Computers and Chemical Engineering 34 (2010) 1137–1154

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

An ontological knowledge-based system for the selection of process monitoring and analysis tools

Ravendra Singh[a], Krist V. Gernaey[b], Rafiqul Gani[a,*]

[a] CAPEC, Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[b] Center for Process Engineering and Technology (PROCESS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 8 December 2009
Received in revised form 6 April 2010
Accepted 12 April 2010
Available online 18 April 2010

Keywords:
Knowledge base
Ontology
Process monitoring
Sensor
PAT
Inference system

ABSTRACT

Efficient process monitoring and analysis tools provide the means for automated supervision and control of manufacturing plants and therefore play an important role in plant safety, process control and assurance of end product quality. The availability of a large number of different process monitoring and analysis tools for a wide range of operations has made their selection a difficult, time consuming and challenging task. Therefore, an efficient and systematic knowledge base coupled with an inference system is necessary to support the optimal selection of process monitoring and analysis tools, satisfying the process and user constraints. A knowledge base consisting of the process knowledge as well as knowledge on measurement methods and tools has been developed. An ontology has been designed for knowledge representation and management. The developed knowledge base has a dual feature. On the one hand, it facilitates the selection of proper monitoring and analysis tools for a given application or process. On the other hand, it permits the identification of potential applications for a given monitoring technique or tool. An efficient inference system based on forward as well as reverse search procedures has been developed to retrieve the data/information stored in the knowledge base.

1. Introduction

For the manufacturing industries, opportunities exist to improve the production process with respect to processing time, plant safety, product quality and production cost by implementation of suitable process monitoring and analysis tools, nowadays commonly known as implementation of a PAT (Process Analytical Technology) system. The design of PAT systems needs to satisfy process and user-specified constraints precisely, and therefore selection of suitable process monitoring and analysis tools for any application is an important task. Several factors need to be taken into account for the selection of process monitoring and analysis tools such as accuracy, precision, operating range, response time, resolution, sensitivity, drift and cost. Availability of an increasingly large number of monitoring tools for a wide range of operations makes the selection procedure a difficult task, and also indicates the need for a proper knowledge management system based on a systematic and efficient representation of the domain knowledge.

A substantial number of papers related to knowledge-based systems (KBSs) have been published since the early 1990s (Choi, 2009; Gebus & Leiviska, 2009; Kandil, Farghal, & Abdel-Aziz, 1992; Lai, 2007; Paton, Goble, & Bechhofer, 2000; Pirró, Mastroianni, & Talia, 2009; Qian, Liang, & Dang, 2009; Wen, Chen, & Chen, 2008). In these papers, emphasis has been given mainly to the collection, management and application of the knowledge using information technology. KBSs have been used extensively to solve problems related to, for example, enterprise performance measurement (Wen et al., 2008), airborne decision support (Howells, Davies, Macauley, & Zancanato, 1999), environmental decision support (Avouris, 1995), HAZOP analysis (Leone, 1996; Rahman, Khan, Veitch, & Amyotte, 2009), safety control (Lee, 2006), and statistical data base management support (Gabor & Barna, 1988) to name a few. Therefore, it can be concluded that there is a wide range of problems where KBSs have proved to be valuable tools (Hendriks, 1999). Much less attention, however, has been paid to the design and development of a knowledge-based system that can be applied systematically for the selection of appropriate process monitoring and analysis techniques/tools.

In general, KBS should include a set of properties, such as, ability to perform inferences, the ability to update its own knowledge base and the ability to control the search space (Hendriks, 1999). As defined by Hendriks (1999), the KBSs are computer systems using a formal mechanism to represent or simulate specific aspects of human knowledge, and to apply these representations in actual problem situations. The solutions of the problems also depend on the type of reasoning systems applied. Rule-based reasoning

* Corresponding author at: Søltofts Plads, Building 229, DK-2800 Kgs. Lyngby, Denmark. Tel.: +45 45252882; fax: +45 45932906.
E-mail address: rag@kt.dtu.dk (R. Gani).

0098-1354/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2010.04.011

(RBR) (Dupuit, Pouet, Thomas, & Bourgois, 2007) and case-based reasoning (CBR) (Burke, MacCarthy, Petrovic, & Qu, 2000; Finnie & Sun, 2003; Patterson, Rooney, Galushka, Dobrynin, & Smirnova, 2008) have emerged as two important and complementary reasoning methodologies in artificial intelligence – AI (Dutta, 1993). The knowledge base needs to be appropriately structured and organized, so that it can be accessed in a straightforward manner through a computer and a robust and efficient inference system can be designed. It should be noted that different types of knowledge bases require different levels of knowledge structuring.

Ontologies have been used extensively by information technologists to systematically represent the knowledge within a domain. The literature contains many, sometimes contradictory definitions of ontology. The definition that is widely accepted is given by Gruber (1995): "*an ontology is an explicit specification of a conceptualization*". Gruber has also proposed the design criteria for ontologies. Ontology should have clarity in defined terms and relations. It should be coherent: that is, it should sanction inferences that are not consistent with the definitions. The ontology should be extendable to increase the range of applications. The encoding bias and the ontological commitment (Rayo, 2007) should be minimal to make it more generic. Basically, according to Gruber (1995), each ontology defines a set of classes, relations, functions, and objects for some domain of discourse, and includes an axiomatization to constrain the interpretation. A knowledge base may use an ontology to specify its structure (entity types and relationships) and classification scheme. An ontology, together with a set of instances of its classes, constitutes a knowledge base. Guarina (1997) also discussed the importance and limitations of some of the definitions of ontology.

In the last decade several ontologies have been developed in different domains for providing solutions to a range of problems. For example, OntoCAPE (Morbach, Yang, & Marquardt, 2007) is developed in the domain of chemical process engineering, KnowBaSICS-M (Bratsas et al., 2007) is developed to manage a class of medical science problems, yOWL (Villanueva-Rosales & Dumontier, 2008) is developed for yeast biologists and CelOWS (Matos, Campos, Braga, & Palazzi, 2010) is developed to manage a class of biological models. However, very few attempts have been made to develop an ontology in the process monitoring and analysis domain.

Eid, Liscano, and Saddik (2006) have proposed an ontology for sensor network data. Ong, Masud, and Eyada (1992) proposed a knowledge-based system for sensor selection based on a group technology coding scheme. The proposed KBS can provide a set of alternatives for sensors that satisfy the operational requirements. The selection of the final sensor is, however, done manually. Shieh, Huber, Fleck, and Ashby (2001) proposed a graphical approach for the selection of sensors. They attempted to locate the sensors in the graph on the basis of their specifications. A pair of sensor specifications – for example, accuracy and operating range – was considered and plotted in a two-dimensional mapping.

The interferences due to the presence of other components in the medium also affect the measurements related to the component of interest. The interferences (cross-sensitivities) affect the selectivity and the sensitivity of the sensors. Brudzewski and Dolecka (1995) took into account the effect of these interferences for the selection of sensors for gas measurements. Barua, Ray, and Sinha (1993) proposed a knowledge-based expert system for transducer selection related to measurements of temperatures, flows and pressures. KBS development attempts for sensor selection thus have a clear focus on the selection of sensors based on sensor specifications. However, process knowledge that needs to be integrated in the selection of sensors for the design of the process monitoring and analysis systems, is typically not taken into consideration (Singh, Gernaey, & Gani, 2010).

In this manuscript the ontology for a knowledge base that can be used by methods and tools for the design of PAT systems is presented. This developed ontological knowledge base has two main purposes: (1) to support the selection of suitable process monitoring and analysis tools for a specific application or process; (2) to support the search for the range of potential applications of a specific monitoring tool. It consists of two main integrated sections. The first section of the knowledge base represents the process knowledge (type of processes, corresponding process points, process variables and actuator candidates) while the second section represents the knowledge on measurement methods and tools (type of variables, corresponding monitoring techniques, monitoring tools, specifications, and suppliers). An inference system based on forward as well as reverse search procedures has been developed to retrieve and manage the data/information stored in the knowledge base.

2. Knowledge base acquisition and reuse

The useful knowledge/data needed for the design of process monitoring and analysis systems is stored in the knowledge base. The procedures for knowledge acquisition and knowledge management are illustrated in Fig. 1. A literature survey, an industrial survey and process experience are the primary sources of knowledge, in this case. The knowledge obtained from these sources has been structured so that it can be used by an inference system and can contribute to the solution of a wide range of problems.

3. Systematic development of the PAT knowledge base

In this paper, ontology has been used to systematically represent the knowledge on process monitoring and analysis systems. Classes, slots, domain of the slots, range of the slots, facets, relations and functions, instances and axioms are the core elements of any ontology. The definition of these terms are available in the literature (Knowledge Systems Laboratory, 1997; Patel-Schneider, Hayes, & Horrocks, 2004; Siricharoen, 2009).

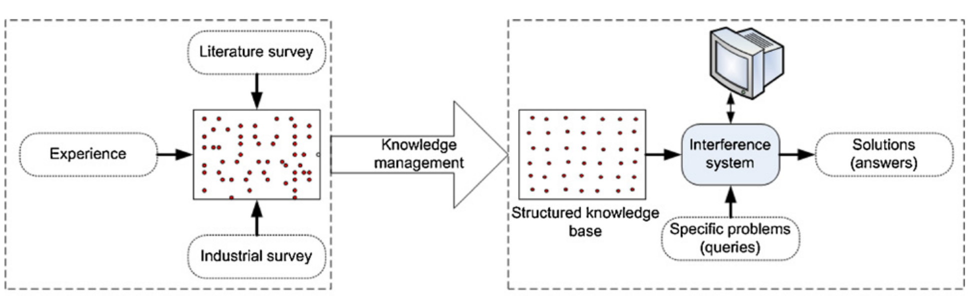

Fig. 1. Knowledge acquisition and reuse.

3.1. Design of ontology

Ontology building is an evolutionary design process consisting of proposing, implementing and refining classes and properties that comprise an ontology (Natalya & McGuinness, 2007). Since 1990, a handful of methodologies have been proposed for the development of the ontologies (Fernandez-Lopez, Gomez-Perez, & Juristo, 1997; Fernandez-Lopez, Gomez-Perez, Pazos-Sierra, & Pazos-Sierra, 1999; Grüninger & Fox, 1995; Kim, Choi, Shin, & Kim, 2007; Lenat & Guha, 1990; Natalya & McGuinness, 2007; Staab, Schnurr, Studer, & Sure, 2001; Swartout, Ramesh, Knight, & Russ, 1997; Uschold, 1996; Uschold & King, 1995). Some of the methodologies available for the design of ontologies have been compared and evaluated by different researchers to identify their advantages and the major limitations (Corcho, Fernandez-Lopez, & Gomez-Perez, 2003; Fernandez-Lopez, 1999). It is observed that several steps are common in most of the methodologies proposed for the design of ontologies, obviously influenced by the lifecycle concept from software engineering (Morbach et al., 2007). It is worth to mention here that the ontology development is an iterative and corrective procedure and there is no straightforward way to develop it (Venkatasubramanian et al., 2006). The existing methodologies for design of an ontology have been reviewed and a systematic procedure for the design of an ontology has evolved as shown in Fig. 2. Most of the steps of the evolved methodology are adapted from Natalya and McGuinness (2007). As shown in Fig. 2, the ontology design procedure involves nine hierarchical steps. The design of ontology is an iterative and corrective procedure (see steps 1–8 in Fig. 2); therefore, the design procedure needs to be repeated until the evolved ontology satisfies the desired requirements. The nature of the methodology is generic and therefore can be employed for the design of any ontology.

Following the methodology shown in Fig. 2, the ontology for process monitoring and analysis systems has been developed (a detailed description is given in Appendix A). The reader may find this detailed description useful for implementation into commercially available ontology editors (e.g. Protégé, http://protege.stanford.edu). The final ontology defines the structure of the knowledge base as shown in Figs. 3 and 4 and described in Section 3.2.

3.2. PAT knowledge base structure

According to the final evolved ontology, the knowledge base is organized into two sections. The first section, the process knowledge base, consists of information related to the process while the second section, the monitoring knowledge base, consists of information related to monitoring methods and tools. The structure of the process knowledge base is shown in Fig. 3, while the structure of the monitoring knowledge base is shown in Fig. 4.

3.2.1. Process knowledge section

The type of processes, the process points, the process variables and the actuator candidates are the main process knowledge categories. They are also called the classes of Section 1 of the knowledge base, as shown in Fig. 3. The first process knowledge class consists of a list of processes. The set of independent processes considered, include, cheese manufacturing process, fermentation process and tablet manufacturing process, which are the instances of this class. The second class of the process knowledge section consists of the list of process points. The instances of this class (process points) are classified in terms of groups and each group is linked to the corresponding process from the first class. For example, a group of process points $\overline{PP_1}$ is linked with corresponding process P_1 (see Fig. 3). The process points, corresponding to the process equipments of the existing processes, are identified directly from the process flowsheet. The third class of the process knowledge section consists of the process variables. The instances of this class (process variables) are classified in terms of groups, and each group is linked to the corresponding process point from the second class. The variables (instances of the third class) are also linked to the instances of the variable type class, which is also called a connecting object, as shown in Fig. 3. The instances of the variable type class (connecting object) are on the one hand linked to the process variables (see Fig. 3) and on the other hand linked to the monitoring techniques (see Fig. 4). These connecting objects therefore play a crucial role in the knowledge base, since they provide the link between the process knowledge section and the section containing the monitoring knowledge. Ensuring a consistent predefined product quality in the process also depends on the successful implementation of the control system. The availability of on-line monitoring tools and the identification of the proper actuators are two important prerequisites for successful implementation of the control system. A class named actuator candidates is therefore also created in the process knowledge section. The instances of this class are the potential actuator candidates (manipulated variables) of the considered process variables.

3.2.2. Monitoring knowledge section

The variable types, monitoring techniques, monitoring tools, specifications and suppliers (of monitoring tools) are the main classes of the monitoring knowledge section (see Fig. 4). The first class, variable types, contains the independent set of variables that need different monitoring techniques. The next class, monitoring techniques, contains the list of monitoring techniques available for each variable type. These monitoring techniques have been classified into different groups on the basis of their potential applications and each group has been linked with the corresponding measurement (monitoring variable). One of the main difficulties in product quality control is to identify the appropriate monitoring tools for measurement of the process variables so that they can satisfy the user and process constraints (e.g. economical and controllability constraints). Therefore a class named monitoring tools has been created in this section of the knowledge base. For a specific measured variable that is measured by a specific technique, many monitoring tools may be appropriate. Therefore all the monitoring tools available for measurement of a specific variable are placed in different groups based on the specific techniques used by the monitoring tools, and each group is linked to the corresponding monitoring technique on the level above. Each monitoring tool is also linked to the corresponding supplier. Since the selection of the monitoring tools for a specific application is one of the important applications of the knowledge base, a set of selection criteria are needed. Therefore, a set of specifications for the monitoring tools are also included in the monitoring knowledge section. Examples of these specifications are the accuracy, precision, operating range, response time, sensitivity, resolution, drift, operating temperature range, and cost (see Appendix A for the definition of these specifications). These specifications have a numerical value and can be considered as the leaves (end-points) of the knowledge tree.

The representation of the specifications of the monitoring tools in the knowledge base is shown in Fig. 5, where, the individual specifications (e.g. accuracy, precision, operating range, response time, etc.) are the sub-classes of the class named 'specifications'. These are actually the forms for each end-point of the knowledge tree. Values of the specifications, units and references form the individuals/instances of these sub-classes. Each specification has a numerical value (type: numeric), corresponding unit (type: character string) and information about the source of data/reference (type: mixed). All the references are placed in a

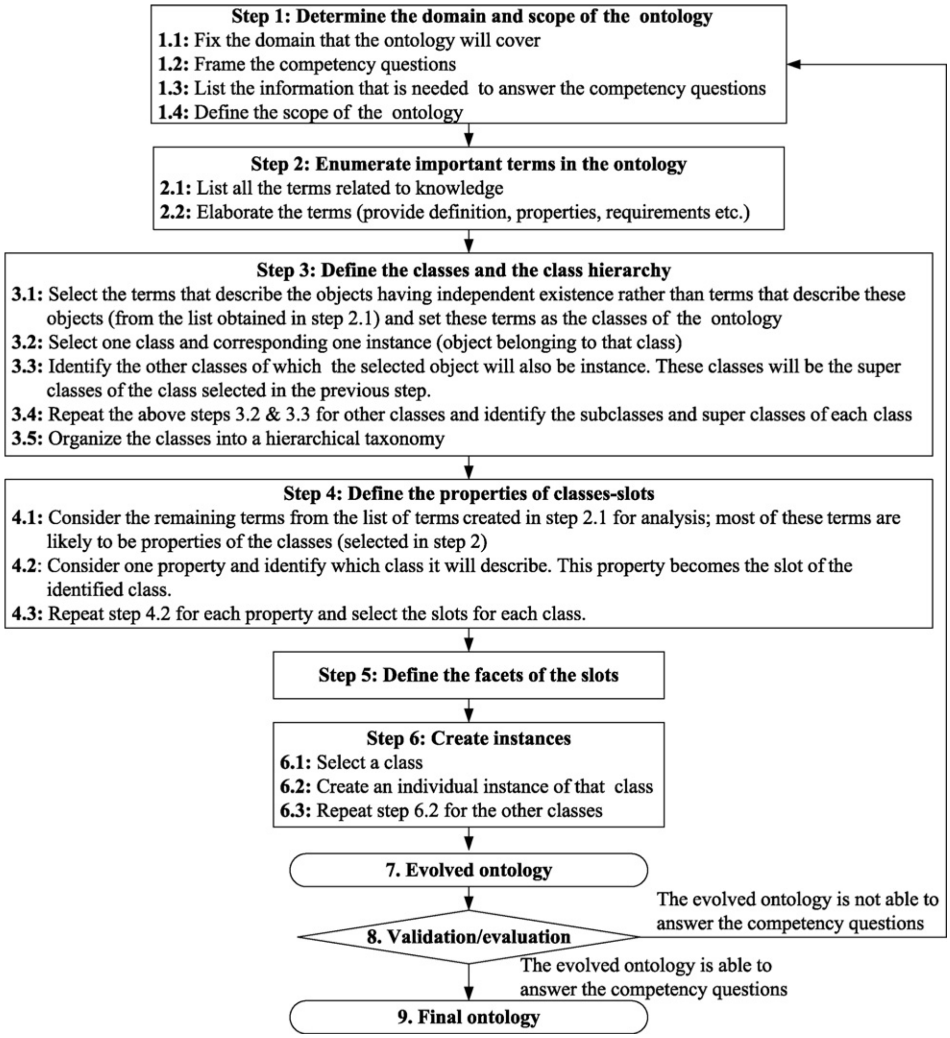

Fig. 2. Methodology for design and development of ontology.

separate sub-section with a unique reference number. The monitoring knowledge section contains these reference numbers and is linked to the original references through the references sub-section. As shown in Fig. 5, two of the specifications, operating range (minimum and maximum limiting value that a monitoring tool can measure) and the operating temperature range (minimum and maximum limiting operating temperature within which a monitoring tool can be used), have two parts such that their numerical data can be considered during the selection process.

3.3. Main contents of the PAT knowledge base

The main contents of the knowledge base are summarized in Table 1. The type of knowledge (process or monitoring) that the knowledge base contains is listed in the first column of the table. The second column of the table relates these classes with the corresponding instances that these classes contain. For instance, the main class ‘processes’ contains a list of process attributes (P_1 to P_m) where P_i is the ith process present in the knowledge base. A list of attributes (instances) for different classes is also given in the last column of Table 1. Table 1 also shows how the instances of one class are linked to the instances of other class. For instance $\overline{PP_1}$ is the ith attribute of the vector of process points $\overline{PP}$, corresponding to the process points of process P_i. Similarly, $\bar{V}_{j,i}$ is the jth attribute of the vector of variables V_i corresponding to the ith process.

4. Extension of the PAT knowledge base

The structure of the knowledge base is generic and can be easily extended to increase the range of applications and/or to accommodate newly developed process monitoring and analysis tools. The range of applications of the knowledge base can be increased through vertical extension, and it can be made more predictable (i.e. provide more information) through horizontal extension. Horizontal extension means adding more columns of data and is achieved by simply adding more classes in the knowledge base. Vertical extension means adding more rows of data and is achieved by simply adding more instances to the existing classes. Horizontal extension increases the diversification of the instances of the knowledge base. Addition of new specifications for a monitoring tool is an example of horizontal extension while the addition of new processes in the knowledge base is an example of the vertical extension of the knowledge base. As the vertical extension needs more work, a systematic procedure has been developed (see Appendix B).

5. Current status of the PAT knowledge base

Often statistics of the knowledge base are needed to determine the range of applications and the capability of the knowledge base. The current PAT knowledge base statistics are shown in Figs. 6 and 7 (see the numbers between brackets). The process knowledge sec-

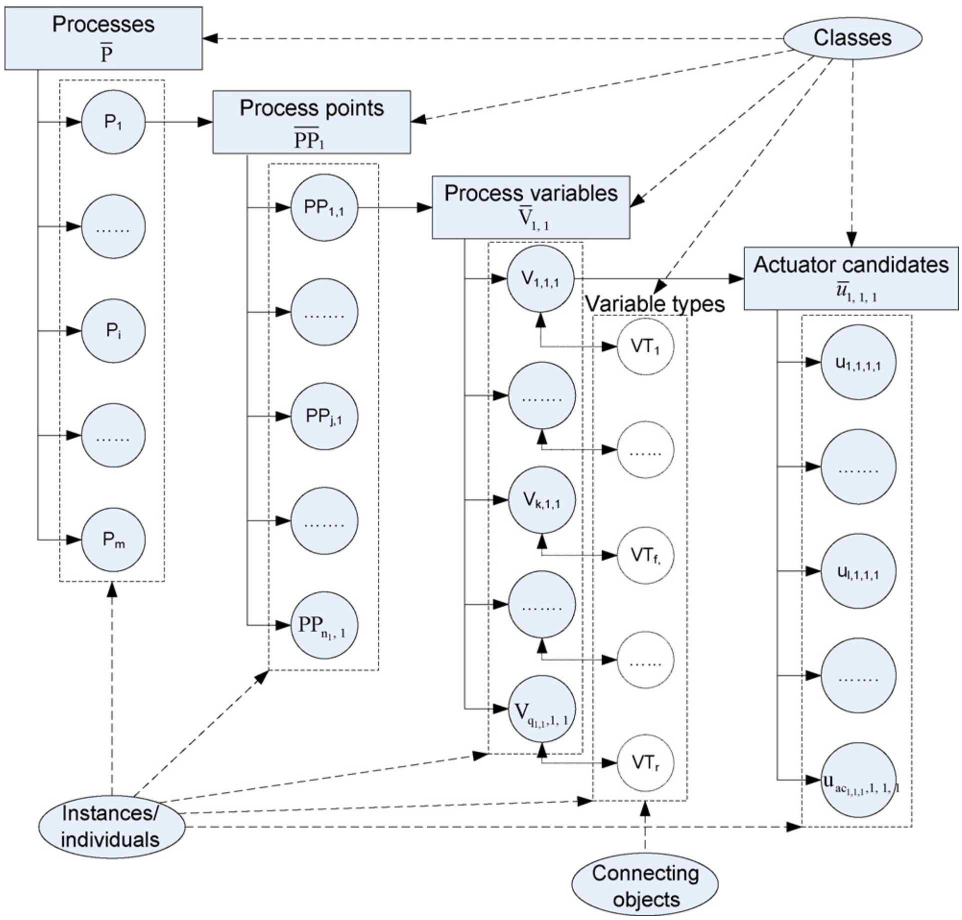

Fig. 3. Structure of the process knowledge section.

Table 1
Main classes of the PAT knowledge base and corresponding instances.

Main classes	Relation with individuals/instances	Object descriptions
Processes ($\overline{P}$) Process points ($\overline{PP}$)	$\overline{P}=[P_1 \dots P_i \dots P_m]$ $\overline{PP}=[\overline{PP}_1 \dots \overline{PP}_i \dots \overline{PP}_m]$ $\overline{PP}_i=[PP_{1,i} \dots PP_{j,i} \dots PP_{n_i,i}]$	P_i: ith process in the knowledge base $\overline{PP}_i$: process points of ith process $PP_{j,i}$: jth process point of ith process
Process variables ($\overline{V}$)	$\overline{V}=[\overline{V}_1 \dots \overline{V}_i \dots \overline{V}_m]$ $\overline{V}_i=[\overline{V}_{1,i} \dots \overline{V}_{j,i} \dots \overline{V}_{n_i,i}]$ $\overline{V}_{j,i}=[V_{1,j,i} \dots V_{k,j,i} \dots V_{q_{j,i},j,i}]$	$\overline{V}_i$: process variables of ith process $\overline{V}_{j,i}$: variables related to jth process point of ith process $V_{k,j,i}$: kth variable related to jth process point of ith process
Actuator candidates ($\bar{u}$)	$\bar{u}=[\bar{u}_1 \dots \bar{u}_i \dots \bar{u}_m]$ $\bar{u}_i=[\bar{u}_{1,i} \dots \bar{u}_{j,i} \dots \bar{u}_{n_i,i}]$ $\bar{u}_{j,i}=[\bar{u}_{1,j,i} \dots \bar{u}_{k,j,i} \dots \bar{u}_{q_{j,i},j,i}]$ $\bar{u}_{k,j,i}=[u_{1,k,j,i} \dots u_{l,k,j,i} \dots u_{ac_{k,j,i},k,j,i}]$	$\bar{u}_i$: actuator candidates used in ith process $\bar{u}_{j,i}$: actuator candidates related to jth process point of ith process $\bar{u}_{k,j,i}$: actuator candidates of kth variable related to jth process point of ith process
Variable types ($\overline{VT}$)	$\overline{VT}=[VT_1 \dots VT_f \dots VT_r]$	VT_f: fth variable type in the knowledge base (Section 2)
Monitoring techniques ($\overline{MTE}$)	$\overline{MTE}=[\overline{MTE}_1 \dots \overline{MTE}_f \dots \overline{MTE}_r]$ $\overline{MTE}_f=[MTE_{1,f} \dots MTE_{g,f} \dots MTE_{t_f,f}]$	$\overline{MTE}_f$: monitoring techniques for fth variable $MTE_{g,f}$: gth monitoring technique of fth variable
Monitoring tools ($\overline{MT}$)	$\overline{MT}=[\overline{MT}_1 \dots \overline{MT}_f \dots \overline{MT}_r]$ $\overline{MT}_f=[\overline{MT}_{1,f} \dots \overline{MT}_{g,f} \dots \overline{MT}_{t_f,f}]$ $\overline{MT}_{g,f}=[MT_{1,g,f} \dots MT_{h,g,f} \dots MT_{T_{g,f},g,f}]$	$\overline{MT}_f$: monitoring tools for fth variable $\overline{MT}_{g,f}$: monitoring tools based on gth technique for fth variable $MT_{h,g,f}$: hth monitoring tool based on gth technique for fth variable
Suppliers ($\bar{S}$)	$\bar{S}=[\bar{S}_1 \dots \bar{S}_f \dots \bar{S}_r]$ $\bar{S}_f=[\bar{S}_{1,f} \dots \bar{S}_{g,f} \dots \bar{S}_{t_f,f}]$ $\bar{S}_{g,f}=[S_{1,g,f} \dots S_{h,g,f} \dots S_{c_{g,f},g,f}]$	$\bar{S}_f$: suppliers of monitoring tools for fth variable $\bar{S}_{g,f}$: suppliers of monitoring tools based on gth technique for fth variable $S_{h,g,f}$: supplier of hth monitoring tool based on gth technique for fth variable
Specifications ($\overline{SP}$)	$\overline{SP}=[\overline{SP}_1 \dots \overline{SP}_f \dots \overline{SP}_r]$ $\overline{SP}_f=[\overline{SP}_{1,f} \dots \overline{SP}_{g,f} \dots \overline{SP}_{t_f,f}]$ $\overline{SP}_{g,k}=[\overline{SP}_{1,g,f} \dots \overline{SP}_{h,g,f} \dots \overline{SP}_{T_{g,f},g,f}]$ $\overline{SP}_{h,g,f}=$ [accuracy, precision, operating range, response time, sensitivity, drift, resolution, otr, cost, ...]	$\overline{SP}_f$: specifications of monitoring tools for fth variable $\overline{SP}_{g,f}$: specifications of monitoring tools for fth variable based on gth technique $\overline{SP}_{h,g,f}$: specifications of hth monitoring tool based on gth technique for fth variable

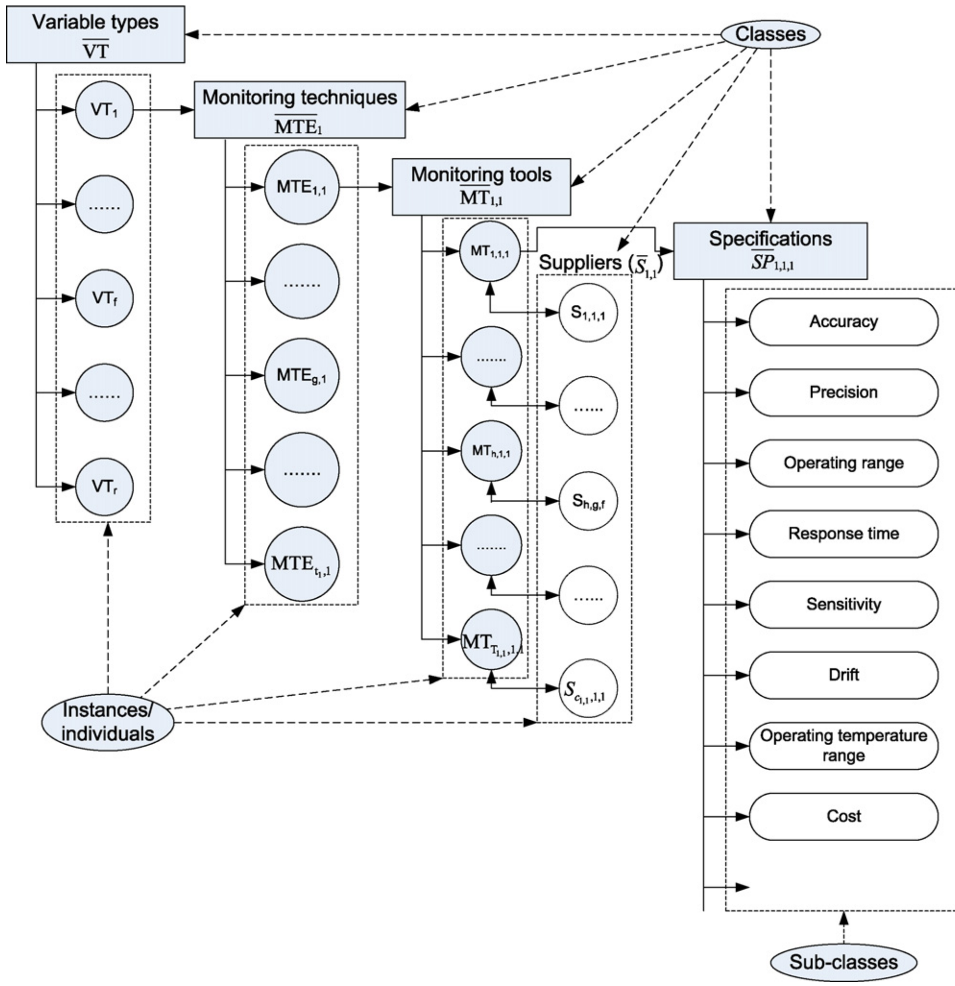

Fig. 4. Structure of the monitoring knowledge section.

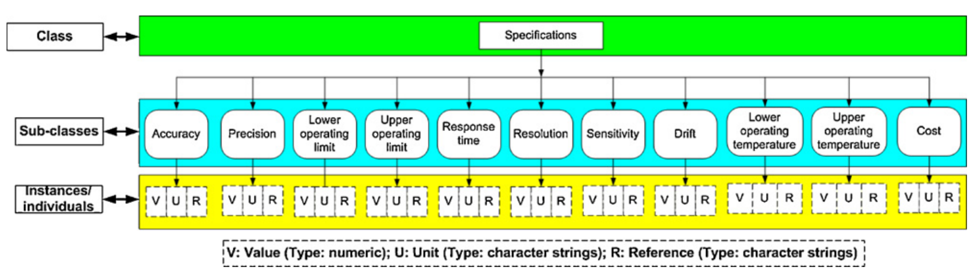

Fig. 5. Systematic representation of the specifications of the monitoring tools.

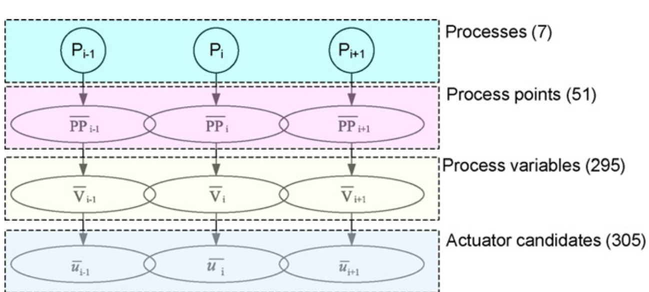

Fig. 6. Process knowledge section statistics (shows also the repetition of the objects).

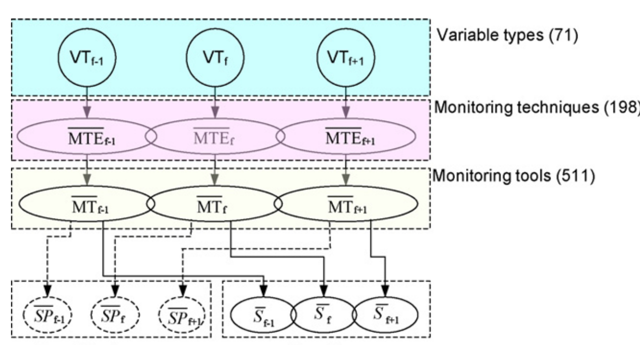

Fig. 7. Monitoring knowledge section statistics (shows also the repetition of the objects).

tion currently includes seven processes (a fermentation process, a pharmaceutical tablet manufacturing process, a crystallization process, an insulin production process, chemical reactions, a cheese manufacturing process and a butter manufacturing process). These processes consist of 51 process points, corresponding to 295 process variables and 305 actuator candidates. The monitoring knowledge section includes 71 independent variable types. For the monitoring of these variables, 198 monitoring techniques and 511 monitoring tools have thus far been added to the knowledge base. Each monitoring tool has 11 specifications. In total the knowledge base contains around 6000 data points. This information has been collected through an intensive literature search combined with an industrial survey covering more than 240 references in total.

The methods for calculating the above statistics are described briefly in Appendix C.

6. Applications of the PAT knowledge base

The developed PAT knowledge base has dual features that provide the selection of suitable monitoring and analysis tools for a given application or process as well as identify the potential applications of a given monitoring technique or tool. An efficient inference system based on forward as well as reverse search procedures has been developed to retrieve the data/information stored in the knowledge base. The forward search procedure (searches from left to right in the knowledge base) is employed to find the appropriate process monitoring and analysis tool for a particular application while the reverse search procedure (searches from right to left in the knowledge base) is employed to identify the potential applications of the monitoring techniques and tools. The forward search engine is particularly useful for the chemical, biochemical, pharmaceutical and food industries where appropriate process monitoring and analysis tools are needed to monitor and control the manufacturing process while the reverse search engine is particularly useful for manufacturers and suppliers of these monitoring tools to identify potential applications and thus potential customers.

This PAT knowledge base may also be part of other software, such as the tool for design and analysis of PAT systems. In this case, the knowledge based together with its search procedures need to be used within the work-flow of the host software.

6.1. Forward search engine

The forward search procedure provides direct access to the knowledge/data stored in the knowledge base, and it is especially designed for the cases where the user would like to explore the available alternatives of process monitoring and analysis systems (e.g. to find out the available monitoring tools and their specifications for a particular process variable). Like the knowledge base, the search engine also consists of two sections that are integrated

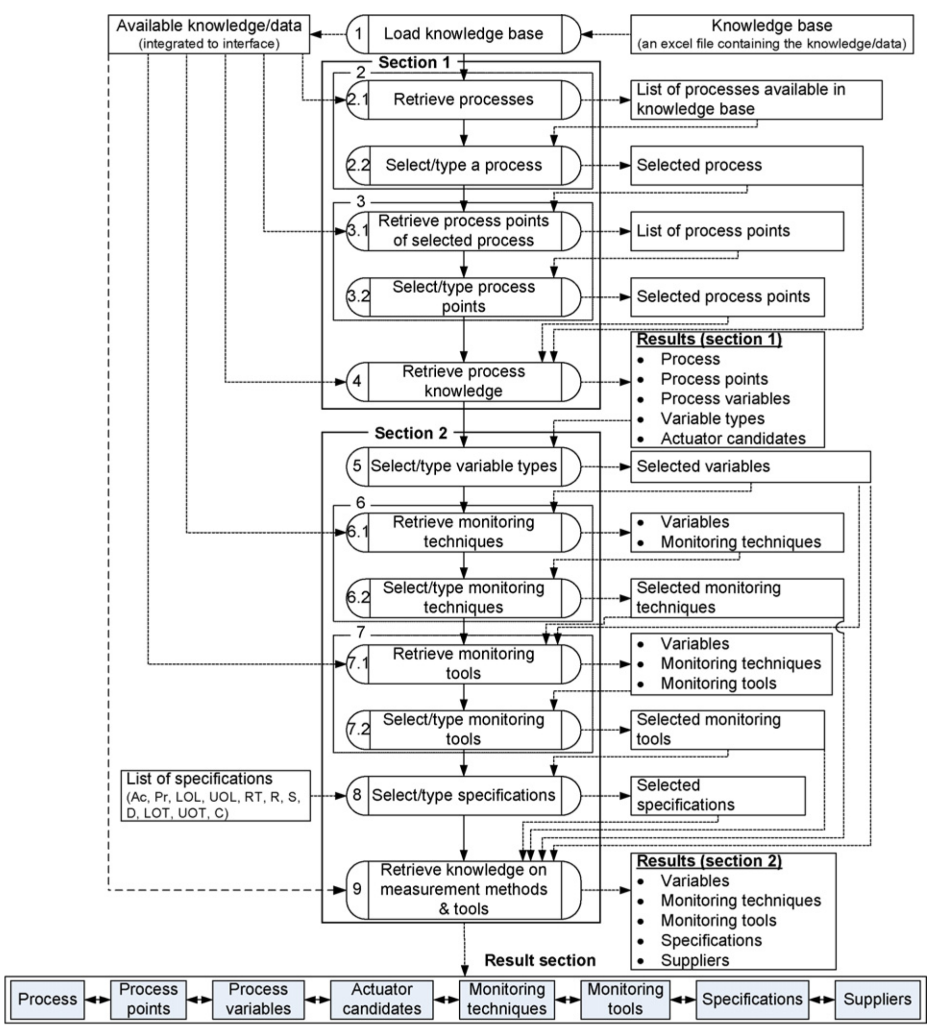

Fig. 8. Activity diagram for the forward search engine.

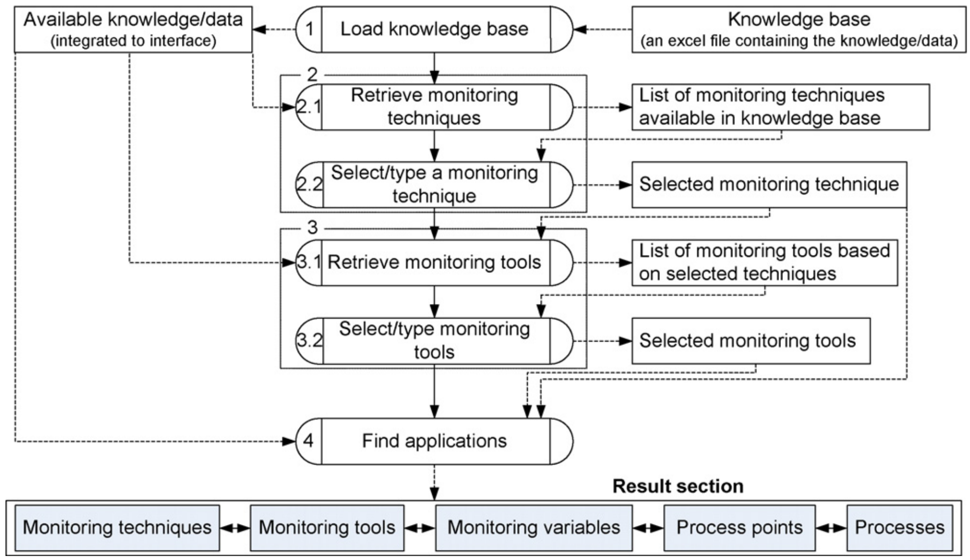

Fig. 9. Activity diagram for the reverse search engine.

with each other. Each section of the search engine may be used as a stand-alone search engine as well as for data retrieval operations for a host software.

The work-flow (solid lines) and data-flow (dashed lines) for the forward search procedure is shown through an activity diagram in Fig. 8. The knowledge base first needs to be integrated with the search system interface or the host software interface before application.

Section 1: The starting point for this section is to retrieve all the processes available in the knowledge base and to create a list of processes. In the next step a process is selected from the created list and the corresponding process points of the selected process are retrieved from the knowledge base. In the next step, the process points are selected and the required information (data) is generated. Note that multiple selections are possible in this step. The generated process knowledge contains a list consisting of the selected process, the corresponding process points, the process variables, the variable types and the actuator candidates. This list is the result of Section 1 of the search engine.

Section 2: The result of Section 1 acts as the input to Section 2 if the user is interested to combine process knowledge with monitoring knowledge. The starting point for this section is to select the variables for which monitoring techniques are needed followed by retrieval of the available monitoring techniques from the knowledge base. Then the monitoring techniques are selected and the corresponding monitoring tools are retrieved followed by the specifications of the selected monitoring tools. The results of this search (selected variables, monitoring techniques, monitoring tools, specifications) act as the input for the final step where the stored numerical values for each tool selection criteria are retrieved. The generated results of Section 2 contain a list of monitoring variables, corresponding monitoring techniques, monitoring tools, specifications and suppliers. The results of Sections 1 and 2 together comprise the final results of the forward search procedure. A user-friendly interface has been developed to retrieve the knowledge/data using the forward search procedure. This user interface is described in detail in Singh et al. (2010).

6.2. Reverse search engine

This option is created to allow data retrieval from the monitoring knowledge section using a reverse search procedure. Through this option the potential application range of a specific monitoring technique/tool can be identified. For example, a manufacturer of near-infrared spectroscopy based monitoring tools could use this feature of the knowledge base to find the potential applications of their product. More specifically, this option of the knowledge base allows to identify for which processes and within the processes at which process point and within the process point for which process variable, a specific monitoring tool can be applied.

The work-flow and data-flow for the reverse search procedure are shown through an activity diagram in Fig. 9. The first step is to integrate the knowledge base with the user interface. The next step is to create a list of monitoring techniques available in the knowledge base and then to select a monitoring technique from the generated list. Then all the monitoring tools related to the selected technique are retrieved from the knowledge base and the monitoring tools for which the applications need to be identified are selected from the generated list of monitoring tools. Based on the above selections, the potential applications of the selected monitoring tools are identified. The final search results contain a list of monitoring techniques, corresponding monitoring tools, monitoring variables, process points and processes. A user-friendly interface has been developed to retrieve the knowledge/data using the reverse search procedure (see example 7.4 and Fig. 12).

As long as both sections of the knowledge base are linked (see Section 3.2) through the elements of the connecting object, there will always be a match between these sections. However, it is also possible for a specific process variable that no monitoring tools are available in the knowledge base. Similarly, it is also possible that a specific monitoring tool cannot be used in any of the processes included in Section 1 of the knowledge base. In such cases the search engines (forward as well as reverse) provide the option to either include the missing information/data manually into the result section of the search engines, and then to add this information/data to the knowledge base, or alternatively to ignore the elements for which the information is not available.

6.3. Automatic selection of suitable process monitoring and analysis tools

The PAT knowledge base can be employed to identify suitable monitoring tools for any process. Typically, the process points where monitoring and analysis equipments need to be placed together with the corresponding monitoring variables are already

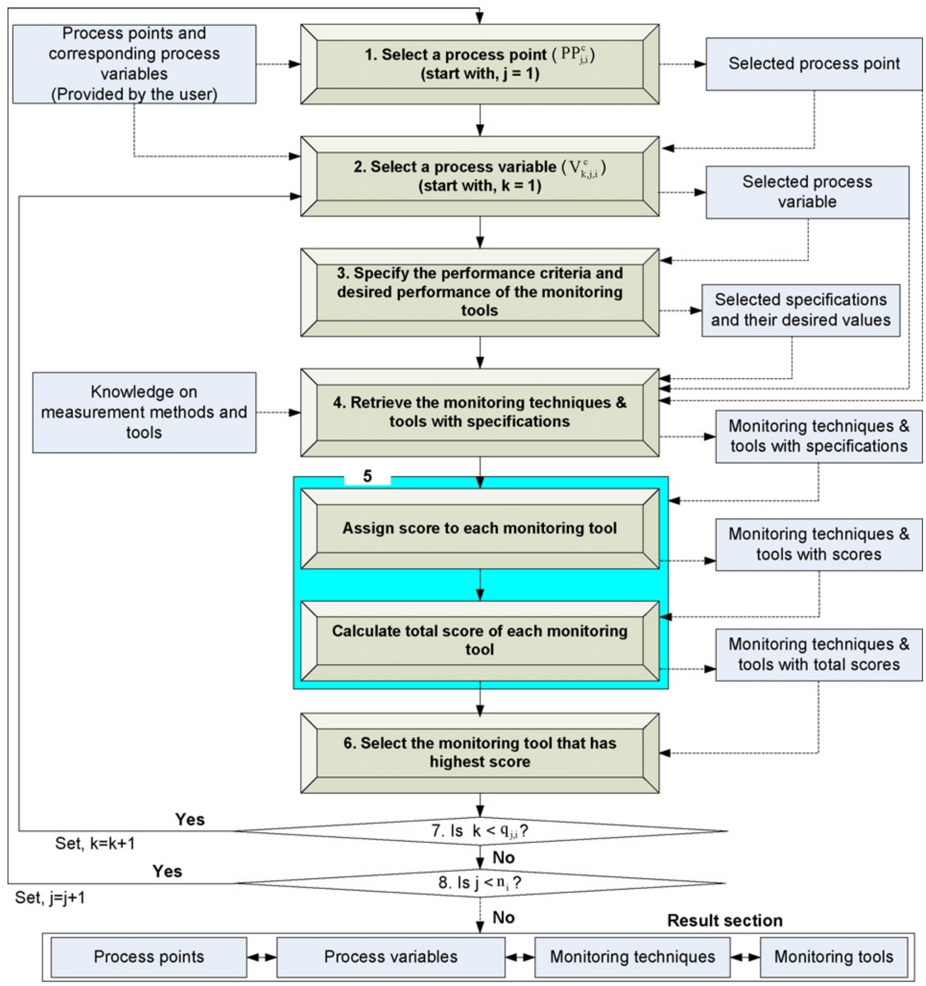

Fig. 10. Activity diagram for the selection of monitoring tools ($q_{j,i}$: number of variables involved with the *j*th process point of the *i*th process; n_i: number of process points in the *i*th process).

known. However, if the process points and the corresponding monitoring variables are not known then they can first be identified through the ICAS-PAT software (Singh et al., 2010) and then from the generated list of process points and the corresponding process variables information on the monitoring tools can be generated from the PAT knowledge base by using the forward search procedure.

The work-flow and data-flow for use of the knowledge base from the ICAS-PAT software for the selection of monitoring tools is shown in Fig. 10. As shown in the figure, a process point is first selected through step 1 (a list of process points and corresponding process variables is provided by the user or generated through ICAS-PAT). A corresponding process variable is selected through step 2, using the selected process point as the input for this search. The selected process variable is the input to step 3 through which the required specifications of the monitoring tools are selected. Moreover, the desired range is provided for the process variable, and will be considered as a criterion to be used in the selection of suitable monitoring tools. The selected specifications, their desired values together with the selected critical process point and critical process variable are the input to step 4 through which the monitoring techniques and tools (with specifications) that satisfied the specified criteria are retrieved from the PAT knowledge base. Note that there is an option to retrieve the monitoring techniques and tools for all the critical process variables through a single run to save the data retrieval time. The results of step 4 form the input for step 5 through which the performance of the monitoring tools is compared. A scoring system has been developed to compare the performance of the monitoring tools. First the monitoring tools are assigned scores according to their specifications – the monitoring tool with the best performance for a specification, e.g. drift, receives the highest score; the monitoring tool with the second best performance receives the second best score, and so on – and then the total score obtained by each monitoring tool is obtained. The algorithm for scoring of monitoring tools has been presented in detail in an earlier paper (Singh et al., 2010). Finally, in step 6, the monitoring tools with the highest score are selected.

7. Case studies highlighting the use of the PAT knowledge based

The application range of the PAT knowledge base is highlighted in detail for a cheese manufacturing process and briefly, for a tablet manufacturing process and a fermentation process. In addition, the use of the reverse search is highlighted for a specified monitoring technique.

7.1. Cheese manufacturing process: food process case study

The application of the knowledge base is highlighted below for a cheese manufacturing process.

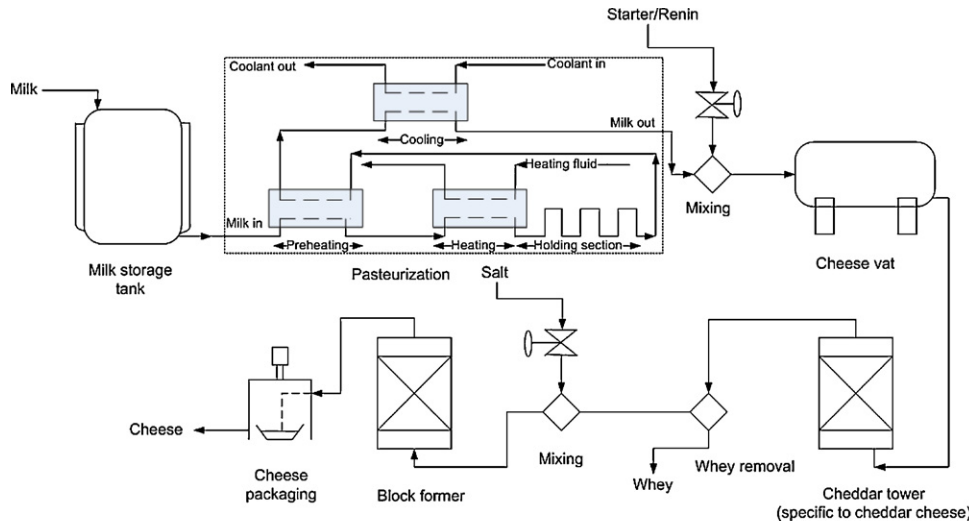

Fig. 11. Process flowsheet for the cheese manufacturing process (adapted from Zaror & Pyle, 1997).

7.1.1. Process description

The cheese manufacturing process flowsheet is shown in Fig. 11, which has been adapted from Zaror and Pyle (1997). Milk is pumped from tankers into a storage tank. The temperature in the storage tank is maintained at around 4 °C. From the storage tank milk flows through a pasteurizer, which is a heat exchanger operating at around 70 °C. The milk then flows through a cooler to reduce its temperature at the inlet to the cheese vats to 30 °C. Lactic acid starter culture and rennin are added in small proportions to the contents of the vats. Next, the formed curd is solidified for 2 h in the cheddar tower. The aqueous whey stream composed of unconsumed proteins, fat, lactose, mineral salts, and water is removed. Next, salts are added to the stream and the cheese is formed into blocks. After block formation, the cheese is sealed, cooled and stored to mature.

7.1.2. Process points and corresponding process variables

The process points and the corresponding process variables for which suitable monitoring tools need to be identified from the knowledge base are listed in Table 2. This information is specified in this example. Alternatively, it can be retrieved from the knowledge base, if not specified.

7.1.3. Selection of process monitoring and analysis tools

The suitable monitoring techniques and tools for each variable of the cheese manufacturing process are identified by following the procedure described in Section 6.3. First the available monitoring techniques and tools that satisfy the user constraints for each variable to be monitored are retrieved from the PAT knowledge base, and then the performances of each monitoring tool are compared to select the final set of monitoring tools. The pasteurization temperature is considered here as an example to illustrate monitoring tool selection. The specified selection criteria (specifications for temperature monitoring tools) and their desired performances (desired target values of specifications) are given in Table 3. These are user-specified values and based on them the available monitoring techniques and the corresponding monitoring tools for temperature measurements are retrieved from the knowledge base (listed in Table 4). The monitoring tools listed in Table 4 are compared on the basis of the specified performance criteria given in Table 3. The scoring system developed by Singh et al. (2010) is used to compare the monitoring tools. The scores obtained by each monitoring tool

Table 2
List of process points and corresponding process variables for the cheese manufacturing process.

Process points	Process variables
Milk storage tank	Temperature Fat Casein protein Whey protein Lactose
Pasteurizer	Pasteurization temperature Outlet milk temperature
Starter/rennin mixer	Starter concentration in milk stream Rennin concentration in milk stream
Cheese vat	Temperature pH Lactic acid concentration Coagulum concentration
Cheddar tower	Lactic acid concentration Curd moisture content
Whey separator	Curd moisture content
Salt mixing equipment	Salt concentration
Block former/cheese press	Hardness Weight Moisture content

Table 3
Selected specifications of monitoring tools for temperature and their desired values (user specific).

Specifications (performance criteria)	Desired value of specifications (desired performance)
Accuracy (Ac)	Ac ≤ 1% of reading
Precision (Pr)	Pr ≤ 0.02 °C
Lower operating limit (LOL)	LOL ≤ 0 °C
Upper operating limit (UOL)	UOL ≥ 100 °C
Response time (RT)	RT (T90) ≤ 30 s
Resolution (*R*)	*R* ≤ 0.1 °C
Sensitivity (*S*)	*S* ≥ 5E−5 sensor output/°C
Drift (*D*)	*D* ≤ 1 °C per year
Lower operating temperature (LOT)	LOT ≤ 0 °C
Upper operating temperature (UOT)	UOT ≥ 80 °C

Table 4
Monitoring tools for pasteurization temperature measurement with specifications. Results obtained from the knowledge base.

Monitoring techniques	Monitoring tools	Ac	Pr	LOR	UOR	RT (T90)	*R*	*S*	*D*	LOT	UOT
Thermocouple	Base Metal Thermocouple Element – TEAB	0.75% of reading [119]	0.02 °C [126]	−270.0 °C	1000.0 °C	100.0 ms [107]	0.1 °C [121]	0.00005 V/°C [123]	1.0 °C per year [119]	−200.0 °C [117]	1750.0 °C [117]
	T Series Standard Temperature Sensor – T4267-K20FGEJ	0.75% of reading [119]	0.02 °C [126]	−270.0 °C	1372.0 °C	100.0 ms [107]	0.1 °C [121]	0.00005 V/°C [123]	1.0 °C per year [119]	−200.0 °C [117]	1750.0 °C [117]
Resistance thermometers (RTD)	Thin Film Platinum RTD – PPG101A1	0.06% of reading	0.001 °C [115]	−50.0 °C	500.0 °C	15,000 ms	0.001 °C [118]	0.39200 Ω/°C [118]	0.05 °C per year [119]	−50.0 °C	500.0 °C
Thermistor	DO-35 Interchangeable Thermistor – 103JL1A	0.40% of reading [11]	0.003 °C [125]	−55.0 °C	300.0 °C	150.0 ms [123]	0.0001 °C [127]	3.92000 Ω/°C [124]	0.027 °C per year [122]	−55.0 °C	80.0 °C
	CWF Series – CWF 3100	1.0% of reading	0.003 °C [125]	−55.0 °C	125.0 °C	15,000 ms	0.0001 °C [127]	3.92000 Ω/°C [124]	0.027 °C per year [122]	−55.0 °C	125.0 °C

Ac: accuracy; Pr: precision; LOR: lower operating limit; UOR: upper operating limit; RT: response time; *R*: resolution; *S*: sensitivity; *D*: drift; LOT: lower operating temperature; UOT: upper operating temperature; ms: milliseconds; [ ] indicates the reference number in the knowledge base.

Table 5
Monitoring tools for pasteurization temperature measurement with scores.

Monitoring techniques	Monitoring tools	Ac	Pr	LOR	UOR	RT (T90)	*R*	*S*	*D*	LOT	UOT	Total scores
Thermocouple	Base Metal Thermocouple Element – TEAB	3	3	5	4	5	3	3	3	5	5	39
	T Series Standard Temperature Sensor – T4267-K20FGEJ	**3**	**3**	**5**	**5**	**5**	**3**	**3**	**3**	**5**	**5**	**40**
Resistance thermometers (RTD)	Thin Film Platinum RTD – PPG101A1	5	5	3	3	3	4	4	4	3	4	38
Thermistor	DO-35 Interchangeable Thermistor – 103JL1A	4	4	4	2	4	5	5	5	4	2	39
	CWF Series – CWF 3100	2	4	4	1	3	5	5	5	4	3	36

Note: Only a few of the monitoring tools corresponding to each monitoring technique are listed here for demonstration purposes. The full table can be obtained from the corresponding author.

Table 6
Process variables and selected monitoring techniques and tools for the cheese manufacturing process.

Process points	Process variables	Monitoring techniques	Monitoring tools
Milk storage tank	Temperature	Thermocouple	T Series Standard Temperature Sensor – T4267-K20FGEJ
	Fat	NIR	KJT550 NIR On-line Composition Analyzer
	Casein protein	NIR	KJT550 NIR On-line Composition Analyzer
	Whey protein	NIR	KJT550 NIR On-line Composition Analyzer
	Lactose	NIR	KJT550 NIR On-line Composition Analyzer
Pasteurizer	Pasteurization temperature	Thermocouple	T Series Standard Temperature Sensor – T4267-K20FGEJ
	Outlet milk temperature	Thermocouple	T Series Standard Temperature Sensor – T4267-K20FGEJ
Starter/rennin mixer	Starter concentration in milk stream	Refractometer	ATR W Series (Automatic Critical angle Refractometer)
	Rennin concentration in milk stream	Refractometer	ATR W Series (Automatic Critical angle Refractometer)
Cheese vat	Temperature	Thermocouple	T Series Standard Temperature Sensor – T4267-K20FGEJ
	pH	Electrochemical sensor	pH Meter – Model 2410
	Lactic acid concentration	Refractometer	ATR W Series (Automatic Critical angle Refractometer)
	Coagulum concentration	Refractometer	ATR W Series (Automatic Critical angle Refractometer)
Cheddar tower	Lactic acid concentration	Refractometer	ATR W Series (Automatic Critical angle Refractometer)
	Curd moisture content	NIR	AvaSpec-NIR256-1.7
Whey separator	Curd moisture content	NIR	AvaSpec-NIR256-1.7
Salt mixing equipment	Salt concentration	Refractometer	ATR W Series (Automatic Critical angle Refractometer)
Block former/cheese press	Hardness	Strain gauge	BBC-Strain Gauge-SS-027
	Weight	Strain gauge	BBC-Strain Gauge-SS-027
	Moisture content	NIR	AvaSpec-NIR256-1.7

are given in Table 5, where, it can be noted (marked in bold) that the "T Series Standard Temperature Sensor" obtained the highest total score and therefore this monitoring tool is selected for temperature monitoring. The selected monitoring technique is a thermocouple. Similarly the monitoring techniques and the corresponding monitoring tools for other process variables are identified as given in Table 6.

7.2. Tablet manufacturing process: pharmaceutical process case study

The detailed description of the PAT system design of a tablet manufacturing process has been presented by Singh et al. (2010). A variable, "tablet coating thickness" is considered here to demonstrate the application of the knowledge base for pharmaceutical

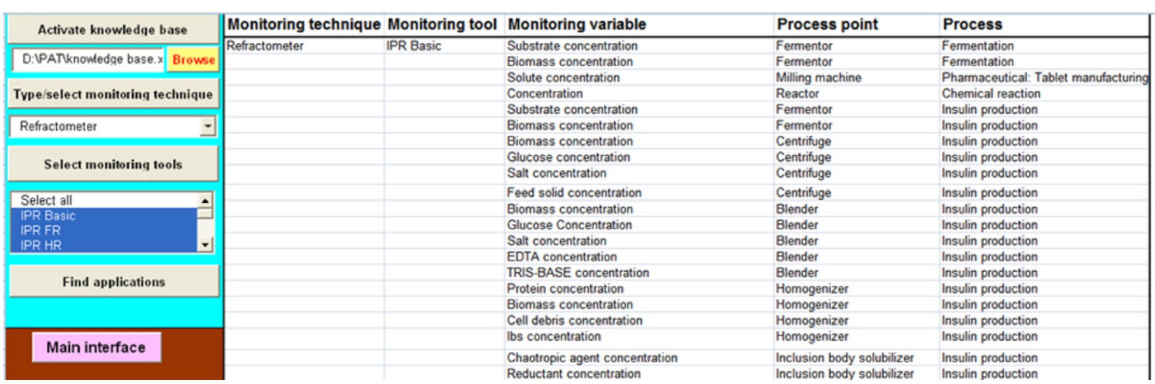

Fig. 12. Reverse search procedure to find applications of monitoring tools, user interface (left) with example of a result (right). Note that this screen shot is not showing all the retrieved information.

Table 7a
Monitoring tools for tablet coating thickness measurement with specifications.[a] Results obtained from the knowledge base (Ac: accuracy; Pr: precision; LOR: lower operating limit (wave length); UOR: upper operating limit (wave length); RT: response time (sampling time); *R*: resolution; *C*: cost (approximate)[b]; nm: nanometer; [ ] indicates the reference number in the knowledge base).

Monitoring techniques	Monitoring tools	Ac	Pr	LOR	UOR	RT	*R*	*C*
NIR	EPP 2000 Fiber Optic Spectrometer-NIR [211]	0.200 nm	0.200 nm	500.00 nm	1200.00 nm	0.0020 s	0.80 nm	1936.00€
	AvaSpec-NIR256-1.7 [210]	0.170 nm	0.170 nm	900.00 nm	1750.00 nm	0.0011 s	50.00 nm	8950.00€
UV	AvaSpec 3648-UA-25-AF [168]	0.140 nm	0.140 nm	200.00 nm	1100.00 nm	N/A	1.40 nm	3156.00€
	F20-UV [300]	2.000 nm	0.100 nm	200.00 nm	1100.00 nm	1.0000 s	N/A	22000.00€
Raman	AvaRaman-532 [170]	N/A	N/A	535.00 nm	752.00 nm	60.0000 s	10.00 cm^{-1}	11900.00€
	AvaRaman-785TEC [170]	N/A	N/A	785.00 nm	930.00 nm	60.0000 s	8.00 cm^{-1}	14500.00€

Note: Only a few of the monitoring tools corresponding to each monitoring technique are listed here for demonstration purposes. The full table can be obtained from the corresponding author.

[a] Specifications (Ac, Pr, LOR, UOR, *R*) of spectroscopic monitoring tools are given in terms of wave length.

[b] The cost of monitoring tools depends on various factors (e.g. suppliers, taxations, manufacturing sites, etc.) and is given here only for demonstration purposes.

Table 7b
Monitoring tools for dissolved oxygen concentration with specifications. Results obtained from the knowledge base (Ac: accuracy; Pr: precision; LOR: lower operating limit; UOR: upper operating limit; RT: response time; *R*: resolution; *D*: drift; LOT: lower operating temperature; UOT: upper operating temperature; [ ] indicates the reference number in the knowledge base).

Monitoring techniques	Monitoring tools	Ac	Pr	LOR	UOR	RT (T90)	*R*	*D*	LOT	UOT
Optical sensor	RDO [185]	1.0% of reading [50]	0.11% of reading [77]	0.0 mg/l [50]	20.0 mg/l [50]	30.0 s [185]	0.001 mg/l [96]	0.019% of reading per week [77]	0.0 °C [106]	50.0 °C [106]
	FOXY Sensor System [182]	0.1% of reading	0.1% of reading	0.0 mg/l	40.7 mg/l [50]	1.0 s	0.001 mg/l [96]	0.019% of reading per week [77]	−44.0 °C	80.0 °C
Electrochemical polarographic sensor	770MAX 368-210 [181]	2.0% of reading	2.0% of reading	0.0 mg/l	5.0 mg/l	60.0 s	0.01 mg/l [109]	1.000% of reading per week [110]	5.0 °C	50.0 °C [106]
	M300 High Performance DO Sensor – 58-037-204 [181]	1.0% of reading	1.0% of reading	0.0 mg/l	10.0 mg/l	90.0 s	0.01 mg/l [109]	1.000% of reading per week [110]	0.0 °C	60.0 °C
Electrochemical galvanic sensor	Dissolved Oxygen Probe – Model 760DOP [174]	0.15% of reading [1]	0.18% of reading [7]	0.0 mg/l	20.0 mg/l	9.0 s [102]	0.10 mg/l	0.110% of reading per week [102]	0.0 °C	50.0 °C [106]
	DirectLine® Series – DL425	5.0% of reading	5.0% of reading	0.0 mg/l	0.2 mg/l	60.0 s	0.001 mg/l	0.110% of reading per week	0.0 °C	50.0 °C

Note: Only a few of the monitoring tools corresponding to each monitoring technique are listed here for demonstration purposes. The full table can be obtained from the corresponding author.

processes. Techniques and tools for "tablet coating thickness" monitoring are retrieved from the knowledge base together with the specifications (type of specifications and their desired target values are user specified) as given in Table 7a. The full list of suitable monitoring techniques and the corresponding monitoring tools for each critical variable in the tablet manufacturing process has been reported by Singh et al. (2010).

7.3. Fermentation process: bioprocess case study

The detailed description of the PAT system design for a fermentation process has been presented in an earlier work (Singh, Gernaey, & Gani, 2009). In this paper, the variable "Dissolved oxygen concentration" is considered to demonstrate the application of the knowledge base for bioprocess operations. Techniques and tools for "Dissolved oxygen concentration" monitoring are retrieved from the knowledge base together with the specifications (type of specifications and their desired values are provided by the user) as given in Table 7b. The full list of suitable monitoring techniques for each critical variable involved in the fermentation process has been reported in Singh et al. (2009).

7.4. Identification of the potential application range of a specified monitoring tool-technique

This example highlights the use of the reverse search procedure to identify the application range of a specified monitoring technique (for example, a refractrometer) from the developed knowledge base. The user interface of the reverse search engine together with the search result is shown in Fig. 12, where on the left the user interface to define the problem is highlighted, while, on the right, the results of the reverse search is highlighted. A specific monitoring technique (refractrometer) and tool (IPR Basic) are selected for which the potential applications need to be identified. It can be noted that according to the PAT knowledge base, potential applications of the refractrometer and the selected monitoring tool are identified in terms of monitoring variables, process points, and processes where they occur.

8. Conclusions

A knowledge base on process monitoring and analysis systems has been developed to support the design of PAT systems. This PAT knowledge base consists of the process knowledge as well as the knowledge on measurement methods and tools. The structure of the PAT knowledge base has been designed such that it allows easy implementation of the inference engine. An ontology has been defined to obtain the structure of the knowledge base and covers a wide range of industrial processes, such as fermentation, pharmaceutical tablet manufacturing, insulin production, cheese manufacturing, butter manufacturing, chemical reaction and a crystallization process. The structure of the PAT knowledge base is generic and can be easily extended to include new processes as well as to include new objects, for example to add a newly developed sensor, within an existing process. Moreover, the structure of the PAT knowledge base can be adapted to build similar knowledge-based systems for other applications, for example, a knowledge base on reaction systems can be built to represent the type of processes, corresponding process points, reactions involved in each process point, favorable reaction conditions, available catalyst and corresponding catalyst properties. This kind of knowledge base can be employed to identify the appropriate catalyst and optimum operating conditions for any reaction. As was demonstrated through the case studies, the PAT knowledge base together with its applying system provide the manufacturing industries an easily accessible and systematic option to compare the salient features of available sensors in order to select the best one that satisfies a specific set of requirements.

Acknowledgement

The PhD project of Ravendra Singh is financed by a PhD scholarship from the Technical University of Denmark.

Appendix A. Design of ontology on process monitoring and analysis systems

Following the methodology shown in Fig. 2, the ontology for process monitoring and analysis systems is developed as follows:

1. Step 1: Definition of the domain and scope of the ontology
   - *Step 1.1 – Define the domain*: The representation of process monitoring and analysis systems is the domain of the ontology.
   - *Step 1.2 – Create a list of competency questions*: Use the competency questions suggested by Natalya and McGuinness (2007). In the process monitoring and analysis domain, the following competency questions may be asked:
     a. What are the process points where monitoring might be needed?
     b. Which process variables should be considered for monitoring?
     c. Which monitoring techniques are available for monitoring of a particular process variable?
     d. Which monitoring tools are available for monitoring of a particular process variable?
     e. Which monitoring tools satisfied the user specifications and constraints for monitoring a particular variable?
     f. What is the cost of a particular monitoring tool?
     g. Who are the suppliers of a particular monitoring tool?
     h. What are the actuator candidates for a particular controlled variable?
     i. What are the applications of a specific monitoring tool?
   - *Step 1.3 – Create a list of information needed to answer the competency questions*: To answer the above mentioned competency questions, the knowledge base needs to have information about the contents: process, process points, process variables, monitoring techniques, monitoring tools, monitoring tools specifications, cost, supplier, actuator candidates.
   - *Step 1.4 – Define scope*: This ontology has to be used for the design of process monitoring and analysis systems (PAT systems).
2. Step 2: Enumeration of the important terms in the ontology
   - *Step 2.1 – Create a list of important terms*: The important terms related to process monitoring and analysis systems are listed in Table 8.
   - *Step 2.2 – Elaborate the important terms*

Table 8
List of the important terms.

Processes, process points, process variables, actuator candidates, variable types, monitoring techniques, monitoring tools, supplier, monitoring tools specifications, accuracy, precision, operating range, response time, resolution, sensitivity, drift, operating temperature range, cost, process points of a process, process variables related to a process point, process variables of a process, actuator candidates of a variable, type of a process variable, monitoring techniques for a variable type, monitoring tools based on a monitoring technique, monitoring tools of a variable type, supplier of a monitoring tool, specifications of a monitoring tool, accuracy of a monitoring tool, precision of a monitoring tool, operating range of a monitoring tool, response time of a monitoring tool, resolution of a monitoring tool, sensitivity of a monitoring tool, drift of a monitoring tool, cost of a monitoring tool

Process: It can be any chemical or biological process, for example, fermentation process, tablet manufacturing process, crystallization process, etc.

Process points: In general the process equipments in the process flowsheet are considered as the process points. Note that each considered process point can have more than one process operation (e.g. the fermentor has more than one operation: charging, fermentation, and discharging).

Process variables: Each process operation is defined by a set of variables. The variables defining each operation related to the considered process points come under this category.

Variable types: Sometimes it happens that a group of process variables can be monitored by the same monitoring tools. So this set of variables can be represented by a single variable type. For instance the three variables, inlet temperature, outlet temperature and process temperature can be placed in the category of the variable temperature because the monitoring tools are the same for these variables.

Actuator candidates: In this category those variables have to be listed which can be the actuator of the considered variable. This list provides the alternatives available for the manipulated variables.

Monitoring techniques: The available monitoring techniques/methods for each variable have to be listed in this category.

Monitoring tools: Each monitoring technique/method comprises a set of monitoring tools.

Monitoring tools specifications: Each monitoring tool has some characteristics which reflect its performance. Some of them are described here:

a. **Accuracy**: The accuracy is defined as, "The ability of a measurement to match the actual value of the quantity being measured" (Haby, 2008). Accuracy of a displayed value is characterized as an uncertainty of a measurement display representing the actual value being measured. It is expressed in terms of how far off any given reading could be from the true value, given in terms of a fixed value (e.g. $\pm 0.5\,^{\circ}$F), a percent of reading (e.g. 0.2% of reading), a fixed value plus a percent of reading or a percent of the instrument's full-scale value. The accuracy of the sensors varies with the type, purpose and manufacturer (Stum, 2006).
b. **Precision**: Precision is defined as, "The ability of a measurement to be consistently reproduced", i.e. it is the ability of repeatability (Haby, 2008). In other words the precision is how close the measured values are to each other. There are several ways to report the precision of results. The simplest is the range (the difference between the highest and lowest results) often reported as a $\pm$deviation from the average. A better way, but one that requires statistical analysis, would be to report the standard deviation. Note that, it is possible for an instrument to have good accuracy and good precision, good accuracy but bad precision, bad accuracy but good precision and bad accuracy and bad precision.
c. **Operating range**: The operating range of a monitoring tool is defined as the minimum and maximum limiting value that it can measure. For example, a given pressure sensor may have a range of 0 to 400 mm Hg (Carr & Brown, 2008).
d. **Response time**: Sensors do not change output state immediately when an input parameter change occurs. Rather, the sensor output will evolve over a period of time before truly reflecting the new state. The response time can be defined as the time required for a sensor output to change from its previous state to a final settled value (when subjected to a step input change) within a tolerance band of the correct new value (Carr & Brown, 2008; Sutherland, 2004). For

example, T90 means time required for a sensor output to reach 90% of the final value, when subjected to a step input change.

e. **Sensitivity**: The sensitivity of the sensor is defined as the change in sensor output per unit change in measured quantity (Sutherland, 2004). Sensors that measure very small changes must have very high sensitivities.

f. **Resolution**: The resolution of the sensor is defined as the smallest change in measurable quantity, which the sensor can detect (Sutherland, 2004). In other words, the smallest detectable incremental change of the input variable that can be detected in the output signal. Resolution can be expressed either as a proportion of the reading (or the full-scale reading) or in absolute terms (Carr & Brown, 2008).

g. **Drift**: Drift is defined as the gradual change in any quantitative characteristic that is supposed to remain constant. It is an undesired change in output over a period of time that is unrelated to the input. It can be due to aging, temperature effects, sensor contamination, etc. (CAPGO, 2007).

h. **Operating temperature range**: It is defined as the temperature range within which the monitoring tools can be used, i.e. the temperature range in which the device will meet performance specifications (Kionix, 2004).

i. **Cost**: It is the price of the monitoring tool. The cost of a specific monitoring tool depends on the supplier and the location of the manufacturing company.

Supplier: It is important to know the manufacturer of the monitoring tools. It should be noted that the same manufacturing companies can produce different monitoring tools (depending on the specifications) for measurement of one variable.

Selection of monitoring tools: The process of finding out the best monitoring tool for a particular application.

3. Step 3: Definition of the classes and their hierarchies
   - *Step 3.1 – Define the classes of the ontology*: The terms that describe the objects having independent existence are identified from Table 8. These terms are the classes of the ontology as listed in Table 9.
   - *Step 3.2*: Select a class from Table 9 and a corresponding object for the selected class (for example, class is process point and object is fermentor).
   - *Step 3.3*: Verify that the selected object (fermentor) is not the instance of any other class.
   - *Step 3.4*: Repeat steps 3.2 and 3.3 until all classes and objects have been identified.
   - *Step 3.5*: Organize the classes in a hierarchical taxonomy as shown in Fig. 13.
4. Step 4: Definition of the domain and range of each class (slot)
   - *Step 4.1 – Create a list of slots*: The slots are process points of a process; process variables related to a process point; process variables of a process; actuator candidates of a variable; type of a process variable; monitoring techniques for a variable type; monitoring tools based on a monitoring technique; monitoring tools for a variable type, supplier of a monitoring tool; specifications of a monitoring tool; accuracy of a monitoring tool; precision of a monitoring tool; operating range of a monitoring tool; response time of a monitoring tool; resolution of a monitoring tool; sensitivity of a monitoring tool; drift of a monitoring tool; and cost of a monitoring tool.
   - *Step 4.2*: For each slot, identify its class as the domain and the instances of the class as the range (for example, the slot 'process points of a process' belongs to the class 'processes' while 'process points' is the instance of the class).
   - *Step 4.3*: Repeat step 4.2 for all slots to create the table of slots, domains and ranges (see Table 10).
5. Step 5: Definition of the facet of the slots

   Two types of values can fill in the slot, character strings and/or numerical values. For example, the slot, 'process points of a process' is filled with the names of process points while the slot, 'accuracy of a monitoring tool' is filled with numerical values. Divide the slots according to facet-1 (top sub-level 1) and facet-2 (top sub-level 2).

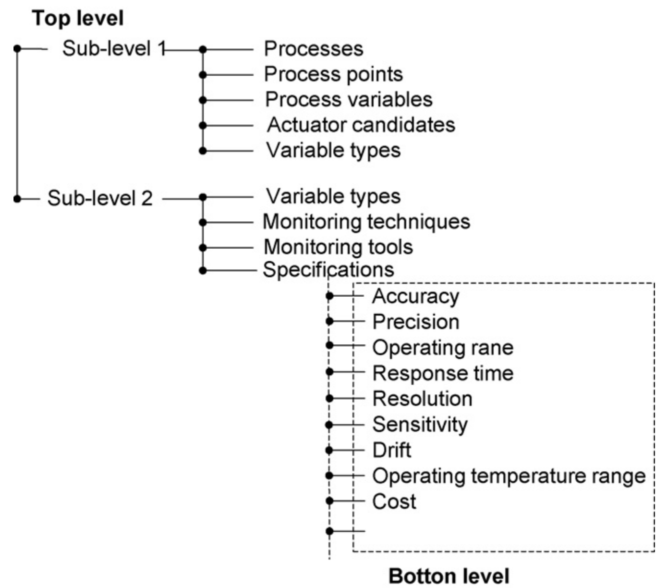

Fig. 13. Hierarchical taxonomy of the classes.

Table 9
List of the classes.

Processes, process points, process variables, actuator candidates, variable types, monitoring techniques, monitoring tools, supplier, specifications, accuracy, precision, operating range, response time, resolution, sensitivity, drift, operating temperature range, cost

Table 10
List of slots and corresponding domains and ranges.

Slots	Domain	Range
Process points of a process	Processes	Process points
Process variables related to a process point	Process points	Process variables
Process variables of a process	Processes	Process variables
Actuator candidates of a variable	Process variables	Actuator candidates
Type of a process variable	Process variables	Variable types
Monitoring techniques for a variable type	Variable types	Monitoring techniques
Monitoring tools based on a monitoring technique	Monitoring techniques	Monitoring tools
Monitoring tools for a variable type	Variable types	Monitoring tools
Supplier of a monitoring tool	Monitoring tools	Suppliers
Specifications of a monitoring tool	Monitoring tools	Specifications
Accuracy of a monitoring tool	Specifications	Accuracy
Precision of a monitoring tool	Specifications	Precision
Operating range of a monitoring tool	Specifications	Operating range
Response time of a monitoring tool	Specifications	Response time
Resolution of a monitoring tool	Specifications	Resolution
Sensitivity of a monitoring tool	Specifications	Sensitivity
Drift of a monitoring tool	Specifications	Drift
Cost of a monitoring tool	Specifications	Cost

6. Step 6: Creation of instances
   - *Step 6.1*: Select a slot and its corresponding domain (for example, domain of class of the slot 'process points of a process' is 'processes').
   - *Step 6.2*: Identify and list all instances of the selected class within the slot (for example, instances of processes are, fermentation process, tablet manufacturing process, etc.).
   - *Step 6.3*: Steps 6.1–6.2 are repeated for all slots and their classes.
7. Step 7: Evolved ontology

   The evolved ontology contains two parts. As shown in Fig. 13 top sub-level 1 (facet-1) evolves into process knowledge (see Fig. 3) while the top sub-level 2 (facet-2) evolves into monitoring techniques and tools knowledge (see Fig. 4).
8. Step 8: Validation/evaluation

   The evolved ontology has been found to be able to provide reliable answers to the competency questions (as framed in step 1.2). The applications of the designed ontology have been highlighted in Sections 6 and 7. Moreover, the developed ontological knowledge-based system has been applied successfully for the design of the process monitoring and analysis system of a fermentation process (Singh et al., 2009) and a tablet manufacturing process (Singh et al., 2010).
9. Step 9: Final ontology

   The final ontology defines the structure of the knowledge base as shown in Figs. 3 and 4 and the amount of data/knowledge stored has been discussed in Section 5.

Appendix B. Extension procedure of the knowledge base

The procedure for extending the PAT knowledge base is highlighted through the work-flow diagram in Fig. 14.

Appendix C.

C.1. Statistics calculation procedure

Table 11 summarizes the calculation procedures that apply to the various objects of the knowledge base. For instance the maximum total number of monitoring tools (T) stored in the knowledge base can be estimated by summing of all monitoring tools available for each technique for each variable (see the row "total number of monitoring tools" in Table 11).

The equations given in Table 11 did not take into account repetition of data in the two sections of the knowledge base. Figs. 6 and 7 show the repetition that can occur in the two sections. As shown in Fig. 7, there is no repetition of objects under the category, 'variable types', but it is quite common that the same monitoring technique can be applied to monitor several variables, i.e. the objects under the category 'monitoring techniques' can be repeated. The repetition of traditional monitoring tools (e.g. temperature sensor) is not common but exceptionally a spectroscopic sensor can be employed to monitor more than one variable simultaneously. In general each monitoring tool has distinct specifications. Also, the same company is often involved in the manufacturing of several types of monitor-

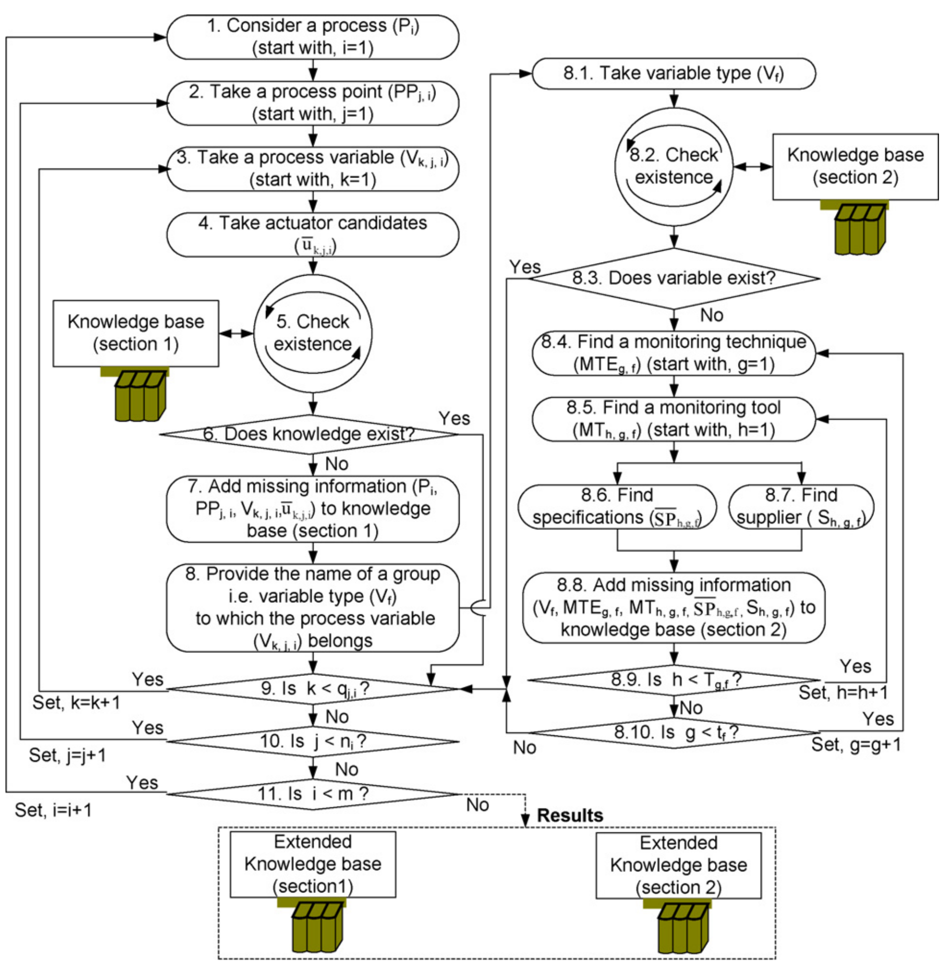

Fig. 14. Procedure for extension of the PAT knowledge base (m: no. of processes added to the knowledge base; n_i: no. of process points of the ith process; $q_{j,i}$: no. of process variables involved in the jth process point of the ith process; t_f: no. of monitoring techniques available for the fth variable type; $T_{g,f}$: no. of monitoring tools based on the gth technique for the fth variable type).

Table 11
Formulas to be applied for calculation of knowledge base statistics.

Number of elements	Calculation
Total number of processes (m)	m
Total number of process points (n)	$n=\sum_{1}^{m} n_i$
Total number of process variables (q)	$q=\sum_{i=1}^{m} q_i=\sum_{i=1}^{m}\sum_{j=1}^{n_i} q_{j,i}$
Total number of actuator candidates (ac)	$ac=\sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{q_{j,i}} ac_{k,j,i}$
Total number of variable types (r)	r
Total number of monitoring techniques (t)	$t=\sum_{f=1}^{r} t_f$
Total number of monitoring tools (T)	$T=\sum_{f=1}^{r}\sum_{g=1}^{t_f} T_{g,f}$
Total number of suppliers (c)	$c=\sum_{f=1}^{r}\sum_{g=1}^{t_f} c_{g,f}$

ing tools, and therefore, the objects under the ‘supplier’ category can be repeated.

References

Avouris, N. M. (1995). Cooperating knowledge-based systems for environmental decision support. *Knowledge-Based Systems*, *8*(1), 39–54.

Barua, A., Ray, S., & Sinha, S. (1993). Transelex: A knowledge-based approach for the selection of transducers using a dynamic database. *Engineering Applications of Artificial Intelligence*, *6*(1), 73–75.

Bratsas, C., Koutkias, V., Kaimakamis, E., Bamidis, P. D., Pangalos, G. I., & Maglaveras, N. (2007). KnowBaSICS-M: An ontology-based system for semantic management of medical problems and computerized algorithmic solutions. *Computer Methods and Programs in Biomedicine*, *88*, 39–51.

Brudzewski, K., & Dolecka, U. (1995). Optimal selection of sensors based on the catalogue data: A vector-space approach to cross-sensitivity effects. *Sensors and Actuators, B*, *23*, 71–73.

Burke, E. K., MacCarthy, B., Petrovic, S., & Qu, R. (2000). Structured cases in case-based reasoning—Re-using and adapting cases for time-tabling problems. *Knowledge-Based Systems*, *13*, 159–165.

CAPGO. (2007). *Sensor glossary*. http://www.capgo.com/Resources/Sensors/SensorGlossary.html

Carr, J. J., & Brown, J. M. (2008). *Sensor terminology* (3rd ed.). National Instruments. http://zone.ni.com/devzone/cda/ph/p/id/227

Choi, J. W. (2009). Architecture of a knowledge based engineering system for weight and cost estimation for a composite airplane structures. *Expert Systems with Applications*, *36*, 10828–10836.

Corcho, O., Fernandez-Lopez, M., & Gomez-Perez, A. (2003). Methodologies, tools and languages for building ontologies. Where is their meeting point? *Data & Knowledge Engineering*, *46*, 41–64.

Dupuit, E., Pouet, M. F., Thomas, O., & Bourgois, J. (2007). Decision support methodology using rule-based reasoning coupled to non-parametric measurement for industrial wastewater network management. *Environmental Modelling & Software*, *22*, 1153–1163.

Dutta, S. (1993). Integrating case- and rule based reasoning. *International Journal of Approximate Reasoning*, *8*, 163–203.

Eid, M., Liscano, R., & Saddik, A. E. (2006). A novel ontology for sensor networks data. In *International conference on computational intelligence for measurement systems and applications*.

Fernandez-Lopez, M. (1999). Overview of methodologies for building ontologies. In *IJCAI99 workshop on ontologies and problem-solving methods: Lessons learned and future trends*.

Fernandez-Lopez, M., Gomez-Perez, A., & Juristo, N. (1997). METHONTOLOGY: From ontological art towards ontological engineering. In *AAAI symposium on ontological engineering*.

Fernandez-Lopez, M., Gomez-Perez, A., Pazos-Sierra, A., & Pazos-Sierra, J. (1999). Building a chemical ontology using METHONTOLOGY and the ontology design environment. *IEEE Intelligent Systems & Their Applications*, *4*(1), 37–46.

Finnie, G., & Sun, Z. (2003). R 5 model for case-based reasoning. *Knowledge-Based Systems*, *16*, 59–65.

Gabor, A., Barna, G. (1988). Knowledge-based system for supporting statistical database management. *1*(5), 279–284.

Gebus, S., & Leiviska, K. (2009). Knowledge acquisition for decision support systems on an electronic assembly line. *Expert Systems with Applications*, *36*, 93–101.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human–Computer Studies*, *43*, 907–928.

Grüninger, M., & Fox, M. S. (1995). Methodology for the design and evaluation of ontologies. In *Workshop on basic ontological issues in knowledge sharing*.

Guarina, N. (1997). Understanding, building and using ontologies. *International Journal of Human–Computer Studies*, *46*, 293–310.

Haby, J. (2008). *What is the difference between accuracy and precision?* USA: Mississippi State University. http://www.theweatherprediction.com/habyhints/246/ (weather prediction website)

Hendriks, P. H. (1999). The organisational impact of knowledge-based systems: A knowledge perspective. *Knowledge-Based Systems*, *12*, 159–169.

Howells, H., Davies, A., Macauley, B., & Zancanato, R. (1999). Large scale knowledge based systems for airborne decision support. *Knowledge-Based Systems*, *12*(1999), 215–222.

Kandil, M. S., Farghal, S. A., & Abdel-Aziz, M. R. (1992). Knowledge base of an expert system for generation expansion planning. *Electric Power Systems Research*, *23*, 59–70.

Kim, J., Choi, B., Shin, H., & Kim, H. (2007). A methodology for constructing of philosophy ontology based on philosophical texts. *Computer Standards & Interfaces*, *29*, 302–315.

Kionix. (2004). *MEMS sensor terminology accelerometers/inclinometers*. http://www.kionix.com/Adobe-Documents/MEMS%20Sensor%20Terminology.pdf

Knowledge Systems Laboratory. (1997). *Report, a glossary of ontology terminology*. USA: Stanford University. http://www-ksl-svc.stanford.edu:5915/doc/frame-editor/glossary-of-terms.html

Lai, L. F. (2007). A knowledge engineering approach to knowledge management. *Information Sciences*, *177*, 4072–4094.

Lee, D. (2006). Knowledge-based system for safety control of damaged ship. *Knowledge-Based Systems*, *19*, 187–191.

Lenat, D. B., & Guha, R. V. (1990). *Building large knowledge-based systems: Representation and inference in the Cyc Project*. Boston: Addison-Wesley.

Leone, H. (1996). A knowledge-based system for hazop studies. *Computers & Chemical Engineering*, *20*(Suppl.), 369–374.

Matos, E. E., Campos, F., Braga, R., & Palazzi, D. (2010). CelOWS: An ontology based framework for the provision of semantic web services related to biological models. *Journal of Biomedical Informatics*, *43*, 125–136.

Morbach, J., Yang, A., & Marquardt, W. (2007). OntoCAPE—A large-scale ontology for chemical process engineering. *Engineering Applications of Artificial Intelligence*, *20*, 147–161.

Natalya, F. N., & McGuinness, D. L. (2007). *Ontology development 101: A guide to creating your first ontology*. Stanford, CA, 94305: Stanford University. http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html

Ong, J. B., Masud, A. S. M., & Eyada, O. K. (1992). Senses: A knowledge-based sensor selection system. *Computers & Industrial Engineering*, *22*(1), 1–8.

Patel-Schneider, P. F., Hayes, P., & Horrocks, I. (2004). *OWL web ontology language semantics and abstract syntax, W3C*. http://www.w3.org/TR/owl-semantics/syntax.html#2.3.1

Paton, N. W., Goble, C. A., & Bechhofer, S. (2000). Knowledge based information integration systems. *Information and Software Technology*, *42*, 299–312.

Patterson, D., Rooney, N., Galushka, M., Dobrynin, V., & Smirnova, E. (2008). SOPHIA-TCBR: A knowledge discovery framework for textual case-based reasoning. *Knowledge-Based Systems*, *21*, 404–414.

Pirró, G., Mastroianni, C., & Talia, D. (2009). A framework for distributed knowledge management: Design and implementation. *Future Generation Computer Systems*, doi:10.1016/j.future.2009.06.004

Qian, Y., Liang, J., & Dang, C. (2009). Knowledge structure, knowledge granulation and knowledge distance in a knowledge base. *International Journal of Approximate Reasoning*, *50*, 174–188.

Rahman, S., Khan, F., Veitch, B., & Amyotte, P. (2009). ExpHAZOPþ: Knowledge-based expert system to conduct automated HAZOP analysis. *Journal of Loss Prevention in the Process Industries*, *22*, 373–380.

Rayo, A. (2007). *Ontological commitment*. http://web.mit.edu/arayo/www/ontcom.pdf

Shieh, J., Huber, J. E., Fleck, N. A., & Ashby, M. F. (2001). The selection of sensors. *Progress in Material Science*, *46*, 461–504.

Singh, R., Gernaey, K. V., & Gani, R. (2009). Model-based computer-aided framework for design of process monitoring and analysis systems. *Computers and Chemical Engineering*, *33*, 22–42.

Singh, R., Gernaey, K. V., & Gani, R. (2010). ICAS-PAT: A software for design, analysis and validation of PAT systems. *Computers and Chemical Engineering*, *34*, 1108–1136.

Siricharoen, W. V. (2009). Ontology modeling and object modeling in software engineering. *International Journal of Software Engineering and Its Applications*, *3*(1), 43–60.

Staab, S., Schnurr, H. P., Studer, R., & Sure, Y. (2001). Knowledge processes and ontologies. *IEEE Intelligent Systems*, *16*(1), 26–34.

Stum, K. (2006). Sensor accuracy and calibration theory and practical application. In *National conference on building commissioning*.

Sutherland, J. W. (2004). *Sensor fundamentals*. http://www.mfg.mtu.edu/cyberman/machtool/machtool/sensors/fundamental.html

Swartout, B., Ramesh, P., Knight, K., & Russ, T. (1997). Toward distributed use of large-scale ontologies. In *AAAI symposium on ontological engineering*.

Uschold, M. (1996). *Building ontologies: Towards a unified methodology*. Cambridge: Expert Systems.

Uschold, M., & King, M. (1995). Towards a methodology for building ontologies. In *IJCAI95 workshop on basic ontological issues in knowledge sharing*.

Venkatasubramanian, V., Zhao, C., Joglekar, G., Jain, A., Hailemariam, L., Suresh, P., et al. (2006). Ontological informatics infrastructure for pharmaceutical product development and manufacturing. *Computers and Chemical Engineering*, *30*, 1482–1496.

Villanueva-Rosales, N., & Dumontier, M. (2008). yOWL: An ontology-driven knowledge base for yeast biologists. *Journal of Biomedical Informatics*, *41*, 779–789.

Wen, W., Chen, Y. H., & Chen, I. C. (2008). A knowledge-based decision support system for measuring enterprise performance. *Knowledge-Based Systems*, *21*, 148–163.

Zaror, C. A., & Pyle, D. L. (1997). Process design: An exercise and simulation examples. In *Chemical engineering for the food industry*. London: Blackie A&P, an imprint of Chapman & Hall. ISBN 0 412 49500 7.