Instrumental Representations of Sensorimotor Control: Representations at Intermediate Level

**Hsi-wen Daniel Liu (hwliu@pu.edu.tw)**
Center for General Education, Providence University
Taichung, Taiwan

**Abstract**

In cognitive science, computation is largely accompanied with representational theory of mind. Yet, it remains unclear whether this companionship also appears in the realm of sensorimotor control. Grush’s (2004) and Pezzulo’s (2008, 2011) account of anticipatory representations provide a limited answer, as they are only suitable for forward models, but not the entire sensorimotor control. Rescorla’s (2016) representational explanation for sensorimotor psychology addresses several intentional states considered in Bayesian inference and optimal modeling. However, the above accounts do not explain how motor commands are produced and chosen in the course of sensorimotor control for maintaining accuracy of goal-achievement. The present paper aims to explain it with a representational account by considering instrumental representations of sensorimotor control, which appear at the intermediate level and are exemplified by motor commands and costs. Such representations do not presume decouplability, as they need to run on-line in the maintenance of accuracy.

**Keywords:** Sensorimotor control; representation; optimal feedback control; Bayesian decision theory.

**Introduction**

Within cognitive science there is a long-standing dispute between different paradigms concerning the role of representation in computation. Classical cognitive science understands cognition in terms of computation over mental representations, and considers the role of cognition to be deriving world-models that provide a database for thinking, planning and problem-solving. Decouplability between representations and their immediate environment is taken as intrinsic to representation. By contrast, a ‘pragmatic turn’ raises the ‘action-oriented paradigm’ that considers cognition to be providing skillful know-how in situated and embodied actions (Engel et. al., 2013). Clark (1997) raises the notion of action-oriented representations, which do not presume decouplability, in his ‘minimal representationalism’. Action-oriented representations, yet, are mostly applied to reactive motor activities, regardless of various models of motor control.

The computational theory of mind is largely accompanied with a representational theory for explaining a variety of faculties, including perception, language, thinking, and problem-solving. However, the role of representation is highly debated for the faculty of motor action. Within models of motor control, the forward model is firmly associated with a role of representation, in the notions of emulating representation or anticipatory representations (Grush, 2004; Pezzulo, 2008, 2011). For those representations, decouplability is claimed to be an intrinsic property. In addition, Bayesian models of motor control are seen as highly associated with a robust notion of mental representation (Rescorla, 2016). Decouplability is also seen as intrinsic to that notion of representation (Haselager et al., 2003). As to other models of motor control, however, the role of representation is unclear. The computational theory of sensorimotor control has been established (Wolpert and Ghahramani, 2000; Franklin and Wolpert, 2011; Orbán and Wolpert, 2011). It is suggested that motor control is conformed to the notion of pragmatic representation, as different from that of semantic representation (Jeannerod, 2006). Yet, questions remain as to what the pragmatic representation is and how it is related to computational models of motor control.

Given that computation of cognition is an explanatory account of cognition (Marr, 1982), an explanation can be found in the computational theory of sensorimotor control, as shown above. The sensorimotor control, as seen in the determination of an appropriate motor command, is considered to be fundamentally a decision process (Körding and Wolpert, 2006). The decision is to choose an appropriate motor command in order to achieve a given goal. The present paper contends that the explanation of the sensorimotor control is to explain why and how the decision could determine an appropriate motor command that turns out to achieve the goal. The explanandum is the way in which the goal is achieved by choosing an appropriate of motor command. The present paper characterizes this way of goal-achievement in terms of the end-means relation. The goal of a motor task is the end, and a chosen motor command is its means. Humans can rationally consider the appropriate means for an end; similarly, the sensorimotor system can choose an appropriate motor command for a given goal. This similarity, yet, is subject to two caveats. Firstly, while in economic decision-making or daily affairs in general, the rationality proceeds at the personal-level, the sensorimotor processes largely operate at subpersonal-level. Secondly, while characterization of human rationality need not be put in terms of probability, the actual performance of sensorimotor control is found to be very close to descriptions made with Bayesian decision theory (Körding and Wolpert, 2006), a theory with probabilistic measures.

Explanation of sensorimotor control can be pursued in anti-representationalist accounts, which hold certain explanatory perspectives (Turvey and Fonseca, 2009). Motor control, as contended in such accounts, is determined by interactions between the neural system, body, and the environment. Different from such accounts, the present
paper preserves an explanatory role for representation in terms of a novel notion of representation—instrumental representation—and regards the above interactions as complementary to computation and representation. Arrangement of that complementarity is for two reasons. One is the overall success of the computational approach to sensorimotor control (Franklin and Wolpert, 2011; Orbán and Wolpert, 2011; Todorov, 2004; Wolpert and Ghahramani, 2000). The other reason is that problems of sensorimotor control, to put it in Clark’s (1997) terms, are not ‘representational-hungry’. The neural system of sensorimotor control has tight interactions with body, and the environment. The notion of instrumental representation differs from the mainstream conceptions of representation in that it does not presume decouplability from the environment. Putting above two reasons together has an implication in the level of instrumental representations. Like that the present account of representation stands between the classic account of representation and those anti-representationalist accounts, instrumental representations can be conceived of as standing at the intermediate level of the mind.

The present paper aims to raise a representational account of the sensorimotor control. This aim is to be achieved by explaining computation of sensorimotor control in terms of representation, on grounds that computation is a way to explain the mind, in the first place. Section two discusses computational explanations of sensorimotor control in terms of end-means relations. Section three specifies the notion of instrumental representations for explaining the sensorimotor control, on grounds of computational explanations of sensorimotor control.

**Computation in the Sensorimotor Control**

The Bayesian decision theory in sensorimotor control holds a computational perspective with two components—estimation of environmental and bodily conditions, on the one hand, and decision made upon motor commands for the most desirable performance, on the other. To put it in an epistemological dichotomy, the former component is to measure environmental and bodily facts, while the latter one is to evaluate motor actions. To put it otherwise, the former is close to perception, while the latter to decision-making.

**Uncertainty**

Measurement of states in the sensorimotor system is affected with various factors of uncertainty, and consequently it cannot be accurate like that we manage to measure the length of an object left on the table with a ruler. Sensory signals of the environment have inherent delays, which affect signals at all stages of sensorimotor system from the afferent (coming-in-from-the-outside) sensory information, to conduction along the neural fibers, together with the complexity of processing (face recognition being longer than motion perception) and ‘slower’ modality (vision being longer than proprioception). It can be said that we ‘live in the past’ by accessing the ‘out-of-date information’ about the common world and our own bodies (Franklin and Wolpert, 2011, pp. 425–6). In addition, the nervous system is corrupted by noise, which also affects sensorimotor control at all stages, from the reception of sensory information, to planning, resulting in variability in movement endpoints. Noise, hence, contaminates our observation of the sensorimotor system internally and externally, by affecting estimation of body states and world conditions (ibid., p. 425). Noise in motor commands, in particular, increases proportionally to the size of their signal (the so-called ‘signal-dependent noise’). Different motor commands would incur different degree of variability in the resulting endpoints. To put it more specifically, the motor performance is subject to speed–accuracy trade-off described by Fitt’s law (Harris and Wolpert, 1998). Noise in the sensorimotor system corrupts not only estimation of internal and external states, on the one hand, but also performance of motor actions, on the other.

Apart from delay and noise, there are more factors of uncertainty residing in the sensorimotor system. Environmental conditions are constantly subject to change, for example, forces imposed upon the arm in the reaching movement (Shadmehr and Mussa-Ivaldi, 1994). It is also uncertain as to which actions or tasks would be beneficial in the real world outside the laboratory (Franklin and Wolpert, 2011). Furthermore, our motor system is non-stationary, for example, the length and weight of our limbs are changing when we grow up, and our muscles are getting stronger with larger forces (ibid.). Those uncertain factors would constantly contaminate information of the sensorimotor system, information which consequently needs to be optimized in order to achieve the given goal.

Given the aforementioned factors of uncertainty, the computation of sensorimotor system needs to optimize its information, in order to find the best resolution in view of goal-achievement. The sensorimotor system should maintain optimal estimation of world and body states, and should conduct optimal choice of motor commands. This need of optimality, in both state estimation and action choice, introduces a version of instrumental rationality that is immanent in the sensorimotor system at the subpersonal-level.

**Decision on Grounds of Utility**

By contrast, another version of instrumental rationality turns up in the sensorimotor system’s course of decision-making for optimal choice of a particular motor command, given a particular goal. Various particular goals can be given to a sensorimotor system, which are ends for the system to seek appropriate means for their fulfillment. Whenever a means is determined, a decision is made for attaining the end. The degree of fulfillment can be evaluated in positive terms, such as benefit, reward, utility or prospect; or alternatively in negative terms, such as loss or cost (Körding and Wolpert, 2006). The degree of fulfillment is a foundational notion, the present paper proposes, for explaining the sensorimotor control. In the realm of sensorimotor control, the utility of a
motor movement is the decision theory's way of evaluating the fulfillment of the goal. The fulfillment to a higher degree would receive the evaluation of a higher utility.

According to the decision theory, the expected utility of an action is defined quantitatively in terms of probability, as follows:

\[ E[\text{Utility}] = \sum_{\text{possible outcomes}} p(\text{outcome} \mid \text{action}) \cdot U(\text{outcome}) \]

where \( p(\text{outcome} \mid \text{action}) \) is the probability of an outcome given an action, and \( U(\text{outcome}) \) is the utility assigned to that outcome. An action is chosen for maximizing the expected value, which is put in terms of utility. A decision made in this way is defined to be a rational choice (Körding and Wolpert, 2006). This is a normative theory that defines the way in which people should behave; that way is put in terms of rationality. In other words, the rationality of action urges that people in their actions should pursue a higher degree of utility. When it is put in the context of sensorimotor control, a utility function evaluates how well a movement is performed. This way of evaluation quantifies, in terms of utility, the total desirability of a chosen movement. In addition, the decision theory is also considered to be a descriptive theory by assuming that people act rationally. The rationality assumed in this theory explains why people behave in the way they do. In fact, empirical findings indicate that the Bayesian decision theory shows successfully how people actually perform their sensorimotor control (ibid.).

The decision theory measures a motor movement, in terms of utility (measuring positively) or cost (measuring negatively), by evaluating its degree of fulfillment, that is, how well it achieves the goal. The end-means relation is assumed in the notion of fulfillment. The utility, in a descriptive term, defines how well a means attains the end.

The measurement of a cost depends on the immediate conditions of all relevant factors in the sensorimotor processing. Specifically, it depends on current states of the sensory system and properties of the to-be-chosen motor commands. The considered states include body states and the environmental conditions, for example, joint angles and velocities, and positions of relevant objects. Properties of motor commands can be measured with different emphases, for example, the jerk, torque change, energy, time, variance, of the to-be-chosen motor commands. The measurement, whatever the emphasis, takes the form of ‘minimize \( X \)’ (Todorov, 2004). It is to minimize the size of relating factors, for example, the energy-to-be-consumed of the motor commands. In other words, the measurement of cost is sensitive to the size of the system’s immediate response to its (bodily and environmental) conditions, which are embodied properties (i.e. jerk, torque change, energy, time, variance, etc.) of the sensorimotor system. To be noted, measuring the cost is not purely an internal matter, but is to be put in real situations, which consist of bodily and environmental conditions.

The Basics of Motor Commands
Choice of motor commands, as above considered, is to be managed after a series of motor commands is organized. The computation of sensorimotor control is required to explain how to transform a higher-level goal into a series of motor commands, which are strictly constrained in the embodied sensorimotor system. This explanation consists of two parts: coordinate transformation and modular structure, both of which assume the end-means relationship. The former, coordinate transformation, is to convert sensory signals of the goal into motor commands. Sensory signals consist of visual information of the object in the goal-state together with the signals relating to the posture of bodily parts (hands, arms, shoulders, head, and eyes). Those signals need to be transformed into a set of motor commands that would bring about the goal-state when they are performed. This task is named to be sensorimotor transformations, which are accomplished with the mapping of a three-layered neural network: from the input layer of posture signals, to the intermediate layer that consists of population codes, to the output layer of the motor command that consists of the change in joint angles needed for the task (Pouget and Snyder, 2000). This is a way of reverse engineering, which is called the inverse model, as its direction of transformation is opposite to the forward model of motor control (Wolpert and Ghahramani, 2000). The sensorimotor transformations and their products can be regarded as means for the end of achieving the given goal.

As for the construction of movements, it remains controversial as to how much of movement might be controlled by modular processes (Zelik et al., 2014). Insofar as modular organization is applicable, various complex motor movements are constructed through flexible combinations of a limited number of modules, in order to simplify computation by reducing degrees of freedom (Jing et al., 2004). In other words, a complex motor command is organized with a combination of motor primitives. A motor command consists of a series of muscle activations for the needed changes of joint angles. With a study in the vertebrate spinal cord, it is shown that a complex motor command is produced by combining a few motor primitives, which are ‘unit burst generators’ organized in the spinal cord. Each burst generator is to control the activation of a small group of synergistic muscles, or motor synergies (Tresch et al., 1999). Such a combination is a modular representation (Musall-Ivaldi, 1999). A motor task is to produce a motor command with an appropriate modular structure. The motor commands produced in the above way are basic elements for the choice of cost function (Wolpert and Ghahramani, 2000). Given that a single motor command has the above modular representation, it would naturally be questioned as to how the series of motor commands leading to the achievement of a goal are organized.

The notion of modular structure appearing in sensorimotor control does not strictly follow Fodor’s (1983) sense of modularity. Firstly, Fodor (1975) argues that
mental modules are combined with a language of thought (LOT, Mentalese). The modular structure of movement generation, however, does not seem to follow the structure of LOT, in the following two aspects. The weighted and graded combination of modules (Jing and Weiss, 2005) does not show the structure of LOT. Furthermore, generation of movements out of modular organization has practical limitations, as is found in generation of movements for diverse locomotor behaviors; sufficient flexibility needs a further basis of coordination (Zelik, 2014). There may be no clear distinction between planning and execution, because coordinated motor movements may emerge out of real-time optimal feedback control (Todorov and Jordan, 2002), as discussed below. Emergence of coordinated movements out of interaction with the environment makes generation of a movement deviant from the modular organization. To summarize, the modular organization is only loosely applicable in sensorimotor control.

Coordination During Execution

The sensorimotor control on the basis of decision, cost and optimality, as aforementioned, can be managed, to a certain degree, in abstraction from the real situations. A version of optimality can be so pursued, by way of open-loop optimization, with detailed planning in advance of execution. The accuracy can be maintained to a certain degree, yet with serious limitations. The application of optimal principles seeks average optimality over previous performance. As it is detached from the real situations, the sensorimotor control is like playing “a prerecorded movement tap” and consequently the given goal is treated like a laboratory task. It would be unable to encounter the trial-to-trial variability in the real situations (Todorov, 2004, p. 2).

Such an abstract way of sensorimotor control is rather like the maintenance of thought, as it can run in abstraction from the real environment. It has the merit of a Popperian creature, that is, planning internally for a best solution of the considered problem before its execution. Decision can be made for a relatively optimal performance. Yet, an important way of sensorimotor control would be completely missing—coordination during execution. Understanding how this is done is a central problem in motor control for nearly 70 years (Todorov and Jordan, 2002).

The coordination during execution presumes optimal feedback control—the optimal control with on-line sensory feedbacks. It does not plan a desired trajectory before execution, but maintain the coordination on-line in response to all the task-specific contingencies in the real situations. Coordination in the sensory system is highly important as such a system is highly redundant, with a high number of ways over the combination of motor activations, and full of a variety of uncertain factors such as noise and delay, as aforementioned. The optimal feedback control produces “continuous trajectory of movement in response to contiguous stream of sensory input (Körding and Wolpert, 2006)” Cost are continuously generated with on-line control of sensory feedbacks. The evaluation of cost is put in terms of ‘cost-to-go’—the continuous and integral summation of costs (Todorov, 2004). The optimal feedback control responds to real situations of the body and the environment, and fully manifests the continuous way of motor decision in fast-changing conditions. This cannot be done in a detached model.

When the optimal feedback control operates as a way of coordination, motor synergies and the achievement of the given goal emerges. It only asserts what to achieve, without dealing with the how question in detailed. After the goal is given, the optimal feedback control can keep on seeking an appropriate resolution because of its coupling with the plant, in a way like the operation in the dynamic systems view. The stages of planning and execution are not separate (Todorov and Jordan, 2002).

The success of the coordination, in the optimal feedback control, relies on a normative property of the end-means relation, which is immanent in the sensorimotor system. That is, the sensorimotor system as a system with the end-means relation would seek appropriate means for its given end. This property of the optimal feedback control can be seen as endowed in evolution. It is a process that makes possible the emergence of coordination in the sensorimotor system. After it encounters its contingencies, including the fast-changing environment (with body) with various uncertain factors, the sensorimotor system operates like the way present in the dynamic systems view. In that dynamic relations, motor synergies will eventually emerge.

Instrumental Representations

Representations in the sensorimotor system are generally divided into three types: end, means, and cost. The goals in the sensorimotor system are regarded as ends, and motor commands are means for attaining those ends. An end represents a world state that is to be attained. A common element in the latter two types is the end-means relation: a means can attain its end, but therein lies a certain cost; in this sense, the means and the cost are called ‘instrumental representations’. The means are first-order representations, which represent ways to attain their relating ends. Costs, by contrast, are second-order representations that represent prospects of the relating means in the processes of attaining their ends. Furthermore, those three types of representations hold different foundations of representation qua representation. An end represents a to-be-attained future state, in which the end refers to the to-be-attained future state on grounds of similarity. By contrast, the means represents the to-be-attained end in the way that the end is to be attained through the means. The cost, in addition, evaluates the prospect of a means in the way to attain its end. Finally, whether a means represents an end successfully, is measured in terms of accuracy (as opposed to truth), that is, accuracy with which the means attains the end. As accuracy can be measured with various degree, misrepresentation is subject to different degree.
Representations as Stand-Ins

Representation, generally speaking, is something R that stands in a system S for something else E. That is, R is a surrogate in the system S for E. Representation sometimes serves as a Poppian creature: something that can run internally in a system before it is actually carried out. Representation in this sense simulates what will actually happen. It (R) is a surrogate of E’s actual performance. Based on Cummins (1996), a surrogate in this sense refers to its target with the information of its content. A city map refers to the city streets according to structures of the map. A map user can simulate a feasible route in the map without actually walking in the street. Further, in order to account for the wits in the sensorimotor control, the notion of representation can be extended from the predicative relation to the end-means relations, insofar as they bring about a certain target (the end) with recourse to a certain content (the means). That is, the predicative representation Rb in a descriptive relation describes E. Rb refers to the target of E, and the content of Rb describes E. By contrast, the instrumental representation Rm in an instrumental (that is, end-means) relation brings about E, and the content of Rm guides the system S to reach the goal-state E. Furthermore, a different surrogate Ri in the system S would be likely to bring about a different state, as opposed to E. Rm and Ri, hence, are alternative means generated in the system S, alternatives which can be compared for a higher degree of prospect P to bring about the end-state E. The prospect P is a second-order instrumental representation, as it evaluates the degree in which certain means would bring about the end. Rm and Ri are exemplified in the instrumental system by motor commands, and the prospect P is exemplified by costs. Thus, instrumental representations are genuine representations because they stand in a system (S) for something else (E) that they bring about. The instrumental representations stand in the sensorimotor system, rather than merely serving as physical components of a mechanism, because they always have alternatives to be chosen in their way to bring about an end. The above profile will be discussed more specifically below.

The decouplability between representations and their environments is but a particular case for a system capable of generating alternatives. The system produces something else apart from a fixed representation. In a dark night, as I encounter a distal horse (a target) I may consider it to be a cow; in a darker night, I may even feel that it is like a unicorn. For achieving a given goal, sensorimotor representations produce various motor commands that are all likely to achieve the goal. The inverse model of a sensorimotor system transforms a single goal-state G into various motor commands, which would all be likely, to a certain degree, to achieve the same goal G. The sensorimotor system always has alternatives to be chosen, in a decision for an optimal motor command, as indicated by the redundancy present in the musculoskeletal system. This is unlike a physical causal relation, such as knocking a group of billiard balls with a single ball, where a move in a particular circumstance will determine a single result. Alternatives, as aforementioned, are made by the system S in a non-physical connection, when the system S encounters a fixed condition (e.g. encountering a horse, or given a particular goal-state). In the sensorimotor system, the inverse model generates various motor commands in a non-physical connection, which is non-physical insofar as it is computational. The mechanism on the basis of Bayesian decision theory, in addition, makes a choice among redundant motor commands for optimality with a lowest cost. The choice from alternatives justifies that the sensorimotor control is not a ‘merely physical’ device.

Decoding of sensorimotor representations in the sensorimotor control is grounded on the use of those representations in the way to achieving the goal. Therein, the pragmatic dimension of sensorimotor control is considered in terms of end-means relations. The use consists of estimation over environmental conditions for applying them and choice between them, as manifest in the application of Bayesian decision theory in sensorimotor control (Körding and Wolpert, 2006). The generated motor commands in the inverse model, in addition, are made with alternatives, which are available for choice in their use dedicated to the achievement of the goal.

The motor commands are genuine representations because their way of bringing about goal-achievement is internally rich. Based on Cummins’ (1996) notion of representation, the goal-state is the target while the content in use is the information employed, serving as guidance of the instrumental control, for achieving the goal. Specifically, the estimation of environmental conditions in the Bayesian model of sensorimotor control presumes the need of achieving the goal, and so is the model of optimality. Thus, the sensorimotor representations are internally rich, even compared to the classic representations, which are dubbed as representation-hungry. For example, the Bayesian inference, in order to estimate external conditions out of noise, needs to take account of priors, that is, previous experiential outcomes. This makes the Bayesian model of the sensorimotor control even no less emphasized on internal wits than the computation related to the classic theory of representation. In addition, the decision made in relation to cost, as discussed previously, making a choice from various alternative motor commands. Furthermore, the on-line measurement of cost adds more wits on the top of the computation with open-loop optimization. The instrumental representation based on alternatives of action command, together with their accompanying estimation and choice, provides computation with internal alternatives, evaluation and inference. Such a way of representation is internally rich.

Representations can serve as stand-ins of a system without being predicative. Representations are stand-ins for the existence of certain states, or for those states’ activities. The former relates to representational production and the latter representational consumption. As considered above, the stand-ins can be instrumental, and consequently need
not be built with the function of imagination, on grounds of which counterfactual representations are possible. Imagination is surely a characteristic of human cognition. Cognition, however, can have other characteristics, for example, instrumental allocation, that is, arrangement of end-means relations. The end-means relation, for a given end, need not consist of a single string of causal chain, as it can produce alternative means for the same end. Those alternatives can be evaluated with different degree of prospects for attainment of the end. The choice from alternatives justifies the cognitive bearing of the instrumental representations.

Before concluding, it should be noted that instrumental representations can have a combinatorial structure only in a loose sense. Instrumental representations of sensorimotor control do not follow the LOT, basically because its modular organization is only loosely applicable, as discussed in a previous section. As a consequence, the combinatorial structure—with which mental representations can be generated recursively and systematically from primitive states—would not be generally salient in the realm of sensorimotor control. In particular, the costs of motor commands are continuously generated in on-line feedback control, as manifest in ‘cost-to-go’—the continuous and integral summation of costs, as aforementioned. With this way of computing costs, the consequently chosen motor commands can only have modular structures (if there are) in a loose sense.

Conclusions

Computation of sensorimotor control employs instrumental representations—representations with end-means relations—as exemplified by motor commands and costs. Motor commands represent ways to achieve the goal, and costs represent prospects of goal-achievement. They are intermediate-level representations, because the computation of motor commands does not rely on reactive machinery, and because they appear at the sub-personal level. Although they have modular structures, sensorimotor representations are initiated continuously and connected integrally. In order to maintain accuracy of goal-achievement, the sensorimotor system needs on-line incorporation of sensory feedbacks, and consequently sensorimotor representations cannot be detached from the body and the environment.

References


