Predictive and Anticipatory Computing

Mihai Nadin
antÉ—Institute for Research in Anticipatory Systems, University of Texas at Dallas, Richardson, Texas, U.S.A.

Abstract

Predictive computation is the path that leads from reaction-based forms of computing to anticipatory forms of computing. It reflects the long-term goal of understanding and emulating the performance of the living. Prediction is informed by the past; probability underscores the gamut of predictive types. Anticipation implies awareness of past, present, and future. It underlies evolution and thus holds promise for realizing artificial adaptive systems. The possibility space is where anticipation is substantiated. Improved understanding of biological processes, paralleled by vastly scaled-up and diversified computer performance, has triggered a variety of engineering successes (mobile computing, robotics, artificial intelligence, the Internet-of-Everything, etc.). These have made predictive and anticipatory computing part of the new ecology of computation.

INTRODUCTION

All science is about dynamics, i.e., how everything changes. What it takes to understand and adapt to change is best embodied in the functioning of the living. Therefore, it is not surprising that knowledge of life processes guide the effort to provide nature-like means and methods for dealing with change, moreover for predicting it. Science expressed computationally integrates life-inspired knowledge, as well as faster and more diverse processing of data pertinent to change. Integration of life science and technological performance is a prerequisite for both predictive and anticipatory computing. Thus the goals pursued herein are

1. To present efficient processing models that describe the various levels at which the future state of a system can be effectively represented;
2. To address specific forms through which predictive computation is performed; and
3. To define progress toward anticipatory computation.

Therefore, the path from algorithmic computation, through which predictive procedures are performed, to the non-algorithmic, as a medium for anticipatory expression, will be framed in theoretical and practical terms. Among the examples of predictive performance to be featured are deep-learning-based human-level performance in language (degree of competence) and in playing games in new computational forms. These are driven by future states. The same applies to all kinds of control mechanisms, especially those embodied in robots and predictive control procedures. Consideration will be given to predictions upon which diagnostics in a variety of domains (e.g., medicine, risk assessment, ecology) are issued by systems that deploy mobile computation. Sensor fusion, multilayered and recurrent neural networks (RNNs), as well as cloud computing, underlie specific predictive performance. It will become evident that appropriate data (small or big) is a prerequisite for all predictive endeavors. In this context, scaling from individuals to communities, with the challenge of an integrated ecology of billions of devices, is the natural progression from discrete forms of predictive computation to decentralized peer-to-peer interactions on the Internet-of-Everything (IoE).

DEFINING THE CONCEPTS

To predict (from the Latin *prae*: before and *dicere*: to say) means to state something about a sequence: what follows in time, and in space, words or expression in language ("the story's ending"), degree or significance. Predictions can be time-independent (extenders), pertinent to simultaneous occurrences (portents), or can infer from data describing a previous state or the current state of the world to a future state. Anticipation (from the Latin *ante*: ahead and *capere*: to understand) means an action (avoiding danger, reaching a goal) informed by a possible future state. Neither prediction nor anticipation invites prescience or psychic understandings. The premise of predictive or anticipatory performance is the perception of reality. Data about it, acquired through sensors, as well as generated within the subject, drive the predictive effort or inform anticipatory action. Prior to the advent of digital computation, predictive and anticipatory goals were pursued within the defining gnoseological metaphor: the world as hydraulic, pneumatic, mechanical, steam-powered, etc.

THE FUTURE IN COMPUTATIONAL TERMS

Predictive and anticipatory computations are inspired by living processes. The premise that biological

phenomena—brain activity, for instance—are an outcome of some sort of computation led computer science to adopt means and methods characteristic of the domain of the living. Currently, this trend is widening, given the interest in applying predictive and anticipatory computing to medicine (Fig. 1).

Some of these biologically inspired computations proved successful in approaching problems otherwise difficult, if at all possible, to address. It should be noted that anticipatory computing corresponds to a holistic understanding of the living. Rosen (pp. 202–203)[1] ascertained that the living is not reducible to the machine; Dubois[2] searched for a compromise in advancing the distinction between weak and strong anticipation; Nadin[3] developed the model of soft machines with variable configurations; Pezzulo et al.[4] initiated the European effort (MindRACES) to design machines with predictive capabilities. In April 2000, *Computer* (an IEEE publication) headlined configurable computation. DARPA allied itself with the Decade of the Brain (1990–2000) through the Augmented Cognition Program (AugCog). It had a very clear goal: to build cognitively aware computational systems. Two directions reflected in the current innovation impetus in predictive computation merge: 1) the design of new computers, sensors, control processes, and communication devices, and their seamless integration and 2) the focus on biological processes, of medical significance, in particular. The program suggested accessing the individual's cognitive state in real time. This was the premise for leveraging cognitive processes and thus integrating humans and machines. Machines adapting to users have to detect their real-time cognitive states, facilitate changes in the cognitive states, and prepare for autonomous cognitive state manipulation. The project is still in progress.

Quite in the same spirit, DARPA (Grand Challenge 2004, and the Robotics Challenge 2012–2014) continues to stimulate research in engineering projects (driverless car, robots for extreme situations).

To date, preoccupation with anticipation-like computations has taken a limited number of distinct forms: variable configurations, precomputation (what prior to computation were look-up tables; Han et al., p. 159[5]), parallel computation at different speeds (to achieve the goal of "faster-than-real-time" models), reverse computation, neural networks, learning, and deep learning procedures, among others. Extrapolation is also pursued as a predictive avenue (similar to how humans think). The convenience of predicting patterns of individual behavior by using sensor data associated with short-term decisions is probably the most frequently utilized method for predictive computation.

REACTION VS. ANTICIPATION

The increased interest in computations of predictive and anticipatory nature entails the need to frame them in the knowledge context in which they are anchored. In an anticipatory system, the current state depends not only on the past state but also on possible future states:

$$\text{Current state} = f(\text{past state, current state, possible future state})$$

In its standard Turing condition, the computer is a deterministic machine: the past state determines the current state. By its nature, it cannot carry out anticipatory computation: the machine performing it would have to simultaneously be in two different states. For predictive computation to take
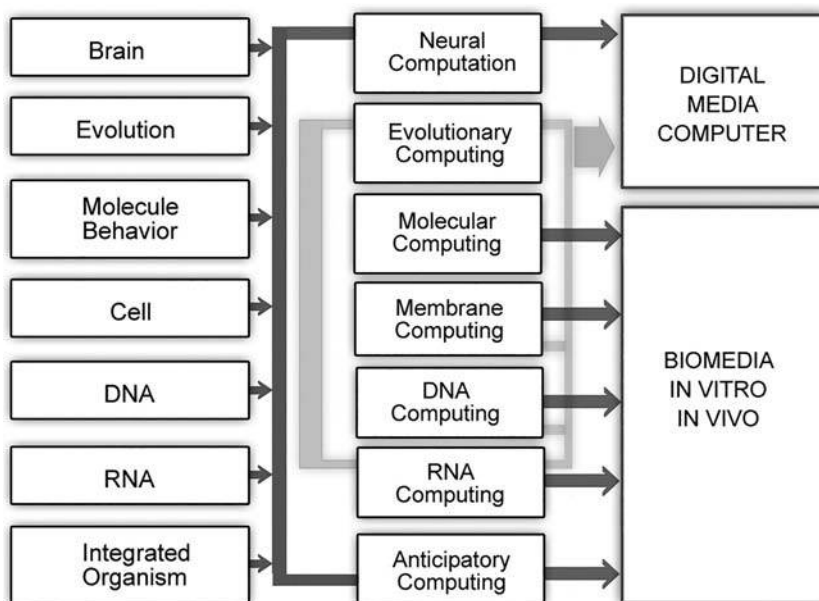

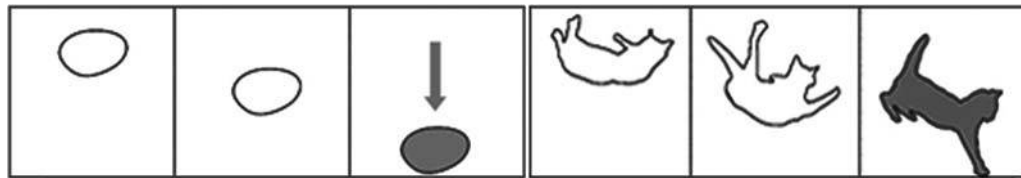**Fig. 1** Biologically inspired forms of computation.

Fig. 2 Not all falls are the same, but all are subject to gravity.

place, data from experience constitutes the premise for a probabilistic description:

Current state $= f$(past state, current state)

Future state $= f$(probability description of successive state changes)

The augmented cognition program led to a simple realization: computation of future states can be performed in a variety of ways.

Achievements in the area of predictive computation and, related to it, in neural network-based deep reinforcement learning, which compete with human-level control, are indicative of the influence of brain research on computing. On the hybrid platform of mobile computation, machine learning affords the connection of data and meaning (e.g., position, activity, possible purpose, in other words: what, where, why). It produces information pertinent to the context.

The living, like the physical, is embodied in matter; hence, the reactive dynamics is unavoidable. However, the physical dynamics—i.e., how change of matter takes place—of what is alive is complemented by the anticipatory dynamics—how the living changes ahead of what might cause its future condition. Newton's laws, like all laws anchored in the deterministic cause-and-effect sequence (past → present → future), preempt the goal-driven changes ahead of material causes. The living reacts, but at the same time it continuously prepares for change. Adaptivity over shorter or longer intervals is the specific expression of this interplay. It also explains the long-term stability of the living. Extrapolation takes prior knowledge as the premise for inference to new situations.

From the perspective of physics, the following would appear as unavoidable: A stone and a cat of equal weight fall (Fig. 2), regardless of the moment in time, and even regardless of the measuring process, acceding to Newton's law of gravity. But the stone falls "passively"—along the path of the gravitational force. The cat's fall is suggestive of anticipation. Predictive computation describing the stone's fall—always the same—makes available to the user an efficient description of all variables involved (stone's position at each moment, speed, impact of the fall, etc.). Anticipation guides the cat's fall—never the same. The equation pertinent to the fall of the stone still applies; but to describe the unique manner in which a cat

falls, additional information describing the cat's condition is needed. Sensors can provide such information. By extension, sensors can, for example, help persons (athletes, firefighters, as well as the aging) who need to mitigate the consequences of falling.

The significance of these distinctions becomes evident when we consider how predictive computation reflects them. The laws of physics (e.g., determine the stone's position and speed during the entire process) provide all that is needed to make a prediction. Navigation systems (providing arrival time, for instance) integrate this kind of predictive computation. For example, inspired by the cat's fall, Apple, Inc. patented a method for controlling the accidental fall of the iPhone on its fragile screen. Sensors activate the iPhone's vibration motor in order to change the angle of the fall in midair to avoid landing on the phone's monitor (Fig. 3).

The patent is an example of engineering inspired by anticipation in the living. The purpose is to change the device's behavior (anticipation is always expressed in action). Knowledge from physics and from observations of the living are integrated in the design.

WAYS OF CONSIDERING THE FUTURE

Predictions can take many forms of expression. Understanding the difference between guessing, expectation, forecasting, etc. allows for defining the relation between computation in the physical substratum and in the living substratum. As a physical entity, a machine is subject to the laws of physics (descriptions of how things change over time). A machine cannot anticipate the outcome of
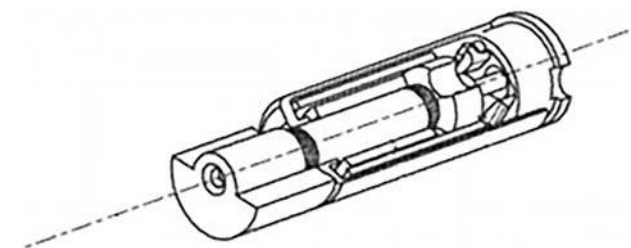

Fig. 3 Statistical analysis of the fall, by comparing gathered data against other information stored in device memory, serves as a trigger to activate the spin, using the vibration motor, and change the phone's center of gravity (cf. patent application).

its functioning. If it could, it would choose the future state according to a dynamical characteristic of the living (evolution), and not according to that of physical phenomena (the minimum principle). A machine, as opposed to a living medium of calculations, is reducible to its parts (the structure of matter down to its finest details). Nothing living is reducible to parts without giving up exactly the definitory characteristic: self-dynamics. Each part of a living entity is of a complexity similar to that of the entity from which it was reduced. From guessing and expectation to prediction and planning as human endeavors, we infer that reaction and anticipation are not reciprocally exclusive, but rather intertwined.

Guessing and Conjecture

To *guess* is to select from what might happen—a sequence of clearly defined options—on account of various experiences: one's own; of others; or based on unrelated patterns (the so-called "lucky throw" of a coin or dice, for example).

> **Guess → selection from a well-defined set of choices**

$(P_{os}(p_i)), p_i \geq N(C(c_i)), c_i \geq N$

Guessing a number from 1 to 100 involves the need to reduce the space of choices ("Is it greater than 50?"). Other guesses involve the processing of information not directly related to the correct answer (cues). When patterns emerge, there is learning in guessing: the next attempt integrates the observation of related or unrelated information. Conjecture is the informed guess; so is extrapolation. These associative actions are the cognitive ingredient most connected to guessing and, at the same time, to learning. Ad hoc associative schemes correspond to responses of the human frontal cortex to surprising events.[6] The dorsolateral prefrontal cortex contributes to the adjustment of inferential learning. Associative relationships that lead to learning are based on the action of discriminating the degree (strength) of interrelation. Fuzzy sets are the appropriate mathematical perspective for describing such interrelations.

Expectation

Expectation does not entail choosing, but rather an evaluation of the outcome of some open-ended process. Several sources of information pertinent to forming an expectation are weighed against each other: Would the dinner guest like pizza? Red wine? What appears *most* probable out of all that is probable gets the highest evaluation. For example, expectations associated with experiments are usually intended to confirm a hypothesis or someone else's results. If the outcome is judged as negative, then avoiding it is the basis for action. Farming prior to the integration of

digital information in agricultural production was often in the realm of expectation. Predictive computation pairs the methods of agronomics and the means of climatology. For example, data from Next Generation Radar (Nexrad) and distributed computing are integrated in a product (by Climate Corporation) that assists farmers in mitigating weather-related risk. An expectation machine (such as Climate Corporation) is actually a learning procedure that attaches weights (some subjective) to choices from the limited set of possibilities. The reactive component dominates the anticipatory.

> **Expectation → evaluation of outcome based on incomplete knowledge (from a limited set of probabilities)**

$P(p_1, p_2, \ldots, p_n)$

Experts absorbed in data patterns more relevant to past performance than applicable to future developments can be given as examples: physicians, economists, politicians, educators, and gamblers. These have in common the perception of random and non-random events. Statistically significant deviations from the expected lead to beliefs that translate into actions. The result of a blood test (for example) is an expectation map. Physicians interpret deviations in respect to expected values (blood glucose, cholesterol, vitamin D, creatinine), and automatic procedures (comparison with average values) trigger warnings. False expectations (of personal or group significance) are the outcome of skewed evaluations.

Forecasting

To ascertain that something will happen in advance of the actual occurrence—prediction (it will rain)—and to cast in advance—forecast—(chance of rain tomorrow) might seem more similar than they actually are. A computer program for predicting weather could process historic data—weather patterns over a long time period—or associate them with the most recent sequence. The aim is to come up with an acceptable global prediction for a season, year, or even a decade. In contrast, a forecasting model would be local and specific.

> **Forecast → infer from past data-based predictions to the future under involvement of self-generated data**

$\mathcal{F}$ (predictions, self-generated data)

Forecasting implies an estimation of what, from among few possibilities, might happen. The process of estimation can be based on "common knowledge" (e.g., black wooly caterpillars mean a harsh winter); on time series; on data

from cross-sectional observation (the differences among those in a sample); or on longitudinal data (same subject observed over a long time). Forecasting is domain specific. It involves data harvested outside the living system, as well as data that the living themselves generate (informed by incomplete knowledge or simplified models). The interplay of initial conditions (internal and external dynamics, linearity and non-linearity, to name a few factors), that is, the interplay of reaction and anticipation, is what makes or breaks a forecast. Sometimes forecast is a recursive inference from the state of the system at a certain moment in time, to a succeeding state.

PREDICTION

Causality, as the primary, but not exclusive, source of predictive power is rarely explicit. Prediction—explicit or implicit—expresses the degree of ignorance: what is not known about change. Bernoulli, the father of probability theory, pointed out that uncertainty is the shadow projected by each prediction. Therefore, it is representative of the limits of understanding whatever is predicted. In some cases, the prediction is fed back into what the process is supposed to predict: e.g., how a certain political decision will affect society; how an economic mechanism will affect the market; how technological innovation will affect education. As a result, a self-referential loop is created. The outcome is nothing more than what is inputted as prediction. Those who predict are not always fully aware of the circularity inherent in the process.

Prediction $\rightarrow$ inference based on probability

$$\mathcal{P}(\text{frequency}, \text{ignorance}, \text{belief})$$

$$F:\mathcal{D} \rightarrow \mathcal{X}(\mathcal{D}) \quad \text{from an initial state (D) to state (x)}$$

Bayes-inspired prediction is driven by a hypothesis: You "know" the answer, or at least part of it (your best guess). The conditional probability of a disease, given a set of findings, is where physicians start from (whether or not they are aware of Bayes). Predictions of election results, of weather patterns, of the outcome of sports competitions rely on similar assumptions. Prediction as a process that describes the outcome of action–reaction dynamics can be usefully affected by experiential evaluations. Predictive computations based on aggregating individual guesses ("crowdsourcing") are deployed in, for example, market research and political consultancy (e.g., Nate Silver's FiveThirty-Eight Political Calculus). DARPA's Policy Analysis Market (PAM) generalized from futures markets to predicting hostilities (probability of overt action or terrorist activity). Dissemination of accurate aggregated information was supposed to address national security concerns. The speed and accuracy of the market pinpointed what went wrong in the 1986 crash of the *Challenger*. PAM was quashed in

its preliminary phase, but the model it developed was not abandoned. Quite a large segment of predictive computation geared toward risk assessment rides on the model DARPA advanced in 2003. In some robotics applications, as well as in artificial intelligence (AI) procedures, various kinds of "auctions" take place: the adequate choice for the task is rewarded with resources to carry it out.

THE PROBABILITY SPACE AND ANTICIPATION

There are also predictions driven, to an extent larger than the Bayesian state of belief or the futures market model, by anticipatory processes, also involving the probability space. Facial expression as a predictor is another example of Bayesian probability-based inferences. Ekman and Rosenberg[7] have shown that the "language" of facial expression speaks of facts to happen before they are even initiated. The Facial Action Coding System (FACS), which is a taxonomy of facial expression (and the associated emotional semantics), inspired Rana El Kaliouby's prediction model by computationally interpreting the language of faces.

Predictions made according to known methods, such as time series analysis and linear predictors theory, capture the reaction component of human action.[8] Mechanisms, as embodiments of determinism, rarely fail. Perpetual calendar watches are a good example. And when they do, it is always for reasons independent of the mechanism's structure. Sensor-based acquisition of data provides in algorithmic computation the simuli of learning through experience. Evidently, the focus is on relationships as a substratum for deriving instructions pertinent to the present and of relevance to the future. Ignorance, which is what probabilities describe, is fought with plenty of data. The typology of predictions (linear, non-linear, statistical inference, stochastic, etc.) corresponds to the different perspectives from which change and its outcome are considered. At the processing level, extraction of knowledge from data makes available criteria for choices (such as those in spatial navigation, playing games, choosing among options, etc.).

Learning and Deep Learning

For learning (prerequisite to prediction and to anticipation) to come about, representations of the dynamic process have to be generated. Some will correspond to the immediateness of the evolving phenomena—what the next state will be; how the phenomena will evolve over time—others involve deeper levels of understanding. Whether in medicine, economy, politics, military actions, urban policy, or education, predictions or anticipations emerge on account of considerations regarding cascading amounts of data. The ever-increasing amount of sensors deployed can be considered as the source of this data, provided that sensor fusion is achieved and the aggregate data can be associated with meaning.

Sensors, inspired by the simplistic understanding of senses as discrete (five senses instead of the sensorial continuum), provide data about the physical world. They can be passive (such as a video camera or microphone) or active (the radar sensor sends a signal and measures the interaction between the signal and the immediate context). Engineering requirements of predictive processes stimulated the design of sensors for representing the external world (exteroperceptive), the state of the subject (proprioceptive), and the state of the system itself (interoceptive). As such, sensors do not entail predictivity, but are necessary conditions for achieving it. Integrated sensors generate high-level, multidimensional representations. Their interpretation, by individuals or intelligent agents, emulates the machine model of neuronal activity. As a consequence, we end up with algorithmic computation, extremely efficient in terms of generalizing from past to present. The so-called deep Q-network (DQN) agent, which has as output "human-level control" performance (in playing games, but applicable as well to other choice-making situations, such as robotic performance), is the embodiment of prediction based on reinforcement learning.[9] Computational power, such as a dedicated supercomputer designed for deep learning, enabled the training of larger models on increased amounts of data. As a result, visual object recognition is accelerated and predictive performance increased.[10]

In studying learning and selective attention, Dayan et al.[11] refer to reward mechanisms in the Kalman filter model (more experience leads to higher certainty). For any process in progress—e.g., moving a vehicle, recalling a detail in an image, thinking something out—there are, from the perspective of the Kalman filter, two distinct phases: 1) predict and 2) update. The filter is a recursive operation that estimates the state of a linear dynamic system. In physical entities, the space of observable parameters is smaller than that of describing the degrees of freedom defining the internal state. In the living, the situation is reversed. Learning, for instance, triggers expectations that turn out to be a measure of how much the deterministic *instinct* (*culture*) takes over the more complex model that accounts for both reaction and anticipation in the dynamics of the living.

Self-Awareness, Intentionality, Planning

A plan is the expression of understanding actions in relation to their consequences. In many ways, predictive computation is planning: goals (catch a flight), means to attain these goals (get a taxi for the airport), and the time sequence for achieving them (robots embody this understanding). A plan is a timeline, and a script for interactions indexed to the timeline. To what we call *understanding*—a prerequisite of anticipatory action—belong goals, the means, the underlying structure of the endeavor (tasks assumed by one person, by several, the nature of their relation, etc.), a sense of progression in time, and awareness of consequences, i.e., a sense of value. In every plan, from the

most primitive to the utmost complex, the goal is associated with the reality for which the plan provides a description (a theory), which is called *configuration space*.

Planning sets the limits within which adaptive processes are allowed. Each plan is in effect an expression of learning in action, and of the need to adapt to circumstances far from remaining the same. The robot learns how to distinguish among objects before acting upon some of them. Processes with anticipatory, predictive, and forecasting characteristics are described through:

> **Control → function of (past state, current state, future state) system**
>
> **Adaptivity → circumstances related to goals**

Knowledge of future states is a matter of possibilistic distributions:

$$r: \cup \to [0, 1]$$

in which ∪ defines the large space of values a variable can take. The function $\mathcal{R}$ is actually a fuzzy restriction associated with the variable X:

$$\mathcal{R}(X) = F$$

It is associated with a possibility distribution Π_x. Nothing is probable unless it is possible. Not every possible value becomes probable.

Functioning under continuously changing conditions means that control mechanisms will have to reflect the dynamics of the activity (Fig. 4). This is not possible without learning. If the automated part (everything involving the change of the physical can be automated) can be combined with human performance (expressed in behavior features), an architecture can be attained, one reflecting the hybrid nature of plan-driven human activities that feed values into the sensors. On the basis of these values, the system is reconfigured under the control of the dynamic model continuously refreshed in accordance with the behavior of the world. Learning results in the process of successive refreshment of data. Effectors act upon the world as a control procedure. This architecture is operationally different from that of the Google DeepMind Group (the new *Go Game* world champion). In the DeepMind, convolutional neural networks are used to appropriate the parameters that guide the action. The Q-network agent is nothing other than a reduction of anticipation to prediction. Indexed behavior features and the methods for extracting regularities characteristic of their behavior are connected. Learning ensues from adapting to new circumstances (i.e., change). This was the premise for the adaptive automobile, which first associated with it the notion of anticipatory computing.[12] If the generic diagram of the hybrid control mechanisms endowed with learning conjures associations with the smartphones of our time (i.e., mobile computing), it is not by accident. It also points to robot architectures (Fig. 5).
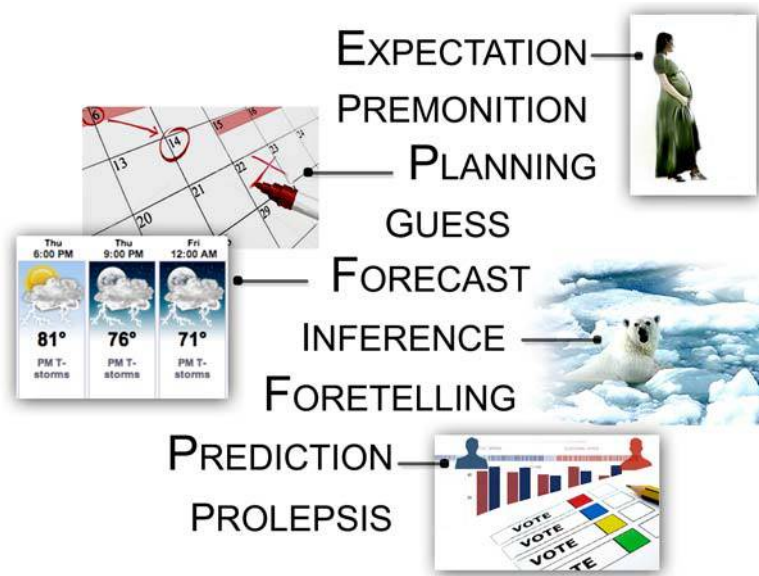
Fig. 4 Ways of considering the future.

A variety of predictive and anticipatory computing means and methods along the line of the conceptual foundation sketched so far were submitted for intellectual property protection. Patents were issued for "touch event anticipation in a computing device" (Microsoft, 2010), following Echostar Technologies' use of anticipatory preprocessing and Qualcomm's gesture-based commands. Mind Meld™ focused on speech recognition: *Show your users the right answer even before they finish speaking*. Before them, the prediction was focused on typing (the *autocomplete* routine automatically finished the word). Currently, it is SwiftKey™ that learns from the user's patterns of expression and predicts the type of message to be formulated. Cover™ will show on the phone screen apps appropriate to the context (at work, home, jogging, etc.). Apple, Inc. applied a proactive control method in conceiving the "connected home" (with the focus on interoperability). Wearable computing (from clothes, to rings, to earrings) integrates predictive devices for many applications (health, appliance maintenance, security, etc.). Songza™

(partnering with the Weather Channel) recommends music based on the weather. Amazon patented "anticipatory package shipping." More areas of predictive applications are inspiring a variety of theoretic and practical contributions. Robotics leads by far in this direction. The common denominator is the convergence of modeling, machine learning, and the more recent deep learning, and data mining. Predictive analytics is becoming more precise, and less dependent on access to large data sets.

PROBABILITY AND POSSIBILITY

In reference to various forms of computation that facilitate forecasting, prediction, planning, and even some anticipatory processes, it should be clear that regardless of the medium in which probability-based computing is attempted, what defines this kind of calculation is the *processing* of probabilities. Probability values can be inputted to a large array and processed according to a functional description. A probability distribution describes past events and takes the form of a statistical set of data. In this data domain, inductions (from some sample to a larger class) or deductions (from some principle to concrete instantiations), or both, serve as operations based upon which we infer from the given to the future. The predictive path can lead to anticipation. From regularities associated with larger classes of observed phenomena, the process leads to singularities. The inference is based on abduction, which is history dependent. Indeed, new ideas associated with hypotheses (another name for abduction) are not predictions, but an expression of anticipation (Fig. 6).

The interplay of probability and possibility is yet another option. This is relevant in view of the fact that information —i.e., data associated with meaning that results from being referenced to the knowledge it affords or is based
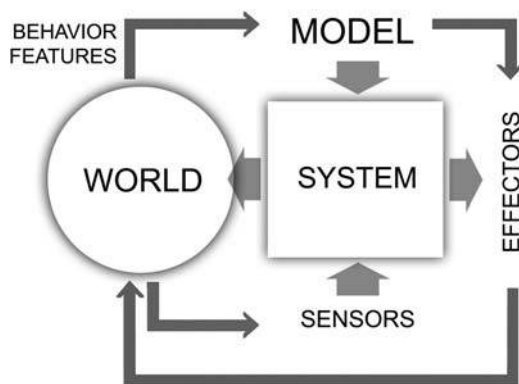

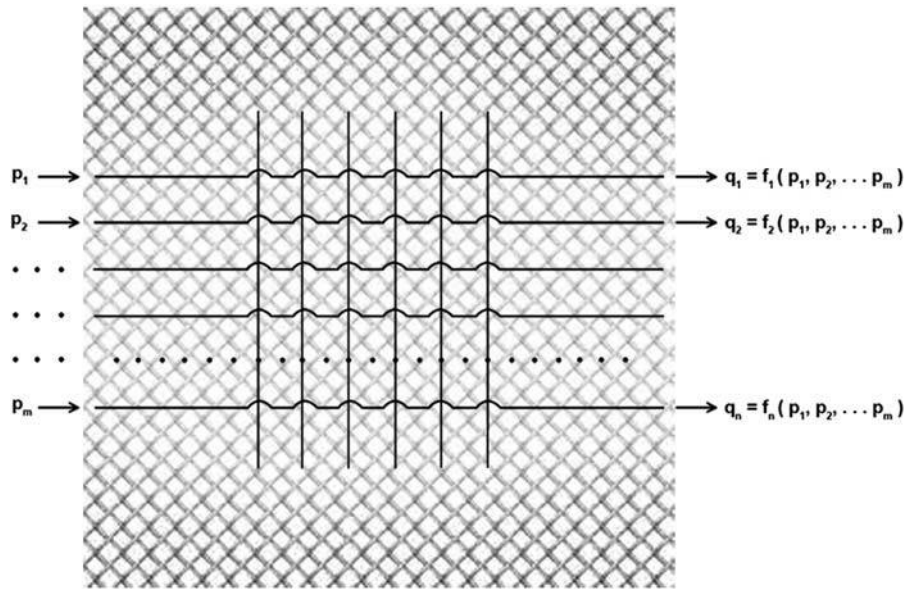**Fig. 5** Generic diagram of a hybrid control mechanism endowed with learning.

Fig. 6   Probability computer: the input values are probabilities of events. The integration of many probability streams makes possible dynamic modeling.

upon—can be associated with probability distributions (of limited scope within the [0,1] interval), or with the infinite space of possibilities corresponding to the nature of open-ended systems. Zadeh[13] and others took note of the fact that in Shannon's data-transmission theory (misleadingly called "information" theory), information is equated with a reduction in entropy, not with form (not morphology). He understood this reduction to be the source of associating information with probability. Possibilistic information, orthogonal to the probabilistic (one cannot be derived from the other) refers to the distribution of meaning associated with a membership function. In more illustrative terms (suggested by Chin-Liang Chang[14]), possibility corresponds to the answers to the question, "Can it happen?" (regarding an event). Probability (here limited to frequency,

which, as we have seen, is one view of it) would be the answer to, "How often?" (Clearly, frequency, underlying probability, and the conceivable, as the expression of possibility, are not interdependent Fig. 7.)

One particular form of anticipative evaluation can be computing *perceptions*.[15] The underlying logic for this is fuzzy logic, in which a qualifier (e.g., young, heavy, tall) is defined as a matter of degree. Anticipation from a psychological viewpoint is the result of processing perceptions, most of the time not in a sequential but in a configurational manner (in parallel). For instance, facial expression is an expression of anticipation (like/dislike, etc., expressed autonomously) based on perception. Soundscapes are yet another example (often of interest to virtual reality applications).
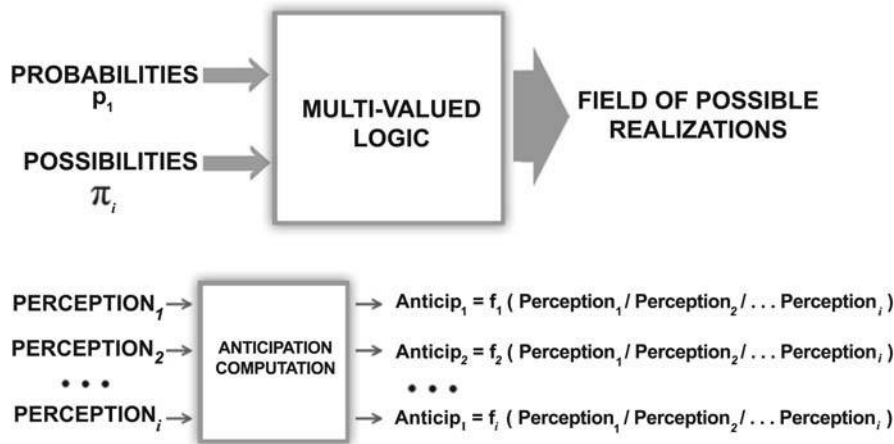

Fig. 7   Computing with probabilities and possibilities, computing with perceptions.

ANTICIPATION IMPROVES PREDICTION

If we could improve such predictions by accounting for the role of anticipation—the possible future state influencing, if not determining, the current state—science would be in a better position to deal with life-threatening occurrences (strokes, sudden cardiac death, diabetic shock, epileptic seizure, etc.[16]). Learning (especially deep reinforcement learning) about such occurrences in ways transcending their appearance and probability is one possible avenue. Things are not different in the many and varied attempts undertaken in predictions concerning the environment (for example, climate change), education, and market functioning.

Predictions, explicit or implicit, are a prerequisite of forecasting. The etymology points to a pragmatics that involves randomness—as in casting. Under certain circumstances, predictions can refer to the past (more precisely, to validation after the fact). This is the case for data fitted to a hypothesis. In other cases, what is measured as "noise" was treated as data. Procedures for effectively distinguishing between noise and data are slow in coming, and usually involve elements that cannot be easily identified. In medicine, where the qualifiers "symptomatic" vs. "non-symptomatic" are applied in order to distinguish between data and noise, this occurs to the detriment of predictive performance. Extrapolation can be erroneous.

In general, theories are advanced and tested against the description given in the form of data. Predictions pertinent to previous change (i.e., descriptions of the change) are comparable to descriptions geared to future change. In regard to the past, one can continue to improve the description (fitting the data to a theory) until some pattern is eventually discerned and false knowledge discarded.

THE PREDICTION MACHINE

Predictions are based on the explanatory models (explicit or not) adopted. Forecasts, even when delivered without explanation, are interpretive. They contain an answer to the question behind the forecasted phenomenon. A good predictive model can be turned into a machine.

Guesses, expectations, predictions, and forecasts—in other words, learning in a broad sense—co-affect human actions and affect pragmatics. Each of them, in a different way, partakes in shaping actions. Their interplay makes up a very difficult array of factors impossible to escape, but even more difficult to account for in detail. Mutually reinforcing guesses, expectations, predictions, and forecasts, corresponding to a course of events for which there are effective descriptions, allow, but do not guarantee, successful actions. Occasionally, they appear to cancel each other out, and thus undermine the action, or negatively affect its outcome. Learning and unlearning (which is different from forgetting) probably need to be approached together. Indeterminacy can be experienced as well. It

corresponds to descriptions of events for which we have no knowledge or for which we have insufficient information and experience. They can also correspond to events that by their nature seem to be ill-defined. The living, in all its varied embodiments, reacts and anticipates.

MACHINES FOR CALCULATIONS

Calculations seem the best path toward inferring from the present to the future. However, the nature of the calculations affects the outcome. But why automate? Leibniz provided a short answer: "… it is unworthy of excellent men to lose hours like slaves in the labor of calculation which could be safely relegated to anyone else if machines were used." This was written 12 years after he built (in 1673) a hand-cranked machine that could perform arithmetic operations (Fig. 8). Today, mathematics is automated, and thus every form of activity with a mathematical foundation, or which can be mathematically described, benefits from high-efficiency data processing. Still, the question of *why machines for calculation* does not go away, especially in view of realizations that have to do with a widely accepted epistemological premise: mathematics is the way to acquire knowledge. The reasonable (or unreasonable) effectiveness of mathematics in physics confirms the assumption. It also raises question regarding the representation of data pertinent to the living.

Wigner[17] contrasts the "miracle of the appropriateness of the language of mathematics for the formulation of the "laws of physics" to the "more difficult" task of establishing a "theory of the phenomena of consciousness, or of biology." Adamant on the subject, Gelfand and Tsetlin[18] went so far as to state, "There is only one thing which is more unreasonable than the unreasonable effectiveness of mathematics in physics, and this the unreasonable ineffectiveness of mathematics in biology." Predictive computation and more so anticipatory computation are subject to such considerations regarding the role of mathematics and probably the need for alternatives to it.

ANALOG AND DIGITAL; ALGORITHMIC AND NON-ALGORITHMIC

In the context of interest in machines of all kinds (for conducting wars, for successful wagers, for calculating the position of stars, for navigation, for making things, for dealing with dangerous situations, etc.), the theoretic machine called *automaton* was the most promising. For a while, what happened in the box (how the gears moved in Leibniz's machine, for example) and what rules were applied—which is the same as saying which algorithm was used—was not subject to questioning. Once the model of the neuron—more precisely, its deterministic reduction—was adopted, a discussion on the nature of computation was triggered (Fig. 9).
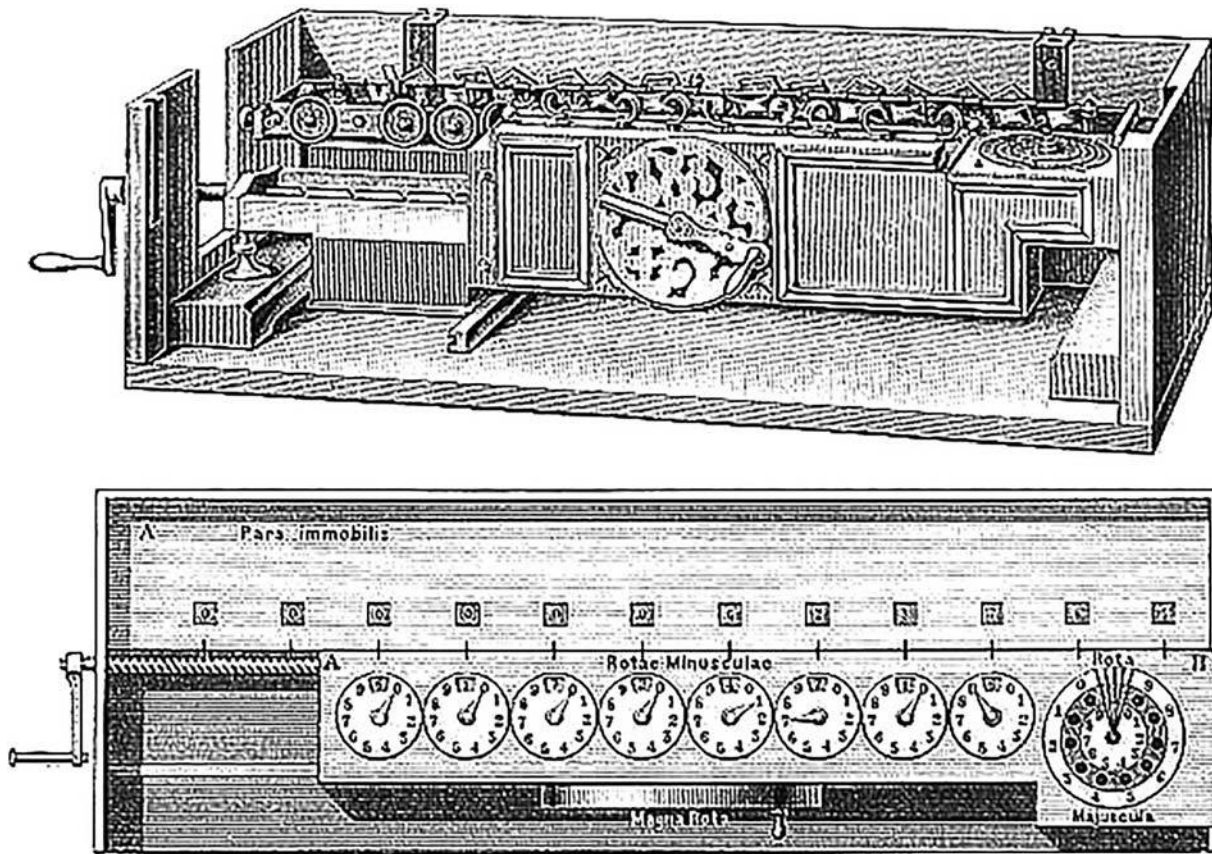
Fig. 8   Leibniz machine: Algorithm in hardware.

It is important to understand that, for the neuron, input values are no longer given and that in the calculation scheme of neuronal networks, the *machine* is "taught" (through training) what it has to do. This applies from the simplest initial applications of the idea (the McCulloch and Pitts model) to the most recent DQN that combines reinforcement learning in association with deep neural networks (in the case of mimicking feed-forward processing, in early visual cortex; Hubel and Wiesel[19]).

Evidently, the subject of interest remains the distinction between reaction-based processes—the theoretic machine has an input, a number of inner states, and an

Fig. 9   A reductionist neuron-based model.

output that is the outcome of the calculation—and predictive performance.

TURING'S MACHINES AND PREDICTION

Hilbert's conjecture that mathematical theories from propositional calculus could be decided—*Entscheidung* is the German word for decision, as in proven true-or-false—by logical methods performed automatically was rejected. Turing provided the mathematical proof that machines cannot do what mathematicians perform as a matter of routine: develop mathematical statements and validate them. Nevertheless, the insight into what machines can do, which we gain from Turing's analysis, is extremely important. Turing[20,21] stated, "A man provided with paper and pencil and rubber, and subject to strict discipline, is in effect a universal machine." At a different juncture, he added: "disciplined but unintelligent."[22] Gödel would add, "Mind, in its use, is not static, but constantly developing."[23] This is where prediction and anticipation are couched. "Strict discipline" means: *following instructions*. Instructions are what, by consensus, became the algorithm. Intelligence at work often means shortcuts, new ways for performing an operation, and even a possible wrong
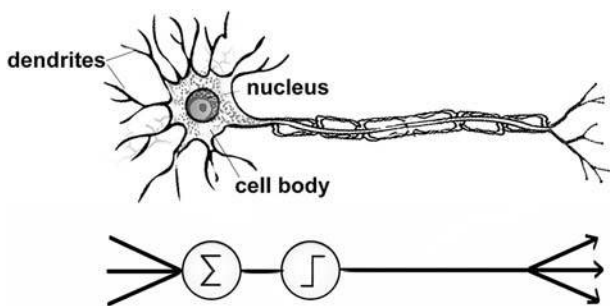
decision. Therefore, non-algorithmic means are not subject to predefined rules, but rather discovered as the process advances, as predictions are made.

Automatic Machines

A-machines, as Turing labeled them, can carry out any computation that is based on complete instructions; that is, they are algorithmic. The machine's behavior is predetermined; it also depends on the time context: whatever can be fully described as a function of something else with a limited amount of representations (numbers, words, symbols, etc.) can be "measured," i.e., completed on an algorithmic machine. The algorithm is the description. With the *a-machine*, a new science is established: the knowledge domain of decidable descriptions of problems. It ascertains that there is a machine that can effectively measure all processes— physical or biological, reactive or anticipatory—as long as they are represented through a computational function.

Choice, Oracle, and Interactive Machines

In the same paper,[22] Turing suggested different kinds of computation (without developing them). Choice machines, i.e., *c-machines*, involve the action of an external operator. Even less defined is the *o-machine* (the oracle machine advanced in 1939), which is endowed with the ability to query an external entity while executing its operations. The *c-machine* entrusts the human being with the ability to interact on-the-fly with a computation process. The *o-machine* is rather something like a knowledge base, a set subject to queries, and thus used to validate the computation in progress. The oracle's dynamics is associated with sets. Through the *c-machine* and the *o-machine*, the reductionist *a-machine* is opened up. Interactions are made possible— some interactions with a living agent, others with a knowledge representation limited to its semantic dimension. Interactions are future-oriented queries. Furthermore, Turing diversifies the family of his machines with the *n-machine*, unorganized machine (of two different types), leading to what is known today as neural networks computation (the B-type *n-machines* having a finite number of neurons), which is different in nature from the algorithmic machine.

There is one more detail regarding Turing's attempt to define a test for making the distinction between computation-based intelligence and the human intelligence possible: human intelligence corresponds to the anticipatory nature of the living. Therefore, to distinguish between machine and human intelligence (the famous "Turing test") is quite instructive for our understanding of anticipation.

Interactivity

Let us not lose sight of interactivity, of which Turing was aware, since on the one hand Turing computation is captive to the reductionist–deterministic premises within which only the reaction component of interactivity is expressed, and, on the other, since interaction computing[24] is not reducible to algorithmic computation. The most recent developments in the areas of robotics, quantum computation, evolutionary computation, and even more so in terms of computational ubiquity (in mobile computing and wearables associated with sensory capabilities), represent a grounding for the numerous interrogations compressed in the question: Is anticipatory computation possible? Moreover, the "Internet of Everything" (IoE) clearly points to a stage in computation that integrates reactive and anticipatory dimensions.

IS ANTICIPATORY COMPUTATION POSSIBLE?

Anticipation comes to expression within entities, the description of which is undecidable. The criterion for this distinction is derived from Gödel's[25] notion of the undecidable (the first incompleteness theorem originally appeared as Theorem VI): entities of complex nature, or processes characterized as complex, cannot be fully and consistently described.[26] The living is undecidable.[26] Anticipation pertains to change, i.e., to the sense of the future that the living has. Quantum processes transcend the predictable; they are non-deterministic. Consequently, their descriptions entail the stochastic (the aim), which is one possible way to describe non-deterministic processes. To the extent that such quantum-based computers are embodied in machines, one cannot expect them to output the same result all the time (Fig. 10).
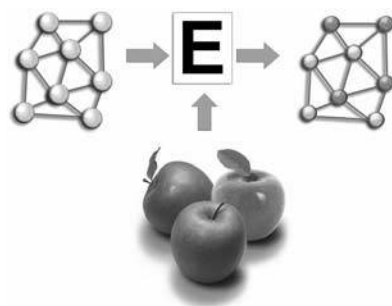

Fig. 10 Quantum computation used in image recognition: apples and a moving car.

Rather, such a computer has no registers or memory locations, and therefore to execute an instruction means to generate samples from a distribution. There is a collection of qubit values—a qubit being a variable defined over the interval {0,1}. A certain minimum value has to be reached. Currently, the art of programming is to affect weights and strengths that influence the process analyzed. Instructions are not deterministic; the results have a probabilistic nature.

Predictive calculations are in one form or another inferences from data pertinent to a time of reference (t_0) to data of the same phenomenon (or phenomena) at a later time ($t_1 > t_0$). Phenomena characteristic of the physical can be precisely described. Therefore, even if non-linearity is considered (a great deal of what happens in physical reality is described through non-linear dependencies), the inference is never of a higher order of complication than that of the process of change itself. In quantum phenomena, the luxury of assuming that precise measurements are possible is no longer available. Even the attempt to predict a future state affects the dynamics, i.e., the outcome. It is important to understand not only how sensitive the process is to initial conditions, but also how the attempt to describe the path of change is affected in the process.

In computations inspired by theories of evolution or genetics, the situation is somehow different. Without exception, such theories have been shaped by the determinism of physics. Therefore, they can only reproduce the epistemological premise. But the "computations" we experience in reality—the life of bacteria, plants, animals, etc.—are not congruent with those of the incomplete models of physics upon which they are based. Just one example: motoric expression (underlying the movement of humans and animals) might be regarded as an outcome of computation—an extrapolation practiced much too often. But in doing so, the complexity of the process is reduced. Even simple movements are indeterminate (there are many degrees of freedom). Motion control (the subject through which Bernstein[27] introduced his concept of anticipation), decision-making, navigation, and autonomic behavior have informed engineering efforts that extend from endowing robots with capabilities comparable to those of humans to conceiving predictive systems whose performance exceeds that of their creators (for example, IBM's *Big Blue*).

CURRENT APPLICATIONS AND FUTURE POSSIBILITIES

Robotics

If the origin of a word has any practical significance to our understanding of how it is used, then *robot* tells the story of *machines supposed to work* (*robota* being the Russian word that inspired the Czech Karel Čapek to coin the term). Therefore, like human beings, they ought to have predictive

capabilities: when you hit a nail with a hammer, your arm seems to know what will happen. From the many subjects of robotics, only predictive and anticipatory aspects, as they relate to computation, will be discussed here.

The predictive abilities of robots pose major computational challenges. In the living, the world, in its incessant change, appears as relatively stable. Motor control relies on rich sensor feedback and feed-forward processes. Guiding a robot (toward a target) is not trivial, given the fact of ambiguity: How far is the target? How fast is it moving? In which direction? What is relevant data and what is noise? Extremely varied sensory feedback as a requirement similar to that of the living is a prerequisite, but not a sufficient, condition. The living does not passively receive data; it also contributes predictive assessments—feed forward—ahead of sensor feedback. This is why robot designers provide a forward model together with feedback. The forward (prediction of how the robot moves) and inverse (how to achieve the desired speed) kinematics are connected to path planning. The uncertainty of the real world has to be addressed predictively: advancing on a flat surface is different from moving while avoiding obstacles (Fig. 11).

Intelligent decisions require data from the environment. Therefore, sensors of all kinds are deployed. To make sense of the data, the need for sensor fusion becomes critical. The multitude of sensory channels and the variety of data formats suggested the need for effective fusion procedures. As was pointed out,[28,29] the position of arms, legs, fingers, etc. correspond to sensory information from skin, joints, muscles, tendons, eyes, ears, nostrils, and tongue. Redundancy, which in other fields is considered a shortcoming
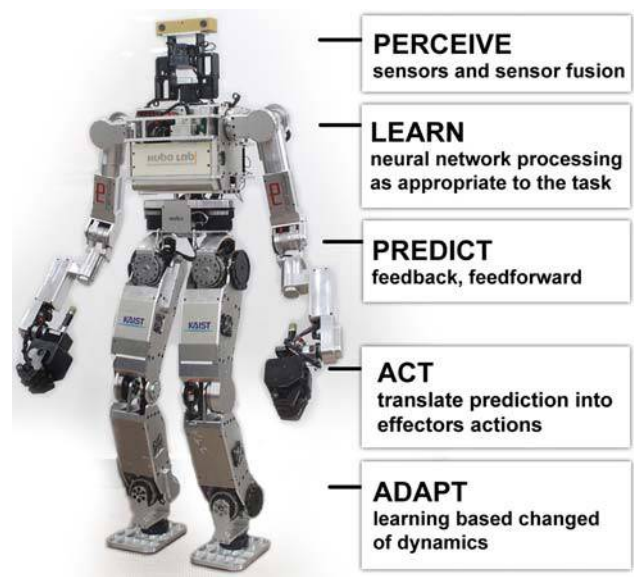

Fig. 11 Prediction permeates the dynamics of robots. The robot displayed serves only as an illustration. It was an entry from KAIST (Korea Institute for Science and Technology) in the DARPA Robotics Challenge Competition (2014).

(costly in terms of performance) helps eliminate errors due to inconsistencies or sensor data loss and due to compensation of variances. The technology embodied in neurorobots endowed with predictive and partial anticipatory properties (e.g., "Don't perform an action if the outcome will be harmful") integrates RNNs, multilayered networks, Kalman filters (for sensor fusion), and, most recently, deep learning architectures for distinguishing among images, sounds, etc., and for context awareness.[30] Robots require awareness of their state and of the surroundings in order to behave in a predictive manner. (The same holds for wearable computers.)

MOBILE COMPUTING: AN UNEXPECTED ALTERNATIVE

Mobile computing, which actually is the outgrowth of cellular telephony, offers an interesting alternative. From the initial computer–telephone integration (CTI) to its many current embodiments (tablets, notebooks and netbooks, smartphones, watches, Google Glass™, etc.), mobile computing evolved into a new form of computation. First and foremost, it is interactive: somehow between the *c-machine* and *o-machine* envisaged by Turing. The computer *sine qua non* telephone is also the locus of sensor interactions. In other words, we have a computer that is a telephone in the first place, but actually a video scanner with quite a number of functions in addition to communication. On a head-mounted wearable device, such as the Google Glass, one can identify a touchpad, a see-through display, a camera, a microphone, a magnetometer, and sensors to characterize cardiovascular and respiratory activity, as well as other unobtrusive sensors. For instance, subtle movements of the head are associated with respiration patterns and heart activity. The result is close to ballistocardiography, an efficient diagnostic method.

The Mobile Paradigm and Anticipatory Computing

The integration of a variety of sensors from which data supporting rich interactions originate is the most difficult challenge. But no predictive performance is possible without such integration (Fig. 12).

Distinct levels of processing are dedicated to logical inferences with the purpose of minimizing processing. Anticipation is expressed in action pertinent to change (*adapt* or *avoid* are specific actions that everyone is aware of). It seems trivial that under stress, anticipation is affected. It is less trivial to detect the degree and the level of stress from motoric expression (abrupt moves, for instance) or from the speech data. A utility, such as StressSense™, delivers useful information, which is further associated with blood pressure, heart rhythm, and possibly electromyography (EMG), the results of which can assist the individual in mitigating danger. The spelling out of specific procedures—such as the Gaussian mixture models (GMM) for distinguishing between stressed and neutral pitch—is probably of interest to those technically versed, but less so for the idea discussed.

El Kaliouby developed a similar facility for reading facial expression. This facility makes available information on attention—the most coveted currency in the world of computer-supported interactions. Initially, the MindReader was merely making predictions under the guidance of a Bayesian model of probability inferences. Currently, the focus is more on associating emotional states and future choices. The system can be integrated into mobile devices, Google Glass or into the Apple Watch™. The description of physical processes (cause-and-effect sequence) and that of the living process, with its anticipatory characteristics, fuse into one effective model. This is a dynamic model, subject to further change as learning takes place and adaptive features come into play.

In the physical realm, data determine the process.[31] For instance, in machine learning, the structure of classifiers—simple, layered, complicated—is partially relevant. What counts is the training data, because once identified as information pertinent to a certain action, the training data will guide the performance. However, the curse of dimensionality does not spare mobile computing. Data sets scale exponentially with the expectation of more features.

At this time in the evolution of computation, the focus is changing from data processing to understanding the meaning. This is no longer the age of human computers or of computers imitating them for the purpose of calculating

Fig. 12 Sensor integration with the purpose of facilitating rich interactions.

the position of stars, or helping the military to hit targets with their cannons. Routine computation (ledger, databases, and the like) is complemented by intelligent control procedures. Self-driving cars, boats, or airplanes follow the smart rockets (and everything else that the military commissioned to scientists), still within the spirit of DARPA's Augmented Cognition Program. It is easy to imagine that the DQN will soon give place to even higher-performing means and methods that outperform not only the algorithms of games, but also the spectacular intelligent weapons.

INTEGRATED COMPUTING ENVIRONMENT

In the area of mobile computation, the meeting of many computational processes, some digital, some analog (more precisely, some manner of signal processing), is the most significant aspect. Signal processing, neural network computation, telemetry, and algorithmic computation are seamlessly integrated. The aspect pertinent to anticipation is specifically this integration. Sensor fusion proved critical in robotics as well.

In this sense, we need to distinguish between actions initiated purposely by the user (taking a photo or capturing a video sequence) and actions triggered automatically by the behavior of the person carrying the device (sensing of emotional state, evaluating proximity, issuing orientation cues pertinent to navigation). It is not only the *a-machine* on board (the computer integrated in the "smartphone"), but the mobile sensing connected to various forms of machine learning based on neuronal networks and the richness of interactions facilitated, which make up more than an algorithmic machine. The execution of mobile applications using cloud resources qualifies this as an encompassing information processing environment. Taken independently, what is made available is a ubiquitous calculation environment. In this ever-expanding calculation environment, we encounter context sensing, which neither the desktop machine nor any other computer provides, or considers relevant for their performance.

In mobile computation, motion tracking, object recognition, interpretation of data, and the attempt to extract meaning—all part of the calculation environment—are conducive to a variety of inferences. This is an embodied interactive medium, not a black box for calculations transcending the immediate. The model of the future, still rudimentarily limited to predictable events, reflects an "awareness" of location, of weather, of some environmental conditions, or of a person's posture or position. A pragmatic dimension can be associated with the interpreted *c-* and *o-machines*: "What does the user want to do?"—find a theater or a bar, take a train, reserve a ticket, dictate a text, initiate a video conference, etc. Inferring usage (not far from guessing the user's intentions) might still be rudimentary. Associated with learning and distribution of data over the cloud, inference allows for better guessing, forecast, and

prediction, and becomes a component of the *sui generis* continuous planning process. The interconnectedness between the human and the device is extended to the interconnectedness over the network, i.e., cloud. Using statistical data from modeling, machine learning, and data mining, predictive analytics makes choices available (as predictions). Regression techniques are used for representing interactions among variables.

From a technological perspective, what counts in this environment is the goal of reaching close-to-real-time requirements. For this, a number of methods are used: sampling (instead of reaching a holistic view, focus on what might be more important in the context), load shedding (do less without compromising performance), sketching, aggregation, and the like. A new category of algorithms, dedicated to producing approximations and choosing granularity based on significance, is developed for facilitating the highest interaction at the lowest cost (in terms of computation).

It is quite possible that newer generations of such integrated devices will avoid the centralized model in favor of a distributed block chain process. Once issues of trust (of extreme interest in the context of vulnerability) are redefined among those who make up a network of reciprocal interest, anticipation and resilience will bind. The main reason to expend effort in dealing with a few aspects of this new level of computation is that it embodies the possibility of anticipatory computing.

In the evolution from portable wireless phones to what today is called a "smartphone," these interactive mobile computing devices "learned" how to distinguish commuting, resting, driving, jogging, or sleeping, and even how to differentiate between the enthusiasm of scoring in a game and the angry reaction (game-related or not). Once the activity (current or intended) is identified, predictions can be made regarding their unfolding. A short description for suggesting the level of technological performance will help in further seeing how integration of capabilities scales to levels comparable to those of anticipatory performance. From GPS connection (and thus access to various dynamic knowledge bases), to sensors (accelerometer, gyroscope, etc.), to communication protocols (facilitating WiFi, Bluetooth, near-field communication), everything is in place to locate the user, the device, the interconnected subjects, and the actions meaningful within the context. Multi-core processors, large memories (not the infinite Turing machine tape but, by extension, to the cloud, close to what an infinite memory could be associated with), and high performance input and output devices (cameras, microphones, touch screen, temperature sensitive surfaces) work in concert in order to support the generation of a user profile that captures stable, as well as changing, aspects (identity and dynamic profile). Models connect raw sensed data in order to interface (the ambient interface) the subject in the world and the mobile station. Information harvested by a variety of sensors (multimodal level of sensing) is subject to

disambiguitization. It is exactly in respect to this procedure of reducing ambiguity that the mobile device distinguishes between the motorics of running, walking, climbing stairs, or doing something else (still within a limited scope). The attempts to deploy physical therapy, and keep a record of it, based on the mobile device rely on this level. The habit component compounds "historical" data—useful when the power supply has to be protected from exhaustion. Actions performed on a routine basis do not have to be recomputed. Other such strategies are deployed in using the GPS facility (path tracking, but only as the device moves, i.e., the user is on a bike, on a car, train, etc.). Overall, the focus is on the *minima* (approximate representations). Instead of geolocation proper, infer location from data (as recorded in the person's calendar: restaurant, doctor, meeting, etc.).

There is no need for excessive precision in the performance of most of the mobiles. (This is why sampling, load shedding, aggregations, etc. are used.) Nevertheless, the user taking advantage of the on-the-fly translation of a phone/video conversation easily makes up the missing details (where sketching is important), or corrects the sentence. Images are also subject to such corrections.

COMMUNITY SIMILARITY NETWORKS: HOW DOES THE BLOCK CHAIN MODEL SCALE UP TO PREDICTIVE COMPUTING?

The anticipation potential of interactive mobile devices is significant in the context of the tendency to scale from individuals to communities. Autonomous Decentralized Peer-to-Peer Telemetry (the ADEPT concept that IBM devised in partnership with Samsung) integrates proof-of-work (related to functioning) and proof-of-stake (related to reciprocal trust) in order to secure transactions. Mobile computation becomes part of the ecology of billions of devices, some endowed with processing capabilities and others destined to help in the acquisition of significant data. Each device—phone, objects in the environment, tagged individuals and flora and fauna—can autonomously maintain itself. Devices signal operational problems to each other and retrieve software updates as these become available, or order some as needed. There is also a barter level for data (each party might want to know ahead of time "What's behind the wall?"), or for energy ("Can I use some of your power?"). There is no central authority; therefore, one of the major vulnerabilities of digital networks supporting algorithmic computation is eliminated. Peer-to-Peer facilitates the establishment of dynamic communities: they come into existence as necessary and cease to be when their reason for being no longer exists.

This is an example of decentralized consensus building in which the aggregate choices continuously refresh the distributed cryptographically encoded ledger shared by those pursuing a common goal. Harnessing block chain

technology, such as tools for smart contracts that Ethereum (Canada) is developing, the community similarity networks (CSNs) reactivate DARPA's PAM. The outcome of real-world events (e.g., medical treatment, political conflicts, financial transactions, diplomatic negotiations) is subject to a chain prediction.

So-called CSNs associate users sharing in similar behavior. A large user base (such as the Turing *o-machine* would suggest) constitutes over time an ecosystem of devices. Fitbit™ (a digital armband) already generates data associated with physical activities (e.g., exercise, rest, diet) and prompted the spontaneous aggregation of data for improved interpretations. A variety of similar contraptions (the "chip in the shoe," heart monitors, hearing- or seeing-aid devices) also generates data, inviting a new understanding of its meaning. The Apple Watch, the Google Glass, or any other integrating artifact scales way further as a health-monitoring station. One can envision real-time physiological monitoring diagnostic devices. Their deployment could result in a new situation where medical conditions are identified and treated before they become symptomatic. The Precision Medicine Initiative (PMI) is heading in this direction. This could be rewritten to apply to economic processes and transactions, to political life, to art, and to education. The emphasis is on before, which characterizes anticipation.

Some researchers (Pejovic and Musolesi,[32] among others) advance the idea of behavior change interventions from an anticipatory perspective. Indeed, behavior change (in terms of diet, exercise, respiration, posture, etc.) informed by a "smart" device is action, and anticipation is always expressed in action. Few realize that posture (affecting health in many ways) depends a lot on respiration. Upon inspiration (breathing in), the torso is deflected backward, and the pelvis forward. It is the other way around during expiration (breathing out). Anticipation is at work in maintaining proper posture as an integrative process. Behavior change interventions could become effective if this understanding is shared with the individual assisted by integrated mobile device facilities. Similar projects could be expected for behavior guidance on the economy, social and political life, education, art, etc. Indeed, instead of reactive algorithmic remedies to crises (stock market crash, bursting of economic bubbles, inadequate educational policies, ill-advised social policies, etc.), anticipatory capabilities could be embedded in new forms of computation.

THE NEXT FRONTIER: COMPUTATION AS UTILITY

Engineering still dominates predictive computation endeavors. It is quite possible that computer scientists will formalize hypotheses originating from the current trial-and-error phase. From data-rich, theory-poor status (similar to the knowledge domain of anticipation), the state of the art in predictive and anticipatory engineering will progress to
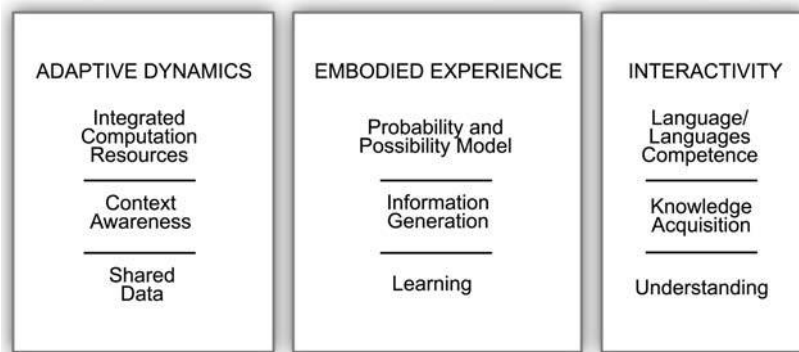
Modeling—Predictive

Fig. 13 Adaptive dynamics, embodied experience, and rich interactivity are premises for anticipatory performance.

the discovery of new principles and, quite probably, alternative forms of computation. Progress in predictive computation technology, sometimes confusingly branded as anticipatory, is promising. The next frontier is anticipatory computation as a hybrid human–machine interactive entity. For a computation to qualify as anticipatory, it would have to be couched in the complexity corresponding to the domain of the living. Specifically, complexity corresponds to variety and intensity of forms of interaction, not to material or structural characteristics of the system. The interaction of the mono-cell (the simplest form of the living) with the environment by far exceeds that of any kind of machine. This interactive potential explains the characteristics of the living. According to the foundations of anticipatory systems, the following are necessary, but not sufficient, conditions for anticipatory computation:

- Self-organization (variable configurations)
- Multiplicity of outcome
- Learning: performance depends on the historic record
- Abductive logic
- Internal states that can affect themselves through recursive loops
- Generation of information corresponding to history, context, goal
- Effective ontology engineering
- Operation in the non-deterministic realm
- Open-endedness

In practical terms, anticipatory computing would have to be embodied (in effective agents, robots, artifacts, etc.) in order to be expressed in action. A possible configuration would have to integrate adaptive properties, an efficient expression of experience, and, most important, unlimited interaction modalities (integrating language, image, sound, and all possible means of representation of somatosensory relevance) (Fig. 13).

In view of the newly acquired awareness of decentralized interaction structures—i.e., pragmatic dimensions of computation—it can be expected that computation as a utility, not as an application, would be part of the complex process of forming, expressing, and acting in anticipation. Achieving an adaptive open system is most important.

Real progress in understanding where the journey into anticipatory computing should take us is to be expected in the years to come as anticipatory processes themselves are better understood.

ACKNOWLEDGMENTS

REFERENCES

1. Rosen, R. *Life, Itself. A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*; Columbia University Press: New York, 1991.
2. Dubois, D.M. Computing anticipatory systems with incursion and hyperincursion, CASYS, AIP Conference Proceedings; New York, 1998, 437, 3–29.
3. Nadin, M. *Mind—Anticipation and Chaos*; Belser: Stuttgart/Zurich, 1991.
4. Pezzulo, G. MindRACES: From Reactive to Anticipatory Cognitive Embodied Systems, 2004–2007. 2008 Report. http://www.mindraces.org/index_html/view?searchterm= pezzulo (accessed May 2015).
5. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*; Elsevier: Amsterdam, 2011.
6. Fletcher, P.C.; Anderson, J.M.; Shanks, D.R.; Honey, R.; Carpenter, T.A.; Donovan, T.; Papadakis, N.; Bullmore, E.T. Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. Nat. Neurosci. **2001**, *4* (10), 1043–1048.
7. Ekman, P.; Rosenberg, E.L. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*, 2nd Ed.; Oxford University Press: New York, 2005.

8. Arsham, H. *Time Series Analysis and Forecasting Techniques*, 2002. http://ubmail.ubalt.edu/~harsham/stat-data/opre330.htm (accessed February 2015).

9. Minh, V.; Kavukculoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; Petersen, S.; Beattle, C.; Sadik, A.; Antonoglu, I.; King, H.; Kumaran, D.; Wierstra, D.; Legg, S.; Hassabis, D. Human-level control through deep reinforcement learning. Nature **2015**, *518*, 529–533. http://www.nature.com/search?journals=nature%2Cnews&q=volodymyr&shunter=1424698400067 (accessed February 2015).

10. Ren, W.; Yan, S.; Shan, Y.; Sun, G.; Dang, Q. *Deep Image: Scaling up Image Recognition. arXiv:1501.02876 S[cs.CV]* (accessed May 2015).

11. Dayan, P.; Kakade, S.; Montague, P.R. Learning and selective attention. Nat. Neurosci. **2000**, *3*, 1218–1223.

12. Nadin, M. Anticipatory mechanisms for the automobile Lecture presented at AUDI Headquarters. **2003**. http://www.nadin.ws/wp-content/uploads/2007/06/audi190203.pdf (accessed March 15, 2015). See also: The car that ages with you. http://www.nadin.ws/archives/2295 (accessed June 23, 2015)

13. Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst. **1978**, *1*, 3–28.

14. Chang, C-L. Fuzzy sets and pattern recognition (Doctoral Thesis). University of California: Berkeley. **1967**.

15. Zadeh, L.A. Foreword. In *Anticipation—The End Is Where We Start From*; Nadin, M. Ed; Lars Müller: Baden, 2003.

16. Nicolelis, M.A.L.; Lebedev, M.A. Principles of neural ensemble physiology underlying the operation of brain–machine interfaces. Nat. Rev. Neurosci. **2009**, *10* (7), 530–540.

17. Wigner, E. The unreasonable effectiveness of mathematics in the natural sciences. Richard Courant Lecture in Mathematical Sciences delivered at New York University, May 11, 1959. Commun. Pure Appl. Math. **1960**, *13*, 1–14.

18. Gelfand, I.M.; Tsetlin, M.L. Mathematical modeling of mechanisms of the central nervous system. In *Models of the Structural–Functional Organization of Certain Biological Systems*; Gelfand, I.M., Ed., Beard, C.R., Trans.; MIT Press: Cambridge, MA, 1971.

19. Hubel, D.H.; Wiesel, T.N. Receptive fields of single neurones in the cat's striate cortex. J. Physiol. **1963**, *165* (3) 559–568.

20. Turing, A.M. *Intelligent Machinery [Technical Report]*. National Physical Laboratory: Teddington, 1948.

21. Copeland, B.J. Ed., *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life Plus the Secrets of Enigma*; Oxford University Press: Oxford, 2004.

22. Turing, A.M. Programmers' Handbook for Manchester electronic computer. Mark II, Computing Machine Laboratory, Manchester University: Manchester, England, **1951**.

23. Gödel, K. *Some Remarks on the Undecidability Results, Collected Works II, 305–306*; Oxford University Press: Oxford, 1972.

24. Eberbach, E.; Goldin, D.; Wegner, P. Turing's ideas and models of computation. In *Alan Turing. Life and Legacy of a Great Thinker*; Teuscher, C. Ed.; Springer: Berlin, Heidelberg, 2004.

25. Gödel, K. Über formal unentscheidbare Sätze der Principa Mathematica und verwandte Systeme. Monatshefte für Mathematik und Physik **1931**, *38*, 173–198.

26. Nadin, M. G-complexity, quantum computation and anticipatory processes. Comput. Commun. Collab. **2014**, *2* (1), 16–34.

27. Bernstein, N.A. *O Postroenii Dvizenii [On the Construction of Movements]*; Medgiz: Moscow, 1947.

28. Makin, T.R.; Holmes, N.P.; Ehrsson, H.H. On the other hand: dummy hands and peripersonal space. Behav. Brain Res. **2008**, *191*, 1–10.

29. Nadin, M. Variability by another name: "repetition without repetition." In *Anticipation: Learning from the Past. The Russian/Soviet Contributions to the Science of Anticipation*; Springer (Cognitive Systems Monographs): Switzerland, 2015; 329–336.

30. Schilling, M.; Cruse, H. What's next: recruitment of a grounded predictive body model for planning a robot's actions. Front. Psychol. **2012**, *3*, October, 1–19.

31. Landauer, R. The physical nature of information. Phys. Lett. A **1996**, *217*, 188–193.

32. Pejovic, V.; Musolesi, M. Anticipatory mobile computing: a survey of the state of the art and research challenges. ACM Computing Surveys, *V*, (N) Association for Computing Machinery: New York, **2014**. http://arxiv/pdf/1306.2356.pdf (accessed February 2015).

Modeling—Predictive