Where Does Compositionality Come From?

Mark Steedman
School of Informatics
University of Edinburgh
2 Buccleuch Place
Edinburgh EH8 9LW, United Kingdom

Abstract

This paper builds on the insight of Lashley (1951) and Miller, Galanter, & Pribram (1960) that action and motor planning mechanisms provide a basis for all serially ordered compositional systems, including language and reasoning. It reinterprets this observation in terms of modern AI formalisms for planning, showing that both the syntactic apparatus that projects lexical meanings onto sentences and the neural mechanisms that are known to be implicated in both language behavior and motor planning reflect exactly the same primitive combinatory operations. The paper then considers some neurocomputational mechanisms that can be applied to modeling this system, and the relation of the compositionality property to such mechanisms.

Introduction

Compositionality of the kind universally assumed in natural language behavior demands a number of properties of representations that have been difficult to reeconcile with neural-computational mechanisms. One is that composition inherently requires recursion. Another is the type-token distinction. Yet another is abstraction. The paper argues that the origins of such properties should be sought in pre-linguistic motor planning mechansisms, and that the two main mechanisms of neural network learning—multi-layer perceptrons and associative nets—support the basic building blocks of compositional systems.

Planning with LDEC

Basic Dynamic Logic

Dynamic Logic (Harel 1984) offers a perspicuous notation for reasoning about dynamic systems such as computer programs:

(1) $n \geq 0 \Rightarrow [\alpha](y = F(n))$

—in a state where n is greater than or equal to 0, running α necessarily results in a state where y is set to a value $F(n)$

(2) $n \geq 0 \Rightarrow \langle\alpha\rangle(y = F(n))$

—running α in a state where n is greater than or equal to 0 possibly results in a state where y is set to a value $F(n)$

Such logics typically include the following dynamic axiom, in which the operator ; is *sequence*, an operation related to functional composition of functions of type *situation* → *situation*:

(3) $[\alpha][\beta]P \Rightarrow [\alpha;\beta]P$

Composition is one of the most primitive *combinators*, or operations combining functions, which Curry & Feys (1958) call **B**, writing the above sequence $\alpha;\beta$ as **B**$\alpha\beta$, where

(4) $\mathbf{B}\alpha\beta \equiv \lambda s.\alpha(\beta(s))$

The Linear Dynamic Event Calculus

To adapt dynamic logic for purposes of planning, in order to avoid the representational and inferential aspects of the frame problem, we need to distinguish linear logical implication ($\multimap$ Girard 1987) from standard intuitionistic implications ($\Rightarrow$). Thus to express such (oversimplified) facts as that if a door is shut, and you push it, you get to a state where it is open, or that if you are inside, and you go through, you are outside, we write:

(5) a. $shut(x) \multimap [push(x)]open(x)$
    b. $in \multimap [go\text{-}through(x)]out$

This rule makes the fluents $shut(x)$ and $in(y)$ resources that are consumed by inference, in the sense that they necessarily no longer hold in the state that results, although any other fluents that hold in the original state but are not involved in the linear inference continue to hold.

By contrast, if we want to express the (again, oversimplified) fact that if something is a door, it is possible to push it, or if something is a door and its open, then its possible to go through it, we write:

(6) a. $door(x) \Rightarrow possible(push(x))$
    b. $door(x) \wedge open(x)$
        $\Rightarrow possible(go\text{-}through(x))$

The predicate *possible* is not the same notion of possibility as the dynamic logical modality $\langle\alpha\rangle$. $possible(\alpha)$ means that the preconditions of α are fulfilled. In fact our rules make no use of $\langle\alpha\rangle$. For planning purposes, all actions are assumed to be deterministic. Of course, the world is not like that but we must deal with nondeterminism in a different way.

The Linear Dynamic Event Calculus (LDEC) made up of deterministic linear rules like the above expressing the state-

changing effects of instantaneous[1] actions, and intuitionistic rules expressing the preconditions on those actions taking place makes a very transparent notation for STRIPS-like knowledge for constructing plans, via the following axiom expressing the transitivity property of the *possible* relation:

(7) $\models possible(\alpha) \wedge [\alpha]possible(\beta) \Rightarrow possible(\alpha;\beta)$

This says that if you are in a state where its possible to α and its a state in which actually α-ing gets you to a state where its possible to β, you are in a state where it is possible to α then β.

It supports a simple planner in which starting from the world (8) in which I am *in*, and the door is *shut* and stating the goal (9) meaning "find a possible series of actions that will get me *out*," can be made to automatically deliver a constructive proof that one such plan is (10):

(8) $in \wedge door(d) \wedge shut(d)$

(9) $possible(\alpha) \wedge [\alpha]out$

(10) $\alpha = push(d); go\text{-}through(d).$

The situation that results from executing this plan in the start situation (8) is one in which the following conjunction of facts is directly represented by the database:

(11) $out(me) \wedge door(d) \wedge open(d)$

This calculus is explored in Steedman (2002b).

Formalizing Affordance

Such a calculus offers a simple way to formalize the notion of "affordance". The affordances of objects can be directly defined in terms of STRIPS/LDEC preconditions like (6).

The affordances of doors in our running example are pushing and going through:

(12) $affordances(door) = \left\{ \begin{array}{l} push \\ go\text{-}through \end{array} \right\}$

This is what is needed to support the Reactive, Forward-Chaining, Affordance-Based planning that is characteristic of primates and other animals.

The Gibsonian affordance-based door-schema can then in turn be defined as a function mapping doors into (second-order) functions from their affordances like pushing and going-through to their results:

(13) $door' = \lambda x_{door}.\lambda p_{affordances(door)}.px$

The operation of turning an object of a given type into a function over those functions that apply to objects of that type is another primitive combinator called **T** or *type raising*, so (13) can be rewritten $door' = \lambda x_{door}.\mathbf{T}x$, where

(14) $\mathbf{T}a \equiv \lambda p.p(a)$

Plans and the Structure of Language Behavior

Interestingly, the combinators **B** and **T** also show up as primitives in the theory of syntax proposed as Combinatory Categorial Grammar (CCG, Steedman 2000). According to this theory, all language specific information resides in the lexicon. For example, the fact that the English transitive clause exhibits SVO word-order is captured as follows in the lexical category for transitive verbs, which comprises a syntactic type (to the left of the colon) and a semantic interpretation (to the right):

(15) like $:= (S\backslash NP)/NP : like'$

The following Combinatory Rules combine such functor categories with arguments:[2]

$$\frac{X/Y:f \quad Y:g}{X:f(g)} > \qquad \frac{Y:g \quad X\backslash Y:f}{X:f(g)} <$$

$$\frac{X/Y:f \quad Y/Z:g}{X/Z:\lambda z.f(g(z))} >\mathbf{B} \qquad \frac{Y\backslash Z:g \quad X\backslash Y:f}{X\backslash Z:\lambda z.f(g(z))} <\mathbf{B}$$

$$\frac{X/Y:f \quad Y\backslash Z:g}{X\backslash Z:\lambda z.f(g(z))} >\mathbf{B}_\times \qquad \frac{Y/Z:g \quad X\backslash Y:f}{X/Z:\lambda z.f(g(z))} <\mathbf{B}_\times$$

Crucially, all arguments are type-raised via the lexicon:

$$\frac{X:x}{T/(T\backslash X):\lambda f.f(x)} >\mathbf{T} \qquad \frac{X:x}{T\backslash(T/X):\lambda f.f(x)} <\mathbf{T}$$

—where T and X are variables over syntactic types.

The combination of Type-raising and Composition allows lexical categories to "project" semantic predicate-argument relations onto "unbounded" constructions such as relative clauses and coordinate structures that have appeared to require powerful grammatical Transformational Rules, as in (16) and (17).

These examples exemplify a number of crosslinguistic universal generalizations that the theory captures (Steedman 2000; Baldridge 2002), and it has been successfully applied to wide-coverage statistical parsing using the Penn Wall Street Journal treebank (Hockenmaier & Steedman 2002; Clark & Curran 2004).

Neural and Computational Theories

Quite a lot is known about this system in neurocomputational terms.

The primate cytoarchitectonic homolog of area 44 or Broca's area in humans, F5, has been shown by single cell recording to include "Mirror Neurons" that fire not only to specific goal oriented actions such as reaching and grasping, but also (with exquisite specificity) to the sight of another animal performing the same goal-oriented action (Rizzolatti, Fogassi, & Gallese 2002). If the animal knows that the goal is not contextually valid, or if the other animals gaze is not consistent, the mere sight of appropriate motion is not enough to fire the mirror neuron.

Interestingly, other neurons in F5 fire only to the animals own actions, and/or fire to visual presentation of the object involved (Rizzolatti, Fogassi, & Gallese 2001; Miall 2003).

This system has usually been interpreted in terms of recognition, understanding, and imitation of the actions of other animals (Gallese *et al.* 1996). However, it seems likely that

---

[1]Durative or interval events are represented by the instantaneous state changes that initiate and terminate them.

[2]Combination of categories by rules is further restricted by a system of features distinguishing slash types that is omitted here—see Baldridge (2002).

(16)

$$
\begin{array}{cccccccc}
\text{a} & \text{man} & \text{who} & \text{I} & \text{think} & \text{you} & \text{like} & \text{arrived}\\
\hline
(T_1/(T_1\backslash NP))/N & N & (N\backslash N)/(S\backslash NP) & T_2/(T_2\backslash NP) & (S\backslash NP)/S & T_3/(_3T\backslash NP) & (S\backslash NP)/NP & S\backslash NP
\end{array}
$$

think + you: S/S $>\mathbf{B}$

you + like: S/NP $>\mathbf{B}$

(S/S) + (S/NP): S/NP $>\mathbf{B}$

who + ...: $N\backslash N$ $>$

man + $N\backslash N$: N $<$

a + N: $T_1/(T_1\backslash NP)$ $>$

$\cdots$: S $>$

(17)

$$
\begin{array}{cccccccc}
\text{give} & \text{a teacher} & \text{an apple} & \text{and} & \text{a policeman} & \text{a flower}\\
\hline
(VP/NP)/NP & T_1\backslash(T_1/NP) & T_2\backslash(T_2/NP) & CONJ & T_3\backslash(T_3/NP) & T_4\backslash(T_4/NP)
\end{array}
$$

a teacher: $<\mathbf{T}$; an apple: $<\mathbf{T}$; a policeman: $<\mathbf{T}$; a flower: $<\mathbf{T}$

$T_2\backslash((T_2/NP)/NP)$ $<\mathbf{B}$

$T_4\backslash((T_4/NP)/NP)$ $<\mathbf{B}$

$T_6\backslash((T_6/NP)/NP)$ $<\Phi>$

VP $<$

such understanding is founded on an even more basic capability for planning the animal's own actions, of the kind proposed above. In particular, it seems likely that the purely motor-sensitive neurons of F5 are closely related to rules of the LDEC type, aka TOTE units or operants, and that the visual object-related neurons are related to the apparatus that associates objects with the actions that they afford (Miall 2003:2135). The interest of the mirror neurons themselves is then that their generalization over participant identities makes them necessarily symbolic representations, distinct from both efferent motor activity and afferent pure perception These units also look as though they would map very directly onto *verbs*, whether we think of these as case-frames (Rizzolatti & Arbib 1998), dependency structures (Pulvermüller 2002) or CCG lexical categories, as above. In CCG, of course, such lexical items constitute the entire language-specific grammar.

In terms of the theory of evolution of the language faculty, it is striking that this entire system is prelinguistic, rather than language-specific, and can therefore be investigated in primates and other animals other than humans. We need to know more about F5 in primates, specifically in relation to tool use. Are there "affordance" neurons that fire both to use and appearance of tools? Study of the regions adjacent to F5, (e.g. F4 which has spatially located action units Rizzolatti, Fogassi, & Gallese 2002) and pathways to and from the cerebellum (Miall 2003) which executes and monitors them, are likely to be important. We also need to understand how the planning process exploits units in F5, particularly the role of the limbic system, and the way in which repeated construction of a plan can lead to compilation as a long-term memory.

In addition, we need neurocomputational and machine-learning theories of how symbolic units of the kind found in F5 can be induced from sensory-motor input. Much of this system seems to be highly localized, rather than parallel-distributed. (However, mechanisms like Simply Recurrent Networks (SRN, Elman 1990) may well be appropriate for the process of compilation of repeated plans into compound actions and episodic memories, as opposed to novel plan construction and natural language understanding—cf. Steedman 1999.)

The computational character of the cortico-cerebellar-hippocampal sensory motor system is fairly well understood since Marr (1969)—see Gluck & Myers 2000. Perceptron-like reinforcement learning conditional on the intended goal state of LDEC-like operants seems to offer a mechanism for the neocortex and cerebellum. Associative networks similarly offer a mechanism for the hippocampus in its latent learning aspects, in neural network systems like those in .
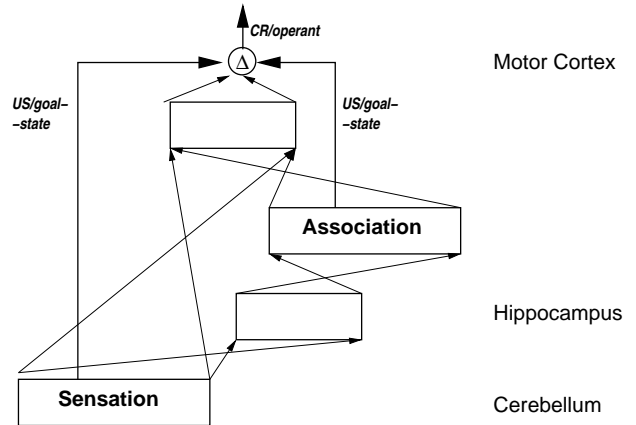

Figure 1: The basic Cerebellar-Hippocampo-Cortical dual-path circuit.

The neural pathways to and from the motor cortex remain less clear (Daskalakis *et al.* 2004). It is likely that several levels of plan representation mediate (Wolpert, Doya, & Kawato 2003). The process of abstracting over complete action representations needed to specify the verb/affordance-like units of F5 seems to be an open problem in neural computational terms.

Where Does Recursion Come From?

The neurological, developmental, and evolutionary conspiracy between serial ordering for motor plans and language suggests that the composition and type raising are prelinguistic primitives that we share with some animals. Yet apes show no sign of being able to acquire productive syntax. What more is needed to support the language faculty?

One candidate is modal and propositional attitude concepts—that is, functions over propositional entities. (We have so far glossed over the important fact that plans compose actions of type *state → state*, whereas syntax composes functions of type *proposition → proposition*.) These induce true recursion in conceptual structures and grammar via the grounded lexicon. There is no evidence that apes can attain the kind of theory of other minds that is required to support such concepts. Perhaps this is *all* they lack (Premack & Premack 1983;Tomasello 1999; Steedman 2002a,b; Hauser, Chomsky, & Fitch 2002).

This suggests that we need to know much more about the development of propositional attitude concepts in human infants, and their relation to planning and tool use around Piagetian sensory-motor developmental stage 6.

Conclusion

Compositionality appears on this account to be a very general property of simple sensory-motor planning systems. The units of compositional sensory motor plans appear to be learnable by known neural computational mechanisms, and can be observed by single-cell recording. While processes of plan formation and plan execution are well understood in symbolic terms, their translation into neurocomputational architectures remains an exciting challenge.

References

Baldridge, J. 2002. *Lexically Specified Derivational Control in Combinatory Categorial Grammar*. Ph.D. Dissertation, University of Edinburgh.

Clark, S., and Curran, J. R. 2004. Parsing the WSJ using CCG and log-linear models. In *Proceedings of the 42nd Meeting of the ACL*, 104–111. (submitted).

Curry, H. B., and Feys, R. 1958. *Combinatory Logic: Vol. I*. Amsterdam: North Holland.

Daskalakis, Z.; Paradiso, G.; Christensen, B.; Fitzgerald, P.; Gunraj, C.; and Chen, R. 2004. Exploring the connectivity between the cerebellum and motor cortex in humans. *Journal of Physiology* 557:689–700.

Elman, J. 1990. Finding structure in time. *Cognitive Science* 14:179–211.

Gallese, V.; Fadiga, L.; Fogassi, L.; and Rizzolatti, G. 1996. Action recognition in the premotor cortex. *Brain* 119:593–609.

Girard, J.-Y. 1987. Linear logic. *Theoretical Computer Science* 50:1–102.

Gluck, M., and Myers, C. 2000. *Gateway to Memory*. Cambridge MA: MIT Press.

Harel, D. 1984. Dynamic logic. In Gabbay, D., and Guenthner, F., eds., *Handbook of Philosophical Logic*, volume II. Dordrecht: Reidel. 497–604.

Hauser, M.; Chomsky, N.; and Fitch, W. T. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298:1569–1579.

Hockenmaier, J., and Steedman, M. 2002. Generative models for statistical parsing with Combinatory Categorial Grammar. In *Proceedings of the 40th Meeting of the ACL*, 335–342.

Lashley, K. 1951. The problem of serial order in behavior. In Jeffress, L., ed., *Cerebral Mechanisms in Behavior*. New York: Wiley. 112–136. reprinted in Saporta (1961).

Marr, D. 1969. A theory of cerebellar cortex. *Journal of Physiology* 202:437–470. Reprinted in Vaina 1991.

Miall, R. C. 2003. Connecting mirror neurons and forward models. *NeuroReport* 14:2135–2137.

Miller, G.; Galanter, E.; and Pribram, K. 1960. *Plans and the Structure of Behavior*. London: Holt.

Premack, D., and Premack, A. J. 1983. *The Mind of an Ape*. New York, NY: Norton.

Pulvermüller, F. 2002. *The Neuroscience of Language*. Cambridge University Press.

Rizzolatti, G., and Arbib, M. 1998. Language within our grasp. *Trends in Neuroscience* 21:188–194.

Rizzolatti, G.; Fogassi, L.; and Gallese, V. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews: Neuroscience* 2:661–670.

Rizzolatti, G.; Fogassi, L.; and Gallese, V. 2002. Motor and cognitive functions of the ventral premotor cortex. *Current Opinions in Neurobiology* 12:149–154.

Saporta, S., ed. 1961. *Psycholinguistics: A Book of Readings*. New York: Holt Rinehart Winston.

Steedman, M. 1999. Connectionist sentence processing in perspective. *Cognitive Science* 23:615–634.

Steedman, M. 2000. *The Syntactic Process*. Cambridge, MA: MIT Press.

Steedman, M. 2002a. Formalizing affordance. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society, Fairfax VA, August*, 834–839. Mahwah NJ: Lawrence Erlbaum.

Steedman, M. 2002b. Plans, affordances, and combinatory grammar. *Linguistics and Philosophy* 25:723–753.

Tomasello, M. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Vaina, L., ed. 1991. *From Retina to Neocortex: Selected Papers of David Marr*. Boston, MA: Birkhauser.

Wolpert, D.; Doya, K.; and Kawato, M. 2003. A unifying computaitonal framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, B* 358:593–602.