Engineering Applications of Artificial Intelligence 27 (2014) 1–16

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Bridging control and artificial intelligence theories for diagnosis: A survey

L. Travé-Massuyès [a,b,*]

[a] CNRS, LAAS, 7, Avenue du Colonel Roche, F-31031 Toulouse, France
[b] Univ de Toulouse, LAAS, F-31031 Toulouse, France

ARTICLE INFO

Article history:
Received 12 December 2012
Received in revised form
24 September 2013
Accepted 26 September 2013
Available online 26 October 2013

Keywords:
Model based diagnosis
Data based diagnosis
Abstractions
Fault management architectures

ABSTRACT

Diagnosis is the process of identifying or determining the nature and root cause of a failure, problem, or disease from the symptoms resulting from selected measurements, checks or tests. The different facets of this problem and the wide spectrum of classes of systems make it interesting to several communities and require bridging several theories. Diagnosis is actually a functional fragment in fault management architectures and it must smoothly interact with other functions. This paper presents diagnosis as it is understood in the Control and Artificial Intelligence fields, and exemplifies how different theories of these fields can be synergistically integrated to provide better diagnostic solutions and to achieve improved fault management in different environments.

1. Introduction

The goal of diagnosis is to identify the possible causes explaining a set of observed symptoms. A set of concomitant tasks contribute to this goal and the following three tasks are commonly identified:

- *fault detection*, which aims at discriminating normal system states from abnormal ones, i.e. states which result from the presence of a fault,
- *fault isolation*, also called *fault localization*, whose goal is to point at the faulty components of the system,
- *fault identification*, whose output is the type of fault and possibly the model of the system impacted by this fault.

Faced with the diversity of systems and different views of the above problems, several scientific communities have addressed these tasks and contributed with a large spectrum of methods. The Signal Processing, Control and Artificial Intelligence (AI) communities are leading actors in this field.

Diagnosis is carried out from the signals that permit efficient fault detection toward the upper levels of supervision that call for qualitative interpretations. Proposing relevant abstractions to interpret the available signals is hence a key issue.

Signal processing provides specific contributions in the form of statistic algorithms for detecting changes in signals, hence detecting faults. This remains out of the scope of this paper and has been surveyed in several reference books and papers (Basseville, 1988; Basseville and Nikiforov, 1993; Basseville et al., 2004; Fillatre and Nikiforov, 2007; Fouladirad et al., 2008).

Interfaces between continuous signals and their abstract interpretations, in symbolic or event-based form, implement qualitative interpretations of the signals that are required for supervision. To do that, discrete formalisms borrowed from AI find a natural link with continuous models from the Control community. These two communities have their own model-based diagnosis track:

- the FDI (Fault Detection and Isolation) track, whose foundations are based on engineering disciplines, such as control theory and statistical decision making,
- the DX (Diagnosis) track, whose foundations are derived from the fields of logic, combinatorial optimization, search and complexity analysis.

In the last decade, there has been a growing number of researchers in both communities who have tried to understand and bridge FDI and DX approaches to build better, more robust and effective diagnostic systems.

Data-based diagnosis approaches based on machine learning techniques, such as pattern recognition (Fukunaga, 1990; Denoeux et al., 1997), are present in both the Control and AI communities and complement model-based approaches to provide solutions to a variety of diagnostic problems where the difficulty arises from the scarce nature of the instrumentation or, conversely, from the massive amounts of data to be interpreted to extract hidden knowledge. Interesting bridges also arise when we consider data-based and model-based approaches.

* Correspondence address: CNRS, LAAS, 7, Avenue du Colonel Roche, F-31031 Toulouse, France. Tel.: +33 561336302.
E-mail address: louise@laas.fr

0952-1976/$ - see front matter
http://dx.doi.org/10.1016/j.engappai.2013.09.018

Other bridges can be found when considering that diagnosis is not a goal per se but a component in fault management architectures. It takes part in the solutions produced for tasks such as failure-mode-and-effects analysis, sensor placement, on-board recovery, condition monitoring, maintenance, repair and therapy planning, and prognosis. The contribution of diagnosis in such architectures requires close links with decision tasks such as control and planning and calls for innovative integrations.

In this paper, different facets of diagnosis investigated in the Control or the AI fields are discussed. While Venkatasubramanian et al. (2003a, 2003b, 2003c) provide three interesting surveys of the different approaches that exist in these fields, this paper aims at reporting the works that integrate approaches of both sides, hence creating "bridges". In particular, the concepts and results of the FDI and DX tracks are put in correspondence and the lessons learned from this comparative analysis are pointed out. Causal model-based diagnosis is presented as a typical example of integration of FDI and DX theories, in which fault detection is implemented along an FDI approach and fault isolation calls for the logical DX framework. Hybrid model-based diagnosis is then used to illustrate several interesting bridges, in particular, how FDI estimation schemes can be combined with search algorithms rooted in AI to achieve hybrid state tracking efficiently. Combining FDI estimation filters with the logical DX theory is also illustrated. Finally, it is shown that the hybrid model-based diagnosis problem can also find a solution by combining the FDI approach with the so-called diagnoser approach of the discrete event systems (DES) field. Subsequently, learning the models that support diagnosis reasoning is shown to be a rich field for bridging theories. It has been intensively investigated for continuous model identification for which regression analysis is essential. Learning discrete event models calls for other bridges, which are illustrated with chronicle learning. Finally, diagnosis is discussed in relation with theories that participate to provide global solutions to fault management problems. On one hand, autonomous architectures exemplify the integration of diagnosis, control and planning. On the other hand, it is shown how diagnosis can enhance prognosis in condition maintenance architectures.

The paper is organized as follows. After the Introduction section, Section 2 first presents a brief overview of the approaches proposed by the FDI and DX model-based diagnosis communities. Although quite commonplace, this overview is necessary because it provides the basic concepts and principles that form the foundations of any diagnosis method. It is followed by the comparison of the concepts and techniques used by these communities and the lessons learned from this comparative analysis. Section 3 is concerned with the trends that integrate and take advantage of techniques from both sides, in particular causal model based diagnosis in Section 3.1 and diagnosis of hybrid systems in Section 3.2. Section 4 then raises the problem of obtaining the models supporting diagnosis reasoning and discusses bridges that can contribute to learning them in an automated manner. Section 5 widens the scope of diagnosis and is concerned with diagnosis as a component of fault management architectures, discussing the several links with control and planning. Finally, Section 6 concludes the paper.

2. DX and FDI model-based diagnosis bridge

The FDI and DX streams both consider the diagnosis problem from a *system* point of view, which results in significant overlaps. Even the name of the two tracks is the same: *Model-Based Diagnosis* (MBD).

The diagnosis principles are the same, although each community has developed its own concepts and methods, guided by different modeling paradigms and solvers. FDI relies on analytical models, linear algebra, and non-linear system theory whereas DX takes its foundations in logic. In the 2000s, catalyzed by the "*Bridging AI and Control Engineering model-based diagnosis approaches*" group, known as the BRIDGE (2000) group, within the European Network of Excellence MONET II (MONET), and its French counterpart, the "*Intégration de Méthodes Alliant Automatique et IA*" group, known as the IMALAIA group, supported by GDR MACS (2000), GDR-I3 (2000), as well as AFIA, there were more and more researchers who tried to understand and synergistically integrate methods from the two tracks to propose more efficient diagnostic solutions. This collaboration launched several events:

- a BRIDGE Workshop in 2001 in the framework of DX'01, the *12th International workshop on Principles of Diagnosis*, Sansicario, Via Lattea, Italy, 5–9 Mars 2001 (DX'01, 2001).
- the co-location, in Washington DC (USA), of the two main events of the FDI and the DX community, namely the *IFAC International Symposium on Fault Detection, Supervision and Safety for Technical Processes* SAFEPROCESS'03 and the *International Workshop on Principles of Diagnosis* DX'03, including a BRIDGE Workshop in the form of a join day.

These events were followed by the publication of a special issue of the IEEE SMC Transactions, Part B, on the topic "*Diagnosis of Complex Systems: Bridging the methodologies of the FDI and DX Communities*" in 2004 (Biswas et al., 2004). The *BRIDGE track* was launched and paved its way until today. Other events followed like the two invited sessions "*AI methods for Model-based Diagnosis*" and "*Bridge between Control Theory and AI methods for Model-based Diagnosis*", recently organized in the framework of the 7*th IFAC International Symposium on Fault Detection, Supervision and Safety of Technical Processes* SAFAPROCESS'09, Barcelona, Spain, 30 July–3 August 2009.

The next subsections first summarize the foundations of the FDI and DX approaches, then proceed with a comparative analysis that allows us to draw some practical assessments in the form of lessons learned. The lessons summarize the respective strengths and weaknesses of the two approaches and provide the guidelines that drive the proposals combining the two approaches.

2.1. Brief overview of FDI approaches

The detection and diagnosis methods of the FDI community rely on behavioral models that establish the constraints between system inputs $u \in U$ and outputs $y \in Y$, gathered in the set of measurable variables Z, and the system internal states defining the set of unknown variables X. The variables $z \in Z$ and $x \in X$ are functions of time. The typical model may be formulated in the temporal domain, then known as a *state-space model*:

$$\begin{aligned} BM &: \; dx/dt = f(x(t), u(t), \theta) \\ OM &: \; y(t) = g(x(t), u(t), \theta). \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^{n_x}$ is the state vector, $u(t) \in \mathbb{R}^{n_u}$ is the input vector and $y(t) \in \mathbb{R}^{n_p}$ is the output vector. $\theta \in \mathbb{R}^{n_\theta}$ is a constant parameter vector. The components of f and g are real functions over $\mathbb{R}$. BM is the behavioral model and OM is the observation model. The whole system model is noted $SM(z, x)$, like in Krysander et al. (2008), and assumed noise-free. The equations of $SM(z, x)$ may be associated to components but this information is not represented explicitly. The models can also be formulated in the frequency domain, for instance in the form of transfer functions in the linear case.

Models are used in three families of methods:

- the methods based on *parameter estimation* that focus on the value of parameters as representing physical features of the system,