pick/put-by-light/voice/vision systems) for guiding the human operator during the assembly process. Their role is to provide real-time instructions for the human operator to minimize the cognitive complexity of performing the task. For doing this, the key challenge is to reliably monitor the operations workflow and to capture the current state of the semi-finished product. In the SmartFactoryKL living lab we have adopted a modular and low cost solution, based on augmented reality and advanced sensor technology to support the manual assembly process. A complete technical description of this solution is given in [19]. Here we synthesize only the necessary details to explain the manual assembly process as an instantiation of an ACPS production system.

The workstation supports the production of customized key-finder products in a „one-piece-flow“. These products are coming from the upstream workstation (labelled “3” in Fig. 2a) with casing-covers that have a RFID-tag glued on it which stores the abstract processing plan and the current status of the product. When a new product
instance is arriving (or available in the intermediate buffer between the workstations), the operator uses the RFID reader (labelled “5” in Fig. 2b) to examine the information for the required assembly operation. This information is translated into a workflow of operational tasks that are visualized in form of virtual instructions, to guide the user step-by-step through the manual assembly process. There are two key complementary aspects for correlating the user’s actions and displaying the correct assembly instructions:

- **to provide recommendations for the next operational task** (augmented reality) in respect to: 1) the current state of the assembling process, and 2) the attentional state of the human operator. It presumes the workflow identification for the first case, while for the second one the detection of relevant objects from the environment (in operator’s eyes) for which to generate and display the augmented assembly instructions to the worker. This is done through an augmented reality visualization of the operational task on a computer screen (an Android tablet with a built in camera, labelled “3” in Fig. 2b), which can be oriented towards the worker’s interest area (the so-called “attentional capabilities” [20]). It includes the relevant representations for both: the human operator (i.e. hands) and the physical workplace (i.e. toolsets, parts). Due to the limited visualization capabilities of the tablet, the augmented assembly instructions are replicated on a public display (labelled “6” in Fig. 2b). Anyway, these two displays may be easily replaced with a pair of virtual reality glasses (i.e. Google glass) to better reflect the spirit of this work.

- **to provide the contextual data** (virtual reality) in which the above mentioned instructions are appropriately being tracked by the operator in order to avoid the “automation surprises” [21], i.e. the consistency between the computational system behaviour (on which basis the system is providing the instructions for the next assembly task) and operator’s mental model (on which basis the operator believes the computational system is giving the recommendations by using his/her own observations). This is realized using a real-time three dimensional (3D) virtualization of the work carried out inside the workstation that is displayed on a separate screen (labelled “2” in Fig. 2b). In other words it supports the feedback loop of what is being interpreted by the computational system (i.e. the assembly workflow) and the operator’s work. As in the first case, the visual representation includes both: the human operator and the physical workplace.

Because the above mentioned components of the virtual instructions are communicated in different ways, they are denoted as **augmented reality** (AR) and **virtual reality** (VR) instructions. Consequently, to engineer the above mentioned complex visualizations - synchronized in real-time with the current state of the assembling workflow (visualised in AR) and the operator’s movements (visualized in VR) - two computer vision systems are concurrently used:

- **an overhead 2D static camera** (labelled “7” in Fig. 2b) to monitor the changes in the physical environment by detecting and tracking the relevant objects and tools used by the worker in the manual assembly task. To
minimize the computational power for this task, the computer vision capability is deliberately restricted to the relevant objects (i.e. tools, parts) from the physical environment that are used by the worker in the manual assembly task (everything else is modelled in virtual reality) and labelled with different and easy to recognize markers (labelled “4” in Fig. 2b).

- a Microsoft Kinect for Windows (labelled “1” in Fig. 2b) to monitor the worker and eliminates the need to explicitly confirm the completion of a certain task from the assembly workflow. To increase the detection accuracy, the monitoring of operator’s movements is restricted to what is relevant for the manual assembly task, i.e. tracking the hands.

The data from these computer vision systems are synchronized and integrated by the computational systems to generate the virtual instructions for the worker (either in AR or in VR; see above). The computational system practically integrates the physical (i.e. the workspace detected by the overhead 2D static camera) and the social (i.e. the human operator detected by the Microsoft Kinect for Windows) space. For the VR instructions, there is a real-time 3D rendering module which updates the worker’s avatar and the other objects according to the data received from the static 2D camera and the Kinect©. For the AR instructions, there is a workflow recognition system that implements a “finite automata” with pre and post conditions to detect the sequence of the assembling operations. Contrary to the previous case, it use the data from the 2D static camera and Kinect© to detect the interactions between the worker and objects for recognizing the pre and post conditions in the assembling workflow. The recognition is pattern based and presumes the assessment of real-time data against baseline data obtained during an offline training phase or extracted from Product Lifecycle Management (PLM) data. It includes also algorithms for filtering the operator’s actions to what is considered to be relevant for the assembly (i.e. eliminate short-time interactions with the objects and tools from the worktable, etc.). A complete description of the algorithms and technologies used to implement these modules are given in [19].

4. Manual assembly workstation as an anthropocentric cyber-physical system

In this section the manual assembly workstation will be depicted as an ACPS instantiation of the production system type from ACPA4SF. The instantiation is represented by the description of its structure and the key interactions among the composite components. For simplicity reasons, the interaction with the other ACPS types is deliberately omitted, being presented in [15].

4.1. The ACPS structure for the manual assembly workstation

![Fig. 3. The ACPS structure of the manual assembly workstation from the SmartFactoryKL demonstrator.](image)
The manual assembly workstation reassembles in a clear way the ACPS structure defined in section 2. It includes all the composite ACPS components (Fig. 3): the human operator, the workstation (as an aggregation of PCs), and the AR/VR application (as an aggregation of CCs).

These components are interacting with each other to accomplish the goal of the workstation, i.e. to assemble the key-finders. These interactions are either natural (e.g. between the operator and the physical toolsets and parts from the workstation) or deliberately supported (e.g. between the workstation/operator and the AR/VR system) via special adaptors (e.g. markers/displays) and engineered services (e.g. video cameras, movement recognition, etc.).

Fig. 4 summarizes the interactions described in the previous section among the ACPS components. For simplicity reasons it does not depicts many other relevant interactions, such as: 1) the interactions with the other ACPS types from ACPA4SF (e.g., it needs interactions with the ACPS product design to provide a complete description of the assembly operation; see [15] for a detailed description of this issue), and 2) the interactions that belong to its continuous development and deployment (an ACPS is never created fully formed, but gradually evolves as new requirements are discovered; for example in the manual assembly workstation case, initially just the AR module has been implemented, driving later on to the development of the VR and workflow recognition module). If the first sort of interactions are related to interoperability and goal-achievement tasks of the ACPS (i.e. to assembly the key-finder), the later ones concern the interactions which make the ACPS design intractable for the classical methods of functional decomposition. For example, the interaction protocol from Fig. 4 does not consider some relevant key performance criteria for any production system (i.e. efficiency, energy consumption, development of worker’s skills, etc.) that shape the messages between the ACPS components. Consequently, it should be viewed as a small part from a more extended and realistic one.

4.2. Continuous adaptation loops among the ACPS components

Although not necessary evident, besides the general considerations presented above, the interaction protocol
among the ACPS components includes their continuous adaptation loops. In the manual assembly workstation, there are multiple adaptation loops among the ACPS components, such as:

- **HC-PC.** As mentioned, the AR instructions are the output of a workflow recognition system that implements a “finite automata” to detect the sequence of the assembling operations. The recognition is pattern based and presumes the assessment of real-time data against baseline data obtained during an offline training phase. In the current implementation this training is done offline beforehand and presumes the collection of object interaction sequences and timing data from a typical assembly scenario. These data are very dependent on the physical configuration of the workstation and the current development level of the operator’s skills. When considerable changes in these two issues are happening, the training phase is necessarily repeated. Of course, more advanced features (i.e. machine learning, teaching by example, etc.) could be easily added to automate and simplify this task, but they are irrelevant in respect to the continuous adaptation loop that truly exists between the HC and PC.

- **PC-CC.** The PC is reflected in the CC by using computer vision methods to recognize the relevant movable objects (i.e. tools, parts) for the manual assembly task. These objects are labelled with different and easy to recognize markers for the computer vision system. Anyway, their recognition presumes calibrations (i.e. on different axes and scaling), that should be repeated whenever the available sensors (i.e. video cameras) are moved. Moreover, when the required toolset is changing, or the stable objects that are considered to be irrelevant for the manual assembly operation are modified (i.e. unmarked ones), will need interventions in the CC to accurately reflect the PC. As in the previous case, these continuous adaptation loops between the PC and CC may be further automated using advanced computer vision techniques for objects identification and classification with the aim of semantic technologies.

- **CC-HC.** In this case we face similar concern as in the PC-CC adaptation loop (see above), applying recognition of hands movement instead of objects. Note that it considers the adaptation of CC to HC. Since in the manual assembly workstation the PC does not have any automated component (i.e. mechatronic devices, sensors, actuators, etc.), it will not play an active role in reflecting the other ACPS components. That is way in the previous analysis of the continuous adaptation loop between the HC-PC and CC-PC we considered only the adaptation to the PC and not conversely. For the CC-HC relationship, we should consider also the adaptation of HC to CC as both of them are active components. At this point in time we do not have relevant longitudinal studies to highlight this issue, but is almost evident if we acknowledge the widely-accepted theoretical framework of adaptive structuration theory [17].

5. Summary

The paper presented an instantiation of our previous ACPS reference model where the humans are not just its users, but elements of the system affecting its overall behaviour. Besides the classical control component, it integrates the physical and human ones as well. This is due to the real need to accommodate the on-going researches from the SmartFactoryKL facility (e.g. augmented reality, mobile interaction technology, virtual training of human operators) where the human is not merely playing a simple and clear role inside the control-loop, but is becoming a composite factor in a highly automated system (“man-in-the-mesh”). Even if it is engineered with state-of-the-art techniques and automation components, the manual assembly workstation presents all the key features of an ACPS, i.e. integrality, sociability, locality, irreversibility, adaptivity, and autonomy. These characteristics should be considered over the entire life-cycle of an ACPS, not only in its normal operational phase. Therefore the paper regards as being “highly automated” the development of operator’s skills for a manufacturing task, and not the task per se, i.e. assembling the key-finders – a task that can be easily automated and consequently irrelevant for this work. The basic assumption is that in an increasingly complex and dynamic environment, where advanced automations may be subsequently engineered (e.g. joint task executing, mixed-initiative hybrid interactions, human computation, etc.), the traditional qualification processes will gradually converge towards in-situ assistance (i.e. learning by doing).
Acknowledgements

This work is partially supported by the “CyProS” project (http://www.projekt-cypros.de/, grant number 02PJ2461), founded by the German Federal Ministry of Education and Research (BMBF) within the “Research for Tomorrow’s Production” Framework, and VISTRA-project (www.vistra-project.eu), which is co-funded by the 7th Framework Programme of the European Union.

References

[1] NIST. Foundations for Innovation in Cyber-Physical Systems, Workshop Report, Available at: http://www.nist.gov/el/upload/CPS-WorkshopReport-1-30-13-Final.pdf; 2013.
[2] Zühlke D. SmartFactory - From vision to reality in factory technologies. Proceeding of 17th International Federation of Automatic Control World Congress, South Korea: IFAC; 2008, p. 82–89.
[3] Rajkumar R. CPS briefing, Pittsburgh: Carnegie Mellon University; 2007.
[4] Lee AE. Cyber Physical Systems: Design Challenges. Int. Symposium Object/Component/Service-Oriented Real-Time Distributed Computing. Orlando: IEEE; 2008, p. 363-369.
[5] Zamfirescu CB, Pirvu BC, Schlick J, Zuehlke D. Preliminary Insides for an Anthropocentric Cyber-physical Reference Architecture of the Smart Factory. Studies in Informatics and Control 2013;22:269-278.
[6] Zhuge H. Interactive semantics, Artificial Intelligence 2010;174:190–204.
[7] NASA. Human system integration, Available at: http://human-factors.arc.nasa.gov/; 2013.
[8] Poslad S. Ubiquitous Computing Smart Devices, Smart Environments and Smart Interaction. Chippenham: Wiley; 2009.
[9] Licklider JCR. Man-computer symbiosis. IRE Transactions on Human Factors in Electronics 1960; HFE-1: 4-11.
[10] Fitts PM. Human engineering for an effective air navigation and traffic control system, Columbus: Ohio State University Foundation Report; 1951.
[11] Parasuraman R, Sheridan T, Wickens C. A Model for Types and Levels of Human Interaction with Automation. IEEE Transactions on Systems, Man and Cybernetics 2000;30:286-297.
[12] Ashby WR. Requisite Variety and its implications for the control of complex systems. Cybernetica 1958;1:83-99.
[13] Scerbo MW. Adaptive automation. In: Karwowski W, editors. International encyclopedia of ergonomics and human factors. London: Taylor and Francis; 2001.p. 1077-1079.
[14] Engelbart DC. Augmenting human intellect: a conceptual framework. AFOSR-3233 Summary Report. Menlo Park: Stanford Research Institute; 1962.
[15] Zamfirescu CB, Pirvu BC, Loskyll M, Zuehlke D. Do Not Cancel My Race with Cyber-Physical Systems. Proceedings of the 19th World Congress of the International Federation of Automatic Control. Cape Town: IFAC; 2014. (to be published).
[16] Hutchins E. Cognition in the wild, Cambridge: MIT Press; 1995.
[17] DeSanctis G, Poole MS. Capturing the complexity in advanced technology use: Adaptive structuration theory. Organization Science 1994;5:121-147.
[18] Pfeifer R, Bongard JC. How the Body Shapes the Way We Think. A New View of Intelligence, Cambridge: MIT Press; 2007.
[19] Gorecky D, Campos R., Chakravarthy H., Dabelow R, Schlick J, Zühlke D. Mastering Mass Customization – A Concept For Advanced, Human-Centered Assembly. Academic Journal of Manufacturing Engineering 2013;11:62 -67.
[20] Helgason HP, Thörisson KR, Attention Capabilities for AI Systems. Proceedings of the 9th International Conference on Informatics in Control, Automation and Robotics. Rome: SCITEPRESS; 2012, p. 281-286.
[21] Salvendy G. Handbook of Human Factors and Ergonomics, 4th ed. Hoboken: John Wiley and Sons; 2012.