The Autonomous Systems Laboratory
Universidad Politécnica de Madrid

Intelligence, Control and the Artificial Mind

Bits of history and ongoing future

Ricardo Sanz

ASLab A-2010-004 v 1.0 Final
2010/01/24

Abstract

Artificial intelligence and cognitive science must look at the world of industrial process control to find the technological reifications of the concept of mind.

1 Motivation

Artificial Intelligence (AI) seems to be at an impasse. The old vision of AI that started as the search for the computer-based technology of the artificial mind is not delivering. The excessive initial hype opened the door to ample criticisms after the failure to fulfill some bold predictions. In a sense, *cognitive systems* research has recently replaced AI as the forefront of this research programme. A new name for the same set of objectives just to elude the tagging as failure. But the problem of the AI research programme may not be in the methods but in the naïve conceptualizations that have driven and are still driving this research.

Indeed, AI has not been a failure. Many AI technologies have been demonstrated and are routinely used with enormous success in many domains. From credit card authentication to nozzle design and language understanding. But, beyond the many focused applications of concrete AI technologies, the big objective of artificial intelligence is also an ongoing success. However, the realization of an artificial intelligence is not to be found in the domain of robotics —still in its infancy— but in the uncontroversially materialistic and practical world of industrial process plants. The challenges posed today by these complex technical systems set the proper stage for continuing the pursue of the old dream of AI: the artificial mind. In this context, for example, current research topics include perception, understanding, self and consciousness. Not for human-like robots —something that would be arguably obvious— but for plainly alien systems like refineries or electrical infrastructures.

2 Intelligent Control Memories

Intelligent control (IC) started as a process of technological immersion of AI in the world of control systems. In the case of process control systems Stock (1988); Boullart et al. (1993) the availability of reusable inference engines led to the implementation of expert systems exploiting the knowledge of human operators. In a first approach the implemented systems were only usable as decision support systems for humans but with the development of real-time expert system shells it was possible to use the inference engines to implement closed-loop real-time controllers. At the same time, the developments in fuzzy logic and the associated fuzzy control technology enabled the construction of systems embracing vagueness with better results in control than those obtained with other mechanisms for dealing with uncertainty —like bayesian or necessity-possibility frameworks. The same can be said concerning neural network technology and its intrinsic learning capabilities.

What was expected in IC was a systematic engineering path to the construction of automated operators, exploiting the knowledge from human operators and the deep plant knowledge from process engineers.

From simple fuzzy rule-based systems at the lowest level to complex model-based reasoners at the strategic control level, AI technology has provided very effective mechanisms for improving controller competence in special circumstances (see Figure 1). The many claimed capabilities of the different AI methods were seen to provide major improvements in all the scales of the control hierarchy, and the capability of learning of non-linear action mechanisms —e.g. using neural networks, adaptive fuzzy controllers or genetic algorithms— was one of the key contributions of IC to the field of automatic control Sanz and Galán (1990); Sanz et al. (1991). However, the degree of predictability of the AI based controllers wasn't as good as desired. This obviously limited their use in safety critical applications but also raised justifiable criticisms (e.g. when an expert system demonstrated brittleness —e.g. suffering the so called cliff effect— or a genetic algorithm evolved truly stupid control rules). At the same time, the ad-hoc approach that is used in most cases renders systems that are far from offering the very needed property of *robust autonomy*.

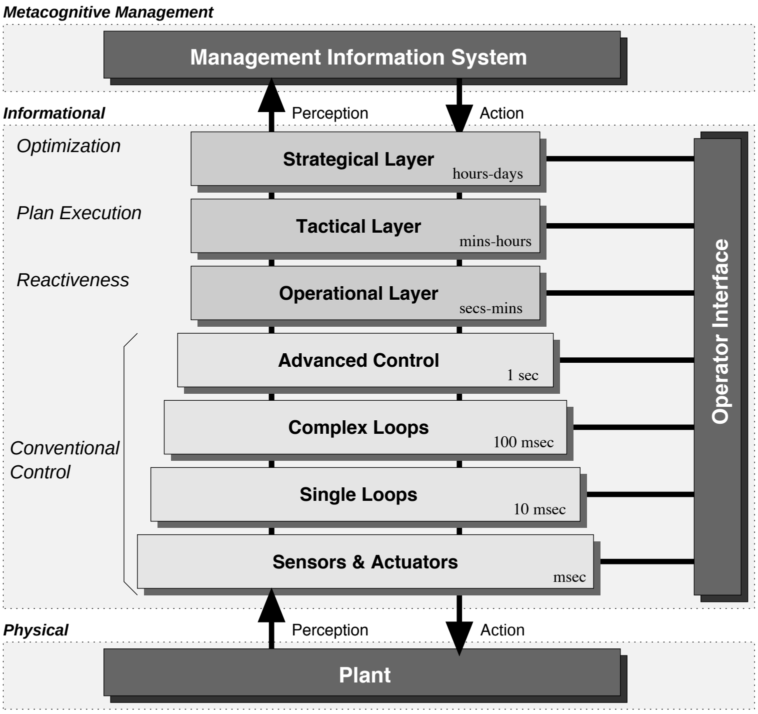

Figure 1: Typical functional layering in a complex industrial process control system.

3 Derailing IC

The discipline of intelligent control quickly become a tool-driven endeavor instead of a problem-driven discipline. The research community grouped around the specific technologies that basically continue to be rule-based systems, artificial neural networks, fuzzy control and evolutionary programming; now classic sub-fields of the *soft computing* worldBonissone (1997); Yen (2006).

However, if we analyze the original motivations, we can see that the control focus on AI was more a natural move than just a search for individual technologies Sanz et al. (1999a,c); Sanz (2000). It was *natural* because the control and AI communities were *originally* in search of the very same objective: the technology of the artificial mind. In the case of AI, this was done as imitation of the human mind; in the case of control, this was done using the methods of physics for any kind of body that was targeted. This common objective was much more clear in the past; so clear that indeed AI and automatic control were just offspring of cybernetics.

4 In Search of the Artificial Mind

Obviously the many approaches of the AI panorama haven't rendered the promised artificial mind as sought Sanz and López (2000); Sanz et al. (2000). Neither has the domain of automatic control so deeply trapped in the limited mathematics of linear systems. The clearest example is perhaps the humanoid robotics field; where body dynamical control is achieving high levels of performance in bipedal walking, cognitive architecture is still very far from offering the minimal glimpse of a real human mindSanz et al. (1999b). The pursue of the complete human-like mind was never an objective in the field of intelligent control. Only some atomic human capabilities were sought to improve localized control systems performance. The many successes of AI in control notwithstanding, at the very heart we still feel the lack of some technical capabilities to engineering some critical human competence in handling abnormal situations Sanz et al. (2000). No matter the apparently different objectives, in both cases —humanoid robots and intelligent controllers— we feel the need of going beyond what we are able to do today and search the seemingly missing essence of mind.

This search for the very essence of mind has been indeed a major pursuit in different fields —philosophy, neuroscience, psychology, robotics, etc.— that have converged into a single programmatic discipline: cognitive science.

This is a widely heterogeneous community due to the many different backgrounds, research practices and personal research sub-objectives. However, the emergence of a unified theory of mind is perceivable in the convergence of the different theoretical models coming from the different domains. This unified vision is so powerful that is providing a way for trying to formalize old age conundrums as perception, knowledge, thought or even consciousness.

The intelligent control community tried to mimic concrete human thought processes in search for competence. The fragility of the realized systems claimed however for a new foundation that was not going to be found in the so called new AI or in postmodern robotics. Cognitive science, on the other side, is lost in to the labyrinth and micro details of the human mind and brain.

Our research has lead us to the conclusion that the only viable strategy to eliminate brittleness and increase mission-level resilience is to make systems epistemologically robust at the mission level Gómez et al. (2007), so we can move the responsibility for real-time cognitive behavior from us engineers to the systems themselves during runtime. And to do this we need what many think is the ultimate human trait: self-consciousnessSanz et al. (2007); Meystel and Sanz (2002). This is what we are trying to do with the development of the SOUL cognitive architecture for robust autonomy. At the end, we expect these conceptually rigorous artificial minds to be theoretical cornerstones of a new science of mind.

One of the critical elements in this approach is the *epistemic control loop* (see Fig-

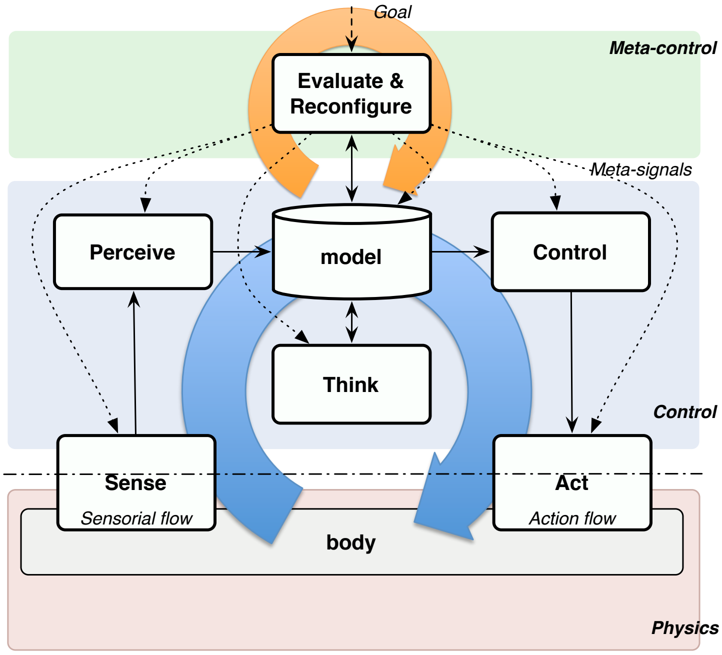

Figure 2: General atomic structure of the epistemic control loop pattern to be pervasively implemented across the control hierarchy.

ure 2) the basic design pattern for embedding intelligence pervasively into the system. A highly robust autonomous system will not only realize a hierarchical federation of cognitive control loops but also a transversal metacognitive competence that will render the necessary self-awareness for achieving full autonomy. The artificial mind is coming.

5 Acknowledgements

The author would like to acknowledge the support coming from the European Community *Framework Programmes* (grant agreements HINT, DIXIT, ICEA and HUMANOBS) and the Spanish *Plan Nacional de I+D* (grant agreements MITA, M3 and C3).

References

Bonissone, P. (1997). Soft computing: the convergence of emerging reasoning technologies. *Journal of Research in Soft Computing*, 1(1).

Boullart, L., Krijgsman, A., and Vingerhoeds, R. A., editors (1993). *Application of Artificial Intelligence in Process Control*. Pergamon Press.

Gómez, J., Galán, R., Sanz, R., and López, I. (2007). Naturalized epistemology for autonomous systems. In *Proceedings KNEW 2007*.

Meystel, A. and Sanz, R. (2002). Self-identity of complex control systems. In *Proceedings of the Workshop on Theoretical Fundamentals of Intelligent Systems*, Durham (NC), USA.

Sanz, R. (2000). Agents for complex control systems. In Samad, T. and Weyrauch, J., editors, *Automation, Control, and Complexity: New Developments and Directions*, chapter 10, pages 171–190. John Wiley and Sons, Chichester, UK.

Sanz, R., Alarcón, I., Segarra, M. J., de Antonio, A., and Clavijo, J. A. (1999a). Progressive domain focalization in intelligent control systems. *Control Engineering Practice*, 7(5):665–671.

Sanz, R. and Galán, R. (1990). A cooperative architecture for intelligent process control. In *Proc. of the International Working Conference on Cooperating Knowledge Based Systems, CKBS 2000*, pages 58–59, Keele, England.

Sanz, R., Jimenez, A., Galan, R., Matia, F., and Puente, E. A. (1991). Intelligent process control: the CONEX architecture. In Tzafestas, S. G., editor, *Engineering systems with intelligence: concepts, tools and applications*, pages 407–413. Kluwer Academic Publishers, Norwell, MA, USA.

Sanz, R. and López, I. (2000). Minds, MIPS and structural feedback. In *Performance Metrics for Intelligent Systems, PerMIS '2000*, Gaithersburg, USA.

Sanz, R., López, I., Rodríguez, M., and Hernández, C. (2007). Principles for consciousness in integrated cognitive control. *Neural Networks*, 20(9):938–946.

Sanz, R., Matía, F., and Galán, S. (2000). Fridges, elephants and the meaning of autonomy and intelligence. In *IEEE International Symposium on Intelligent Control, ISIC'2000*, Patras, Greece.

Sanz, R., Matía, F., and Puente, E. A. (1999b). The ICa approach to intelligent autonomous systems. In Tzafestas, S., editor, *Advances in Autonomous Intelligent Systems*, Microprocessor-Based and Intelligent Systems Engineering, chapter 4, pages 71–92. Kluwer Academic Publishers, Dordretch, NL.

Sanz, R., Segarra, M. J., de Antonio, A., Matía, F., Jiménez, A., and Galán, R. (1999c). Design patterns in intelligent control systems. In *Proceedings of IFAC 14th World Congress*, Beijing, China.

Stock, M. J. (1988). *AI in Process Control*. McGraw-Hill, Inc., New York, NY, USA.

Yen, G. G. (2006). Editor's remarks - a new page unfolded: Ieee computational intelligence magazine. *IEEE Computational Intelligence Magazine*, 1(1):2–2.

Ricardo Sanz —electrical engineer, PhD in AI And robotics— is the coordinator of the *Autonomous Systems Laboratory* (ASLab), a multidisciplinary research group focused in the development of theoretical models and software technology for cross-domain implementation of autonomous systems. Find more about Ricardo Sanz at `www.rsanz.com` and about ASLab at `www.aslab.org`.

This brief article has been prepared for the PerAda Magazine.

Title: Intelligence, Control and the Artificial Mind
Subtitle: Bits of history and ongoing future
Author: Ricardo Sanz
Date: 2010/01/24
Reference: A-2010-004 v 1.0 Final
URL: http://www.aslab.org/documents/controlled/ASLAB-A-2010-004.pdf

Autonomous Systems Laboratory

Universidad Politécnica de Madrid
c/José Gutiérrez Abascal, 2
Madrid 28006 (Spain)