Modeling and Specifying Requirements for Cyber-Physical Systems

L. Ordinez, Member, IEEE, O. Alimenti, Member, IEEE, E. Rinland, M. Gómez and J. Marchetti

Abstract— Cyber-physical systems have to do with a strong interaction between the physical world and a computing system, which should be transparent to an external observer. In this paper, an approach for modeling and specifying requirements for these systems is presented. It is called 3D Approach and it is based on the well known 4-Variable Model. The proposed extension intends to tackle the particularities involved in the development of cyber-physical systems by a multidisciplinary team and consequently it emphasizes a practical and applied point of view. The 3D Approach came out as the result of systematizing lessons learned in the design of a biodigester, which allows the anaerobic digestion of biomass for the production of biogas.

Keywords— Requirements modeling, requirements specification, cyber-physical system, biogas, biodigester.

I. MOTIVATION

THE CULTIVATED onion area in the Patagonian Protected Region varies annually between 8500 and 12000 hectares. The protected region area has also a great manufacturing capacity because it has skilled work force and infrastructure to process large volumes of onions: every year at least 70 packhouses are enabled, most of them in the Colorado River valley in Buenos Aires province. According to statistics of Fresh Onions Certification program carried out by Zoophytosanitary Patagonian Foundation Barrier (FunBaPa) [1], during the first five months of 2008, onion exports from that region reached 156410 tons. The problem of the final disposal of residues left in fields, such as those from packhouses, is a major concern for farmers and residents of Patagonian Protected Region.

Biodigesters are devices that enhance organic matter degradation trough anaerobic digestion [2]. This technology allows a sustainable production of methane gas, which could be produced from organic household, industrial waste or from crops grown for that purpose. However, anaerobic digestion process is extremely sensitive. This aspect makes necessary to control different variables in order to achieve not only efficiency but for the process to actually work. In this sense, it was decided to model the biodigester and anaerobic digestion process, along with an electronic system for monitoring and control of it, as a cyber-physical system.

Figure 1. View of a field with onion wastes in the region of Villarino.

Cyber-physical systems represent the intersection between the computational world and the physical environment [3]. Particularly, the computational part of CPS is constituted by several areas that form the basis of this new research discipline. Thus, this new field takes concepts from the mentioned areas, but it adapts them to inherently involve the physical environment. Some characteristics of CPS are [4]: 1) Physical world modeling, 2) Control theory, 3) Communication, 4) Human interaction, 5) Real-time capabilities, 6) Real-time operating system support and 7) Electronic hardware. Note that these features will be taken into account in the proposed method.

In this paper, a requirements modeling and specification approach for CPS is presented. The approach is based on the well known 4-Variable Model, but it introduces some extensions and adaptations to be used in CPS. The proposed approach is centered in the environmental aspects and the way in which they are affected by the computing system. This is essential when dealing with CPS, since the vision of the complete system as a domain co-managed by the physical laws and the computer system, defines a CPS. Another aspect to consider is the inherent communication between the different components that constitute the computer system. The previously mentioned characteristics and a clear identification of the software entities involved, are the ones in which this paper focuses to adapt the mentioned approach. It is worth noting that the proposed adaptation is based on matters that came out from practical experience. Although emphasis is placed on a practical vision, this is not detrimental to achieve the formal rigor of a scientific work.

Based on the previous paragraph, the rest of the paper is organized as follows: Section II presents the case study that conducted this research; in Section III the adaptation and extension performed is exposed; a discussion of the proposed
approach is done in Section V; finally, conclusions and future works are drawn in Section VI.

II. CASE STUDY DESCRIPTION

Anaerobic digestion is a controlled biological degradation process that can be used to treat several organic wastes and recover bioenergy in the form of biogas. This process consists of a series of bacterial transformations that convert organic compounds mainly to methane (CH4), carbon dioxide (CO2), and new bacterial cells as end products of the process. These events are commonly considered as a three-stage process [5], [6]. The first stage involves the hydrolysis of solids that results in the production of soluble organic compounds (volatile acids and alcohols). The second stage, acetogenesis, involves the conversion of the volatile acids and alcohols to substrates such as acetic acid or acetate (CH3COOH) and hydrogen gas that can be used by methane-forming bacteria (MFB). The third stage, methanogenesis, involves the production of CH4 and CO2 by MFB [6].

The anaerobic digestion process proceeds efficiently if the degradation rates of all three stages are equal. The groups of bacteria responsible of the stages mentioned before, work in sequence, with the products of one group serving as the substrates of another group [6]. There are four groups of bacteria involved in this process. Hydrolytic bacteria break down complex organic molecules present in wastes into simple sugars, amino acids and fatty acids. Next, acidogenic bacteria convert these products into organic acids, which are then degraded by acetogenic bacteria to H2, CO2 and to a huge amount of acetate. Finally, MFB produce biogas from acetic acid, H2 and CO2 [7].

III. EXTENDING THE MODEL: A 3D APPROACH

In this section, the 4-Variable Model, proposed by Parnas and Madey [8] (The original method proposed by Parnas and Madey was then extended and adapted to many types of systems. The works related to the here proposed are [9], [10], [11], [12]), is extended to cope with the characteristics commonly found in CPS (A list can be found in [13], [14]). Note that in the following description, some concepts remain the same as in the original one; however, they are presented again to remain consistent.

A. General Scheme

The proposed extension, called 3D Approach, is depicted in Fig. 2. This new model maintains the same basic sets of variables and their relationships (i.e., MON, CON, INPUT and OUTPUT; and NAT, REQ, IN, OUT and SOF, respectively). However, it expands them to several planes, each of which constituted by a particular case of the original 4-Variable Model. Consequently, forming a three-dimensional body. The key idea behind this is to express the different physical laws involved in a process by means of simpler ones. In addition, this viewpoint allows to establish relationships between control laws (and hence between software requirements) and to have a clearer vision of the system at the time of implementing it.

From a geometric point of view, in this new scheme, the sets MON, CON, INPUT and OUTPUT are not just corners in a rectangle, but edges. Additionally, relationships REQ, IN, OUT and SOF are no longer lines, but planes. Fig. 3 shows those planes. This new formulation eases the separation of concerns in the sense of splitting the different components of the systems in terms of their particular functionalities. This is, the separation of concerns is done on the basis of the purposes of the system, which are expressed as control laws. Moreover, within each plane, a subset of related lines can be grouped together, constituting a subgroup of relationships. Despite pursuing a distinct purpose, each relationship in this subgroup of relationships share a common aspect or characteristic. They are related in some way. This feature will then be exploited to formally establish communications within the system.

Another important aspect that comes out from the 3D Approach, and that is related to the previously mentioned separation of concerns, is that of working with a Real-Time Operating System (RTOS). A fundamental issue to have into account when working with an RTOS is the identification of the tasks that compose the system. Here, the questions that should be answered are: 1) How to establish the functionalities of a system? 2) Which functionalities of the system are implemented through which tasks? 3) How to determine a criteria to perform such implementation?
(i.e., each individual 4-Variable Model) will be implemented as a real-time task. More specifically, each SOF requirement is implemented by means of a real-time task. The concept of implementing a requirement as a real-time task is not new and can be also found in [15], [16] for different kinds of systems.

B. The 3D Approach

Following, the variables, sets and relationships of the 3D Approach will be defined. As mentioned in the introduction of this work, the aim is not to make rigorous demonstrations of the approach, but to present it from a practical point of view. However, certain degree of formalism is necessary in order to make it clear, understandable and unambiguous.

Definition 1 (Monitored variable). A monitored variable is a physical quantity, that is essential to be measured and that may vary over time. Formally, a monitored variable $j$ is expressed as $m_j$ or equivalently $m_j(t)$, where the $t$ represents the time-dependent aspect.

Definition 2 (Controlled variable). A controlled variable is a physical quantity that may vary over time. Formally, a controlled variable $j$ is expressed as $c_j$ or equivalently $c_j(t)$, where the $t$ represents the time-dependent aspect.

Definition 3 (Requirement relationship). The relationship REQ expresses how certain controlled variables are to respond to changes in certain monitored ones. Note that in order to consider the variables actually involved in the relationship, REQ, must be expressible as a control law.

Definition 4 (Set of Monitored Variables). A set of monitored variables $MON_i$ is composed only by those monitored variables $m_j$ that are involved in the relationship REQ.

Definition 5 (Set of Controlled Variables). A set of controlled variables $CON_i$ is composed only by those controlled variables $c_j$ that are involved in the relationship REQ.

Definition 6 (Natural relationship). The relationship NAT expresses the physical law that states how certain controlled variables are related to certain monitored ones.

The above definitions are still the same ones as proposed by Parnas and Madey and are limited to one vertical plane. However, as mentioned earlier, the 3D Approach considers multiple planes composed by individual 4-Variable Models. In this respect, the following definitions extend the 2D scheme to the new 3D one.

Definition 7 (Edge of Monitored Variables). The Edge of Monitored Variables MON is constituted by all the sets of monitored variables $MON_i$. Formally, $MON = \bigcup MON_i$.

Definition 8 (Edge of Controlled Variables). The Edge of Controlled Variables CON is constituted by all the sets of controlled variables $CON_i$. Formally, $CON = \bigcup CON_i$.

Definition 9 (Plane of Requirements). The plane of requirements REQ is composed by every requirement relationship REQ in the system. Formally, $REQ = \bigcup REQ_i$.

Definition 10 (Plane of Natural Laws). The plane of natural laws NAT is composed by every natural law $NAT_i$ in the system. Formally, $NAT = \bigcup NAT_i$.

A first implication of the previous definitions has to do with hardware devices. Once determined the monitored and controlled variables involved in a requirement, and consequently the edges of variables $MON$ and $CON$ and the plane of requirements $REQ$, an obvious next step is to establish which kinds of sensors and actuators are needed. At this point, the important aspect is to identify the kind of device needed and not to go into any depth about its characteristics. This identification is a simple example of an aspect found in CPS that is considering the electronic hardware at the same time of modeling the physical system.

The previous definitions are the ones more related to the physical environment. Next, those closer to the computing system are exposed.

Definition 11 (Input Variable). An input variable is a finite, binary representation of a monitored variable. Formally, an input variable $j$ is expressed as $i_j$ or equivalently $i_j(t)$, where the $t$ represents its time-dependence.

Definition 12 (Output Variable). An output variable is a finite, binary representation of a controlled variable. Formally, an output variable $j$ is expressed as $o_j$ or equivalently $o_j(t)$, where the $t$ represents its time-dependence.

Definition 13 (Set of Input Variables). A set of input variables $INPUT_i$ is composed only by the input variables $i_j$ that represent the monitored variables $m_j$ in the set of monitored variables $MON_i$.

Definition 14 (Set of Output Variables). A set of output variables $OUTPUT_i$ is composed only by the output variables $o_j$ that represent the controlled variables $c_j$ in the set of controlled variables $CON_i$.

Definition 15 (Edge of Input Variables). The edge of input variables $INPUT$ is constituted by all the sets of input variables $INPUT_i$. Formally, $INPUT = \bigcup INPUT_i$.

Definition 16 (Edge of Output Variables). The edge of output variables $OUTPUT$ is constituted by all the sets of output variables $OUTPUT_i$. Formally, $OUTPUT = \bigcup OUTPUT_i$.

Definition 17 (Software Requirement relationship). The relationship $SOFi$ expresses the corresponding requirement relationship in terms of input and output variables.

Definition 18 (Plane of Software Requirements). The plane of software requirements $SOFi$ is composed by every software requirement relationship $SOFi$ in the system. Formally, $SOFi = \bigcup SOFi$.

The previous definitions set up the variables, edges and planes that are closer to the physical environment and the computing system, respectively. Following, the definitions concerning the interface between the two worlds are presented.

Definition 19 (In Relationship). The relationship $IN_i$ expresses the mapping of monitored variables to input variables. Formally, $IN_i : MON_i \rightarrow INPUT_i$.

Definition 20 (Out Relationship). The relationship $OUT_i$ expresses the mapping of output variables to controlled variables. Formally, $OUT_i : OUTPUT_i \rightarrow CONTROL_i$.

Definition 21 (In Plane). The plane $IN_i$ is constituted by every $IN_i$ relationship in the system. Formally, $IN_i = \bigcup IN_i$.

Definition 22 (Out Plane). The plane $OUT_i$ is constituted by every $OUT_i$ relationship in the system. Formally, $OUT_i = \bigcup OUT_i$.

The above relations define how the physical world is mapped to the computer one, and vice versa. These interfaces are implemented through hardware devices that may be
remote or even be constituted as a computer system themselves. Thus, in summary, it can be seen how the proposed approach includes not only the physical world and the computer system, but also the interfaces between them.

Following, a series of results are formally presented. They state a way of specifying requirements in the sense of providing an agreement, between the customer and the developer, on what is being developed. These results reinforce the practical vision devised along this paper. In this sense, it is noteworthy that the proofs exposed are actually sketches of proofs.

**Theorem 1** (Task Identification). Each software requirement \( SOFi \in SOF \), represents a task \( \tau_i \) (in the sense established by real-time systems theory [17]) in the system.

*Proof:* A software requirement is the counterpart in the computational world of a physical requirement. This is, a requirement is a control law that states how controlled variables are to respond to changes in monitored ones. In this line of reasoning, each original 4-Variable slice of the 3D cube, can be seen as a particular requirement along with some (more close to) implementation details. With this and considering the architecture proposed in [18] for real-time control systems, a real-time task is a representation of a software requirement.

**Definition 23.** The set of all tasks \( \tau_i \) constitute the software implementation of a cyber-physical system \( \Psi \). Formally, \( \Psi(n) = \tau_i | i = 1 \ldots n \)

**Lemma 1** (Requirements’ Relationship). If a monitored (controlled) variable \( m_j(c_j) \) belongs to two or more different sets of monitored (controlled) variables \( MON_i \) (CON\(_i\)) and \( MON_k \) (CON\(_k\)), then requirements \( REQ_i \) and \( REQ_k \) are said to relate to each other, expressed as \( REQ_i \approx REQ_k \). Formally, if \( MON_i \cap MON_k \neq \emptyset \Rightarrow REQ_i \approx REQ_k \), analogously for \( CON_i \) and \( CON_k \).

*Proof:* Since requirements are, in fact, control laws expressing a desired relationship in terms of monitored and controlled variables, it comes out evidently that if the same variable monitored or controlled) appears in two or more different laws, then those laws are related to each other by means of the shared variable.

**Corollary 2.1** (Independent Scheduling). If \( \forall REQ_i \neq REQ_k \), \( \cap REQ_i \cap REQ_k = \emptyset \) then the set of tasks that constitute the cyber-physical system \( \Psi \) can be scheduled by a real-time scheduling policy \( P \) for independent tasks.

**Corollary 2.2** (Tasks’ Communication). If \( \exists REQ_i \approx REQ_k \), then the associated tasks \( \tau_i \) and \( \tau_k \) respectively, have some kind of communication between them.

**Corollary 2.3** (Resource Sharing Scheduling). If, in the cyber-physical system \( \Psi \), two or more tasks have a communication between them, the complete set \( \Psi \) has to be scheduled by a real-time scheduling policy able to handle shared resources.

With all, the results obtained in this section cope with the characteristics of CPS enumerated in the introduction of this paper. In particular, physical world modeling is achieved through MON and CON edges and NAT plane; control theory is included in REQ plane by means of the desired relationships between controlled and monitored variables; real-time aspects are considered in the very definition of input and output variables; communication concerns and real-time operating system support, are dealt by Theorem III.1 and Theorem III.2 with its corollaries; finally, some electronic hardware issues are addressed when defining the IN and OUT planes.

### IV. Practical Application

In this section, the application to the case study will be presented.

#### A. The Anaerobic Digestion Process

The first step in most of the fermentation process is called hydrolysis. Here, the complex bio-materials are converted into soluble compounds that are to be hydrolyzed to monomers. The hydrolysis is performed by hydrolytic bacteria [2].

To carry out the hydrolysis process is necessary to have optimum operating conditions for growth and/or survival of microorganisms. Acetogenic bacteria grow in a very close relationship with the MFB depending on one of the other. To achieve this result, we need specific ranges of pH, temperature, pressure and concentrations.

The MFB are anaerobic and are extremely sensitive to changes in alkalinity, pH and temperature. Therefore, the operating conditions within the digester should be periodically monitored and maintained within optimum ranges. Besides the controls mentioned above, other conditions should be monitored to keep them operating within the optimal range of development. These conditions are gas composition, hydraulic retention time (HRT), solids retention time (SRT), total solids (TS), volatile solids (VS) and concentration of volatile fatty acids (VFA).

#### B. Applying the 3D Approach

Based on the previous description of the anaerobic digestion process, the following physical variables were particularly identified in this work.
Once determined the physical variables, different requirements relationships are established. It is worth noting that sometimes the variables determine the relationships and some others in reverse.

\[
\text{REQ1} = \{\text{MON1} = \{\text{temp}, \text{pH}, \text{ch4}\}, \text{CON1} = \{\text{temp}\}\}
\]

\[
\text{REQ2} = \{\text{MON2} = \{\text{temp}, \text{orp}, \text{ch4}\}, \text{CON2} = \{\text{hrt}, \text{srt}\}\}
\]

\[
\text{REQ3} = \{\text{MON3} = \{\text{ts}, \text{vs}\}, \text{CON3} = \{\text{hrt}, \text{srt}\}\}
\]

\[
\text{REQ4} = \{\text{MON4} = \{\text{temp}, \text{ch4}\}, \text{CON4} = \{\text{hrt}, \text{srt}, \text{temp}\}\}
\]

\[
\text{REQ} = \{\text{REQ1}, \text{REQ2}, \text{REQ3}, \text{REQ4}\}
\]

According to the previously stated definition of the requirement relationships, each of them should be expressed as a control law. However, only their inputs and outputs are shown, since the inclusion of them would be out of the paper’s scope. Moreover, the interactions between the control laws is a topic not easily approachable and that requires the physicochemical modeling of the anaerobic digestion process [19].

In the computing system side, the identified variables are named with a prefix indicating whether they are inputs or output (i.e., in_ or out_).

\[
\text{INPUT} = \{\text{in_temp}, \text{in_ch4}, \text{in_pH}, \text{in_orp}, \text{in_ts}, \text{in_vs}\}
\]

\[
\text{OUTPUT} = \{\text{out_temp}, \text{out_hrt}, \text{out_srt}\}
\]

Following, the planes that serve as interface between the physical and computational world are summarized. Beyond the mapping, the devices used in each case are also shown.

\[
\text{IN} = \{
\begin{align*}
\text{temp} & \rightarrow \text{in_temp} \quad (\text{Temperature sensor, LM35}) \\
\text{ch4} & \rightarrow \text{in_ch4} \quad (\text{Gas sensor, TGS-813, MQ5}) \\
\text{pH} & \rightarrow \text{in_pH} \quad (\text{pH electrode}) \\
\text{orp} & \rightarrow \text{in_orp} \quad (\text{redox potential electrode}) \\
\text{ts} & \rightarrow \text{in_ts} \quad (\text{laboratory equipment}) \\
\text{vs} & \rightarrow \text{in_vs} \quad (\text{laboratory equipment})
\end{align*}
\}
\]

\[
\text{CON} = \{
\begin{align*}
\text{temp} & \rightarrow \text{out_temp} \quad (\text{Heating resistance}) \\
\text{hrt} & \rightarrow \text{out_hrt} \quad (\text{Solenoid}) \\
\text{srt} & \rightarrow \text{out_srt} \quad (\text{Solenoid})
\end{align*}
\}
\]

With all the necessary variables and relationships defined, a sketch of the biodigester to be built is shown in Fig. 4.

**V. FURTHER DISCUSSION**

In this section, three aspects common to any requirement engineering approach are discussed. The analysis is centered in topics that go beyond technical features, emphasizing operating matters.

In the first place, an aspect that was previously mentioned, but that deserves to be recalled, is the one related to the choice of the 4-Variable Model as the basis of the one proposed in this paper to deal with CPS. The kinds of systems aimed to be handled with the 3D Approach have a strong interaction with the physical environment and are thought to be functioning without any human intervention. This reinforces the characteristic stated in the introduction about a CPS, concerning that an external observer could not distinguish whether the system behaves in a certain way because of the physical laws, the computing system or both. With this in mind and analyzing the case of the biodigester that led the process, it came out naturally that, the chosen method had to be capable of expressing the particularities of both worlds and their interfaces. And an excellent alternative was the one proposed by Parnas and Madey, since that approach is able to express the characteristics of a system in terms of control laws (i.e., physical laws that show how outputs are to respond to changes in inputs).

![Figure 4. Sketch of the biodigester.](image-url)
requirement engineering approach. In this paper a rough view of how the 3D Approach should be used and documented was given. However, that presentation was very sketchy and it lacked of a well-defined structure or template that allows its application in a bigger and more complex system. Nonetheless, for small-sized systems and even mid-sized ones the scheme used here is usable. On the other side, an agile way of documenting requirements is through the usage of a graphical language. In this respect, it is interesting to consider SysML along with a formal notation (for example, [22] or [23]) to express the 3D Approach in a way that eases the translation from requirements to design to implementation.

VI. CONCLUSIONS AND FUTURE WORK

Cyber-physical systems is an emerging research discipline that settle its foundations in well-known ones such as control theory, real-time embedded systems, communication networks, software engineering and electronics. Nevertheless the key aspect of this new discipline is its very close relationship with the physical world. This imposes the necessity to deal with both worlds in parallel when developing an application.

In this paper, some outcomes from the practical experience of designing a CPS where exposed. The chosen application was a biodigester, which is a device that allows the anaerobic digestion of organic matter. The anaerobic digestion has many advantages among which stands out the generation of biogas and fertilizer. The biodigester was meant to work in an autonomous way, with almost any human intervention. Consequently, the requirements elicited were related to physicochemical characteristics of the anaerobic digestion process. In order to model and analyze those requirements, the 4-Variable Model was adopted. However, it had to be extended to cope with the particularities of CPS. This led the development of a new model, called the 3D Approach. The most relevant aspect of it is the separation of concerns based on the physical laws that governs the process. In this manner, the features of a CPS can be easily distinguished and tackled.

The approach proposed in this paper is the result of lessons learned from practical experience. In this sense, future works have to do with formalizing the method and systematically describing its usage.

REFERENCES


Leonardo Ordinez received the Computer Systems Engineering degree and Ph.D. in Engineering degree from Universidad Nacional del Sur, Argentina, in 2006 and 2010, respectively. He is Teaching Assistant at the same university and Assistant Professional at CONICET, Argentina. His main interests are in the field of embedded and cyber-physical systems, specially dealing with requirements and system design. He has authored several papers and is currently involved in different multidisciplinary projects in the areas of precision agriculture and bioenergy.

Omar Alimenti received the Electrical Engineering degree and the M.S. in Engineering degree from Universidad Nacional del Sur, Argentina, in 1979 and 1999 respectively. He is in the teaching profession since 1980 in subjects of computer architectures. He is currently
Full Professor. He has authored numerous publications in the specialty of Real Time Systems and Microprocessors and is director of numerous scientific research projects and technology transfer.
Maria Emilia Rinland obtained her Biology degree in 2009 from Universidad Nacional del Sur, Bahía Blanca, Argentina. She is currently doing her Ph.D. in Agronomy with a fellowship granted by CONICET. She is working in the microbiological aspects of the anaerobic biodegradation of onion waste for biogas production.

Marisa Anahi Gómez received a Ph.D. from Université de Bourgogne (France). Her current position is Associate Professor of Universidad Nacional del Sur. Courses: Agricultural Microbiology, Biology of Procarriotas and Microbial Ecology in Agroecosystems. She is director of research projects focused on basic understanding on biodiversity of soil microbial communities in agricultural systems and the biodegradation of agrochemicals and organic waste.

Jorge Marchetti received his Bachelor degree from Universidad Nacional del Sur in 2003. He also got from the same university his Ph.D. in Chemical Engineering (biodiesel production) in 2008 and his second Ph.D. in Physics (hydrogen storage) in 2009. He was appointed post doc in Norwegian University of Natural Science and Technology from 2008-2010. He has worked on several different projects and done cooperative work with USA, Norway, Denmark, and Sweden. He is the author of 2 books, 5 chapter contributions and more than 25 publications in international well known peer review journals.