10

Coordination

What It Is and Why We Need It

Christoph von der Malsburg

Abstract

Trying to apply our everyday concept of coordination to the brain raises a number of fundamental questions: What is the nature and meaning of local brain states that are to be brought together? On what grounds are they to be coactivated and connected? What is the nature of meaningful structural relationships, and how does the brain learn them? What is the role of focal attention? How does the brain assess its current level of coordination? How do brain states address goals? What is the nature of our environment's statistics, and how is it captured by the brain? What mechanisms endow brain dynamics with a tendency to fall into coordinated states? Some of these questions seem to be difficult to address within the current experimental paradigm.

Introduction

What is coordination? There are many domains in which coordination plays a central role: from the preparation of a meal or setting of the dinner table to the writing of a literary novel or musical composition. The aesthetic feelings conjured up by a work of art in our mind are often no more than a reflection of the level of coordination that takes place between our sensations. The establishment and running of a company or the organization of a conference requires structural elements to relate to each other in a meaningful way: they must fit together to form a whole if the result is to function well. The coordination inherent in institutions or manufactured structures reflects necessarily that which is present in their creators' brains. Thus, any insight that we can glean from coordinated artifacts is relevant for understanding coordination in the brain.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

What Is Coordination in the Brain?

Coordination means putting things together that belong together. Each term in this sentence raises questions, which I will address in turn. Throughout this essay, I will focus more on posing questions than on providing answers. Because things in our brain[1] are assembled on a slow timescale of learning as well as on a fast timescale of thinking, it is important for us to heed both.

What Are the Things That Must Be Put Together?

The brain has a great variety of modalities or subsystems, corresponding to internal and external senses, to motor control, to memory, to emotions, to behavioral control, and more. Minsky (1988) spoke of the "Society of Mind." Superficial modalities are formed by wiring with the sensory, motor, or humoral periphery and are thus defined (onto-)genetically. More centrally, there is considerable plasticity, at least in the cortical system, as shown by many cases of neurological recovery (Weiller and Chollet 1994). Different modalities focus on different reflections of phenomena and contain patterns that are similar, so that they can be mapped and stacked onto each other. This pervasive phenomenon of cortical localization requires a natural measure of similarity between neural patterns, the need for which we will encounter repeatedly later.

Each subsystem of the brain is able to create a large variety of alternate activity patterns, of which at any given time only one, or a few, can be clearly and unambiguously expressed. For much of the time, neural activity in a subsystem expresses an ambiguous superposition of different states, which is reduced in stages under the influence of signal exchange. This reduction must be coordinated, so that patterns which belong together are coactivated. Such reduction of uncertainty is, for example, modeled in probabilistic formulations (Pearl 1988; Bishop 2006).

What Does It Mean to Belong Together?

Neural patterns belong together because they are generated by stimuli that are statistically linked in the environment (often, a common cause is responsible), or because they successfully interact to attain goals or to exercise and develop capabilities playfully. "Belonging together" is thus both defined in a passive, recording mode and in an active, creative mode. The creative mode is responsible for the brain's ability to handle new situations and has been grossly neglected in the neuroscientific literature.

---

[1] I am adopting the attitude of Spinoza and am viewing brain and mind as two sides of the same coin. When using electrode or microscope, we observe the brain. Psychophysics or introspection, by contrast, lets us view the mind.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

How Does the System Learn What Belongs Together?

The initial motive for patterns to be related is similarity. Different modalities, by contrast, speak different languages and contain neural patterns that are not similar. How, then, can the brain distinguish which patterns from different modalities belong together? The only means of establishing *de novo* pattern association is simultaneity; that is, significant correlation in time. Unfortunately, however, it is not useful to associate all neural patterns in the brain with each other merely on the basis of simultaneity, as is implied in associative memory models (e.g., Hopfield 1982). To an overwhelming extent, pattern simultaneity in the real world is accidental and not of lasting value. If synaptic plasticity had the form of indiscriminate stickiness, the brain would soon be cluttered with myriad connections. It is thus important to single out associations that are significant.

The natural definition of pattern significance—recurrence—suggests a strategy that has actually been adopted by much of the artificial neural network literature (Haykin 1994); namely, sifting through the input for those patterns that appear repeatedly with statistical significance. Unfortunately, this strategy fails for input fields of any realistic size, because the number of patterns which must be tracked is simply too large, and literal repetition of a given pattern is too unlikely. The system, therefore, needs to possess a similarity measure and powerful prejudices concerning the nature of significant patterns. Yet, what is the nature of these prejudices?

Once different modalities have accumulated a sufficient mass of pattern associations between them, they can use general laws of composition to generate creatively novel, modality-spanning composite patterns. An important issue is the nature of these laws of composition. A typical (or perhaps *the* typical) law may be that overlapping patterns in one modality must be associated with overlapping patterns in another.

The Detection of Significant Patterns by Focal Attention

Why is the information content of attention limited? Does this represent an imperfection of the brain, or is this even functionally significant? Focal attention powerfully (though not exclusively) restricts learning to a small subset of active neurons at any one time (for further discussion, see Jiménez 2003). Key questions focus on how this restriction is expressed and what effects this has on learning. One proposal is based on gamma rhythms (Fell et al. 2003) and on the ensuing concentration of neural spikes into narrow temporal windows, as a boost to synaptic plasticity. This reduces the input and the memory domains to small sectors (as modeled in Jacobs et al. 1991). The restriction addresses a fundamental problem of present-day models of learning (i.e., the scaling-up to realistically large input and storage domains) by restricting system modification to the narrow focus of attention. Accordingly, the reason

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

why informational content of attention is limited results not from an imperfect mechanism but from the necessity to preclude confusion.

What is the informational basis on which focus of attention is formed? Attention is very much at the core of the problem of coordination: it has to bring together sets of patterns that belong together. If it only unites patterns that have already been associated in the past—in passive mode, so to speak—it does not aid the problem of learning. For that, it has to find new associations in a creative manner.

External events, signaled by temporally isolated sensory signals (which are emphasized by the filter properties of our senses), can bring together specific patterns that light up simultaneously in different modalities. Patterns can also be brought together on the basis of abstract properties with which they are tagged. For example, if the senses can be focused on the same point in space, for which the colliculus superior seems to be well equipped (at least in the mouse; see Dräger and Hubel 1975), patterns aroused from that point can be associated. The general Gestalt laws (Koffka 1935; Crick 1994) can define significant patterns as figures set apart from a background. An object moving against a static background can, for example, be made to stand out as the focus of attention on the basis of common motion (Spelke 1998). Thus, the Gestalt laws formulate abstract properties (common motion, color, stereo depth, spatial grouping, good form, and edge continuity), which help to tag novel patterns as significant.

A statistical definition of pattern significance is not sufficient; a biological definition is also required. Important classes of patterns or events must be genetically defined to inform an individual on what is required for success in life. Such a definition must be laid down in some abstract fashion so that concrete occurrences can be recognized and selected. Ethologists have described many such schemata, such as the facial schema with which the human infant is born, the definition of mother goose for the gosling, or a red dot on the beak of the seagull to indicate to the chick the source of food (Toates 1980). Upon recognition of the releaser for an innate cue, attention is focused on the recognized stimulus, which is then separated from the ground; an appropriate action is induced (e.g., an orienting reflex or grasping); and an appropriate sector of memory is selected for modification by the stimulus. This sequence of events may be referred to as schema-based learning.

Important questions remain, however, regarding the technical implementation of these processes.

How Is Coordination Evaluated?

People have a keen sense of the level of coordination that goes on in their brains. Sometimes we feel distracted or confused; we cannot make up our mind or feel that something is awry, or not quite right. Other times, we experience

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

the sensation of being sharply focused, highly concentrated, or fully conscious of a situation. Then there are those precious moments, when we suddenly feel that we have it; everything has fallen into place and we shout Eureka!

To a large extent, aesthetic pleasure is due to the level of coordination generated in our brain by the object of our attention. It is not conscious insight into the structure of the work of art that we experience, but rather some direct feeling of the level of coordination in our brain. This *Aha!* effect, this falling into a state of organization, is what Gestaltists refer to as reorganization of the perceptual field and insight (Köhler 1925).

Which structures in our brain are responsible for the evaluation of this measure of coordination, and what is the nature of this measure? Obviously, it cannot be attributed to some superior intellectual entity whose insight into the subject matter serves as judge. Rather, it must be some signal that can be "mechanically" generated and globally evaluated. The essence of it may be the level of nontrivial agreement of independent signals at convergence points, evaluated over the whole brain.

How Is Purpose Defined, Enforced, and Achieved?

In an active mode, the brain must coordinate patterns to achieve its purposes. Here, the central issue is that purpose (e.g., I am hungry and am looking for food) as well as the generation of neural patterns that serve to achieve that purpose (opening the fridge, or calling the pizza delivery service) are defined at very different levels of detail, and generally in different parts of the brain. A newborn possesses, presumably in the midbrain, the schematic definition of a set of fundamental goals. These form a hierarchy, the honing of which keeps us busy over much of the course of our life. Goals are activated either spontaneously or in response to some stimulus, like "danger" or "thirst." Goals tend to be mutually exclusive and come equipped with powerful mechanisms of enforcement. Complex tasks require the attainment of goals and subgoals in hierarchical fashion, and there are profound questions concerning the nature, establishment, and implementation of goals in our brain.

Behavioral patterns usually have a number of functional components, each of which has a range of possible role fillers. In the looking-for-food scenario, relevant roles include possible foodstuffs, sources of food (e.g., the refrigerator, delivery service), and potential modes of acquisition. In a specific situation with a concrete goal, the system must select the appropriate role fillers which will interact functionally to attain the goal. How is this type of coordination achieved through the interaction between a goal schema that contains a set of role descriptions, the possible role fillers that have the ability to combine appropriately, and the sensory patterns that describe the situation? In addition, we need to know the way in which behavioral schemata and goal descriptions are implemented neurally, the mechanisms by which these schemata are triggered

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

and prioritized, the reward mechanisms by which the achievement of goals is evaluated, the mechanisms by which the activity of the brain is biased in the direction of goal fulfillment, and the mechanisms by which, over the course of our life, goal schemata are elaborated in richer and richer ways in terms of detailed sensory and motor patterns.

How Are the Environment's Statistics to Be Captured?

The brain receives signals from the environment over many millions of fibers and, in turn, influences the environment through multiple output fibers. All our brain can ever know and learn is contained in the statistics of these activity patterns. From all possible combinations of individual neural input or output signals—a space of vast volume—only a minute subvolume is ever realized in terms of actual signals. Exhaustive recording of global activity patterns is not possible, nor would it make sense as no sensorimotor activity pattern ever has a chance to recur. Only by extracting significant subpatterns, by ordering them in groups of similar patterns, and by developing schemata for their arrangement is it possible to capture the environment's statistics and cope with the ever-changing, novel situations that humans encounter daily. This requires a prejudice to define significant patterns for extraction; it demands a general similarity measure by which these patterns are to be grouped; and it requires a preestablished format for the representation of pattern arrangements.

One is caught, so to speak, between a rock and a hard place. If the prejudices are too weak, the system is overwhelmed by variance that cannot be captured in a realistic finite system. If the prejudices are too narrow, the reality of the environment may be missed (the bias-variance dilemma; Geman et al. 1992). Another indication that the system's prejudices must be tuned to the environment are the no-free-lunch theorems (Wolpert and Macready 1997), according to which any learning or optimization mechanism can be totally vitiated by an environment that does not fit its a priori assumptions. In summary, the brain needs powerful a priori assumptions, and these must fit the actual environment! What, then, are these a priori assumptions?

What Is the Nature of Our Environment's Statistics?

Of all the questions, this is probably the most crucial, since, as argued, the mode of operation of the brain must be tuned to the environment. In fact, the brain must coordinate with the environment.

Some important aspects of sensory pattern statistics result from the media through which they are transmitted. The visual medium, for example, is the optical radiation field captured by the eye, and the patterns that appear on our retinae are shaped by the laws of reflection and propagation of light, geometry,

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

and motion. These transformations need to be inverted for the brain to decipher the structure of the patterns that are there. The laws of transformation are to a large extent independent of the environmental patterns themselves. The first sensory stages of the brain can reduce the complexity of the input patterns tremendously by inverting these transformations, thus reducing large sets of patterns to invariance classes (Wiskott 2006).

What is the regularity of the world beyond that? What are the repeating patterns? Or, rather, in what general format do they appear? According to Kant's analysis, we come equipped with "categories" (i.e., a priori structures that permit us to absorb information). Among these he counted space, time, and causality. We take for granted that repeating spatial and temporal patterns and causal sequences of events play an important role in our environment. It is a wide-open question, however, as to how whole scenes are to be decomposed to find repeating patterns, and how the general rules of composition by which our environment generates its configurations in ever-new ways are formulated. Coordination, to relate back to our theme, is the ability to create internal scenes that capture the reality of the environment. The brain's task, then, is to extract environmental patterns and their relations, together with a measure of likelihood for their relevance, so as to acquire the ability to complement partial information in a given scene with additional details familiar from the past, to generate a more complete description of the scene. Our challenge is to second-guess the general form—the architecture—on the basis of which this is possible.

What Is the Nature of Structural Relationships?

The patterns that we experience never repeat precisely. When recording a novel pattern, it is thus important to be able to define a spectrum of other patterns that are similar to it. This implies a similarity measure, or some definition of the likelihood that a sensory pattern is to be identified with a stored pattern. If properly constructed, the stored structure and the similarity measure can decide to a high statistical significance whether a perceived pattern is an accidental arrangement of elements, or whether it is the same pattern repeating itself.

What is this similarity measure? The simplest idea of pattern recognition is template matching, where a rigid pattern, the template, is moved over an image to find an identical fit. The "motion" takes care of the invariance aspect if it includes all possible transformations (e.g., translation, scaling, and rotation). Template matching has long since fallen out of favor because identical fits are never found in real images. A first step toward solving the problem is to dismantle the "template" and endow the resulting pieces with flexible relationships so that distortion can be addressed. Thereafter, pieces of the model must be replaced with statistical models of possible variants. A version of this is the leading mechanism of face recognition, as described, for example, by Wolfrum

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

et al. (2008). Finally, the pieces themselves can then be replaced by composite models to create a hierarchical structure.

What has been described here for vision applies analogously to other modalities. Motor patterns (including speech) form patterns within patterns, each being a role filler, each permitting a range of variants, the whole put together flexibly to permit continuous time warping and, of course, further nesting. Hidden Markov models (Rabiner and Juang 1986) capture essential aspects of this. I contend that this kind of architecture applies to all modalities of the brain individually, and the brain as a whole.

Out of these considerations arises a picture according to which mental objects are hierarchical graph structures, with concrete patterns arranged in spatial and temporal relationships to each other. Graphs are, in general, embedded in or linked to more abstract graphs (which, due to their abstractness, are called schemata), whose nodes refer to classes of exchangeable subpatterns, and whose links describe permitted relations. Recognition is the process by which abstract graphs are mapped to concrete patterns homeomorphically (i.e., with equivalent parts mapping to each other under preservation of relations). Coordination is the process by which concrete brain states are generated under the guidance of abstract descriptions (including formulation of goals) and of sensory input. Usually, several abstract schemata conspire to create a detailed description.

Important questions include: How is the repertoire of neural behavior tuned to the construction of such hierarchical descriptions? How are hierarchical descriptions developed in the brain on the basis of experience? How can this architecture be described in concrete mathematical terms?

How Are Neural Patterns Put Together?

The brain is endowed with an architecture that tends to fall into globally ordered patterns, structured accordingly to the world in which we live. This ability can be likened to the process of crystallization, in which constituent atoms or molecules create global order out of local interactions, by exerting their preferences according to the shape of the local environment. For crystallization to be initiated, a seed (or minimal structure) is required such that further molecules quickly find a niche into which to fall.

In the case of the brain, the constituent elements are neurons, and knowing how their behavioral repertoire is structured, so as to favor global order, is crucial. Some aspects of this are already emerging. Outgrowing processes are guided by chemical or electrical signals so as to favor ordered connectivity patterns, as exemplified by the ontogenesis of retinotopy (Goodhill 2007), whereas intrinsic plasticity (Butko and Triesch 2007) regulates the duty cycle of cellular activity.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

If the behavioral repertoire of neurons (or of a collection of neural types) were known to any degree of precision, it should be possible to simulate on the computer the growth of ordered connectivity and activity states. Thus far, this venture has had some level of success, especially in modeling the ontogenesis of ordered connectivity structures. However, modeling the generation of neural states that can be interpreted as mental objects is still out of reach. Thus, questions remain regarding the repertoire of neural behaviors, and, possibly equally important, the equivalent of seed structures with which the process of coordination is initiated.

Concluding Remarks

To coordinate "stuff" in our mind, even while performing routine tasks, we regularly have to put things together that have previously never been assembled. These associations are creative acts that are not imposed on us by our environment. In recognizing an object, for example, we apply a schematic description of this type of object to a concrete image, thereby associating abstract features with concrete instantiations. When we manipulate the object or describe it verbally, this enables us to relate directly to the specific character of the feature instances. Certainly, our brain contains the neural pathways to connect what is to be connected; however, these pathways are embedded, or are drowned, in a virtual continuum of others which, at present, are irrelevant but are all required and useful when their time comes. Although I intended simply to pose questions in this essay, I could not do so without interjecting my conviction that the task of our brain, in any given situation, is not just to activate a subset of all neurons, but to select a tiny subset from the vast numbers of physical connections. Synapses outnumber neurons, and the task of selecting connections is larger than that of selecting neurons by a factor of about ten thousand. If we ignore this task, we are, in my view, ignoring 99.99% of the information in our brain's state.

If indeed the brain, in its rapid state changes, is mainly concerned with selecting structured connectivity patterns, and if we need to study and understand synaptic dynamics in addition to neural dynamics to bridge the chasm between mind and brain, then we have a problem. Experimental technology is highly developed to study neural dynamics and, to a lesser extent, static or slowly changing connectivity. It can even record short-term modification of synaptic effects for individual connections, but the imaging of whole, rapidly changing, connectivity patterns is presently beyond our imagination. When Ludwig Boltzmann first established statistical mechanics, he was ridiculed by his colleagues Ernst Mach, Wilhelm Ostwald, and others for his atomistic ideas, and it took three decades until experiments made the reality of atoms and molecules concrete enough to convince the community. Must the neurosciences also wait for decades for the necessary revolution? Dedicating years of effort

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

to experimental exploration is too risky if it is not supported by community convictions, but the community will not be convinced without focused efforts. Let us hope that this vicious circle can be broken with the help of concrete computer models of cognitive functions whose demonstrable success rests on outlandish physiological assumptions.

Acknowledgment

I acknowledge support received from the BMBF project Bernstein, FKZ 01GQ0840.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.