AAAI Technical Report SS-12-02
Designing Intelligent Robots: Reintegrating AI

Evolutionary Language Games as a Paradigm for Integrated AI Research

Luc Steels[1,2]
[1] ICREA-Institute for Evolutionary Biology (UPF-CSIC) Barcelona and [2] Sony Computer Science Lab - Paris

Abstract

Evolutionary language games are a way to study how perceptions, concepts, and language can emerge in populations of situated embodied agents, driven by the needs of communication and the properties of the environment. Evolutionary language games are currently being investigated using physical robots and this then requires that the full cycle of processing activities from physical robotic embodiment to sensory-motor processing, visual perception and action, conceptualization, and language processing are all integrated in a single system. This contribution reports on a large-scale long term effort to experiment with evolutionary language games and discusses major results achieved so far.

Introduction

Over the past decade, my research group has been involved in large-scale efforts towards deeply integrated AI that involve consideration of the full suite of processes needed for flexible open-ended language interaction on autonomous robots (Steels and Hild 2012). The notion of an evolutionary language game acts as the glue that helps us tie many aspects of intelligence together including multi-agent aspects. A language game is a situated embodied interaction between, at least, two autonomous robots drawn from a population (see Figure 1). It involves a shared goal, for example the speaker wants to draw the attention of the hearer to an object in the world. Both have to perceive and conceptualize their world, the speaker has to plan what to say and transform it into language. The hearer has to parse, reconstruct the meaning and enact it. This in itself poses already enormous challenges for integrated AI.

But our research group wants to go one step further. We seek to understand how the enormously complex systems needed at all these levels can emerge. Concretely, we do not provide the needed feature extraction and pattern recognition primitives, preprogram the ontologies needed for conceptualization, nor supply the lexicons and grammars for parsing and production. We do not even provide the agents with corpus data so that they can learn from existing human verbal interactions. Instead the challenge is to provide the necessary machinery for learning AND invention, so that all these structures emerge under the pressure of being successful in the language game. This clearly ups the ante, particularly because the self-organization of a communication systems is inherently a multi-agent problem. The adoption by one agent of a new concept and a new word or grammatical construction to express it can only be effective if it is also adopted by other members of the population.

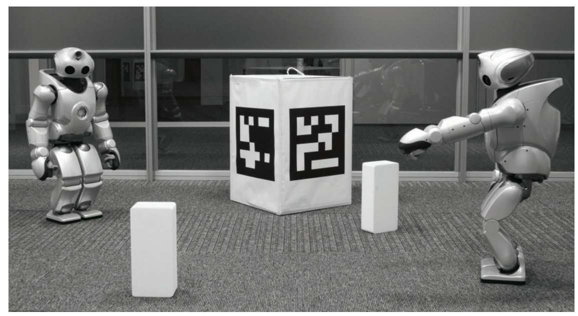

Figure 1: Experimental setup for a language game played by two humanoid robots. This setup was designed to do experiments in the emergence of spatial language. The speaker has to draw attention to one of the objects in the environment by describing it and the hearer signals understanding by point to the object.

Our research program has initially focused on very basic language functions, such as naming objects with proper names or color terms, and naming bodily actions. It then gradually extended to more complex domains, such as naming of spatial relations and using grammar to express landmarks and perspectives, and handling temporal language including tense, aspect, and modality. We now have reached a level where we can attack questions on how grammatical functions, constituent structure and case grammar can emerge, always grounded in the sensory-motor interactions of robots. A representative set of these experiments is discussed in (Steels, 2012a).

Conceptualization

The big issue for integrated AI is to bridge the gap between the continuous dynamics of sensory-motor intelligence and the discrete levels of symbolic reasoning and language. We argue that there is no simple quick fix that will bridge this gap. Instead, we propose a sophisticated layer, called the conceptualization layer, that will achieve this task. This layer contains semantic structures as shown in Figure 2 (from (Gerasymova and Spranger 2012)). The nodes in these structures are cognitive operations that operate either over discrete structures (for example sets of objects) or over continuous data from the sensory-motor layer. They are implemented as constraints that try to fill in or constrain as much as possible their slots represented as variables (written as names with question marks). The cognitive operations are connected into a network which is executed using data flow activation as in constraint propagation systems. We have implemented the necessary mechanisms so that agents can come up with a network to satisfy a particular communicative goal in the context, to chunk networks and store them for later use, and to match networks against each other so that partial networks can be expanded with chunks already in memory.

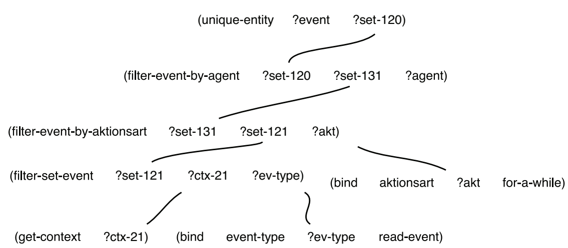

Figure 2: Example of a semantic structure for the sentence "Who has been reading for a while?" One operation is `filter-set-event` which will filter the recognized events in the context in order to find the ones that match with a particular event-type. Another operation is `bind` which anchors a variable to a constant, e.g. bind the `?event-type` to a `read-event`.

Other Integration Principles

We use of course many additional principles to achieve integration. One of them is **reversability**. Any process in the complete chain from sensing to language can work in both directions. For example, all grammar rules are written in such a way that they can be used both for parsing (going from form features to semantics) and for producing (going from semantics back to form) and we have developed a new formalism called Fluid Construction Grammar to achieve this (Steels 2012b). Similarly, the semantic component can either come up with meanings that would achieve a communicative goal, or conversely interpret meanings derived by the parser using the world model derived from sensing. The sensory-motor system either works in a bottom-up way analyzing signals to form internal world models, or in a top-down way predicting what the world is going to look like based on expectations. This reservability has numerous advantages: (1) It aids to cope with noise, errorful input, incomplete fragments, etc. because if one layer of analysis is failing, another can make up for it. (2) It aids in learning because lacking knowledge in one component (e.g. concepts or grammatical rules) can be acquired based on strong predictions coming from other layers (e.g. syntactic or semantic bootstrapping).

Another principle, well known from earlier AI research is **reflective computation**. Agents are operating on two levels. There is a routine level that is able to deal with routine cases, often by pulling out of memory ready-made solutions. There is also a meta-layer that is constantly running diagnostics, for example to detect unknown words, unexpressable meanings, ungrammatical constructions, wrong reactions to a phrase, etc. When certain problems are detected, repairs come into action that try to handle the problem, for example, perform additional sensory-motor activities, trigger additional inference to guess meaning, expand an existing word to deal with a new case, etc. (Beuls, Wellens, and van Trijp 2012).

Conclusion

Evolutionary language games have proven an enormous challenge and fruitful framework to push integrated AI, not only routine perceptual, conceptual and linguistic processing but also learning and multi-agent coordination. In the presentation I will give an overview of some of the experiments that we have already carried out, with concrete examples of evolutionary language games in the domain of spatial language, descriptions of events including participant roles, temporal language, and action language.

Acknowledgement Research activities described here have mainly been carried out at the AI laboratory of the University of Brussels, the Sony Computer Science Laboratory in Paris, and the Neurorobotics Research Laboratory of Humboldt University in Berlin. Research was partly funded by the European FP7 projects ECAgents and ALEAR.

References

Beuls, K., P. Wellens and R. van Trijp 2012. Diagnostics and Repairs in FCG. In: Steels, L. (2012c) (ed) Computational Issues in Fluid Construction Grammar. Springer Verlag, Berlin.

Steels, L. ed. 2012a. Experiments in Cultural Language Evolution. John Benjamins Pub. Co., Amsterdam.

Steels, L., ed. 2012b. Design Patterns in Fluid Construction Grammar. John Benjamins Pub. Co., Amsterdam.

Steels, L. and M. Hild, ed. 2012. Language Grounding in Robots. Springer Verlag, Berlin.

Gerasymova, K. and M. Spranger 2012. Learning Temporal Language: A case study for Russian aspect. In: Steels, L. and M. Hild (2012) (eds) Grounding Language in Robots Springer-Verlag, Berlin.