ELSEVIER

Artificial Intelligence 101 (1998) 337–339

Artificial Intelligence

Commentary

Schema-based learning

Fernando J. Corbacho [1]

Departamento de Ingeniería Informiática, Universidad Autónoma de Madrid, 28049 Madrid, Spain

Received 26 January 1998; received in revised form 2 March 1998

Schema-based learning (SBL) builds on the schema theory of *The Metaphorical Brain* 2 (TMB2) in a number of ways. In SBL we incorporate not only schemas which provide units of interaction of the organism with its environment, but also predictive schemas reflecting expectations about those interactions. These internal models exist at all levels of granularity forming a hierarchical network of schemas. Schemas as generic specifications go beyond any specific implementation (neural networks in particular). Thus, the definition of schemas must be general enough so that it applies to a whole variety of possible implementations. There have been several attempts at formalizing schemas. Lyons and Arbib [7] formalized schemas as port automata and Corbacho and Arbib [4] extended this formalization by including activity variables for each port automaton along with the corresponding dynamics. Since schemas must be general by definition we claim the "power" comes from the operations at this schema level. For instance, SBL includes schema *instantiation*, schema *assertion* by distributed competitive and cooperative dynamics over the schemas activity variables, schema *tuning* and schema *construction* (where predictive schemas play a very important role in deciding when to construct a new schema).

Arbib asserts that "schema theory attempts to bridge between structure and function at the highest level". A schema may refer to either a set of regularities in observed behavior or the internal mechanisms which give rise to such regularities. In analyzing a living system we mostly refer to the first. In designing (synthesizing) a system we must provide the second, i.e., provide internal mechanisms. The ultimate challenge is to match both. Hanson argues that "one might posit different types of schemas, ones perhaps that vary in abstractness or perhaps have some domain or topic dependency ...", and indeed SBL introduces several types of schemas: perceptual, motor, sensorimotor, goal and predictive schemas. Current extensions on progress also include autonomous construction of incrementally abstract schemas grounded on more primitive sensorimotor schemas, much in the way Piaget [9] described the appearance of more abstract schemas along the

[1] Email: corbacho@ii.uam.es.

PII: S0004-3702(98)00029-0

developmental stages in children. These different types of schemas also exist at all levels of granularity. Moreover, where Hanson argues that "what seems to be avoided here are serious issues of Generalization ..." we respond that a schema by definition represents what is stable and generalizable over variability—becoming precise through parameterization and adaptation. The "bottom" or "inner" schema generalization mechanisms depend on the particular implementation of the schema (e.g., artificial neural networks, fuzzy logic, and so on) and hence do not belong at the schema level. This also reflects the fact that there can be generalization at several levels.

Where much connectionism focuses on single networks being trained by a single learning rule, SBL stresses (i) composition besides decomposition, and (ii) integration of learning paradigms/aspects.

(i) For instance, Jordan and Jacobs [6] assert that "The ME architecture solves complex function approximation problems by allocating different modules to different regions of input space. ... If we now inquire about the internal structure of a module, however, we see that the same argument can be repeated. Perhaps it is better to split a module into simpler submodules ...". Hence, Jordan and Jacobs emphasize decomposition ... and do not provide any dynamics to autonomously combine (in run time) structures to form new ones. This is contrary to SBL which specifically emphasizes composition. That is, incremental construction of new schemas based on the current stock of schemas. In this regard much work in connectionism has ignored both built-in structure and principles of organization. After all, much learning happens in already functioning systems. Learning from scratch is not adequate for learning in large, structured domains.

(ii) SBL does not reject other learning methods but actually integrates different aspects of several of them. For instance, reinforcement learning (Barto [2]) is employed in the learning of goal schemas, error learning is employed in schema tuning, and unsupervised learning (Willshaw and von der Malsburg [12]; von der Malsburg [11]) is used to construct topological mappings among previously unrelated spaces. Nevertheless, SBL goes beyond any of these single paradigms. For instance schema tuning in SBL would be the analogous to error learning in a neural network (e.g., Rumelhart et al. [10]), whereas SBL schema construction goes beyond that by autonomous construction of new schemas and involves higher level processes than local synaptic plasticity in neural networks. Corbacho et al. [5] provide an example of schema construction in a lesion study (in the motor system of frogs) where a new schema has to be constructed to efficiently restore a pattern of interaction disrupted by the lesion. In this particular case no simple adaptation rule would have solved the problem unless part of the specific solution had initially been coded implicitly by the designer.

We agree with Hanson that "the subtle interplay of perception and memory [is] a key element of schema acquisition and consequent usage", yet Hanson fails to include the action component to close the action-perception cycle as well as the resultant feedback signals which result from interacting with the environment. In this respect SBL performs *Coherence Maximization* by which predictive schemas can guide learning (as well as acting) by structuring the credit assignment space. No matter how schemas are implemented (neural nets are just one example—i.e., schemas are beyond neural nets), fundamental to SBL is a set of Principles of Organization. Among them we have:

- Coherence maximization: maximize the congruence between the result of an interaction and the expectations for that interaction. This includes the learning of *cause-effect* relations for previously unrelated events.
- Performance maximization: related to the achievement of goals which in turn are related to the reduction of internal drives.
- Topology: "closer" elements tend to be more related (this is important to bias the construction of a particular set of relations within the (usually) large combinatorial relational space).

The schemas' output elements are defined as patterns as they usually have internal structure. The ability to compare patterns is fundamental, for instance to check whether a pattern is getting closer or further from a "goal" pattern (learned in a previously successful interaction). Hence the notion of distance is fundamental and hence we use metric spaces for the schema's output spaces.

Due to space limitations we have only listed three of the main SBL principles of organization. Several authors have listed sets of Principles of Organization for the design of intelligent systems (Arbib and Liaw [1]; Pfeifer [8]; Corbacho and Arbib [4]). SBL integrates many of these inter-related principles of organization in an unifying way.

References

[1] M.A. Arbib and J.-S. Liaw, Sensorimotor transformations in the worlds of frogs and robots, Artificial Intelligence 72 (1995) 53–79.

[2] A.G. Barto, Reinforcement learning, in: M.A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, A Bradford Book/The MIT Press, Cambridge, MA, 1995, pp. 804–809.

[3] R.A.Brooks, A robust layered control system for a mobile robot, IEEE J. Robotics and Automation RA-2 (1986) 14–23.

[4] F. Corbacho and M.A. Arbib, Schema-based learning: towards a theory of organization for adaptive autonomous agents, in: Proceedings First International Conference on Autonomous Agents, Marina del Rey, CA, 1997.

[5] F. Corbacho, K.C. Nishikawa, A. Weerasuriya, J.-S. Liaw and M.A. Arbib, A schema-based learning model of adaptable and flexible prey-catching II: the adaptable architecture, in preparation.

[6] M.I. Jordan and R.A. Jacobs, Modular and hierarchical learning systems, in: M.A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, A Bradford Book/The MIT Press, Cambridge, MA, 1995, pp. 579–582.

[7] D.M. Lyons and M.A. Arbib, A formal model of distributed computation for schema-based robot control, IEEE J. Robotics and Automation 5 (1989) 280–293.

[8] R. Pfeifer, Building "fungus eaters": design principles of autonomous agents, in: Proceedings Fourth International Conference on Simulation of Adaptive Behavior, Cape Cod, MA, 1996.

[9] J. Piaget, The Construction of Reality in the Child, Ballantine, New York, 1954.

[10] D.E. Rumelhart, G.E. Hinton and R.J. Williams, Learning internal representations by error propagation, in: D.E. Rumelhart, J. McClelland, and the PDP Research Group (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1, A Bradford Book/The MIT Press, Cambridge, MA, 1986, pp. 318–362.

[11] C. von der Malsburg, Network self-organization, in: S.F. Zornetzer, J.L. Davis and C. Lau (Eds.), An Introduction to Neural and Electronic Networks, Academic Press, San Diego, CA, 1990, pp. 421–432.

[12] D.J. Willshaw and D. von der Malsburg, How patterned neural connections can be set up by self-organization, Proc. Royal Soc. London, Ser. B 194 (1976) 431–445.