Time, Language and Action - A Unified Long-Term Memory Model for Sensory-Motor Chains and Word Schemata

by Fabian Chersi, Marcello Ferro, Giovanni Pezzulo and Vito Pirrelli

Action and language are known to be organized as closely-related brain subsystems. An Italian CNR project implemented a computational neural model where the ability to form chains of goal-directed actions and chains of linguistic units relies on a unified memory architecture obeying the same organizing principles.

Recent advances in cognitive psychology and neuroscience emphasize that action and language are not organized as insulated brain subsystems. Rather, language processing elicits perceptual and motor processes/responses that are tightly coupled with the referents of what is heard or read. Rizzolatti and Arbib (1998) proposed that linguistic abilities developed phylogenetically on top of action control abilities, on the basis of a common brain substrate where the mirror neuron system plays a key role. Accordingly, area F5 of the monkey brain (where mirror neurons are located) is a precursor of human Broca's area (devoted to language processing), and language could have inherited the "grammatical" and combinatorial structure of actions. An Italian CNR project is currently exploring the related hypothesis that the ability to form chains of goal-directed actions (ie, action sequences leading to a distal result) and chains of linguistic units (eg, sequences of phonemes, morphemes or words forming a sentence) may rely on the same neural architecture obeying a common pool of organizing principles.

Motor chains and lexical chains

Fogassi and colleagues (2005) have shown that motor and mirror neurons in the monkey inferior parietal lobule code single motor acts (eg "reaching" or "grasping") belonging to an action sequence and that their discharge reflects the intended goal of the whole action (eg "grasping to eat" versus "grasping to place"). On this empirical basis, it has been hypothesized (Fogassi et al. 2005; Chersi et al. 2005) that this brain area contains highly ordered neural structures, where each goal-directed action sequence is represented by a separate chain of pools of neurons. Elements in one chain are not interchangeable with elements of other chains even if they code the same motor

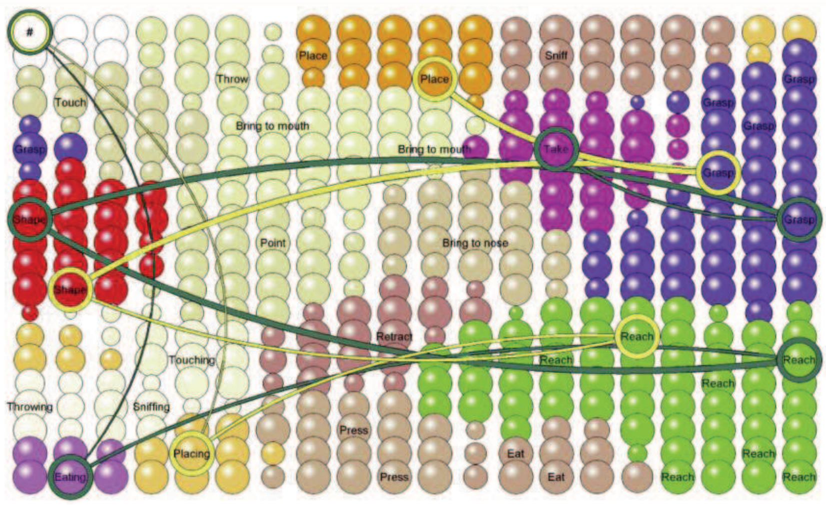

Figure 1: A 400 node T²HSOM trained on sequences of goal-action patterns. Topological clustering is highlighted by colour shades, showing specialization of different areas for different goals (eg "Eating", "Placing", "Throwing" etc., bottom left corner) and actions (eg "Reach", "Shape", "Bring to mouth" etc.). The temporal response of the map for two input patterns (<#, Eating, Reach, Shape, Grasp, Take> and <#, Placing, Reach, Shape, Grasp, Place>) is shown through circles (highlighting BMUs) and clockwise oriented arcs, representing temporal transitions between consecutively-activated BMUs, ie chains of goal-directed actions. The "#" symbol is the "start of sequence" marker.

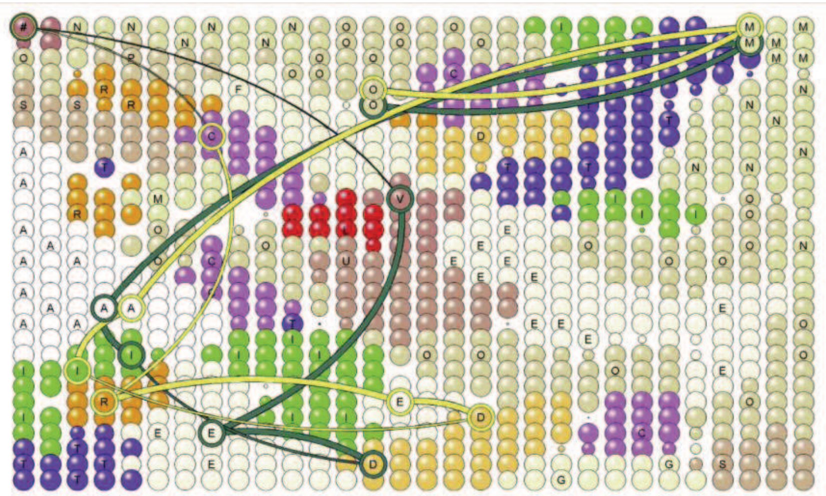

Figure 2: A 900 node T²HSOM trained on Italian verb forms. BMU chains are shown for "vediamo" ("we see") and "crediamo" ("we believe"). Although the two forms share the common ending "-iamo", the fact that the roots "cred-" and "ved-" are different produces the activation of BMU chains independently running through the map at a short topological distance (green and yellow trajectories). The "#" symbol is the "start of sequence" marker. The result shows node sensitivity to morphological structure.

act. Execution and recognition of an action is achieved through the activation of the appropriate chain and thus to the pre-selection of specific neurons.

Results from joint behavioural and functional neuro-imaging studies on the mental lexicon demonstrate the existence of a whole-word level of brain coding (Baayen 2007). Word forms are stored in full, organized into hierarchically-structured chains of sub-lexical units (eg letters or phonological segments), where units in one lexical chain are coded differently from the same units in another lexical chain. Whole-word memory structures account for i) development of dedicated chains of linguistic units, enhancing predictive linguistic behaviour (Ferro et al. 2010); ii) frequency-based competition between inflected forms of a word (eg “bring” and “bringing”) (Pirrelli et al., in press); iii) simultaneous activation of false morphological friends (eg “broth” and “brother”).

The analogy between action and word memory structures persuaded us to investigate the hypothesis that they can both be served by the same memory mechanisms for serial order, modelled as Topological Temporal Hebbian Self-Organizing Maps (T^2HSOMs, Ferro et al. 2010). T^2HSOMs are time-sensitive SOMs (Kohonen 2002, Koutnik 2007) whose nodes are fully connected through an add-on weighted temporal Hebbian layer. Upon presentation of a stimulus, all map nodes are activated synchronously, with the most highly-activated node (or Best Matching Unit, BMU) winning the competition. Through training, nodes are made more sensitive to particular classes of stimuli occurring in specific spatio-temporal contexts, with inter-node Hebbian connections being attuned to transition probabilities between temporally adjacent stimuli, thus affording predictive processing.

Results and future developments

Figures 1 and 2 illustrate chains of BMUs in two T^2HSOMs activated by action chains and word forms respectively. The effect is achieved with a “predictive drive”, making the network maximize prediction accuracy in perception, and effortless memory access of order information in production (note that the same network supports both perception and production). As a result, highly-ordered neural structures emerge as a response to repeated action patterns and word schemata.

Besides unravelling some fundamental mechanisms underlying the processing of time-ordered series, the model shows that apparently unrelated evidence on the neural coding of motor chains and word schemata is accounted for by the dynamic interaction of common principles of topological self-organization and time-bound prediction. This dynamic is key to modelling pervasive aspects of synchronization of multimodal sequences in both linguistic (eg reading) and extra-linguistic (eg visuomotor coordination) tasks.

Link:
http://www.ilc.cnr.it/dylanlab

Please contact:
Marcello Ferro
ILC-CNR, Italy
E-mail: marcello.ferro@cnr.it

Challenges for the Design of Intelligent and Multimodal Cognitive Systems

by José Rouillard, Jean-Claude Tarby, Xavier Le Pallec and Raphaël Marvie

With the MINY (Multimodality Is Nice for You!) project, our goal is to propose some novel possibilities to take many modalities of interaction into account. Using a model-driven engineering approach we present some suggestions in order to tackle the challenges around the design of intelligent and multimodal cognitive systems.

Technology's evolution is an unstoppable process. Consider the regular release of new devices such as smart-phones or the multi-touch tabletop: each new version is more powerful and more interconnected than the previous one. Home automation is an example of improved communication: Washing machines can run Android and be remotely driven by a smartphone or computer. While such interaction is easy to implement, most of these systems offer a single modality of interaction: the Wii only support movement based interactions. Games are at the more complex end of the scale, relying on two modalities such as Mouse / Keyboard and voice.

Multimodality is the ability to combine different modalities of interaction (voice, gesture, touch, etc) as input and/or output, such as the historical "Put that there" from Bolt in 1980. Our goal is the design and implementation of intelligent multimodal systems. By intelligent, we mean the ability to make decisions, to request additional information from outside (eg the user or other applications), and to learn (from mistakes, from the user actions, etc).

Our approach is top-down and is very concerned about the heterogeneity of the material covered. It begins with the specification of tasks that the system can achieve. Then we choose the best suited materials to enable the realization of these tasks. This allows code generation supporting the interaction with the system and associated materials (eg X10 home automation, but also smartphones, webcams etc).

To support this top-down approach, we use a model-driven engineering approach (MDE). A first reason for this choice is that we are designing applications for various domains (such as home automation, botany, tourism). While the modalities are always the same, their implementations change from one application to another.