An Analysis of the Ideomotor Principle and TOTE

Giovanni Pezzulo[1], Gianluca Baldassarre[1], Martin V. Butz[2], Cristiano Castelfranchi[1], and Joachim Hoffmann[2]

[1] ISTC-CNR, Via S. Martino della Battaglia, 44 - 00185 Rome, Italy
{giovanni.pezzulo,gianluca.baldassarre,cristiano.castelfranchi}@istc.cnr.it
[2] University of Wurzburg, Rontgenring 11 - 97070 Wurzburg, Germany
{mbutz,hoffmann}@psychologie.uni-wuerzburg.de

Abstract. What does it mean for a system to be goal oriented? In this paper we investigate how goals are represented and how they activate actions. We review the main philosophical and psychological assumptions about the ideomotor principle and we compare it with the TOTE model in cybernetics. We also present three computational architectures that implement goal orientedness, discussing their main peculiarities and limitations with respect to the ideomotor principle and TOTE.
Keywords: Ideomotor principle (IMP), TOTE, teleonomy, goal, feedback, anticipation, search, reaching

1 Introduction

Intelligence of complex organisms such as humans and other apes resides in the capacity to solve a problem by working on internal representations of problems, that is by acting upon "images" or "mental models" corresponding to word states on the basis of simulated actions ("reasoning"). Recently, many converging evidences in psychology and neurobiology indicate a crucial role of anticipatory representations for many cognitive functionalities such as visual and motor control [13]. As suggested by the discovery of mirror neurons [34], representations are often action-related and are thus grounded on the representations subserving the motor system. Barsalou [2] as well as Grush [11] try to provide unitary accounts of these phenomena proposing *perceptual symbol systems* and an *emulation* theory of cognition. Moreover, anticipatory functionalities are starting to be explored from a conceptual point of view [6] as well as from a computational point of view [7, 5, 41].

Several of these anticipatory mechanisms can be dated back in their origin for decades if not centuries. The *ideomotor principle (IMP)*, which was proposed multiple times during the 19th century [12, 19], hypothesizes a bidirectional action-effect linkage in which the desired (perceptual) effect triggers the execution of the action that previously caused that effect. The *test operate test exit (TOTE)* model of cybernetics [24] proposes goal-oriented action control.

The first aim of this paper is to provide a comprehensive introduction to both of these concepts and to highlight their similarities, differences, and drawbacks in explaining anticipatory goal-oriented behavior (sections 2-4).

The second goal of the paper is to analyze, at an abstract level, three computational architectures which implement several features of IMP and TOTE in distinct ways (Section 5). This analysis is intended to exemplify and clarify the principles underlying the IMP and the TOTE, and to provide a starting point for future research in the investigation of anticipatory goal-oriented behavioral mechanisms. A final discussion concludes the paper with an outlook of the next, most imminent challenges (Sec. 6).

2 The Ideomotor Principle

According to the IMP [14, 17, 19], *action planning takes place in terms of anticipated features of the intended goal.* [10] underlines the role of anticipation in action selection: *a current response is selected on the basis of its own anticipated sensory feedback.* The *Theory of Event Coding* [18] proposes a common coding in perception and action, suggesting that the motor system plays an important role in perception, cognition and the representation of goals. The theory focuses on learning action-effect relations which are used to reverse the linear stage theory of human performance (from stimulus to response) afforded by the sensorimotor view. Neurobiological evidence for common mechanisms in perception and action are reported in [20, 33]. In this respect, Gallese [9] suggests that *the goal is represented as a goal-state, namely, as a successfully terminated action pattern.*

In the ideomotor view, in a sense, *causality, as present in the real world, is reversed in the inner world.* A mental representation of the intended effect of an action is the cause of the action: here it is not the action that produces the effect, but the effect that produces the action. [25, par. 21.5] describes an "automatic mechanism" realizing this principle (see Fig 1): when the features of, say, an apple are endogenously activated, an automatic mechanism is oriented toward (seeing or grasping) apples teleonomically.

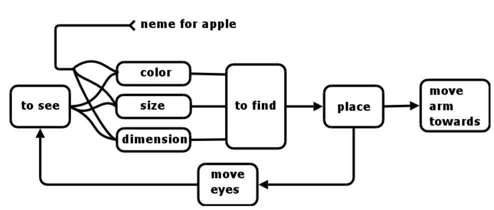

Fig. 1. The "automatic mechanism" in [25, par. 21.5].

The main constituents of the IMP The comparison of the presentations of the IMP by these various authors allows identifying three main constituents of the principle itself. These form the core of the principle and abstract over minor details and different emphases stressed by the various authors. The three constituents are now analyzed in detail (see Fig 1).

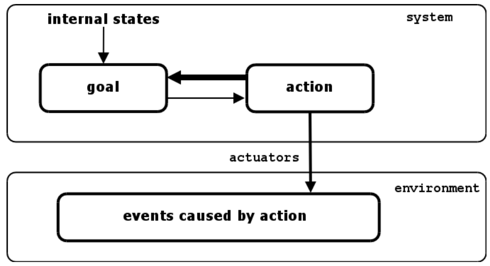

Fig. 2. A scheme that represents the main features of the IMP. Thin arrows represent information flows, whereas the bold arrow represents the direction of the internal association between goal and action following the causality. See text for further explanations.

- *Perceptual-like coding of goals.* An important characteristic of the IMP is that it has been developed within a vision of intelligence seen as closely related to the sensorimotor cycle (for an example from the psychology literature see [22], whereas for an example from embodied artificial intelligence see [11]). As a consequence, the authors proposing the IMP usually stress the fact that the system's internal representations of goals are similar, or the same, as the internal representations activated by perception. This feature of the principle has also an important "corollary": the source of goals is usually assumed to be experience, that is, goals tend to correspond to previously perceived (abstracted) states.
- *Learning of action-effect relations.* Another important constituent of the principle is that experience allows the system to create associations between the execution of actions (e.g., due to exploration, "motor babbling", etc.) and the perceived consequences resulting from it. This requires a learning process that is based on the co-occurrence of actions and their effects observed in the environment [17]).
- *Goals are used to select actions.* Another core constituent of the principle is the fact that the system exploits the learned association between actions and the resulting perceived states of the world to select actions. According to [10, pag. 93]: *For the ideo-motor mechanism, a fundamentally different state of affairs is proposed in which a current response is selected on the basis of its own anticipated sensory feedback.* The idea is that the activation of the representation of a previously experienced state allows the system to select the action that led to it. When this process occurs, the representation of the state assumes the function of goal both because it has an anticipatory nature with respect to the states that the environment will assume in the future, and because it guides behavior so that the environment more likely assumes such states.

It is important to note that the selection of actions with this process requires an "inversion" of the direction of the previously learned action-effect association, from "actions → resulting states" to "resulting states → actions". This inversion is particularly important because it implies that the system passes from the causal association that links the two elements, as resulting from experience, to the teleonomic association between them, as needed to guide behavior. It is only *after* the inversion that the effects can be used as goal states by the system.

3 TOTE and cybernetic principles

TOTE was introduced by [24] as the basic unit of behavior, as opposed to the stimulus-response principle. TOTE was inspired by cybernetics [35], that however focused on homeostatic control and not on *goals*. In a TOTE unit firstly a goal is tested to see if it has been achieved; if not, an operation is executed until the test on the goal's achievement is successful. One of the examples of a TOTE unit is a plan for hammering a nail; in this case, the test consists in verifying if the nail's head touches the surface and the operation consists in hitting the nail. In this case, the representation used for the test is in sensory format, and the operation is always the same, even if the TOTE cycle can involve many steps. TOTE units can be composed and used hierarchically for achieving more complex goals, also including any kind of representation for the test and any kind of action. The TOTE also inspired many subsequent theories such as the General Problem Solver (GPS) [26].

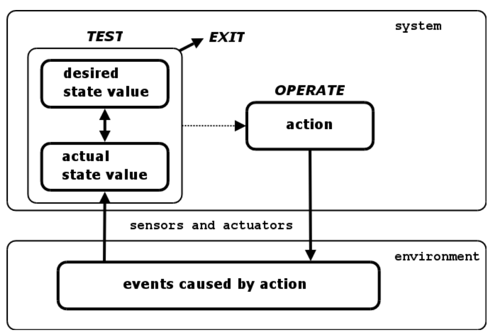

Fig. 3. A scheme that represents the main features of the TOTE. Words in *Italics* represent the main processes composing the principle. Thin arrows represent information flows. The double-headed arrow represents a process of comparison between the desired and the actual state value. The dashed arrow represents the fact that an action is selected and executed in the case the *Test* fails, but not how it is selected. The bold arrow represents a switch in the sequence of processes implemented by the system. See text for further explanations.

The main constituents of the TOTE The three main constituents of the TOTE (see Fig. 3) are now analyzed in detail on the basis of the comparison of the various positions of the authors just mentioned.

– *Test.* A first fundamental constituent of the principle is the internal representation of the desired value(s) of the state of the environment. The representation of this value is a key element of the *Test* sub-process composing the principle. This is the sub-process through which the system repeatedly checks if the current state of the environment matches the goal.
– *Abstract goal.* The desired state value of the system, that is the goal, can be abstract. Indeed, the TOTE is underspecified with this respect, and the literature has used several different types of encodings for goals, from perceptive-like encodings to more abstract symbolic ones. The principle can manage this type of goals as the Test sub-process can be as complex as needed, from simply matching two pattern to a more sophisticated process of logical comparison of several features. This (possibly) abstract nature of the definition of goals has also an important implication on the origin of goals themselves, which can derive from previous experience but also from other sources such as other systems (communication or external setting) and "imagination" processes.
– *Multiple steps.* An important aspect of the TOTE is the fact that it is naturally suited to implement a course of action formed by multiple steps, as suggested by the repetition of the Test sub-process in its acronym. Sensory feedback is also used for chaining actions.

4 Comparison of IMP and TOTE

From the descriptions of IMP and TOTE of the previous sections, it should be apparent that the frameworks referring to them specify rather general behavioral and learning principles. Thus, designing an artificial adaptive learning system according to the principles of the IMP and/or TOTE requires to integrate the two frameworks with many implementation details. To compare the guidelines that IMP and TOTE give for the implementation of an artificial adaptive learning system, we now successively look at the different aspects that need to be integrated. We start by considering goal selection and representation, then analyze action selection and initialization, action execution, and context dependencies. We close by discussing how learning may be integrated into the two principles.

4.1 Goal Origin and Selection

For goals or effects to trigger actions, goals need to be generated and selected in the first place. However, neither one of the two approaches gives suggestions on how such a goal selection process might be structured. Certainly, strong links with motivational and emotional mechanisms might be called into play to tackle this problem. For example undesired low values of variables controlled homeostatically may trigger a goal that previously caused the variable to increase in value (e.g., empty stomach leads to the search and consumption of food). However, literature on IMP simply assumes that some events internal to the system eventually trigger the (re-)activation of an internal representation of action-consequences, that hence assume the function of a pursued goal, without specifying the mechanisms that might lead to this. On the other side, the literature on TOTE tends to generically assume that goals derive from experience or that they originate form outside the system (e.g., other intelligent systems, other module, con-specifics, etc.). Thus, how a goal generation and selection system could be implemented lies outside the scope of IMP and TOTE.

4.2 Goal Representation

Regardless of how goals are selected, goals may be represented in multiple ways. In the IMP, goal representations are encoded perceptually. As a consequence anything that can be perceived might give origin to a goal representation. These goal representations can then trigger bidirectionally linked action codes or action programs that previously led to the activated goal representation.

The IMP does not specify which perceptual goal representations may be possible and how concrete or general they might be. However, two aspects are usually emphasized: the role of experience in the formation of potential goals and the perceptual basis of goal representations. With these restrictions in mind it seems hard to generate some kinds of abstract goals within the IMP, in particular, goals that are defined in terms of qualitative or quantitative comparisons, such as: “find the biggest object in the scene”, or “find the farthest object”. In fact, in these cases the goal cannot be a template or a prototype to be matched, but requires complex processes such as “find an object, store it in memory, find a second object, compare it with the previous one”. Thus, IMP can only apply, if such abstracted, generalized, relational representations are generated from the more basic, perceptual codes.

An interesting additional problem arises in the IMP in that it does not specify how the system may distinguish between actual current perceptual input and current perceptual goal activation. The IMP postulates that the goal is represented in the same format as the percept, generated from the sensation of the state corresponding to it. With this respect, authors usually claim that the physical machinery used to represent the goals and the one used at the higher levels of perceptual processing are the same (e.g. [34]). This raises a problem: how does the system distinguish between the activation of the representation corresponding to a pursued goal and the activation of it caused by the perception of the corresponding state? This information is clearly needed by the system to control actions and essentially act in an anticipatory fashion rather than reactively.

TOTE on the other hand explicitly supposes “abstract” goal representations. The system designer is left to define the possible levels of abstraction. This gives TOTE much more freedom—goals could even be perceptually specified but also other encodings may be used. Even when abstract encodings are used, TOTE needs to be perceptually grounded since the “Test” of the mechanism needs to compare goals with environmental states, which can only be derived from the perceptual input. Thus, differently than IMP, TOTE stresses the importance of abstract goal representations but its goals’ representations need ultimately to be grounded in perceptual input as it needs to test if they have been achieved.

4.3 Action Selection and Initialization

Once a goal representation is invoked, the next question arises: how the corresponding motor program or action is selected and triggered. Both principles remain silent on when the invocation of a goal actually triggers an action, assuming that this is always the case. However, in an actual cognitive system it can be expected that the invocation of a goal representation may not always lead to an actual action trigger, for example when the goal is currently not achievable or too hard to achieve.

The IMP mostly stresses that the perceptual goal representations directly trigger actions or motor programs that previously led to that goal. In contrast to TOTE, though, the IMP does not specify how long this goal is pursued. In particular, it does not specify what happens if the selected goal is already achieved, nor it specifies how the system checks if the currently pursued goal has been achieved. This information is important for the successive correct selection of actions depending on the fact that the pursued goal has been achieved or not. On the contrary, TOTE contemplates an explicit test, applied repeatedly, that allows the system to check when the selected goal has been achieved.

On the other side, whereas IMP suggests the existence of bidirectional links between goal representations and motor programs or actions that achieve them, TOTE is silent on how specific actions are triggered on the basis of the activated goal. For example, the origin of the knowledge needed to select the suitable actions in correspondence to goals is not specified. This is in line with the fact that the literature on TOTE tends to overlook the role that learning and experience might have in goal directed behavior. Given this underspecification, the models working on the basis of TOTE have adopted various solutions. The most common solution (such as the General Problem Solver) assumes that the controlled state is quantitative and continuous, and uses a mechanism that selects and executes actions so as to diminish the difference between the current and the desired values of the state itself. [30] makes explicit that there is a representation of a causal and instrumental link between the actions and the resulting consequences.

4.4 Action Execution

Similar to the testing stage in action selection and triggering, TOTE performs a test on the goal's achievement also during action execution. The IMP remains silent on how the testing process may be realized. TOTE is thus an explicit closed loop framework, which by definition takes the initial state and feedback into account. However, it does not specify if the system should only check for the final goal or also for intermediate perceptual feedback, as suggested for example in the emulation framework of Grush [11] or also in the closed-loop theory by Adams [1]. Moreover, the authors of TOTE do not make any specific assumptions about the specific mechanisms used for control, such as the overall architecture of the system (e.g., hierarchical, modular, etc.). Also the IMP stays silent with regard to the question on how the execution of the "selected" action is carried out, in particular whether or not feedback is used for control. Finally, neither framework concretizes possible distinctions between different types of perceptual feedback such as proprioceptive versus exteroceptive feedback.

4.5 Context Dependence

Both approaches do not make any presuppositions on how goal selection and action selection may be dependent on the current context. The IMP approach considers merely the relation between the desired goal and the "action" to reach it without taking into account that the required action almost always depends on the given initial state of the system. Although modulations of action-effect links are certainly imaginable dependent on currently available contextual information, these are not specified in any form. Also TOTE is silent on this issue as the link between activated goal and corresponding operation is not specified. However, TOTE is context dependent at least in a sense: it explicitly takes the current state into account in order to determine the action.

4.6 Learning

The IMP presupposes the learning of action-effect associations that have a bidirectional nature, contrasting the view that the learning of "forward models" and of "inverse models" are separated learning mechanisms. However, it is less clear how such bidirectional learning is accomplished. If one assumes that the connections between actions and effects are mutually formed by a Hebb-like mechanism (what fires together wires together), one has to face the problem that sensory and motor parameters have to be represented in a way that allows to "wire" different values from both sides with each other. This assumption leads thus to the "common code" hypothesis [31], which is an additional claim. So, in short, learning is to be taken for granted as the IMP does not provide an answer to the question how learning should be organized or structured.

TOTE stays completely silent on how operator modes for specific goals might be learned. Indeed, TOTE does not consider learning at all but rather expects that the system designer creates appropriate operator modules for the goals that may be selected.

In general, both systems remain very underspecified with respect to important issues related to learning. For example, both frameworks do not address important challenges such as learning generalizations over control programs or facing the problem that goals may be achieved in multiple ways. This underspecification with respect to learning remains one of the most crucial challenges for the application of both frameworks.

4.7 Goal Orientedness

We distinguish between three kinds of teleonomic mechanisms:

1. *stimulus determined*, in which some results are reached thanks to learned regularities, without any explicit representation;
2. *goal determined*, in which there is an explicit representation of the expected effect which also triggers an action, via previously learned action-effect links[3].
3. *goal driven*, in which there is an explicit representation of the goal (not only the expectation) having the function to evaluate the current state and to activate an action, given there is a mismatch.

We consider the IMP of the second kind and the TOTE of the third kind, which is a sub-case of the second one. The main difference is that in the IMP the goal is causally reached but not pursued as such. In other terms in the IMP is functionally able to reach a state which is represented in an anticipatory way, but the state is not treated as a goal that is something motivating and to be realized.

On the contrary, the TOTE is goal driven: in fact, there is an explicit goal representation which serves to evaluate the world (in particular, to be matched against the current state). The test is both a trigger for action and a stopping condition; more precisely, the mismatch (indicating the degree of non realization of a goal) serves to select and trigger the rule whose expectation minimizes the discrepancy. Differently from the IMP, the TOTE "knows" if/when a goal is achieved. Another related point is that in the IMP desired results (motivating the action) are not distinguished from expected results of actions, the latter including the former.

The comparison has shown that both, IMP and TOTE, are rather underspecified under many aspects. Whereas TOTE stresses the test-operate cycle, the IMP stresses the linkage between action and contingently experienced effects and the reversal thereof to realize goal-oriented action triggering. With this regards, it seems possible that both principles might be combined into a unique system whose goals are perceptually (but possibly very abstractly) represented, and in which these perceptual goal representations trigger the associated action commands. The triggered goal may then be continuously compared to the current perceptual input enabling the recognition of current goal achievement. To realize this, goal-related perceptual codes need to be distinguished from actual perceptual codes, by, for example, a tag-based mechanism, a difference-based representation, or a simple duplication of perceptual codes. As an example of such a combination, Fig. 4 indicates how the TOTE can exploit action-effect rules as in the IMP, still retaining the test component and using the mismatch for selection and triggering. Of course, the functioning of many processes such as matching, selection and triggering are left unspecified here, because they can be realized in different ways. This is why in the next Section some computational systems are presented that provide concrete examples of possible models that can be obtained by merging different aspects drawn from both frameworks.

[3] As discussed in Sec. 2, an effect can be used as a goal state because there is an "inversion" of the direction of the previously learned action-effect association. For this reason, this mechanism can be conceptually divided into two consecutive steps.

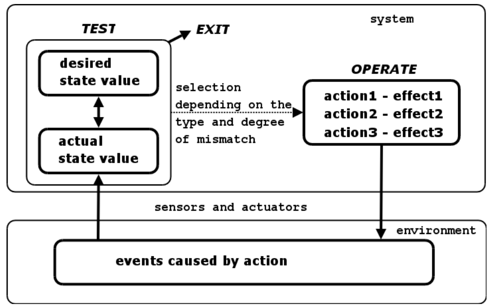

Fig. 4. An example of model integrating some functionalities of both IMP and TOTE. Actions, as in the TOTE, are selected and triggered by the mismatch produced by the test. The action-effect rules are the same used in the IMP.

5 Implementations of IMP and TOTE in Artificial Systems

After having analyzed the IMP and TOTE at a theoretical level, this section reviews and discusses some computational models, presented in detail elsewhere, that on one side represent concrete implementations of some important features of such frameworks, and on the other side offer concrete answers to the issues left open by both frameworks.

5.1 Case Study I: An Architecture for Visual Search

A hierarchical architecture [29] inspired by the IMP and by the "automatic mechanism" in [25] is tested in a *Visual Search* task [40]. The goal is to find the red T in a picture containing also many distractors, i.e. green Ts and red Ls. The system can not see all the picture at once, but has a movable spotlight with three concentric spaces having good, mild and bad resolution.

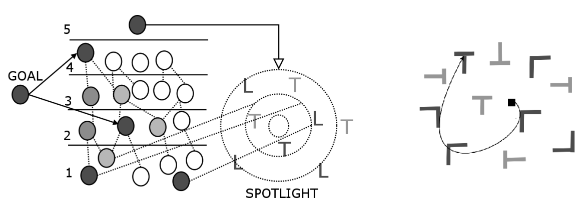

Fig. 5. *Left*: the components of the simulation: the goal, the spotlight and the modules, whose layers are numbered. Light and dark nodes represent more or less active modules. Modules learn to predict the activity level of some modules in the lower layer, which they receive in input (dotted lines). *Right*: a sample trajectory in the visual field, starting from the center (red letters are dark Grey, green letters are light Grey).

The visual search task is performed by many feature-specific modules, such as color-detectors and line-detectors, organized hierarchically (see the left part of Fig 5). According to [8, p. 444] search is *matching input descriptions against an internal template of the information needed in current behavior*: each module consists in an input template and a behavior. Modules have a variable level of activation; more active modules can act more often and, as we will see, influence more the overall computation. Modules in layers 1 and 2 obtain an input from a simulated fovea; the other ones have no access to the fovea, but use as input the activation level of some modules in the immediately lower layer (dotted lines in Fig 5). The architecture has five layers:

1. **Full Points Detectors** receive input from portions of the spotlight, e.g. the left corner, and match full or empty points. Modules are more numerous in the inner spotlight than in the central and outer spotlight.
2. **Color Detectors** monitor the activity of Full Points Detectors and recognize if full points have the color they are specialized to find (red or green).
3. **Line Detectors** categorize sequences of points having the same color as lines; they do not store positions and can only find sequences on-the-fly.
4. **Letter Detectors** categorize patterns of lines as Ls or Ts; they are specialized for letters having different orientations.
5. **The Spotlight Mover** is a single module; as explained later, it receives asynchronous motor commands from all the other ones (e.g. *go to the left*) and consequently moves the center of the spotlight.

In the **learning phase**, by interacting with a simulated environment, each module learns *action-expectation pairs*. Modules learn the relations between their actions and their successive perceptions (the activation level of some modules in the lower layers), as in *predictive coding* [32]. In this way they also learn which actions produce successful matching; for example, a line-detector learns that by moving left, right, up or down the fovea its successive pattern matching operation will be successful (i.e. it will find colored points, at least for some steps), while by moving in diagonal its matching will fail; in this way the line-detector implicitly learns the form of a line by learning how to "navigate" images of lines. In a similar way, a T-detector learns how to find Ts by using as inputs the line-detectors . There is also a second kind of learning: modules *evolve links* toward the modules in the lower layer, whose activity they use as input and can successfully predict; for example, T-detectors will link some line-detectors[4]. These top-down, *generative* links are used for spreading activation across the layers.

The **simulation phase** starts by setting a Goal module (e.g. *find the red T*) that spreads activation to the red-detector(s) and the T-detector(s). This introduces a strong goal directed pressure; at the beginning of the task some modules are more active than others and, thanks to the top-down links, activation propagates across the layers. During the search, each module in the layers

[4] By learning different sets of action-prediction rules, modules can also specialize: for example, there can be *vertical lines detectors* and *horizontal lines detectors*

2, 3, and 4 tries to move the spotlight where it anticipates that there is something relevant for its (successive) matching operation, by exploiting their learned action-perception associations. For example, if a red-detector anticipates something red on the left, it tries to move the spotlight there; a green-detector does the opposite (but with much less energy, since it does not receive any activation from the Goal module). Line- and letter- detectors try to move the spotlight for completing their "navigation patterns". Modules which successfully match their expectations (1) gain activation, and thus the possibility to act more often and to spread more energy; and (2) send commands to the Spotlight Mover (such as *move left*); the controller dynamically blends them and the spotlight moves, as illustrated in the right part of Fig 5. In this way the fovea movements are sensitive to both the goal pressures and the more contextually relevant modules, i.e. those producing good expectations, reflecting attunement to actual inputs. The simulation ends when the Goal module receives simultaneous success information by the two modules it controls; this means that the Goal module has only two functions: (1) to start the process by activating the features corresponding to the goal state and (2) to stop the process when the goal is achieved. As reported in [29], this model accounts for many evidences in the Visual Search literature, such as sensitivity to the number of distractors and "pop-out" effects [40].

IMP and TOTE in play. According to the IMP, activity is preceded and driven by an endogenous activation of the anticipated (and desired) goal state. In this case, the goal "find the red T" can be reformulated as "center the fovea in a position in which there is a red T"; and the process starts by pre-activating the features of the desired state, i.e. the modules for searching the color red (red-detector) and the letter T (T-detector); the "finding machine", once activated, can only search for an object having these features. The key element of the model is the fact that modules embed action-expectation rules and are self-fulfilling; when a module is endogenously activated, its effect becomes the *goal* of the system. It is worth noting that this system does not use any map of the environment, but only sensorimotor contingencies [28] and a close coupling between perception and action.

This system can achieve only two kinds of goals: (1) *goal states that were experimented during learning*, such as "find the red T"; and (2) *goal states that are a combination of features*; for example, by combining a green-detector and an L-detector, the system can *find a green L* even if it has never experimented green Ls during learning, but only green Ts and red Ls. On the contrary, this system cannot achieve other kinds of goals such as: (1) *The red T on the left*, since locations are not encoded; (2) *The biggest red T*, since there is no memory of past searches and different Ts can not be compared; (3) *The farthest red T*, since temporal features are not encoded. These goals require a more sophisticated procedure for testing and a more abstract encoding: two of the features of the TOTE. The system uses a feature of TOTE: a stopping condition, consisting in a matching between the goal and the activation level of the corresponding features.

5.2 Case Study II: An Architecture for Reaching

The second system used to illustrate the IMP and the TOTE in play has been used to control a simple 2D two-segment arm involved in solving sequential reaching tasks by reinforcement learning. Here we present only the features of the system useful for the purposes of the paper and refer the reader to [27] for details.

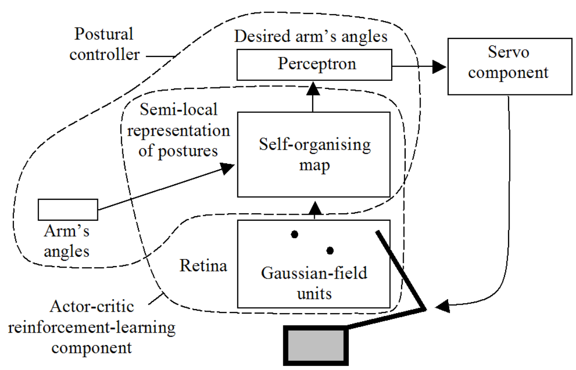

Fig. 6. The architecture of the model of reaching. Rectangular boxes indicate neural layers. Text in boxes indicates the type of neural-network model used. Text near boxes indicates the type of information encoded in the layers. Callouts indicate the two major components of the system. The graph also shows the controlled arm and two targets activating the retina (black dots). See text for further explanations.

The system is mainly formed by two components, a *postural controller* and a *reinforcement-learning component* ("RL component" for short). In a first learning phase, the postural controller learns how to execute sensorimotor primitives that lead the arm to assume certain postures in space. In order to do so, while the system performs random actions (similarly to "motor babbling" in infants, see [23]), the postural controller learns to categorize the perceived arm's angles in a 2D self-organizing map [21]. At the same time a two-layer network is trained, by a supervised learning algorithm [38], to associate the arm's angles (desired output pattern) with the map's representation of them (input pattern). This process allows the system: (a) to develop a population-code representation of sensorimotor primitives within the self-organizing map, encoded in terms of the corresponding "goals" (i.e. postures); (b) to develop weights between the map and the desired arm's angles that allow selecting sensorimotor primitives by suitably activating the corresponding goals within the map.

In a second learning phase, the RL component learns to select primitives to accomplish reward-based reaching-sequence tasks, for example in order to reach two visible dot targets in a precise order (see Fig 6 for an example; the RL component is an "actor-critic model", see [36]). Each time the RL component selects an action (i.e. the achievement of a "desired posture"), the desired arm's angles produced by it are used to perform detailed movements (variations of the arm's angles) through a hardwired servo-component that makes the arm's angles to progressively approach the desired angles (postures): when this happens, control is again passed to the RL component that selects another action.

IMP and TOTE in play. The system has strong relations with both the IMP and TOTE, and in so doing it emphasizes their complementarities. In line with the IMP, in the first phase of learning (motor babbling) the system performs (random) actions, and learns to associate the resulting consequences, in terms of the proprioception of the arm's angles, to them. In the second phase of learning, the system uses the expected consequences of the actions as goals (expected in terms of final postures), to trigger the executions of the actions themselves (by trial-and-error, so as to pursue rewarding states). This feature of the system is in line with two core features of IMP related to learning the action-effect relations and using them in a reverse fashion to select actions. However, a first important departure from the IMP is that the "goals" of the primitives (i.e., the corresponding previously perceived postures), through which the system selects the primitives themselves, are not encoded in a "pure" perceptual-like format, but in terms of more abstract representations generated by the self-organizing map. This might represent a first step towards a more abstract representation of goals in the spirit of TOTE.

A second important departure from the IMP is that the system incorporates a "stop" mechanism on the basis of which, when it achieves the goal for which it selected the corresponding action, control passes again to the RL component. As we have seen, this is a typical feature of TOTE. Note how we had to introduce this "stopping" condition to allow the system to accomplish a task that required the execution of more than one "action" in sequence (two actions in this case).

From an opposite perspective, it is interesting to notice how by using some of the core ideas behind the IMP, the system overcomes some limitations of TOTE. In particular, first it uses experience to create goals' representations and to associate them to actions, a core idea of the IMP. Second, it uses motor babbling to create an association between goals and actions used to achieve them, overcoming TOTE's underspecification about how specific actions are selected in correspondence to a given goal.

5.3 Case Study III: Anticipatory Classifier Systems

The *anticipatory learning classifier system* ACS2 [3] learns anticipatory representations in the form of condition-action-effect schemata, similar to Drescher's schema system [7]. However, ACS2 learns and generalizes these schemata online using an interactive mechanism that is based on Hoffmann's theory of anticipatory behavioral control [14–16] and on genetic generalization [3]. Similar to the described arm-control approach, ACS2 executes some form of motor babbling. It consequently learns a generalized model of the experienced sensory-motor contingencies of the explored environment. In difference to the above system, though, ACS2 learns purely symbolic schema representations, in difference to the dynamically abstracted real-valued sensory information. Generally, though, such an abstraction mechanism might be linked with the ACS2 approach. More importantly, though, ACS2 makes sensory-motor contingencies explicit: The systems learns a complete, but generalized predictive model of the environment.

ACS2 was combined with an online generalizing reinforcement learning mechanism, based on the XCS classifier system [39]. The resulting system, XACS [4], learns a generalized state value function using XCS-based techniques in combination with the model learning techniques of ACS2. Figure 7 sketches the resulting architecture. The reinforcement component is intertwined with the model learning component using the model information for both predictive reinforcement learning and action decision making. For learning, XACS iteratively updates its reinforcement component using a Q-learning-based [37] update mechanisms—testing all possible reachable situations and using the maximum reward value to update the currently corresponding reward value. For action decision making, XACS uses the model to activate all immediately reachable future situations and then uses the reinforcement learning component to decide on which situation to reach and consequently which action to execute. It was also proposed that XACS may be used in conjunction with a motivational module representing different drives. The reinforcement module would then consist of multiple modules that work in parallel, each module influencing decision making according to its current importance [4] (see Figure 7).

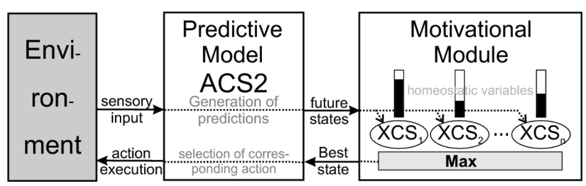

Fig. 7. XACS realizes the IMP in that it selects actions according to their associated perceptual effects. A desired effect is selected using the developed motivational module that is designed to maintain the system in homeostasis. TOTE is realized in that each iteration currently possible effects are compared with currently desired effects.

IMP and TOTE in play. XACS plays a hybrid role being situation-grounded but goal-oriented. In this way, goals that cannot be achieved currently will not have any influence on behavior. Vice-versa, goals that are easily achieved currently will be pursued first. Due to the generalization in the predictive model and in the reinforcement component, abstract generalized goal representations can be reached within differing contexts.

XACS realizes ideomotor principles in that actions are directly linked to their action effects. Initially, XACS learns such schemata by the means of random exploration. Goals are coded using the given perceptual input, which is symbolic. XACS, however, does not start from the goal itself but interactively activates potential goals (that is, future situations), then chooses the currently most desirable one, which finally triggers action execution. In this way, the system is goal-driven—but it is grounded in the current situation.

Goal selection is integrated in XACS by the separate reinforcement component that links to the behavioral component. Thus, XACS proposes a goal selection mechanism realized with reinforcement learning techniques. In difference to TOTE, there is never an explicit test that controls if a goal was reached. This mechanism is implicitly handled by the reinforcement learning component in conjunction with the proposed motivational module. Once a goal is reached, a motivation will become satisfied and thus another drive will control behavior.

6 Conclusions

This paper has investigated the implications of the ideomotor principle (IMP) and the test operate test exit (TOTE) framework for adaptive behavior and action selection. It was shown that the frameworks are actually rather closely related as both stress the importance of goal-oriented action selection. Whereas goals are represented perceptually and are bidirectionally linked to associated actions in the IMP, TOTE emphasizes the interactive cycle of triggering actions by desired goals while iteratively testing if such goals are achieved. Overall the two frameworks enlighten important aspects of the anticipatory nature of goal-driven systems. However, neither of them get concrete enough to pinpoint specific actual implementations.

The three implementations described in the paper not only emphasize the power of the guidelines proposed by the IMP and TOTE, but also represent important attempts to give possible answers to the problems left unresolved by them. With this respect, the paper reviewed three architectures. We now highlight the lessons learned from putting the theoretical principles of IMP and TOTE in practice.

- The first architecture was concerned with a visual search task. It had a "goal node" which contained a test condition (similarly to TOTE) having a sensorimotor encoding (testing two sensorial conditions, color and shape). Like the IMP, action was preceded and triggered by a pre-activation of the desired goal state, but like TOTE this happened as a consequence of a mismatch between the pursued goal and the stimuli. The search proceeded thanks to the learned action-expectation links, which in this architecture were encoded both in the modules, which were procedures that attempt to "self-realize", and the links between them. Interestingly, to allow the architecture to function we had to design a mechanism for which the goal to pursue was selected through an activation with a level above zero (in order to trigger the search) but below the activation achieved when the state corresponding to it was actually achieved through action. In fact, if the pre-activation and the activation had the same level, the test had a positive outcome and the search immediately stopped. In several experiments it was also found that the initial amount of pre-activation led to different response times in finding a solution and could also lead to different search strategies. The interpretation of this was that such pre-activation encoded a measure of *urgency*. The IMP and TOTE do not specify any mechanism to encode quantitative aspects of teleonomic: this is surely an important limitation of the two frameworks pointed out by the attempt to translate them into functioning efficient computational systems.

- The second one is a neural architecture directed to tackle reaching problems with a simple simulated arm. This architecture highly benefited from the suggestion given by the IMP to create the association between goals and actions suitable to tackle them, through experience (learning), and then to use goals to suitably trigger such actions. On the other side, it also highlighted the importance of testing the achievement of goals, similarly to what is suggested by TOTE, to suitably give control either to the reinforcement-learning selector of actions or to the execution of the actions themselves. On the other side the architecture also raised the necessity to have distinct representations of goals to pursue and current states of the system in order to be in the condition of performing such tests.
- The XACS architecture is a more symbolically-based architecture that enables the pursuance of different goals. It implements the IMP directly forming a forward model of its environment, using its forward model to trigger action execution. In TOTE it remains underspecified how goals may emerge and how they may trigger actions. Also the IMP does not specify how desired perceptual states are triggered, nor how the bidirectional sensory-motor knowledge activates appropriate actions. XACS proposes an interlinked process that (1) activates all reachable (currently immediate) future states and (2) selects that action that leads to the currently most desirable one. Multiple goals may thus be active concurrently and the most relevant and most reachable goal will be pursued.

With such a conceptualization and characterization of IMP and TOTE in hand, the next step along this line of research will be to further investigate the many questions left open by the two principles, as well as to further identify the specific advantages and disadvantages stemming from the actual implementations. Hereby, it will be important to use real-world simulations, or actual robotic platforms, both (1) to identify the issues left unresolved, and, at the same time, (2) to crystallize the true potential of the anticipatory principles proposed.

Acknowledgments This work is supported by the EU project **MindRACES**, FP6-511931.

References

1. J. A. Adams. A closed-loop theory of motor learning. *Journal of Motor Behavior*, 3:111–149, 1971.
2. L. W. Barsalou. Perceptual symbol systems. *Behavioral and Brain Sciences.*, 22:577–600, 1999.
3. M. V. Butz. *Anticipatory learning classifier systems.* Kluwer Academic Publishers, Boston, MA, 2002.
4. M. V. Butz, D. E. Goldberg, and K. Tharakunnel. Analysis and improvement of fitness exploitation in xcs: Bounding models, tournament selection, and bilateral accuracy. *Evolutionary Computation*, 11:239–277, 2003.
5. M. V. Butz and J. Hoffmann. Anticipations control behavior: animal behavior in an anticipatory learning classifier system. *Adaptive Behavior*, 10(2):75–96, 2002.
6. M. V. Butz, O. Sigaud, and P. Gérard. Internal models and anticipations in adaptive learning systems. In M. V. Butz, O. Sigaud, and P. Gérard, editors, *Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories, and Systems*, pages 86–109. Springer-Verlag, Berlin Heidelberg, 2003.
7. G. L. Drescher. *Made-Up Minds: A Constructivist Approach to Artificial Intelligence.* MIT Press, Cambridge, MA, 1991.
8. J. Duncan and G. W. Humphreys. Visual search and stimulus similarity. *Psychological Review*, 96:433–458, 1989.
9. V. Gallese and T. Metzinger. Motor ontology: The representational reality of goals, actions, and selves. *Philosophical Psychology*, 13(3):365–388, 2003.
10. A. G. Greenwald. Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review*, 77:73–99, 1970.
11. R. Grush. The emulation theory of representation: motor control, imagery, and perception. *Behav Brain Sci*, 27(3):377–96, Jun 2004.
12. J. Herbart. *Psychologie als Wissenschaft neu gegrundet auf Erfahrung, Metaphysik und Mathematik. Zweiter, analytischer Teil.* August Wilhem Unzer, Koenigsberg, Germany, 1825.
13. G. Hesslow. Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6:242–247, 2002.
14. J. Hoffmann. *Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung und Wahrnehmung [Anticipation and cognition: The function of anticipations in human behavioral control and perception].* Goettingen, Germany, 1993.
15. J. Hoffmann. Anticipatory behavioral control. In M. V. Butz, O. Sigaud, and e. P. G/'erard, editors, *Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories, and Systems*, pages 44–65. Springer-Verlag, Berlin Heidelberg, 2003.
16. J. Hoffmann, C. Stöcker, and W. Kunde. Anticipatory control of actions. *International Journal of Sport and Exercise Psychology*, 2:346–361, 2004.
17. B. Hommel. Planning and representing intentional action. *TheScientificWorld JOURNAL*, 3:593–608, 2003.
18. B. Hommel, J. Musseler, G. Aschersleben, and W. Prinz. The theory of event coding (tec): a framework for perception and action planning. *Behavioral and Brain Science*, 24(5):849–78, 2001.
19. W. James. *The Principles of Psychology.* Dover Publications, New York, 1890.

20. G. Knoblich and W. Prinz. *Higher-order motor disorders*, chapter Linking perception and action: An ideomotor approach., pages 79–104. Oxford University Press., Oxford, UK, 2005.
21. T. Kohonen. *Self-Organizing Maps*. Springer-Verlag, Berlin Heidelberg, New York, 3rd edition, 2001.
22. S. Kosslyn. *Image and Brain: The Resolution of the Imagery Debate*. MIT Press, Cambridge, 1994.
23. A. N. Meltzoff and M. K. Moore. Explaining facial imitation: A theoretical model. *Early Development and Parenting*, 6:179–192, 1997.
24. G. A. Miller, E. Galanter, and K. H. Pribram. *Plans and the Structure of Behavior*. New York, 1960.
25. M. Minsky. *The Society of Mind*. Simon & Schuster, 1988.
26. A. Newell. *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA, 1990.
27. D. Ognibene, A. Rega, and G. Baldassarre. A model of reaching integrating continuous reinforcement learning, accumulator models, and direct inverse modelling. In S. Nolfi, G. Baldassarre, D. Marocco, D. Parisi, R. Calabretta, J. Hallam, and J.-A. Meyer, editors, *From Animals to Animats 9: Proceedings of the Ninth International Conference on the Simulation of Adaptive Behavior (SAB-2006)*, 2006.
28. J. O'Regan and A. Noe. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):883–917, 2001.
29. G. Pezzulo, G. Calvi, D. Lalia, and D. Ognibene. Fuzzy-based schema mechanisms in akira. In M. Mohammadian, editor, *Proceedings of CIMCA'2005*, Vienna, 2005.
30. M. E. Pollack. Plans as complex mental attitudes. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 77–103. MIT Press, Cambridge, MA, 1990.
31. W. Prinz. *Perspectives on imitation: From neuroscience to social science*, volume 1, chapter An ideomotor approach to imitation, pages 141–156. MIT Press, Cambridge, MA, 2005.
32. R. P. Rao and D. H. Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*, 2(1):79–87, January 1999.
33. G. Rizzolatti and M. A. Arbib. Language within our grasp. *Trends in Neurosciences*, 21(5):188–194, 1998.
34. G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 1996.
35. A. Rosenblueth, N. Wiener, and J. Bigelow. Behavior, purpose and teleology. *Philosophy of Science*, 10(1):18–24, 1943.
36. R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge MA, 1998.
37. C. J. C. H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8:279–292, 1992.
38. B. Widrow and M. Hoff. Adaptive switching circuits. In *IRE WESCON Convention Record*, pages 96–104, 1960. Part 4.
39. S. W. Wilson. Classifier fitness based on accuracy. *Evolutionary Computation*, 3:149–175, 1995.
40. J. M. Wolfe. Visual search. In H. Pashler, editor, *Attention*. London, UK: University College London Press, 1996.
41. D. Wolpert and M. Kawato. Multiple paired forward and inverse models for motor control. *Neural Netw*, 11(7-8):1317–1329, Oct 1998.