11

Extending Sensorimotor Contingencies to Cognition

Alexander Maye and Andreas K. Engel

Abstract

An emerging view in cognitive science considers cognition as "enactive" (i.e., skillful activity involving ongoing interactions with the external world). A key premise of this view is that cognition is grounded in the mastery of sensorimotor contingencies (i.e., the ability to predict sensory changes which ensue from one's own action). It is proposed that the learning of sensorimotor contingencies serves basic sensorimotor processing and that it can also be used to establish more complex cognitive capacities, such as object recognition, action planning, or tool use. Recent evidence from robotics and neuroscience supports this claim and suggests that "extended" sensorimotor contingencies might be a viable concept for pragmatic cognitive science.

Introduction

An "action-oriented" approach to cognition holds that cognitive processes are closely intertwined with action and that cognition is best understood as "enactive" (i.e., a form of practice itself). Accordingly, cognition is grounded in a pre-rational understanding of the world—one based on sensorimotor acquisition of skills for real-life situations.

Long before the emergence of modern cognitive science, philosophers emphasized the active nature of perception and the intimate relation between cognition and action. In 1896, the American pragmatist John Dewey formulated an influential sensorimotor approach to perception:

> Upon analysis, we find that we begin not with a sensory stimulus, but with a sensorimotor coordination and that in a certain sense it is the movement which is primary, and the sensation which is secondary, the movement of the body, head and eye muscles determining the quality of what is experienced. In other words, the real beginning is with the act of seeing; it is looking and not a sensation of light.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science," Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18, series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

With striking convergence, the same concept was expressed, more than sixty years later by the French phenomenologist Merleau-Ponty (1962):

> The organism cannot properly be compared to a keyboard on which the external stimuli would play. Since all the movements of the organism are always conditioned by external influences, one can, if one wishes, readily treat behavior as an effect of the milieu. But in the same way, since all the stimulations which the organism receives have in turn been possible only by its preceding movements which have culminated in exposing the receptor organ to external influences, one could also say that behavior is the first cause of all stimulations. Thus the form of the excitant is created by the organism itself.

Merleau-Ponty strongly advocated an anti-representationalist view by emphasizing that the structures of the perceptual world are inseparable from the cognitive agent.

Most motifs of the "pragmatic turn" addressed throughout this volume can be traced back to these two philosophers. Drawing on the pragmatist and the phenomenological traditions, numerous authors have explored the implications of defining cognition as embodied action (Varela et al. 1992; Clark 1998; Noë 2004; Pfeifer and Bongard 2006; Thompson 2007; Engel 2010; Engel et al. 2013).

The notion that cognition can only be understood by considering its inherent action-relatedness is a key ingredient in the sensorimotor contingency (SMC) theory put forth by O'Regan and Noë (2001). Accordingly, an agent's SMCs are constitutive for cognitive processes. In this framework, SMCs are defined as law-like relations between movements and associated changes in sensory inputs that are produced by the agent's actions. "Seeing" (according to the SMC theory) cannot be understood as the processing of an internal visual "representation"; instead, "seeing" corresponds to being engaged in a visual exploratory activity, mediated by knowledge of SMCs. The active nature of sensing has been advocated by other approaches as well. For example, active vision in robotics is often considered to be a sensorimotor approach in which action plays a constitutive role for perception. In the majority of cases, however, the different views captured by the robot's camera at the scene are analyzed without considering the actions involved in bringing about these perspectives. Thus, active vision approaches attempt to compute a veridical representation of a scene by effectively stitching together perceptions from different perspectives. Whereas it has been shown that this improves the reliability of the scene segmentation, the approach still hinges largely on action-ignorant methods for analyzing the individual camera images. The key concept of SMC theory is more radical in the sense that it considers action a necessary component of perception: action does not merely support or interact in some way with perception. Thus, instead of analyzing images individually for their perceptual content, an SMC theory-based approach searches for regularities in

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

the changes between camera images brought about by the specific actions with respect to the objects in the scene.

Although increasing evidence from work in robotics, psychology, and neuroscience support the SMC theory perspective, few attempts have been made to extend these ideas into a more comprehensive framework for cognitive science, and to derive their implications for understanding more complex cognitive capacities. Here, we propose extending the concept of SMC theory and suggest that SMCs be used to define object concepts and action plans, and that the mastery of SMCs could lead to goal-oriented behaviors. Our proposal implies that the notion of SMCs could be expanded into a more generalized concept of action-outcome contingencies, and we use the term "extended" sensorimotor contingencies (eSMCs) to denote this generalized concept.

Limitations of Sensorimotor Contingency Theory

SMC theory provides a fresh approach to explain perceptual awareness as well as a potential alternative to cognitivist theories of consciousness that consider cognition as computation over internal, observer-independent representations. O'Regan and Noë (2001) developed the basic idea and related it primarily to visual awareness, but its pertinence to other sensory modalities is straightforward. In a number of respects, SMC theory reveals limitations that call for further development of the concept as well as its application in empirical research.

Clarification Is Needed

A central claim of SMC theory is that conscious perceptual experience requires attentional exercise to master SMCs (O'Regan and Noë 2001). Building on the intuitive understanding of the term, "mastery" may work to explain the very concept. However, to assess mastery of SMCs in nonhuman agents (in particular, artificial agents), details are needed as to what constitutes mastery and the mechanisms by which an agent can actually achieve it. Similarly, in SMC theory, the process of "attunement" to an environmental feature has not been clearly defined. Attunement can describe the exercise or deployment of already mastered sensorimotor knowledge in a particular context (O'Regan and Noë 2001). However, we do not know whether attunement is a deliberative process which consciously weighs different possible contingencies to be deployed or an automatic process that selects the right set of contingencies in response to a particular situation.

This is complemented by an unclear attitude toward representationalism. O'Regan and Noë (2001) explicitly refute the view of the brain as a world-mirroring device. Perhaps the most interesting proposition of SMC theory is

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

that perceptual experience cannot be equated with the activity of neurons or neuronal populations in specific sensory modalities. Instead, the regularities in the sensorimotor interactions (i.e., the SMCs) give rise to the different qualities of perceptual experience in different sensory modalities. These contingencies, however, need to be acquired, memorized, and maintained. This may create the impression that SMC theory has simply replaced classical concepts of representation with the notion of SMCs, which largely subserve the same role in the overall framework. Even if knowledge of SMCs is declared implicit, by emphasizing the continuous nature of the processes by which an agent attunes itself in its entirety to the environmental structure to fulfill its demands, functions for the discovery of regularities and policies that regulate the plasticity of this knowledge need to be specified.

Notions of Action and Normativity Are Unclear

SMC theory explains perceptual qualities in terms of the regularities in the changes of sensory signals caused by actions. Whereas examples of squashing a sponge or stroking a surface (O'Regan and Noë 2001) may explain the emergence of the corresponding perceptual experience in an intuitive way, closer inspection calls for further elaboration of SMC theory's concept of action. Why can we squash the sponge and stroke the surface and not vice versa? Given that our action space is infinite, why should we select exactly these two movements? How do actions which convey perception get along with actions designed to achieve goals?

These questions suggest that some type of normativity is involved, reflecting the aptitude of each action for achieving perceptual discrimination and other goals. Likewise, actions can succeed or fail in yielding the intended outcome. Consequently, the question becomes: How do such norms arise? Other enactivist approaches (Thompson 2007; Di Paolo et al. 2010) approach this question by pointing out that the body is not only an apparatus that mediates perceptual experience, but that its self-maintenance (homeostasis) constitutes a source of basic norms by distinguishing actions that promote or break self-maintenance.

Another direction of development must be to clarify what constitutes "action." In its original form, SMC theory seems to comprise more than just movements or motor activity, but it does not offer a clear distinction between action and movement. The extended version of SMC theory, which we propose here, interprets "action" as being neither coextensive with that of behavior nor with that of movement (Mead 1938); "action" also includes acts that do not involve any overt movements (e.g., thinking, calculating, imagining, deciding). The description of acts or actions typically makes references to goals, whereas behavior can be described without making any reference to mental states.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Generalizing with Sensorimotor Contingencies

To recognize environments and objects contained therein, current SMCs must be compared to previous experiences drawn from a repository of learned contingencies. The process of recognition does not necessarily involve movement at all, but can rely solely on sensory information to match prior experiences. This allows, for example, quick recognition of one's own cup from the pattern it leaves on the retina, instead of the time-consuming process of reenacting some of the contingencies that were acquired when it was seen for the first time. In addition, after some experience with cups, any new one can be recognized as such without fully exploring the contingencies of this particular instantiation. Both aspects require some form of generalization. Whereas learned contingencies may be the basis for recognizing a particular instance of an object ("my cup"), generalized contingencies allow the recognition of a broader object class. Such a generalization process needs to consider the relevant aspects in the typically high-dimensional sensorimotor interaction patterns and reduce them to a basic set of properties that pertain to the whole object class. Relevance, however, is largely contingent on context. Thus, any generalization schema is likely to operate dynamically. How relevant subsets of eSMCs are determined is currently an open issue.

Extending Timescales

Explanations for the basic concept of SMC theory generally employ a single context: gazing at an object, its haptic exploration, the attunement of an automatically guided missile, or driving a Porsche (O'Regan and Noë 2001). It seems straightforward, however, to assume that there are regularities in sensorimotor interactions that extend across different contexts and over longer timescales. Deployment of contingencies over extended timescales appears necessary for certain actions, such as driving to work or preparing a cake. Going to the movies and growing flowers involve regularities that extend over the course of a day or weeks, respectively. Pursuing a doctoral degree to achieve a fascinating position in science may represent an example that plays out at a lifetime timescale. As we suggest below, the concept of SMC theory might be extended to what we call "intention-related" eSMCs, which capture long-term regularities in action sequences and constitute our conscious experience beyond the timescale of object perception. It may even be interesting to ponder how eSMCs connect individuals on a social level, perhaps on timescales greater than an individual's lifetime.

Disrupted Contingencies and Altered Perception

If SMCs shape perceptual experience, their disruption should yield altered perception. This obvious conclusion needs elaboration and experimental validation.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Changes in body morphology, the skeleton-muscular apparatus, or the transmission characteristic of the sensory organs will activate a learning process for the respective altered contingencies. Conscious experience of altered perception is possible as long as the agent has access to both original and altered sets of contingencies. Without the possibility to reenact the undistorted contingencies, they are likely to be extinguished over time. To address these issues, it would be highly interesting to study perceptual alterations in patients with long-term impairment of the ability to move without concurrent cognitive deficits. Targeted changes of SMCs in virtual environments may offer a promising strategy to augment the experiential quality in such patients.

Extending Sensorimotor Contingency Theory

Our approach explicitly departs from the notion that perception is generated by an internal representation of the outer world. Instead, it accounts for the crucial role of action in the process of developing cognitive capabilities, an idea that is immediately appreciated if one thinks about the development of human infants. Our approach turns the classical view, in which sensory information initiates actions, upside down. The agent initiates actions to receive changes in its sensory stimulation, and it learns associations between them. The initiation of actions seems to emerge from spontaneous behaviors of biological organisms, as in the motor-babbling behavior of infants (Westermann and Miranda 2004; Natale et al. 2007).

Below, we distinguish three types of eSMCs that comprise contingencies at different levels of complexity, from sensorimotor coordination to action-reward contingencies. This suggests a multilevel architecture, where different types of eSMCs are acquired and deployed to implement cognitive processes of increasing complexity. A key hypothesis in our proposal is that a consideration of eSMCs at different scales should help unravel the emergence of cognitive capacities at different levels of complexity and contribute to an understanding of how these may be grounded in basic sensorimotor processes.

Modality-Related Extended Sensorimotor Contingencies

This first type captures the specific changes of the sensory signal in a single modality, depending on the agent's action. Examples include the different perspective distortions that result from eye movements and locomotion, sound pressure profile changes when the head rotates, and the dependence of the force feedback from the force exerted by finger movements. This most basic type of eSMC distinguishes the qualities of sensory experiences in the different sensory channels (e.g., "seeing," "hearing," "touching") and was addressed by SMC theory in its original formulation (O'Regan and Noë 2001). We suggest, however, that the original idea be broadened to account for the full vector of

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

multisensory inputs, together with intrinsically generated normative feedback. This notion is reflected in the computational model of modality-related eSMCs that we describe below.

Object-Related Extended Sensorimotor Contingencies

The next eSMC type relates to the effects on the sensory system that are specific for each object under consideration, and these effects are inherently supramodal. They describe the multisensory impression that an object leaves on a set of actions of the agent. An example is given by the different visual and force feedback signals received when touching a sponge, a piece of cardboard, or a piece of wood. Object-related eSMCs define the object under consideration, and exercising actions from a set of object-specific eSMCs corresponds to the perception of this object. One of the fundamental claims of our approach is that the observed relations between actions and sensory changes are sufficient for recognizing a particular object. Object-specific eSMCs are more numerous and more complex than modality-specific eSMCs.

Intention-Related Extended Sensorimotor Contingencies

The third type denotes a further level of generalization of the concept of SMCs and considers the long-term correlation structure between complex action sequences and the resulting outcomes or rewards, which the agent learns to predict. We propose that intention-related eSMCs capture the consequences of an agent's actions on a more general level as well as on extended timescales. These complex eSMCs include contingencies that are cognitively simulated by the agent and do not relate to factual movement. After learning, intention-related eSMCs could be used to predict whether an action will be rewarding or not and to rank alternatives. At the same time, intention-related eSMCs may provide the basis for action plans that involve several steps to reach an overall goal. In this way, anticipation and anticipatory behavior as well as the sense of agency might be grounded in eSMCs.

Extended Sensorimotor Contingencies in Natural Cognitive Systems

A substantial body of neurobiological findings is compatible with the idea that SMCs are used by the brain, and that different types of contingencies play a key role in natural cognition. Here we review neurobiological evidence that supports the eSMCs concept before considering recent attempts to model and test eSMCs in robot experiments.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Basic Importance

The action-oriented view of cognition advocated here is supported by findings on the role of exploratory activity and sensorimotor interactions for neural development and plasticity. For a long time, developmental processes in the nervous system have been known to be activity dependent. For instance, development of neural circuits in the visual system and acquisition of visuomotor skills critically depend on sensorimotor interactions and active exploration of the environment (Held and Hein 1963). Even in the adult brain, there is considerable plasticity in cortical maps (e.g., in the somatosensory and motor system) that has been shown to depend on action context (Blake et al. 2002). These studies suggest that appropriate action, allowing exercise of relevant eSMCs, is necessary throughout life to stabilize the functional architecture in the respective circuits.

If guidance of action is a dominant function of the brain, one would predict that neuronal response profiles in sensory or association regions should strongly depend on action context. Indeed, clear evidence exists for such an action-relatedness. Activation of visual cortical neurons changes profoundly if self-induced movements are permitted, as compared to passive viewing of stimuli (Niell and Stryker 2010). Gain modulation of neural responses is abundant in the nervous system (Salinas and Sejnowski 2001), demonstrating that sensory activity patterns are always, to a considerable degree, related to or modulated by action. In premotor cortex, the spatial profile of multimodal receptive fields depends on body and limb position (Graziano and Gross 1994). Tactile and visual receptive fields of premotor neurons are in dynamic register and seem anchored to body parts, even if these are moving; this suggests that such polymodal neurons support predictions about expected changes in sensory input (Figure 11.1a). Action-related changes of sensory response properties of polymodal neurons have also been observed in studies involved in tool-use learning (Figure 11.1b) (Maravita and Iriki 2004). It is tempting to speculate that assemblies of such polymodal neurons encode eSMCs.

Important evidence regarding the neural mechanisms of eSMCs comes from research on "corollary discharge" or "reafference signals," which are necessary for an organism to distinguish self-generated sensory changes from those not related to its own action (Crapse and Sommer 2008). Supporting the SMC theory, this research shows that predictions about the sensory outcome of movement are critical for the basic interpretation of sensory inputs. The importance of reafference has been shown in the context of eye movements as well as grasping or reaching movements. Interestingly, similar principles of predicting sensory inputs seem to play a key role in more complex cognitive processes, such as language comprehension or predictions about sequences of abstract stimuli (Schubotz 2007). In all of these cases, activity of motor planning regions seems to be involved in generating the prediction about sensory events, possibly by modulating neural signals in sensory regions. Malfunction

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

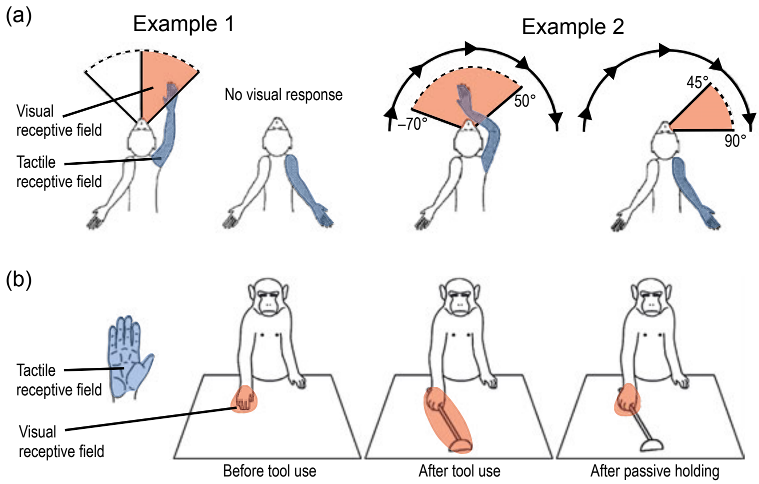

Figure 11.1 Dependence of multisensory receptive fields on motor state and action context. (a) Two examples of neurons recorded from ventral premotor cortex in the monkey. Neurons showed bimodal responses with both tactile (blue) and visual (red) receptive fields. In both examples, the visual response depended on the arm position of the animal. The panel illustrates results from Graziano and Gross (1994). (b) Recording of bimodal intraparietal neurons with tactile (blue) and visual (red) receptive fields. The visual receptive field showed adaptive changes as the animal used a tool to retrieve food, which expanded to include the entire length of the tool. Under control conditions with passive holding of the same tool, this expansion did not take place. The panel illustrates the work of Maravita and Iriki (2004). Reproduced with permission from Engel et al. (2013).

of such modulatory signals and associated disturbance of forward models have been implicated in the pathogenesis of psychiatric disorders such as schizophrenia (Frith et al. 2000b).

Object- and Intention-Related eSMCs in the Human Brain

Our view implies that procedural knowledge is fundamental to acquire object concepts. Thus, storing information about events and objects should generally involve action planning regions. In line with this prediction, neuroimaging studies show that object concepts in semantic memory do not only rely on sensory features alone but, critically, also on the motor properties associated with the object's use (Martin 2007). If subjects are trained to perform functional tasks on certain objects, premotor regions become active during visual perception of these objects. Another intriguing finding is that motor and premotor systems are also active during mental simulation of events (e.g., during mental rotation of objects). Research on the mirror neuron system provides strong support for this view (Rizzolatti and Craighero 2004).

A fruitful approach for investigating the relevance of object-related eSMCs in human cognitive processing is to study the dependence of object recognition

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

on exploratory eye movements during free viewing of images. Using ambiguous images, a recent study showed that eye movements performed prior to conscious object recognition predict the later recognized object identity (Kietzmann et al. 2011). Other studies on natural vision have also shown that eye movements are highly predictive in nature (Hayhoe et al. 2011). These findings suggest that eye movements (like other movements) express the mastery of object- and intention-related eSMCs.

Patients with dysfunctional motor circuits provide the possibility to investigate the functional role and putative mechanisms of eSMCs. Our view predicts that motor dysfunction in such patients should result in a disruption of learned eSMCs and a lack of adaptivity and acquisition of novel eSMCs. This, in turn, should become manifest in altered perceptual processing and altered cognitive capacities. These hypotheses can be tested, for example, in patients with Parkinson disease. In these patients, modified eSMCs could be studied both as a function of dopaminergic medication (with and without levodopa) or as a function of deep brain stimulation in patients who have undergone surgical treatment. Indeed, substantial evidence suggests deficient acquisitions and calibration of eSMCs in patients with Parkinson disease. Thus, altered modality-related eSMCs have been observed in tasks involving rhythmic movement (Gulberti et al. 2015). Alteration of object- and intention-related eSMCs is suggested by difficulties in action naming observed in patients with Parkinson disease (Herrera and Cuetos 2012).

Intention-related eSMCs are also reflected in the "sense of agency," which refers to the experience of oneself as the agent of one's own actions and is, as such, a central constituent for human self-consciousness. Our sense of agency relies heavily on the experience of SMCs or action-effect couplings, but adds to these an additional, more complex level involving intention, experience, and identification. A large body of evidence suggests that the sense of agency depends on the degree of visuomotor congruence or congruence between predicted and actual sensory consequences of an action (David et al. 2008). Only in the case of congruence will an agent register a sensory event as caused by itself; incongruence would lead to the registration of the event as externally caused (Frith et al. 2000b). The more systematic the congruence between action and action effect, the better the agent's capacity to differentiate between self-produced and non-self-produced actions. Thus, we hypothesize that the emergence of the experience of agency is directly related to the mastery of action-effect couplings at the level of intention-related eSMCs.

Extended Sensorimotor Contingencies in Artificial Cognitive Systems

Implementing the concept of eSMCs in artificial agents provides the opportunity to verify the theory and challenge its limits as well as to explore extensions.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

In return, robotics may benefit from the virtues of the eSMCs account of cognition. Learning and using the structure of sensorimotor dependencies could endow a robot with an understanding of its different sensory modalities and an action-based perception of objects in its environment. Among other advantages, the object recognition and manipulation capabilities of eSMCs-controlled robots would be based on the experience that a robot would acquire with an object rather than the experiences of a human programmer, thus allowing the robot to discover relevant features itself and, hence, adapt better to unexpected changes in the environment or sensor failures. This synthetic approach requires casting the theory into computational models that are suitable for controlling robots. In the following, we introduce some of these approaches.

Modality-Related Extended Sensorimotor Contingencies

A general formulation of modality-related eSMCs can be given in computational terms by considering the probability of observing a particular sensory input after executing an action given the previous sensory state. This notion is described mathematically by the conditional probability distribution $P\big(o(t+1)\big|a(t),o(t)\big)$ over future sensory observations $o(t+1)$, when the current sensory state is $o(t)$ and action $a(t)$ is executed. Recently, we used an extension of this idea to study the acquisition of modality- and object-related eSMCs and to show how they can be used to control an agent's behavior (see Figure 11.2; Maye and Engel 2012). The robot observed the dependencies between changes in the accelerations and power consumptions of the motors and actions that led to changes in the movement direction. Using reinforcement learning, the robot preferentially reenacted those patterns which maximized a given utility function and thus learned to move smoothly and in an energy-efficient manner.

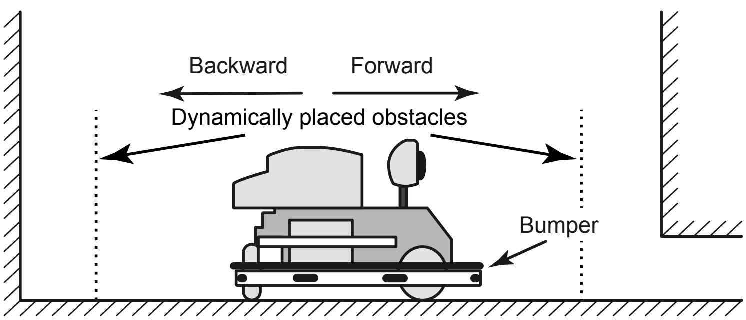

Figure 11.2 Experimental setup to study learning of eSMCs and deploying them for controlling behavior in an artificial agent. In the exploration phase, the robot developed knowledge of the size of its confinement, appropriate movements to escape collisions and energy-efficient movement sequences between the walls. This knowledge allowed the robot to react properly to obstacles that were placed dynamically in the arena. Reproduced with permission from Maye and Engel (2012).

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

It would be interesting to explore to what extent minimalistic systems with a Braitenberg vehicle-like control architecture could enact modality-related eSMCs. This architecture is characterized by a fairly direct coupling between the sensors and actuators of the agent, which is adjusted with respect to a goal function. As an example, O'Regan and Noë (2001) likened this to a tactical missile that is "tuned" to the eSMCs of airplane tracking, a task easily solved by a Braitenberg vehicle. The vehicle has extensive knowledge of possible input-output relations in its domain and uses this information to adjust its behavior to maximize the goal function. Even if the robot malfunctions, it still experiences sensorimotor regularities, but it cannot effectively exploit them to optimize its behavior.

Object-Related Extended Sensorimotor Contingencies

Object recognition poses a key challenge for artificial agents, for which there is currently no general solution. Classical attempts to solve this problem have aimed at recognizing objects in static images (e.g., snapshots from a camera). The pragmatic viewpoint invites us instead to consider vision as a mode of exploration, as an active process of making sense of the visual input. A paradigmatic example for applying this concept to robot vision is given by Bergström et al. (2011), in which a robot arm interacts with objects in a scene to disambiguate camera images of the scene acquired by the robot. A model of the visual appearance of the scene was generated using an expectation-maximization approach and belief propagation. This model produced a weak hypothesis about the centers of potential objects in the scene. The robot arm then pushed one of these centers, and the ensuing motion flow was analyzed for compatibility with the presence of one or more rigid objects. The result was fed back to the appearance model to improve the segmentation.

A similar concept was used by Björkman et al. (2013) to determine the shape categories of household objects. Object shapes were modeled by implicit surfaces that were adjusted by a regression of Gaussian processes. Visual information from a stereo camera was used to initialize the surface models. Regions of high uncertainty were then touched by a robotic hand, and the models were updated using tactile information. If the robot hand touched not only the object, but pushed it across the ground, information in the resulting motion flow could be used to classify the functional properties of the inspected objects and to group them according to their affordances (Högman et al. 2013).

In one of our own studies (Maye and Engel 2011), we addressed two of the aforementioned limitations of SMC theory: action selection and normativity of eSMCs. We developed a scenario in which object-related eSMCs were not only used to recognize objects but also to let the robot show corresponding behavior. We used a Markov model as described for modality-related eSMCs but extended it by attaching to each individual eSMC a utility value which captured the feedback that the robot received in the respective context. By giving

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

positive feedback for correct actions and punishing wrong actions (and neutral feedback if the action was indifferent), the robot was trained to associate the presence of different object types with specific actions (e.g., sorting cans and boxes by pushing them in opposite directions). The robot modeled the utility of these actions, $a(t)$, by observing reward or penalty probabilities conditional on the sensorimotor context, $c(t)$. This was done by counting the number of rewards received in a situation given by sensory observation $s(t + 1)$:

$$p_{reward}\left(s(t+1)\middle|a(t),c(t)\right)=\frac{N_{reward}\left(s(t+1),a(t),c(t)\right)}{N\left(a(t),c(t)\right)}. \tag{11.1}$$

Another histogram was used for computing the probability of punishment, $p_{penalty}$.

Assigning reward probabilities to eSMCs can be regarded as a method for structuring the sensorimotor knowledge that the agent acquires. It establishes relations between eSMCs that have a similar ecological value. Sensorimotor knowledge could be further structured by considering temporal information, for example, by grouping eSMCs that have been observed in the same time interval and which are therefore probably related.

In the terminology of SMC theory, viable actions result when the agent exercises its mastery of eSMCs. However, SMC theory does not suggest a strategy for choosing which set of learned eSMCs to activate. This is why our model uses the value system described above, which structures and weighs eSMCs according to their expected reward. The decision schema (Figure 11.3) constitutes a simple switch from exploration to exploitation: If there are actions that have never been tried before in the current context, one of these actions will be

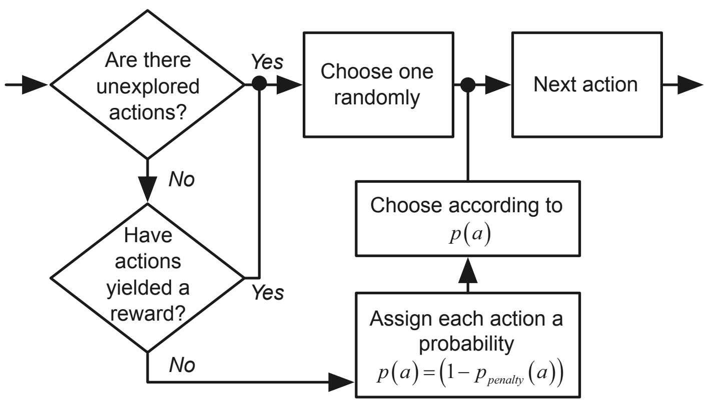

Figure 11.3 Action selection schema that optimizes the robot's behavior based on the observed utility of different eSMCs. Reproduced with permission from Maye and Engel (2011).

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

chosen. Otherwise, an action which most likely will yield a reward is chosen. If only actions which yielded no or negative rewards (i.e., penalties) have been observed in the past, these actions are assigned probabilities according to their least negative effect, and one of them is chosen according to this probability distribution.

Intention-Related Extended Sensorimotor Contingencies

Temporally extended sensorimotor interaction patterns are not only suitable characterizations of the agent's perceptual experience, they also enable the prediction of future experiences and the planning of proper action sequences. The basic idea is to search for partial matches between learned eSMCs and the current sensorimotor context, and then build a probability distribution over future experiences from these eSMCs beyond the point to which they match. We explored this idea in a study where we intentionally introduced a delay between issuing a motor command in a robot and showing the corresponding change in the movement direction (Maye and Engel 2013). This required the agent to predict potential collisions and to adjust its motor commands accordingly.

Taken together, the approach to model eSMCs using Markov processes of different history lengths led us to the hypothesis that the different levels of eSMCs may involve different timescales over which regularities in the sensorimotor patterns are captured (Maye and Engel 2012). Accordingly, from a modeling perspective, there would be no principle differences between the three eSMCs levels introduced here. Rather, these levels would describe different ranges on a continuum of sensorimotor context scales. This would allow eSMCs to be nested, with eSMCs that span shorter timescales forming the building blocks for eSMCs with longer timescales (Figure 11.4).

Intention-related eSMCs can also be considered from the perspective of the related concept of affordances (Gibson 1979). Knowing an object's affordances allows continuous sensorimotor interaction to change the state of the environment in a predictable and goal-directed manner. A primary application scenario for this approach is tool use, which was explored by Sánchez-Fibla et al. (2011) using a robot that learned to push objects of different shapes along a predefined trajectory or to a given target position and orientation. Sánchez-Fibla et al. introduced the concept of affordance gradients, which are object-centered representations of the consequences that an agent's actions may have.

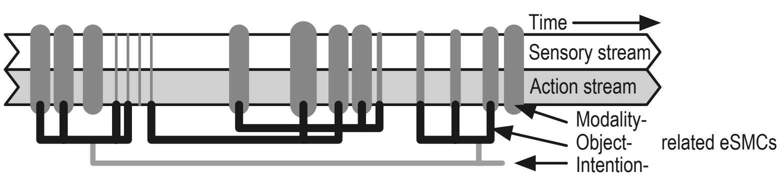

Figure 11.4 Schematic of the timescales that are captured by different types of eSMCs.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

These representations are mathematically formulated by vector fields which allow the robot, by means of interpolation, to make predictions about the consequences for actions that have not been tried before. In the case of push actions, affordance gradients can be formally defined by a triplet consisting of a gradient representing the object's shape and two vector fields describing the displacement and rotation of the object when pushed from a given position and in a given direction. This triplet is learned in an exploration phase in which the robot can try out the consequences of pushing the object from various positions and at different angles. The forward model that the robot acquires in this phase can then be used iteratively to develop an action plan to achieve a goal and to update this plan during execution (Sánchez-Fibla et al. 2011).

Major challenges for eSMCs-based approaches to robot control result from the size of sensorimotor spaces and the complexity of the relations between actions and sensations. Modern robots feature many degrees of freedom and large numbers of sensory channels, composing potentially infinite sensorimotor spaces. This is an issue for any robot control architecture. Squashing action space and quantizing sensory signals are frequent solutions for curbing the sensorimotor space. The second issue, the complexity of sensorimotor relations, is more specific to eSMCs-based approaches. These relations can be simple co-variational or causal; they can span different timescales, be described at different levels of abstraction, and be hierarchically nested. Another challenge may be the development of neurobiologically plausible eSMCs models which would link insights from empirical research with results from synthetic approaches. Corresponding models are only beginning to emerge.

Outlook

The approach that we have pursued here departs from the classical notion that presumed "higher" levels of cognition might be fundamentally different from presumed "basic" levels of sensorimotor coordination. It also rejects the notion that architectures embodying complex cognitive functions (e.g., "higher" processing centers in the human brain) differ in principle from modules for more basic functions. Conceptually, the eSMCs approach diverges from much of what has been characteristic of the classical cognitivist framework, moving toward an enactive, embodied view of cognition and a much more dynamic and holistic view on the underlying architectures.

The central role that this approach assigns to action places it in the family of embodied approaches to cognition. It shares ideas with embodied approaches in cognitive linguistics, autonomous robotics, and ecological psychology as well as with enactivist models. Like SMC theory, enactive approaches to cognition (Varela et al. 1992; Thompson 2007; Di Paolo et al. 2010) conceive of cognitive states as interaction processes of embodied agents and the environment rather than brain states. Enactivism, however, entertains the notion of the

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

body as a self-constructing process that gives rise to norms and the concept of autonomy. The act of striving to maintain autonomous identity can explain why agents have interests in the world and provide a touchstone for developing a concept of sense-making. In this respect, enactivism may be able to provide an answer to the question that SMC theory seems to leave open; namely, what constitutes mastery of eSMCs and how can an agent achieve it?

We suggest that the eSMCs approach may provide new perspectives for robotics. In contrast to top-down architectures used in industrial robot applications, where the relevant knowledge is programmed into the robot, our approach advocates letting the agent acquire knowledge by embedding it in an environment which it can explore and manipulate. A consequent interpretation of the eSMCs approach suggests that the discovery and appropriate utilization of regularities in the robot's sensorimotor interactions with the environment would vest it with a primitive form of cognitive states. Notably, the relevant sensorimotor dependencies comprise properties of the embodiment as well as the environment. As such they can be seen to be a control mechanism for morphological computations (i.e., cognitive functions which are carried out through particular properties or suitable arrangements of body parts) as well as for an extended mind that employs the body and the environment to realize cognitive functions. This places our approach in the perspective of active externalism, as proposed by Clark and Chalmers (1998).

One of the challenges for SMC theory is to understand how abstract concepts might be grounded in sensorimotor processes. Embodied theories of cognitive linguistics consider the human body as a source of meaning, concept articulation, and even reasoning (Gallese and Lakoff 2005). Rather than abstract rules of logic, thought processes employ comparisons to sensorimotor experiences, thereby grounding our understanding of the world and the contents of our verbal communication in the physical reality in which we are embedded. Solving an exercise, for example, is thought to involve patterns used to overcome obstacles, such as considering the problem from different perspectives, selecting an appropriate solution method, progressing along the action plan, or determining the suitability of the result. This idea is compatible with an extended version of SMC theory in which eSMCs define the contents of perceptual experience and constitute the substrate for higher-level cognitive processes. This requires learning mechanisms that enable the cognitive agent to generalize and abstract from sensorimotor interaction patterns.

A highly interesting issue is whether the approach proposed here can be extended even further to ground basic aspects of social cognition. It seems plausible to assume that agents deploy learned action-effect contingencies in social contexts to predict outcomes of their own and other's actions. This would allow an agent to coordinate with actions of other agents and enable effective coupling of agents in social contexts. Accordingly, social interaction may strongly depend on the dynamic coupling of agents, and this interaction dynamics may provide important clues to social cognition. This view shares aspects with the

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

interactionist concept of social cognition proposed by Di Paolo and colleagues (e.g., Di Paolo and De Jaegher 2012). They argue for an extension of an enactivist position and emphasize that sense-making in a social environment occurs in a participatory manner and that central aspects of cognitive performance are inherently relational. Furthermore, our concept agrees well with the joint action model by Knoblich and colleagues (Knoblich and Sebanz 2008), who predict that shared intentionality can arise from joint action.

Finally, it might be interesting to consider implications of our approach for understanding consciousness (cf. Seth et al., this volume). In the original version of SMC theory, a key hypothesis is that SMCs can account for qualia as the basic features of phenomenal experience, and that differences in the qualitative character of perceptual experiences result from differences in the relevant SMCs (O'Regan and Noë 2001). What SMC theory leaves open is how more complex contents of conscious experience can be aggregated, selected, and structured. We believe that it would be worth exploring whether an extended sensorimotor account, as discussed here, might help to account for the structure of conscious experience beyond the basic level of sensory qualities.

Acknowledgments

This work has been supported by the EU (FP7-ICT-270212, ERC-2010-AdG-269716).

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.