Mihai Nadin *Editor*

Anticipation Across Disciplines

Springer

Mihai Nadin   *Editor*

Anticipation Across Disciplines

Never before was anticipation more relevant to the life and activity of humankind than it is today. "It is no overstatement to suggest that humanity's future will be shaped by its capacity to anticipate...." (*Research Agenda for the 21st Century*, National Science Foundation). The sciences and the humanities can no longer risk explaining away the complexity and interactivity that lie at the foundation of life and living. The perspective of the world that anticipation opens justifies the descriptor "the post-Cartesian Revolution."

If anticipation is a valid research domain, what practical relevance can we await? Indeed, anticipation is more than just the latest catch-word in marketing the apps developed by the digital technology industry. Due to spectacular advances in the study of the living, anticipation can claim a legitimate place in current investigations and applications in the sciences and the humanities. Biology, genetics, medicine, as well as politics and cognitive, behavioral, and social sciences, provide rich evidence of anticipatory processes at work. Readers seeking a foundation for anticipation will find in these pages recent outcomes pertinent to plant life, political anticipation, cognitive science, architecture, computation. The authors contributing to this volume frame experimental data in language that can be shared among experts from all fields of endeavor. The major characteristic is the inference from the richness of data to principles and practical consequences.

Anticipation and Computation: Is Anticipatory Computing Possible?

Mihai Nadin

Abstract Anticipation, a definitory characteristic of the living, is expressed in action. It implies awareness of past, present, and future, i.e., of time. Anticipatory processes pertain to the world's dynamics. Anticipation also implies an observation capability, the acquired function of processing what is observed, and the ability to effect change. Computation means processing quantitative distinctions of physical entities and of those that inform the condition and behavior of the living. Autonomic processing is the prerequisite for anticipatory expression. In the physical, processing is reactive; in the living it is autonomic. Automated calculations, inspired by human "computers," are different in nature from those involved in living dynamics. To distinguish between anticipatory and predictive computation is to account for the role of the possible future in dealing with change.

1 Theoretical Considerations

1.1 Preliminaries

The context is clear: it was asserted (Mitchell [1]) that the world to which we belong is the outcome of computation (of quantum nature, in particular, according to Deutsch [2]). Consequently, to understand computation is a prerequisite for evaluating the message of concern regarding the long-term consequences of the

M. Nadin (✉)
anté—Institute for Research in Anticipatory Systems,
University of Texas at Dallas, Richardson, TX, USA
e-mail: nadin@utdallas.edu
URL: http://www.nadin.ws

increasing dependence on a particular machine, i.e., the computer, that humankind is experiencing. The broad view recalled above does not assuage the worry some express. But even if the hypothesis were to prove wrong, the dependence would not go away. In particular, artificial intelligence, together with the associated science of robotics, has prompted messages of doom: "The development of full artificial intelligence (AI) could spell the end of the human race" (Hawking [3]). Such messages are comparable, I hasten to add, to those euphoric forecasts (Kurzweil [4]) announcing an age in which machines will outsmart even those who conceived them. Shannon [5] went as far as to say that "I visualize a time when we will be to robots what dogs are to humans, and I'm rooting for the machines." (As respectful as I am of Shannon, I am not willing to accept the leash.)

Of particular interest in this respect are achievements in the area of predictive computation and, related to it, in neural networks-based deep reinforcement learning competing with human-level control. The most common embodiment of these developments is mobile computing. What used to be wireless telephony became the hybrid platform of algorithmic computation, integrated with a variety of non-algorithmic processes, supported by a vast array of sensors. 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 situation (e.g., a sales chart for a marketing meeting, a simulation for a class in Big Data visualization). Other "feats" make for spectacular headlines: the algorithm for playing video games that plays better than the living player for whom the games were conceived; and the algorithm for understanding language. These transcend Big Blue, a high-performance machine programmed to play chess, and which eventually became a successful contestant on the game show *Jeopardy*, and then a digital doctor. For this purpose, huge resources were made available to be thrown at the problem of beating a world champion (through brute force computation). The more recent claims are for game competence across the gamut of games, regardless of experience; and for understanding questions posed in everyday language, that is, the ability to answer them. High-dimensional sensory inputs (representations of the environment) drive deep network agents able to outperform humans. (In language understanding, deep reinforcement learning is the chosen path.) The video game playing "intelligent" machine knows nothing about the Atari games (of the early "romantic" age of video games). It was tested on 49 of them (Minh et al. [6]). It makes inferences from pixels from the game images and from game scores, that is, how others played. Of course, the game's algorithmic nature itself is congruent with that of the artificial agent driven by high-dimensional sensory inputs.

In reference to understanding language (Weston et al. [7]), proxy tasks are set out in order to facilitate the evaluation of reading comprehension via the mechanism of answering questions. By no coincidence, a particular kind of games (text adventure connected to interactive fiction, Monfort [8]) provides a medium for categorizing various types of questions. Far from being only examples of successful programming and clever methods, these define a new frontier in computation. Implicit in the challenge is the question of whether human performance, anticipatory in nature, can be matched, or even outdone, by algorithmic forms of

computation. Of course, the goal has to be defined as clearly as possible. To understand the significance of all this breakthrough research, we shall first define the underlying concepts involved.

1.2 What Is and What Is not Anticipation—A Question that Does not Go Away

Let us return to the issue of the human being's progressive dependence on computers. Being part of a reality within which everything associated with human existence is, in one way or another, dependent on computers undermines the effort of a neutral evaluation. For example, this present text originates in a word-processing program. In the writing, the author used speech recognition and image processing, and benefited from machine learning-based search for references. The text will, along the academic path of publication (editing, peer review, additional feedback, layout, etc.) be made available—on paper, using digital printing, and in e-formats— to a readership shaped by the experience of computation to the extent that dependencies are established. References will be cross-linked, keywords highlighted; the text will become an easy-to-explore hypertext, all set to be further indexed and eventually fed into a complex network visualization. If the means of expression (language, formulae, images, etc.), communication (sharing), and signification (evaluation of originality, impact, usefulness over time, etc.) were passive, it would not make any difference that this text is not the outcome of orality, or of handwriting on parchment or paper, or of lead-based typography, or of the Gutenberg printing press. But the media involved are *never* neutral. Tools are not passive partakers in the activity. Being used within a culture, they "make" a new content, a new user, a new public—and thus contribute to the change of culture itself.

   This is all the more important as we realize the ubiquity and diversity of computation. We understand that a new human condition is ascertained in the ever-expanding use of digital technology. Humankind might project itself into a way more exciting future than ever. Alternatively, it might wipe itself out (or at least place itself on a degenerative path), and thus eliminate humans from the dynamics of evolution, or at least diminish their influence. Computers that perform better in chess or programs that outperform the human in Atari games (or any other machine-based game), and computers capable of understanding and answering questions, are only indicative of the breadth of the process. What counts in the perspective of time is the depth of the process: how the mind and body change, how human pragmatics is redefined.

1.2.1 Winning, or Changing the Game

Having taken this broader view in order to establish a context, it is time to focus on the terms that frame the question we are trying to answer: anticipation and

computation. Mitchell (or for that matter Wolfram [9] or Zuse [10]) claims that somehow the universe is being deterministically computed on some sort of giant but discrete computer (literally). If indeed all there is is an outcome of deterministic computation, then what is the rationale behind the fact that, in the world as we know it, some entities are alive and some not? Computation, itself grounded in rationality, ought to have an explanation for this, if indeed we are only its outcome (whether as stones, micro-organisms, or individuals who conceived computation). Living entities (from bacteria to the human being) come into existence at some moment in time, unfold in a dynamic driven by survival—including reproduction— and eventually die. As they do so, they join the non-living (water, chemical elements, dissipated energy, etc.), characterized by a dynamic driven by the forces at work on Earth and in the cosmic space Earth occupies (according to descriptions in astrophysics). Of course, sun and wind, humidity, a variety of particles and radiations, as well as interaction (local or galactic), affect stones and rivers, the air, and decomposition of dead organic matter as much as they affect the living. Experimental evidence shows that the dynamics of the living is, moreover, characterized by adaptivity. The non-living does not exhibit adaptivity—at least not at the timescale of those who observe them. What most radically distinguishes life from not-life is the sense of future, i.e., the vector of change. It is goal-driven change that explains why there is life, more so than the rather unsubstantiated, and therefore dubious, affirmation of a universal computation of almost deistic nature.

Between the living and the physical—which is subject to descriptions constituting a body of knowledge known as physics—there is a definite systemic distinction: the living is complex; the physical is complicated. To reproduce here the arguments upon which this epistemological model is based would probably invite attention to a subject different from that pursued herein. Suffice it to say, the criterion used is derived from Gödel's [11] notion of the undecidable: entities of complex nature, or processes characterized as complex, cannot be fully and consistently described. The living is undecidable. This means that interactions implicit in the dynamics of the living cannot be fully and consistently described. As a consequence, no part of a living entity is less complex than the whole to which it belongs—unless it is reduced to a merely physical entity. The famous decerabrated frog experiment (described in almost all physiology books) is illustrative of this thought. The physical is decidable. A fragment of a stone is as much representative for the whole stone as the laws of physics are for the universe. Of course, Gödel referred to descriptions of reality (a nominalist view), to statements about it, to the logic guiding such statements, and, further, to operations upon them. We take the view (resulting in the definition of G-complexity, Nadin [12]), that the decidable/undecidable, as a gnoseological construct, defines states of complementary nature (physics vs. living). Anticipation is associated with the undecidable nature of the living. In the decidable realm, action-reaction entails change.

The most recent attempt to explain the emergence of life (England [13]) returns to the obsessive model of physics inspired by the laws of thermodynamics. If indeed capturing energy from the environment and dissipating that energy as heat were conducive to the restructuring of matter (carbon atoms, in particular), leading

in turn to increased dissipation, the process would not have ended, and more such restructuring would take place. It does not take place on Terra, (i.e., our Earth) and nobody has yet documented it on other planets or cosmic bodies. That physics is fundamentally inadequate for explaining the emergence of life and, further, for explaining it, is a realization that physicists swore to ignore. Anticipation does not contradict the predicaments of physics, but complements them with an understanding of causality that integrates the future. The increased entropy (cf. the Second Law of Thermodynamics) of physical systems explains, to a certain extent, how they change over time. Physical information degrades. Reaction is the only remedy. In the living, we are faced with the evidence of long-term stability of species. Biological information (DNA is an example) is maintained as a condition of life. Entropy does not increase. Anticipation is the expression of the never-ending search for equilibrium in the living, and therefore its definitory characteristic.

It is easier to postulate (such as in the above text) and navigate a clean conceptual universe in which words mean exactly what we want them to mean (cf. Humpty Dumpty in Carroll [14]). To deal with the messy reality of complementarity—the living and the non-living—in which concepts are ill defined, implies awareness beyond the views that shaped civilization after Descartes. In respect to anticipation, such clarity—not only of semantic nature—is essential since the terminology permeates the pragmatic level, in particular in the language domain associated with computation.

When machine performance (of the computer or of any other device) is juxtaposed to that of the human—machines outperforming world champions in chess or game fanatics in 49 Atari games—one has to define the criteria. The same applies for understanding language. On account of anticipation, humans answer questions even before they are posed. Indeed, the knowledge used in such performance is as important as understanding the difference between repetitive tasks and creativity. Algorithms that integrate better prediction models in data-processing characteristic of playing games (or any other form of algorithmic expression) or understanding questions, together with high-speed processing, will outperform the human being to the extent that the "self-driving" car will outperform the human-driven car. But the real problem is "Will they generate new games, or new instances of competitive dynamics? Will they generate, as human do, new language, within which new ideas are seeded?" This is where anticipation comes into the picture. Winning and changing the game are two sides of a coin about to be flipped. Making way for new language is part of the continuous remaking of the individual.

1.2.2 Reaction and Anticipation

Anticipation pertains to change, i.e., to a sense of the future. The image (Fig. 1) is suggestive.

Fig. 1 Ways of considering
the future
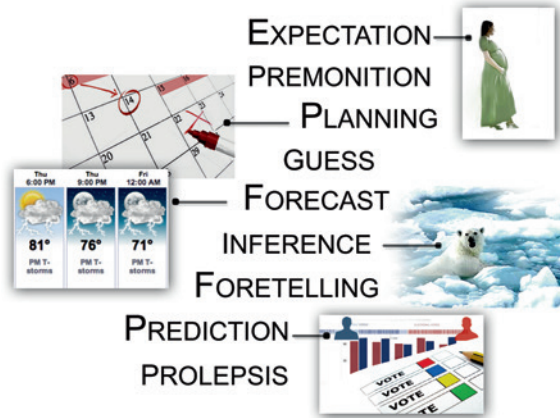
To comment on the particular words would only result in anecdotal evidence. First, let us clarify some of the terms. Foremost in importance is the understanding that the physical is defined through interactions driven exclusively by reaction. The physics of action-reaction, as formulated in Newton's Third Law [15] provides a decidable model:

Lex III: Actioni contrariam semper et æqualem esse reactionem: sive corporum duorum actiones in se mutuo semper esse æquales et in partes contrarias dirigi.

(Translated to English, this reads: Law III: To every action there is always opposed an equal reaction: or the mutual actions of two bodies upon each other are always equal, and directed to contrary parts.)

If body *A* exerts a force **F** on body *B*, then body *B* exerts an equal and opposite force −**F** back on body *A*.

$$\mathbf{F}_{AB} = -\mathbf{F}_{BA} \tag{1}$$

The subscript *AB* indicates that *A* exerts a force on *B*, and *BA* indicates that *B* exerts a force on *A*. The minus sign indicates that the forces are in opposite directions. Often $\mathbf{F}_{AB}$ and $\mathbf{F}_{BA}$ are referred to as the *action force* and the *reaction force*; however, the choice of which is which is completely arbitrary. Elementary particles, as much as physical entities (at the scale of our immediate reality or at cosmic scale), behave as though they follow the law. Within the living, this is no longer the case. Cells, in their infinite diversity, have a dynamics for which the description in Newton's laws are no longer applicable.

Descriptions of entities and processes restricted to the dynamics originating in action-reaction can be fully and non-contradictorily described. As stated above, this applies across the reality of physics—from the micro-universe to cosmic space.
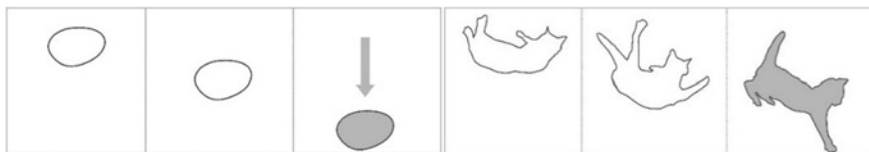
Fig. 2 Not all falls are the same, but all are subject to gravity

1.2.3 The Undecidable "Falling of the Cat"

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. Awareness of change, pertaining to the living, is reactive. The living reacts (to changes in temperature, to stimuli, to other people, etc.), but at the same time, it is also anticipatory (preparedness, as well as foresight, for instance). Adaptivity over shorter or longer intervals is the specific expression of this interplay. It also explains the long-term stability of the living. From a physics perspective, 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. It is expressed in action; and it is meant to preserve life (the cat usually avoids getting hurt).

The equation of the "change"—coordinates (falling from height h) in this case—is straightforward:

$$h = 1/2 g t^2 \tag{2}$$

in which **h** is the falling height, **g** the acceleration due to gravity, and **t** the falling time. If, for example, the height is given by h = 10 m and g = 9.81 m/s^2, the predicted falling time is obtained by inserting these values in (2) and solving for **t**. Introducing the variable **T** for the falling time and the function A_t for the prediction procedure yields the following predicted event, which consists of one element only:

$$A_t(\{10, 9.81\}) = \{1.4278431220270645\} \tag{3}$$

This description omits some variables related to air resistance (object's shape, air density, effect of temperature and humidity, etc.).

The cat falls "actively." The cat's response to falling (even if the fall is acci-dental, i.e., not caused within an experiment) is at least a change in geometry: actively turning (by triggering the motoric) increases the surface, and thus air resistance. The equation pertinent to the fall of the stone still applies, but in a rather approximate way (more approximate than considering friction). The living "fights" gravity. (The metaphor is a mere translation of the fact that nobody likes to fall.) The past (cat's position at the start of the fall), but also the possible future (how and where to land) affect the outcome.

1.2.4   The Living Can Observe the Physical

The fall of the same stone, repeated, from the same position, is captured in the physical law description: air resistance can be precisely accounted for and, even under experimental conditions, maintained. The gravitational field strength (9.8 N upon every 1 kilogram) is a characteristic of the location within the Earth's field of gravity, not a property of the falling stone or cat. The nomothetic (Windelband [16]) corresponds to a description of a phenomenon (or phenomena) characterized as law. The fact that mathematicians extend the nomothetic description to the falling cat is testimony to their incomplete knowledge. It does not include anticipatory dynamics. Indeed, the cat, as opposed to the stone, will not fall the same way twice. Mathematicians, like many others, scientists or not, observed the fact mentioned above.

But since the cat's falling became a mathematical problem, let's take a closer look at what is described. The purpose is simple: that we understand why the "recipe" for calculating the parameters of physical phenomena cannot be extended to predictions of living processes. Prestigious scientists (such as George Gabriel Stokes [1819–1903], James Clark Maxwell [1831–1879], and Etienne Jules Marey [1830–1904]) were tempted to explain the falling of cats, more precisely, how they turn in the air.

For them and their followers, the falling cat problem consists of explaining the physics underlying the common observation of the "cat-righting reflex." To dis-cover how a free-falling cat can turn itself right side up as it falls—no matter which way up it was initially, without violating the law of conservation of angular momentum—is a challenge. There is one limitation: all that counts is that the cat fall on its legs. As a leading mathematician in the falling cat problem puts it:

> Although somewhat amusing, and trivial to pose, the solution of the problem is not as straightforward as its statement would suggest, leading towards surprisingly deep mathe-matical topics, including control theory for nonholonomic systems, optimal motion plan-ning, Lagrangian reductions, differential geometry, and the gauge theory of Yang-Mills fields [17].

Within this study, we will not go into the details of all of the above. In broad strokes: applied differential geometry allows for the approximate description of an object flipping itself right side up, even though its angular momentum is zero.

In order to accomplish that, it changes shape (no stone changes shape in the air). In terms of gauge theory, the shape-space of a principal SO(3)-bundle, and the statement, "Angular momentum equals zero," defines a connection on this bundle. The particular movement of paws and tail conserves the zero angular momentum. The final upright state has the same value. This is the "geometric phase effect," or *monodrony*.

The idea is simple: Let a cat fall; and derive the pertinent knowledge from the experiment. (In 1882, Marey used a chronophotographic gun for this purpose; in our days, motion capture equipment is used.) But this is no longer a reproducible event. It is not the passive fall of a stone—reproducible, of course—but the active fall embodying anticipation. The outcome varies a great deal, not the least from one hour to another, or if the landing topology changes. The stone will never get tired, annoyed, or excited by the exercise. And it will never learn. We shall explain this, in reference to the cat's fall, using images (Figs. 3 and 4).

The cat's shape is given by two angles: (Θ) (ψ).

ψ is the angle between the two halves of the cat's body.

Θ describes the direction of the cat's legs ($\Theta = 0$ when the front and back legs are closest to each other).

A change in Θ corresponds to a rotation of the cat's body around the \spinal axis. Heisenberg's uncertainty relation [20] suggests that, although such descriptions are particularly accurate, we are, in observing the falling of a cat, not isolated viewers, but co-producers of the event. To observe entails influencing the result. The falling of human beings, of consequence as we advance in age, makes it clear that "to know how to fall" (as the cat obviously does) is more than a problem in physics or a mathematical exercise. (No kitten should be subject to such an experiment.)

Just as an aside: inspired by the cat's fall, Apple, Inc. patented a method for controlling the accidental fall of the iPhone on its precious screen. The iPhone's vibration motor (Fig. 5) is programmed to change the angle of the fall in mid-air. This change is based on data from the device's positioning sensors. The patent is, in its own way, an example of engineering inspired by the expression of anticipation in the living. It is based on knowledge from physics (coordinates and center of gravity) and takes predictions based on Newton's laws in order to activate the

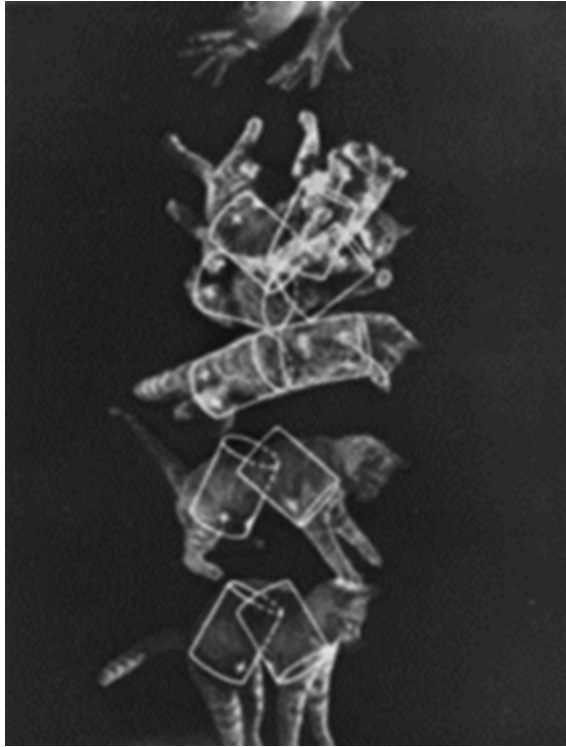**Fig. 3** Representation of a falling cat (Drawing by. E. Kuehne, in Mehta [18, p. 5])
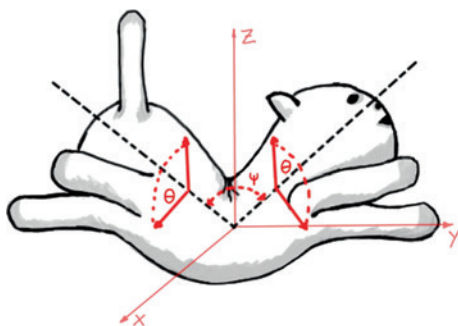
Fig. 4 This image is representative of the Kane-Scher solution [19] to why cats fall on their feet. (Reproduced from [18])
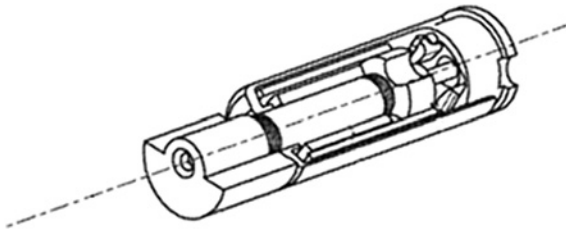

Fig. 5 The vibration motor. Statistical analysis of the fall, by comparing gathered data against other information stored in device memory, serves as trigger to activate the spin and change the phone's center of gravity (cf. patent application)

vibration motor so that the device is turned in the air—pretty much like a cat made out of stone or wood.

With this device, we are in the reaction domain, taking advantage of a good understanding of physical laws.

The unity reaction-anticipation—characteristic of the living—corresponds to a different condition of matter and its change over time. The measurement process, i.e., the observation of the change (the falling cat), influences the outcome. Our watching how the smartphone falls does not affect the process.

Thesis 1: The living can observe the physical.

Actually, as it evolves, it continuously does—because the process is affected by the context. The physical does not have an observation capability. It is rather a stage on which the living performs (while it also reshapes the stage). Perception is nothing more than the process through which awareness of here and now is established.

Thesis 2: Awareness of immediate space and time is the outcome of perception processes.

1.3 Expectation, Guessing, Prediction

We become aware of anticipation when it is successful: falling the "right way," avoiding danger, rewarding creative activities, competence in competition, understanding language, images, and textures, for example. From such activities, we can generalize to processes in which anticipation is sometimes involved, and also to other forms of dealing with the future. From guessing and expectation to prediction and planning as human endeavors, we infer that reaction and anticipation are not reciprocally exclusive, but rather intertwined. Acknowledged in language (and in experiments) are various forms of what is called premonition (of danger, usually), foretelling (mostly associated with the dubious commerce of "seeing into the future"), not to mention curses and blessings, and voodoo. There is no reason to go into these; although for those fixed on the notion of algorithm—description of actions through which a goal is attained—they can be given as examples of processed information to which misinterpretation also belongs. Each of the above-mentioned aspects (including the slippery practices) is a combination of reaction and anticipation. Just as an innocent illustration: The fortuneteller reacts to someone's need for reassurance or comfort, creating the illusion of a successful (or unsuccessful) anticipation, "You will not be awarded the Nobel Prize, but your love life will improve."

1.3.1 Guessing

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

$$(P_{os}(p_i)), p_i \geq N \quad (C(c_i)), c_i \geq N \qquad (4)$$

If you have to guess a number from one to one hundred, the first thing is to reduce the space of choices.

The reaction component ("I know that the person asking the question likes the queen of hearts!") is based on real or construed prior knowledge. For an anticipatory dimension, one would have to combine a wager: "Guess what number I chose" (or what card, or what word) "and you win!" (See Fig. 6). A generalization can be made: when reaction and anticipation are suggested, the outcome will show that some people are better at guessing than others because of heightened perception of cues of all kind.

Reactions are based on the evaluation of the information pertinent to the situation—different when one visits a fortuneteller or a casino from when one guesses the correct answer in a multiple-choice test. In the multiple-choice situation, one infers from the known to the unknown. When patterns emerge, there is learning in guessing: the next attempt integrates the observation of related or unrelated information. This associative action is the cognitive ingredient most connected to guessing and, at the same time, to learning. (We retain patterns and recall them when faced with choices.) The rest is often statistics at work, combined with *ad hoc* associative schemes pertinent to what is possible. (That's where premonition, mentioned above, comes into play.) Let us acknowledge that guessing (as well as learning) is reduced to nil in predictable situations. The anticipation component of guessing is related to the state of the self. From all possible games in the casino, some are more "favorable" at a certain time, something like: "Guess who's knocking at the door," after grandma's voice was heard. Only surprise justifies the effort, even when the result is negative. Recent research of responses of the human frontal cortex to surprising events (Fletcher et al. [21]) points to the relation to learning mentioned above. The dorsolateral prefrontal cortex contributes to the adjustment of inferential learning. Associative relationships (Fig. 7) that lead to learning (also qualified as associative) are based on the action of discriminating the

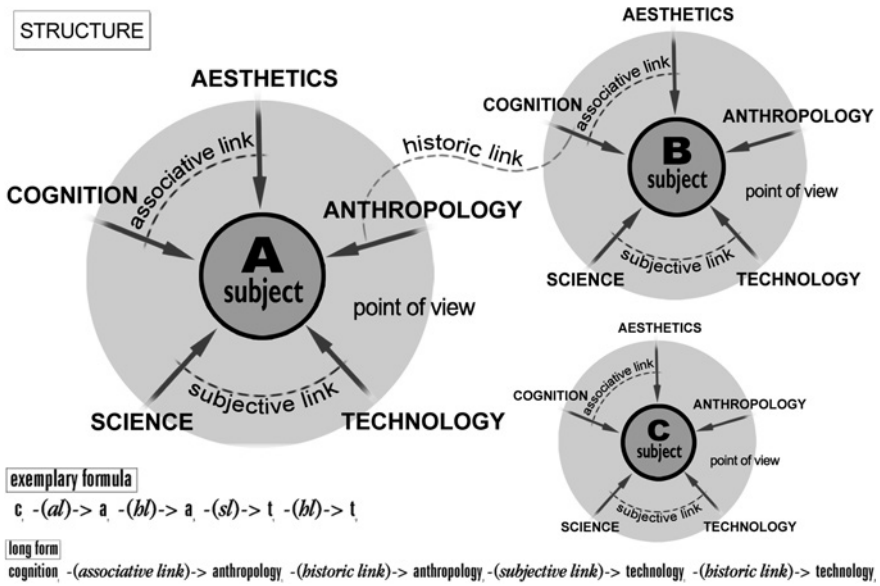

Fig. 6 Two examples of guessing

Fig. 7 Examples of associative relationships (Associative Encyclopedia, Nadin [22])

degree (strength) of interrelation. Of course, fuzzy sets are the appropriate mathematical perspective for describing such interrelations.

Empirical data (statistics, actually) document "better days," i.e., above-average guessing performance. This corresponds to a variety of circumstances: additional information about the process (acquired consciously or, most of the time, through processes "under the radar"), a state of cognitive or sensorial alertness (for whatever reason), or simply a statistical distribution ("lucky"), to name only a few. There is no magic in the exceptional (a "good" day, "bad luck"), but there is quite a bit to consider in terms of the large number of variables involved in the outcome of human actions. The manner in which anticipatory action is intertwined with the reactive is difficult to describe exactly because of the multiplicity of factors involved. If anything, anticipation actually undermines success in guessing, given its non-deterministic nature; it integrates the subjective, the emotional, the spontaneous. A guessing machine—computer or any other type of machine—can automate the guessing knowledge specific to well-defined selection and thus outperform the guessing living (not only human beings are involved in guessing as they face change). Machine learning provides a good basis for such applications. The algorithm for successfully playing computer games is based on data acquired through what is called *deep reinforcement learning*. The algorithm for understanding questions formulated in natural language is based on a *multilinear map* subjected to processing in Memory Networks (MN).

1.3.2 Expectation

In comparison to guessing, expectation does not entail choosing ("Heads or tails?"), but rather an evaluation of the outcome of some open-ended process. An example: A child's expression is informative of what might happen when the child will "hang out" with friends. The parents' evaluation might be difficult, if at all possible, to describe (e.g., "I know what you guys plan to do"), that is, to formalize. In the act of forming an expectation (such as in carrying out experiments), the focus on the reaction component changes from the probable (which number from the set defined?) to the inferred. Several sources of information pertinent to forming an expectation are weighed against each other. What appears most probable out of all that is possible gets the highest evaluation, especially if its outcome is desirable (for instance, pleasant weather preferred over the expectation of rain). Expectations associated with experiments are usually in the area of confirming a hypothesis or someone else's results. If the outcome is judged to be negative, then avoiding it is the basis for action. Again, anticipation—reflected in what is perceived as possible —meets reaction, and information is associated with probable cause. Weather is often expected—inferred from opinion, observation, data—not guessed. So are the outcomes of activities that weather might influence. Agriculture practiced prior to the integration of digital information in agricultural production was often in the realm of the expected. A cornfield is not equally fertile in every spot. Learning how to increase production by extracting data (through GPS-based measurements) pertinent to fertility and applying fertilizers or planting more seeds in certain spots grounds expectation in knowledge—and thus makes it look like an algorithm (a set of rules which, if respected, can yield a result). Based on the evaluation of the outcome, new expectations are generated. Events with a certain regularity prompt patterns of expectation: a wife awaits her husband, a child awaits a parent, a dog awaits its owner, who usually returns from work at a certain time. Such regular events are encountered on many occasions and in many activities.

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

$$P(p_1, p_2, \ldots p_n) \tag{5}$$

An expectation machine is actually a learning procedure that attaches weights (some subjective) to choices from the limited set of possibilities. The reactive component dominates the anticipatory. False expectations (of personal or group significance) are the outcome of skewed evaluations. Expectation and superstition are examples of such evaluations. They are driven more by desire or wishful thinking that tends to falsify the premise (adding self-generated data to the factual incomplete knowledge).

Among the cognitive illusions (Kahneman and Tversky [23] existing in culture are those formed by gamblers (Delfabbro [24], as well as by a professional acting in a state of over confidence. Physicians making inferences based on limited medical tests (Gigerenzer and Gray [25], Sedlmeier and Gigerenzer [26]); coaches captive to the "hot-hand" model (Tversky and Gilovich [27], Miller and Sanjurjo [28]; economists absorbed in data patterns more relevant to the past than applicable to future developments (Hertwig and Ortmann [29]) can be given as examples. These have in common the perception of random and non-random events. Statistically significant deviations from the expected (e.g., the average scoring performance of a gambler in a casino, of a basketball player, of the stock market, etc.) lead to beliefs that translate into actions (a gambler can be refused entry, the basketball player believed to have a "hot-hand" day faces a stronger defense, hot stock market days mean more trades, i.e., more speculation, etc.). Physicians interpret deviations in respect to expected values (blood glucose, cholesterol, vitamin D, creatinine), and automatic procedures (comparison with average values) trigger warnings. What we get after a blood test, for example, is an expectation map. Guessing and expectation, each in its own way, are meant to inform choices or result in decisions. Positive and negative factors weigh in with every option. The integration of biases in making the choice leads to the surprising observation that what some call instinct (choose among options in the absence of identifiable previous knowledge, in common parlance, "gut feeling") can explain successful guesses or actions driven by expectation [30]—such as which direction to take at a fork in the road.

1.3.3 Prediction

Connecting cause and effect, i.e., associating data generalized from statistical observations describing their connection, is the easiest way to characterize 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. Uncertainty is the shadow projected by each prediction (Bernoulli [31]). Therefore, it is representative of the limits of understanding whatever is predicted. In some cases, the prediction is fed back into what we want to predict: how a certain political decision will affect society; how an economic mechanism will affect the market; how technological innovation (let's say multimedia) 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. The impossibility of disconnecting the observer (the subject, in learning) from the observed (the object of learning) is an inherent condition of learning, whether human or machine learning. The constructivist perspective demonstrated the point quite convincingly (von Glasersfeld [32]).

Prediction → inference based on probability

$$\mathcal{P}\,(\text{frequency, ignorance, belief}) \tag{6}$$

$$F : \mathcal{D} \to X(\mathcal{D})$$
from an initial state $(\mathcal{D})$ to state (x)
$$\tag{7}$$

Prediction machines of all kind are deployed in situations in which the outcome is associated with reward/punishment (loss). In particular, Bayes-inspired prediction is driven by a hypothesis: You "know" the answer, or at least part of it (your best guess). Predictions of election results, of weather patterns, of sports competitions are based on such assumptions. Prediction as a process that describes the outcome of action-reaction dynamics can be usefully affected by experiential evaluations.

1.3.4   Future States and the Probability Space

But there are also predictions driven, to an extent larger than the Bayesian state of belief, by anticipatory processes, involving the probability space also. Falling in love at first sight—which is neither guessing nor expectation—is a prediction difficult to make explicit. (It combines rationality and consistency with a subjective perspective such as the above-mentioned "gut feeling.") There is no explicit cause-and-effect connection to uncover, and no frequencies to account for. The future state (the romantic ideal of a great love, or the calculated outcome of an arranged marriage) affects current states as these succeed each other in a sequence of a time often described as "out of this world." We could add the dopamine release during anticipation of a musical experience (Salimpoor et al. [33]). Peak emotional responses to music are different from the experience of winning a computer game, or answering a question. Therefore, it would be inadequate to even consider an algorithm for returning the value of musical experience based on statistical data. Machine-based performance (such as winning games, or understanding questions) corresponds to different domains of computation.

Facial expression as a predictor is yet another example of Bayesian probability-based inferences. In very sophisticated studies (Ekman and Rosenberg [34], Ekman [35]), it was shown that the "language" of facial expression speaks of facts to happen before they are even initiated—which is anticipation in pure form. The Facial Action Coding System (FACS), which is a taxonomy of facial expression (and the associated emotional semantics), inspired Rana El Kalioubi in her work on computationally interpreting the language of faces. For those who "read" the face's expression, i.e., for those who learned the language of facial expression, the emotion predictions based on their own anticipation guides their action. Gladwell [36] describes the case of a Los Angeles policeman who reads on the face of the criminal

holding a gun on him that he will not shoot, leading the officer to avoid shooting the criminal. The expectation—criminal pulls out gun and points it at the policeman pursuing him—and the prediction—this person with the particular facial expression, as studied by the interpreter, will not shoot—collide. So do the probability of being shot and the prediction informed by knowledge otherwise not accounted for.

Descriptions of the relation between expectation and prediction are informative in respect to the mechanisms on which both are based. The various levels at which learning—different in expectation-driven from prediction-based decisions—takes place are not independent of each other. Expectations pertain to more patterned situations: e.g., "I expect to be paid for my work based on our agreement." (The intitial state d, to be hired; the future state x, to be paid for work, depends on initial state d.) The prediction, "Based on the record of your employer, you will be paid," conjures different data. An acceptable description is that the learner extracts regularities or uses innate knowledge. They are often an expression of what in ordinary language is described as stereotype or, in some cases, wishful thinking. However, when the individuals become involved in the activity of predicting (literally, "to say beforehand," i.e., before something happens), they expect the prediction to actually take place. It is no longer a wish, but rather the human desire, expressed in some action, to succeed.

Many activities, from policing the streets to conceiving political reform, urban development, military strategy, educational plans (to name a few areas of practical activity with features with little or nothing in common) are informed by the very competitive "industry" of predictions. Generalizing from the past can take many forms. 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.).

Change means evolution, variability over time. Predictive efforts are focused on understanding sequences: how one step in time is followed by another. However, these efforts focus on what, ultimately, anticipatory processes are: a modeling of the entity for which they are an agency, and the execution of the model in faster than real time speed. The limited deterministic perspective, mechanic in nature, repetitive—i.e., what cause leads to which ensuing effect—affects the understanding of anticipation through a description of predictive mechanisms. Predictions made following known methods (such as time series analysis and linear predictors theory) capture the reaction component of human action (Arsham [37]). The anticipatory component is left out most of the time, as a matter of definition and convenience. Complexity is difficult to recognize, and even more difficult to handle because it corresponds to open-ended systems. Once a predictive hypothesis—let's say every minute the clock mechanism engages the minute hand—is adopted, it defines the

cognitive frame of reference. On a digital display, the predictive hypothesis will be different. Should the predicted behavior of the mechanism somehow not take place, expectation is tested. However, mechanisms, as embodiments of determinism, rarely fail. And when they do, it is always for reasons independent of the mechanism's structure.

1.3.5 Learning and Expectation

Predictions concerning the living are less obliging since interactions are practically infinite. Structure matters, interdependencies are fundamental. It happens at all levels of the living that predictions—what will happen next (immediate or less immediate future)—are either partially correct or not at all. In studying learning and selective attention, Dayan et al. [38] 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; 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, if you prefer) takes over the more complex model that accounts for both reaction and anticipation in the dynamics of the living.

Predictors reflect the desire to understand how change takes place. They express the practical need to deal with change. However, they omit change from the equation of those predicting or subject to prediction. Actions from thoughts, as Nicolelis [39] calls them, account for the self-awareness of change. What is learned supports inferences (statistical or possibilistic); uncertainty results as the competitive resources engaged in the inference are overwritten by unrelated factors. Predictions also capture the interconnectedness of all elements involved in the dynamics of the observed. Learning involves predictions. In this sense, they open access to ways to emulate (or imitate) change.

Expectations have no direct learning component. One cannot learn explicitly how to expect, even accepting that there might be structure in the learning process after an expectation is validated, and in the representation associated with the expectation. Expectations only occasionally produce knowledge: a series of expectations with a certain pattern of success, or failure for that matter. Predictions, even when only marginally successful, support activities such as forecasting—for short or less than short sequences of change—of modeling, and of inference to the characteristics of the observed dynamic entities.

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, the economy, politics, military actions, urban policy, or education, etc., predictions or anticipations emerge on account of considerations regarding cascading amounts of data. Just to genralize, we can consider the ever-increasing amount of sensors deployed as the source of this data. Integrated sensors generate high-level, multi-dimensional 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 agent, which has as output "human-level control" performance (in playing games, but applicable as well to other choice-making situations), is the embodiment of prediction based on reinforcement learning [6].

1.3.6   Interconnectedness

Without the intention of deriving full-fledged conclusions, an example could suggest the interrelated nature of expectation, guessing, and prediction. The painful revelation of the practice of torture associated with the "war on terror" (a very misleading formula) prompted discussions that ranged from the moral, aesthetic, medical, to political, and ultimately focused on how successful torture is in extracting useful information. Data of all kind, from anecdote to statistics (perversely kept by those regimes that for centuries have practiced torture, some methodically, some as circumstances deemed necessary) document both the efficiency of brutal treatment of prisoners and the possibility of collecting misinformation. The process is non-deterministic. Moreover, principles of conduct—some by tacit agreement (what is hateful to you, don't do to others), others codified by the community of nations—associated with extreme treatment of the adversary, set moral borderlines (some less clear than they should be). Still, contrary to this foundation on data and rules, the practice continues. (Those on whose behalf torture is employed tend to find justification for it, since they form the notion that it has served them well.)

Prediction-data show that torture occasionally begets information. A torture information production machine—i.e., a computer, or better yet, a robot, with the applicable moral constraints built in—would decide on a cost-benefit analysis model whether torture should be applied or not. In retrospect, it is evident that guessing would be a weak description of the future: it has the highest margin of error. Expectation would not be much better: a machine does not output expectations since their variability escapes algorithmic descriptions. Predictions, especially in the Bayesian sense, are more effective. According to the Report of the Senate Intelligence Committee on the CIA Counter-Terrorism Program, some cases of torture were doomed from the outset.

Of course, the above description pertains to the macro-level. The interrogator and the interrogated are actually in an anticipatory situation: winning or losing drives their actions. Guessing, expectation, and prediction meld as they do in hide-and-go-seek, in playing tennis, in poker. In view of this observation, the understanding of prediction (implicit in guessing and expectation) takes on new meaning.

Predictions regarding the living, although inappropriate for systematically capturing their anticipatory dimension, are a good indicator of what is lacking when anticipation is ignored. An example: in focusing only on human beings, predictions based on physiological data remain at a primitive stage at best, despite the spectacular progress in technology and in the scientific theory of prediction. Streams of data (from a multitude of sensors) in association with some analytical tool (data-mining, usually) could, of course, help identify where and how the physical component of life is affected by change (aging, environment, medical care, hygiene, alimentation, driving, etc.). The reactive component of what is called "health" is of extreme importance. Clogged arteries, degradation of hearing or of the eyes can be identified with the help of real-time monitoring of blood pressure, hearing, or the macula. But they remain partial indicators. In evaluating change in the condition of the living, of the human being, in particular, what counts are not only the parts under observation, but their interconnectedness, especially of the whole.

Reaction is reductive. Anticipation is a holistic expression. Albeit, if we could improve such predictions by accounting for the role of anticipation—the possible future state influencing, if not determining, the current state—we would be in a better position to deal with life-threatening occurrences (strokes, sudden cardiac death, diabetic shock, epileptic seizure, etc. (Nicolelis and Lebedev [40]). Learning (i.e., 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—the well-known climate change issue, for example—education, market functioning. It is easier, when addressing a given concern, to deal with "recipes" (e.g., reduction of CO_2 emissions as a solution to climate change, with its reductionist focus, to the detriment or exclusion of other variables, either ignored or opportunistically downplayed), than to articulate an anticipatory perspective, holistic by definition.

Unless and until anticipation is acknowledged and appropriate forms of accounting for it are established, the situation will not change drastically. Neither will medical care, environmental policies, political matters, or education change, no matter how consequential their change (if appropriate) could be. Physical processes have well-defined outcomes; living processes have multiple outcomes (some reciprocally antagonistic.) This aspect becomes even clearer when we look at the very important experiences of forecasting and planning. Policies, i.e., social awareness and political action, depend on forecasts and involve responsible planning, liberated from the influence of opportunistic interests.

1.3.7 Forecasting and Planning

Predictions, explicit or implicit, are a prerequisite of forecasting. The etymology points to a pragmatics, one that involves randomness—as in casting. Under certain circumstances, predictions can refer to the past (more precisely, to their validation

after the fact). Take a sequence in time—let's say the San Francisco earthquake of 1906—and try to describe the event (after the fact). In order to do so, the data, as registered by many devices (some local, some remote) and the theory are subjected to interpretations. The so-called Heat-Flow Paradox is a good example. If tectonic plates grind against one another, there should be friction and consequently heat. This is the result of learning from physical phenomena involving friction. Along the well-known San Andreas Fault, geologists (and others) have measured (and keep measuring) every conceivable phenomenon. No heat has been detected. The generalization from knowledge regarding friction alone proved doubtful. Accordingly, in order to maintain the heat dissipation hypothesis as a basis for forecasting, scientists started to consider the composition of the fault. This new learning—extraction of regularities other than those pertaining to friction and heat dissipation—was focused on an aspect of friction initially ignored. A strong fault and a weak fault behave differently under stress, and therefore release different quantities of heat. This is a case in which data is fitted to a hypothesis—heat release resulting from friction. To adapt what was learned to a different context is frequently used in forecasting.

In other cases, as researchers eventually learned, what was measured as "noise" was treated as data. Learning noise patterns is a subject rarely approached. 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. The lawsuit industry has exploited the situation to the extent that medicine is becoming defensive at a prohibitive cost (or overly aggressive, through the variety of surgical interventions, for instance, at an even higher price).

In general, theories are advanced and tested against the description given in the form of data. Regardless, predictions pertinent to previous change (i.e., descriptions of the change) are not unlike 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. (Successive diet plans exemplify how data were frequently fitted to accommodate the pharmaceutical industry's agenda, sometimes to the detriment of patient health.)

To ascertain that something will happen in advance of the actual occurrence—prediction (the weather will change, it will rain)—and to cast in advance—forecast—(tomorrow it will rain) might at first glance seem more similar than they are. A computer program for predicting weather could process historic data: weather patterns over a long time period. It could associate them with the most recent sequence. And in the end, it could come up with an acceptable global prediction for a season, year, or decade. In contrast, a forecasting model would be local and specific. The prediction based on "measuring" the "physical state" of a person (how the "pump," i.e., heart, and "pipes," i.e., blood vessels, are doing, the state of tissue and bone) can be well expressed in such terms as "clean bill of health" or "worrisome heart symptoms." But it can almost never become a forecast: "You will have

a heart attack 351 days from now;" or "In one year and seven hours, you will fall and break your jaw." Or even: "This will be a historic storm" (the prediction, so much off target, of the "Nor'easter" of January 2015).

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

---

$$\mathcal{F}(\text{predictions, self-generated data}) \tag{8}$$

Forecasts are not reducible to the algorithmic machine structure. They involve data we can harvest outside our own system (the sensorial, in the broadest sense). The major difference is that they involve also data that human beings 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.

To summarize: forecasting implies an estimation of what, from among few possibilities, might happen. The process of estimation can be based on "common knowledge" ("Winds from the west never bring rain"); 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). Evidently, forecasting is domain specific. Meteorology practices forecasting as a public service; commerce needs it for adapting to customer variability of demand. Urban planners rely on forecasting in order to optimize municipal dynamics (housing, utilities, traffic, etc.). The latter example suggests a relation between forecasting and planning. How change might affect reality in comparison to how change should affect reality distinguishes forecasts from predictions.

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. "The price of oil will change due to…." You can fill in the blank as the situation prompts: cold winter, pipeline failure, war. "Tomorrow at 11:30 AM it will rain…." because of whatever brings on rain. "There will be a change in government…." "Your baby will be born in the next two hours." A good predictive model can be turned into a machine—something we do quite often, turning into a device the physics or chemistry behind a good prediction: "If you don't watch the heat under the frying pan, the oil in it will catch fire."

Our own existence is one of never-ending change. Implicit in this dynamic condition of the living are:

(a) the impossibility of accurate forecasting, and
(b) the possibility of improving the prediction of physical phenomena, to the extent that we can separate the physical from the living.

Our guesses, expectations, predictions, and forecasts—in other words, our 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 insufficient information and experience, or lack of knowledge. 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. Of course, this applies to every other living form. The reaction-anticipation conjunction defines how effective the living is in dealing with change.

1.4 Self-awareness, Intentionality, and Planning

The human being has a distinct condition in the extraordinarily large realm of the living. It doesn't only play games (the example chosen in advanced research of high levels of control), but also conceives them. It not only understands questions in a given language, but also changes the language according to the human's changing pragmatic condition. Moreover, the human depends on a variety of other forms of living (billions of bacteria, for instance, inhabit the body), but in the larger scheme of things, it acquired a dominant position (not yet challenged by the technology created). In our world, human activity (although often enhanced through science and technology) is, for all practical purposes, the dominant force of change. Humans "are what we do" (the pragmatic foundation of identity, [41, pp. 258ff], [42]). The only identifier of human actions (and of other living entities) is their outcome. This is an instantiation of identity at the same time. The question, "What do you do?" cannot be answered with "I anticipate," followed, or not, by an object, such as "I anticipate that an object will fall," or "I anticipate my wife's arrival," or "I anticipate smelling something that I never experienced before."
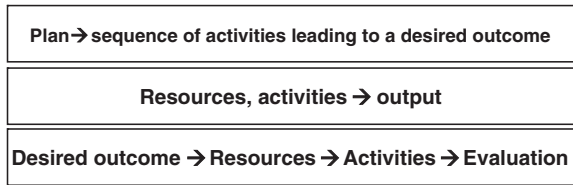
Anticipation is a characteristic of the living, but not a specific action or activity. Humans do not undertake anticipation. The dopamine release in anticipation of high emotional anticipation (associated with sex, eating, music, scientific discovery, for example) is autonomic. Humans are in anticipation. Anticipation is not a specific task. It is the result of a variety of processes. As an outcome, anticipation is expressed through consequences: increased performance (an anticipated tennis serve is returned); danger (such as a speeding car) is avoided; an opportunity (in the

stock market, for instance) is used to advantage. Anticipatory processes are auto-nomic. Implicit in the functioning of the living, such processes result in the pro-active dimension of life. This is where identity originates. Anticipatory processes are defined in contrast to reaction, although they often imply reaction as well. Playing a computer game—with the game "canned on the machine," or competing with someone via the medium of the game or a MMORPG (massively multiplayer online role-playing game)—over the internet can be reactive (with a predictive component) or anticipatory. It can also be random. Characteristic of the deter-ministic sequence of action-reaction defined in physics, reaction is the expression of the living's physical nature. Identity is expressed in the unity of the reactive and proactive dimensions of the human being. It appears as a stable expression, but actually defines change. It is the difference between what we seem to be and what we are becoming as our existence unfolds over time. Identity is affected by, but is not the outcome of, learning.

No matter what humans do, the doing itself—to which explicit and implicit learning belongs—is what defines the unfolding identity. The outcome is the expression of physical and intellectual abilities. It also reflects knowledge and experience. The expression of goals, whether they are specifically spelled out or implicitly assumed, affects the outcome of actions as well. The process through which existence is preserved at the lowest level—as with the phototropic mono-cell and progressing all the way up to the human being—is anticipatory. But at a certain level of life organization and complexity, the preservation drive assumes new forms through which it is realized. Anticipation is the common denominator. However, the concrete aspect of how it is eventually expressed—i.e., through self-awareness, intentionality, or in the activity called "planning"—changes as the interdependence of the processes through which the living unfolds increases.

Anticipation at the level of preserving existence is unreflected. Facial expression in anticipation of an action is a good example here, too. It seems that facial expression is not defined on a cultural level but is species wide (Ekman [35], Gladwell [36]). It is not a learned expression. Individuals can control their facial expression to an extent. However, there is always that one second or less in which control is out of the question. Intentionality is always entangled with awareness—one cannot intend something without awareness, even in vague forms. But this awareness does not automatically make human expressions carry anticipations more than the expression of the rest of the living does. We sweat "sensing" danger even before we are aware of it. The difference is evident on a different level. Humans reach self-awareness; the mind is the subject of knowledge of the mind itself. As such, we eventually recognize that our faces "speak" before we act (or before perspiration starts). They are our forecasts, the majority of them involuntary. Those intent on deciphering facial expression obtain access to some intriguing anticipatory mechanisms, or at least to their expression.

Fig. 8 Goal-driven means: future informed

Plan→ sequence of activities leading to a desired outcome
Resources, activities → output
Desired outcome → Resources → Activities → Evaluation

RESOURCES ⟵ **GOAL** ⟵ OUTCOME
⤷ ACTIVITIES ⟹ OUTPUTS ⬏

Planning (Fig. 8) is more than calculation. A planning machine for integrated activities carried out in an open system over a longer period of time would require real-time adaptive capabilities.

The planning dimension is based on learning capabilities: what road to choose at which time; how long it takes to find the daughter and prepare for the gym; how long will parent and daughter spend at the gym; which is the best way home, assuming that some other activity might be spontaneously chosen. It also implies flexibility, as a form of adapting to new circumstances (the daughter has a lot of homework, for example). "Take me from the University to my daughter's school. After she joins me, take us to the gym. After that, we go home." To prepare for a worst-case situation, one would have to generate possible breakdown timelines and provide contingency measures for each. Various reactive components (which correspond to reactive planning, i.e., how to react) can be effectively described in computational terms. For instance, process planning maps from design (which is an expression of anticipation) to instructions and procedures, some computer-aided (e.g., 3D printing), for effectively making things, or changing things. Operations (deterministic), operation sequences, tooling, and fabrication procedures are described in computer process planning and serve as input for automated activities.

Planning, expressed through policymaking, management, prevention, logistics, and even design, implies the *ante* element—giving advance thought, directing towards something, looking forward, engaging resources (including the self). Moreover, it implies understanding, which resonates with the initial form of the word denoting anticipation: *antecapere*. As such, the activity through which human beings identify themselves as authors of the blueprint of their actions takes place no longer at the object level, but on a meta-level. It is an activity of abstracting future actions from their object. It is also their definition in a cognitive domain not directly associated with sensory input, but rather with understanding, with knowledge. Plans synthesize predictive abilities, forecasting, and modeling (Fig. 9).

A plan is the expression of understanding actions in relation to their consequences. It is what is expressed in goals, in means to attain these goals, as well as in the time sequence for achieving them. A plan is a timeline; it is a script for interactions indexed to the timeline. To what we call understanding belong goals, the means, the underlying structure of the endeavor (tasks assumed by one person,

Fig. 9 Integrated Planning
Process (IIP) as part of the
understanding process
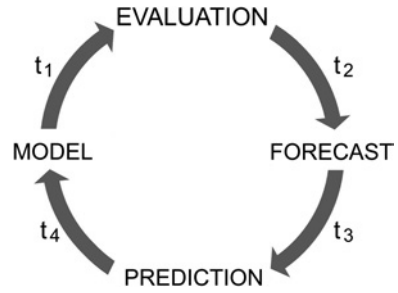
by several, the nature of their relation, etc.), a sense of progression in time,
awareness of consequences, i.e., a sense of value. As such, they report upon the
physical determination of everything people do, and of the anticipatory framework.
In every plan, from the most primitive to the utmost complex, the goal is associated
with the reality for which the plan provides a description (a theory), which is called
configuration space. If it is a scientific plan, such as the exploration of the moon or
the genome project, the plan describes where the "science" actually resides, where
those equations we call descriptions are "located." If it is a political plan, or an
education plan, the configuration space is made up of the people that the plan
intends to engage, and of the means and methods to make it work. Our own
description of the people, like the mathematical equations of science, is relative.
Such description of the configuration space, and, within that space, of the inter-
actions through which people learn from each other are subject to adjustments.

The plan also has to describe the time-space in which the goal pursued will
eventually be embodied. This is a manifold, towards which the dynamics of actions
and interactions (social context) will move those involved. In science, this is the
landing on the moon, or the map of the human gene; it can as well be a new
educational strategy or, in politics, the outcome of equal opportunity policies. The
plan associated with the self-driving automobile taking its user to the daughter's
school, to the gym, etc., is of a different scale, but not fundamentally dissimilar. All
the goals are anticipations projected against the background of understanding
change in the world as an expression of the unity between the dynamics of the
physical and the living. Plans spell out variables to be affected through actions, and
the nature of the interrelationships established in pursuing the plans. Quite often,
plans infer from the past (the reactive component) to the future (proactive com-
ponent). They also project how the future will eventually affect the sequence of
ensuing current states. Planning and self-regulation are related. The inner dynamics
of phenomena and their attractors—the goals to be attained—reflect this intercon-
nectedness. These attractors are the states into which the system will settle—at least
for a while. They are the descriptions of self-organizing processes, their eventual
destination, if we can understand it as a dynamic entity, not the statement of a static
finality. 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.

Processes with anticipatory, predictive, and forecasting characteristics are described through

Control $\rightarrow$ **function of (past state, current state, future state) system**
Adaptivity $\rightarrow$ **circumstances related to goals**

Knowledge of future states is a matter of possibilistic distributions:

$$r : \cup \rightarrow [0,1] \tag{9}$$

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

$$\mathcal{R}(X) = F \tag{10}$$

It is associated with a possibility distribution $\Pi_\mathbf{x}$ (Nadin [43]). Nothing is probable unless it is possible. Not every possible value becomes probable.

The anticipated performance (von Glasersfeld [32]) and the actual performance are usually related. The difference between the pursued goal and the concrete output of the process, together with the reward mechanism, guides the learning component.

Functioning under continuously changing conditions means that control mechanisms will have to reflect the dynamics of the activity (Fig. 10). This is not possible without learning. If we finally combine the automated part (everything involving the change of the physical can be automated) and human performance (expressed in behavior features), we arrive at an architecture that reflects the hybrid nature of plan-driven human activities that feed values into the sensors. Based on 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. If we compare this architecture to that of the Google Deep Mind Group, we notice that the difference is operational. Convolutional neural
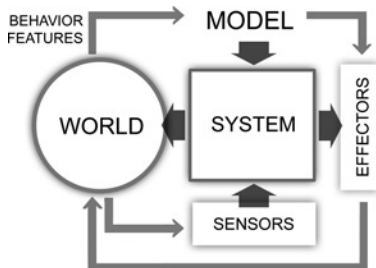

Fig. 10 Generic diagram of a hybrid control mechanism endowed with learning (Nadin [44])

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 (of students in a class, patients, vehicle drivers, airplane pilots, politicians in a power position, computer game choices, etc.) and the methods for extracting regularities characteristic of their behavior are connected. Learning ensues from adapting to new circumstances (i.e., change). The "learning"—classroom, physician's office, car, airplane, management system guiding a prime minister or a secretary of state, successfully playing a game, understanding a question and answering it, etc.—is thus one that combines its own dynamic (modified, evolving knowledge) and that of the persons involved (anticipation included). The suggestion here is that conceiving intelligent classrooms, intelligent schools, intelligent cars and airplanes, intelligent "assistants" for those in power, or intelligent game players is characteristic of an anticipatory perspective. However, the perspective does not automatically translate into proactive activity. Most of the time the system remains reactive. Embodied intelligence and the intelligence of challenged users could augment the perception of time, and thus help mitigate consequences of change for which society is rarely (if ever) prepared. 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 (as we shall see in the second part of this study).

2   Practical Considerations

Pursuing the enticing goal of making everything behave like a machine—and paying the price for it—stands in sharp contrast to a vision of acknowledging the living and its definitory anticipation. One writer put it in quite expressive terms: "Think of the economy as being more like a cat than a washing machine," (Taleb [45]). Evidently, becoming servants to robots, as Shannon cavalierly conceded, goes in the opposite direction. With this note, we are back to the preliminaries to the broader question of whether anticipatory computing is possible.

2.1   The "Why?" Question

Actually, "Why anticipatory computing?" would be a better question than simply questioning its feasibility. The reason for entertaining the question is straightforward: computation, of any nature, is nothing other than counting or measuring. The digital computer, as opposed to the person whose work was to calculate, i.e., to be a "living computer" provides automated calculation. (The first documented use of the word computer dates to 1613; it refers to persons performing calculations, to which we shall return.) It is not surprising that the human associates with calculation certain desired capabilities: the ability to make distinctions (large, small, wide,

narrow), to compare, to proceed in a logical manner, to guide one's activity. From the stones (*calculae*) used yesteryear to describe property, effort, and sequences of all kind to the alphabet of zeroes and ones (or Yes and No) of the new electronic abacus, the change was merely in scale, scope, speed, and variety of calculations, but not in its nature. The reason for calculations, and, implicitly, for measurement, remains the same: to cope with change, to account for it, to impact change.

Considering computer industry claims (e.g., Cigna CompassSM™, the MindMeld™ iPad app, among others), the question of whether anticipatory computing is possible appears to be meaningless. The public and the major users of computation (banks, the military, healthcare, education, the justice system, etc.) are enticed by gadgets supposedly able to perform anticipations. Leaving aside the marketing gags—"We know where you will be on February 2, 2017 and with whom" states the Nostradamus bot—what remains, as we shall see, are computations with predictive features. As respectable as one or the other is, their performance does not have identifiable anticipation features. It is still hard to believe that the computer community, of presumed smart individuals, simply falls prey to the seductive misrepresentation methods characteristic of marketing. This is an example of lack of knowledge, of incompetence. Setting goals (to anticipate) not connected to what is actually offered—to extract information from patterns of behavior with the aim of predicting—does not qualify as competence. To define human-level control in terms of computer game proficiency is as misleading (regardless of the unreserved blessing of being published in *Nature* [6]).

The reason for this extended study of whether anticipatory computing is possible is to set the record straight and inform future work that reflects the understanding of anticipatory processes.

2.1.1 Counting

Stones, or knots on a rope (Fig. 11), are a form of record keeping: so many sheep, so many slaves, so many arrows, whatever; but also number of days, of bricks, of containers (for water, oil, wine, etc.), of anything that is traded. Such measurements translate as the basis for transactions: those entitled to a portion of the exchange (yes, change of ownership, currency of reference related to the days and weeks it took to hunt, process, make, preserve, etc.) will exercise their rights. To know ahead of time what and how things will change always afforded an edge in the economy of survival (as in any subsequent economy, including the transaction economy characteristic of our time). Ahead of time (*ante* is Latin for "before") is ahead of others. Actions ahead of time are anticipatory. They are conducive to higher performance.

Therefore, anticipation as the "sense of the ever-changing context" is co-substantial with the preoccupation of describing change, either in words or in numbers, or, more generally, in any form of representation. Hence: representing change, as image (in the prehistoric cave paintings, for instance), as words trying to describe it, as numbers, equations, visualization, etc. is indicative of anticipation— the never-ending wager against change.

Fig. 11 The Incan *quipu*
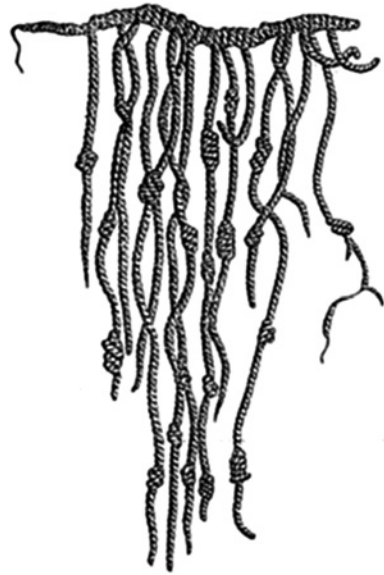
The fact that some descriptions (i.e., representations) are more adequate than others for certain activities is a realization of the nature of observations leading to learning. Where quantitative distinctions are more effective, numbers become more important than images, sounds, or words. With numbers—very much derived from the geometry of the human body (single head; pair of eyes, nostrils, ears; set of fingers and toes; myriad strands of hair)—comes the expectation of capturing change in operations easy to understand and reproduce. Counting emerges as a fundamental cognitive activity. Leibniz went so far as to state that music is the pleasure that the human mind experiences from counting without being aware that it is counting. (Poetry would easily qualify for the same view, so would dance.) This statement can be generalized, although temporal aspects (music unfolding over time/duration/ interval, as expression of rhythm) and spatial aspects are rather complementary. Still, counting involves the most basic forms of perception: the visual and the aural. To count implies the abstraction of the number, but also the abstraction of point, line, surface, and volume. A straight line is a set of adjacent points that can be counted (and the result is the length of the line); a surface is the collection of all lines making it up; and volume is represented by all the elements needed to arrive at it.

Just for illustration purposes: a more than 500-year-old woodcut (*The Allegory of Arithmetic*, Gregor Reisch, 1504). This image is part of the *Margarita Philosophica* describing the mapping from a counting board (for some reason, Pythagoras was chosen as a model) to the emergent written calculation (in which, for some even more obscure reason, Boethius is depicted). In this image, Hindu representation of numbers is used in what emerges as the art and science of mathematics.

Fig. 12 *The Allegory of Arithmetic.* The *abacist* uses the abacus; the *algorist* is involved in formulae

Indeed, the calculating table (counting board) is a machine—conceptual at that stage—and so are the abacus and all the contraptions that make counting, especially of numbers of a different scale than that of the immediate reality, easier, faster, cheaper. In the image (Pythagoras competing with Boethius), an *abacist* (one who knows how to count using an abacus) and an *algorist* (one who calculates using formulae) are apparently competing (Fig. 12). The abacist is a computer, that is, a person who calculates for a living. It is appropriate to point out that there are many other forms of computing, such as counting the elements that make up a volume of liquid, or a mass (of stone, wood). Under these circumstances, counting becomes an analog measurement. With the abacus, after moving the appropriate beads, you only have to align the result, and it will fall in place. This holds even more with the attempt to measure, i.e., to introduce a unit of reference (describing volumes, or weights).

2.2 Measuring

The act of measuring most certainly implies numbers. It also implies the conventions of measuring units, i.e., a shared understanding of what the numbers represent, based on the science behind their definition. Indeed, the numbers as such are data, their meaning results from associating the numbers to the measuring process, such as "under the influence of gravity" (Fig. 13).

Fig. 13 Measurement as
pouring medicine into beaker
(The pharmacist "computes"
the quantities prescribed)

Quantitative distinctions are associated with numbers. Qualitative distinctions
are associated with words, or any other means for representing them (e.g., sounds,
colors, shapes). In the final analysis, there are many relations between quantitative
and qualitative distinctions. Of course, numbers can be represented through words
as well, or visually, or through sounds.

It should be clear at this time that counting numbers (or associating qualities, such
as small, round, soft, smelly, etc.) is a discrete process, while "falling in place"
(measuring, actually) is based on analogies and is a continuous process. This dis-
tinction pretty much defines the numbering procedure—one in which representations
are processed sequentially, according to some rules that correspond to the mapping
from the questions to be answered to what it takes to answer it. Example: If you have
50 sticks of different length, how does one order them from shortest to longest? Of
course, you can "count" the length of each (using a measuring stick that contains the
"counted" points corresponding to the units of measurement) and painstakingly
arrange them in the order requested. Or you can use a "recipe," a set of instructions,
for doing the same, regardless of how many there are and how long each one.

2.2.1 The Early Meaning of Algorithm

Let us recall the calculating table from *Margarita Philosophica* allegory of arith-
metic: to the right, the human "computer," checking each stick, keeping a record of
the length of each, comparing them, etc.; to the left the *algorism* (no spelling error
here)—a person using a counting method by writing numbers in a place-value form
and applying memorized rules to these numbers. Before that, once the scale
changed (say, from tens to hundreds to thousands and tens of thousands), the
representations used, i.e., the symbols, such as Roman numerals, changed. LVI

means 56. A native of Kharazuni (a locality in what today is Uzbekistan) gave his name (Al-Khwarizuni arrived in Latin as *Algorituni*) to a treatise on the number system of the Indians (*Algorituni de Numero Indorum*). Actually, he provided the decimal number system for counting.[1] The fact that the notion of the algorithm, which has characterized the dominant view of computation since Turing, is associated with his name is rather indicative of the search for simple rules in counting. Those implicit in the decimal number system and in the place value form are only an example. Algorithm (on the model of the word logarithm, in French) maintains the connection to *arithmos*, ancient Greek for *number*. Numbers are easier to use in describing purposeful operations, in particular, means and methods for measuring. Such means and methods replace guessing, the raw estimate that experienced traders knew how to make, and which were accepted by all parties involved. The ruler and the scale are "counting" devices; instructions for using them are algorithms.

   Historic accounts are always incomplete. Examples related to the need and desire to make counting, and thus measurement, more like machine operations are usually indicative of the dominant knowledge metaphor of the time (the clock, the pneumatic pump, the steam engine, etc.). The machines of those past times embody algorithms. Leibniz used the label *machine* in describing the rules of differential calculus, which he translated into the mechanical parts (gears) of his machine. Machines such as clocks, water wheels, and even the simple balancing scale (embodying the physics of the lever) were used for various purposes. The balancing scale was used to estimate weight, or the outcome of applying force in order to move things (the most elementary form of change, i.e., in position), or change their appearance. Recalling the various contributions to computation—Blaise Pascal and his Pascaline device; Leibniz and his computer; Schickard and the calculating clock associated with his name; Babbage's analytical engine inspired by the loom, etc.— means to recall the broader view of nature they embody. It also suggests that calculations and measurements can be performed basically in either an analog manner or a digital manner (Fig. 14).

2.2.2 Why Machines for Calculations?

To this question 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.

---

[1]He also wrote the famous *Istikhvay Tarikh al-Yahud* (Extraction of the Jewish Era), a lunar cycle-based calendar for the Jewish holidays. This elaborate calendar, still in use, is a suggestion of his possible Jewish roots. (At the time of his activity, he was a Muslim scholar of the early stages of the creed, after he gave up his Zoroastrian identity). Inspired by the Hebrew *Mishnat ha Middot* (Treatise on Measures), he suggested elements of geometry well aligned with those of Euclid. That was a world of practiced diversity, including diversity of ideas.
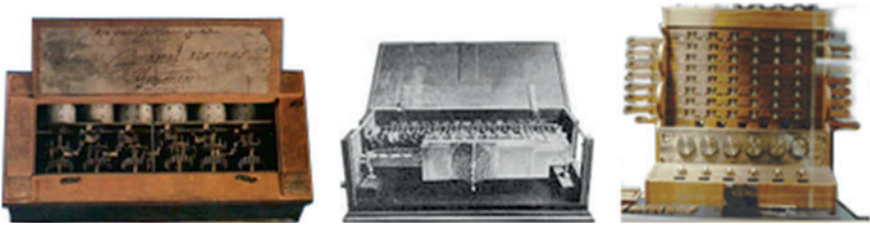
Fig. 14 The Pascaline, Leibniz's machine, Schickard's calculating clock

What he wrote is of more than documentary relevance. Astronomy, and applications related to it, required lots of calculations in his time. 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. Excellent men and women (to paraphrase Leibniz) program machines that process more data and faster, because almost every form of human activity involves calculations. 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 justifies the assumption. But there is no proof for it. Can this hypothesis be "falsified"?

It is at this juncture that our understanding of what computation is and the many forms it takes becomes interesting. Eugen Wigner's article of 1960 [46] 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." Less than shy about the subject, Gelfand and Tsetlin [47] 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." Leibniz would have seconded this formulation. Vellupillai [48] uses the same formulation in respect to economics.

2.3   Analog and Digital: Algorithmic and Non-algorithmic

The analog corresponds to the continuity of phenomena in nature. Pouring water or milk into a measuring cup in order to determine the volume, and thus, indirectly, the weight, is indicative of what analog calculations are. It is counting more than one molecule at a time, obviously. Similarity defines the domain. The lever on a scale automates the counting. It functions in the analog domain: add two or three more ounces until the scale is level, and that is the outcome of the calculation. Of course, to model simple phenomena and scale them is easy. But to reconfigure the analog—from calculating the volume of a liquid to that of a gas or solid—is more difficult.
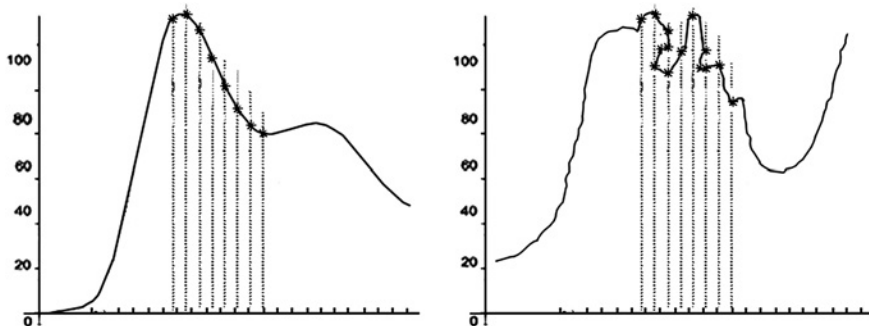
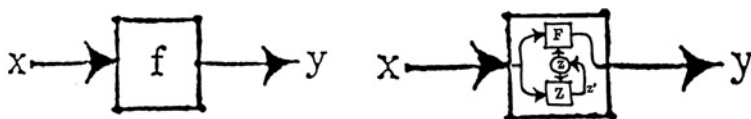**Fig. 15** Sampling: comparisons of rate and data retention in sampling

Fig. 16 Heinz von Foerster: trivial and non-trivial machines (his own drawings)

Moreover, it is hard to distinguish between what is actually processed and the noise that interferes with the data. By way of counter-example, when you count beans in a bag, by hand, you can easily notice the stone among them.

The digital is focused on sampling, on making the continuous discrete. At low rates of sampling, much relevant data is lost. The higher the rate, the better the approximation. But there is a cost involved in higher rates, and there are physical limitations to how fast a sampling machine can go (Fig. 15).

In the context of interest in machines of all kind (for conducting wars, for successful wagers, for calculating the position of stars, for navigation, for making things, 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. Heinz von Foerster took the time to distinguish between trivial and non-trivial machines (Fig. 16).

His distinction proved to be more consequential than initially assumed, once the model of the neuron (more precisely, its deterministic reduction was adopted (Fig. 17).

It is important to understand that input values are no longer a 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 (McCulloch and Pitts, 1943) to the most recent deep Q-network (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 [49]).

Fig. 17 The neuron machine
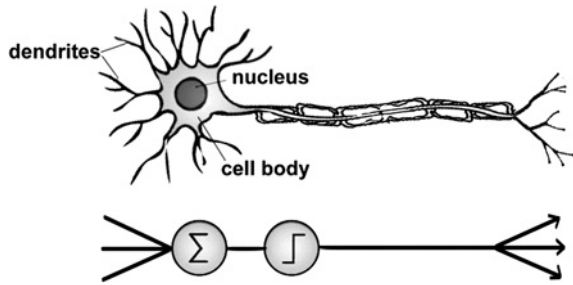
Evidently, the subject of interest remains the distinction between reaction-based processes—the theoretic machine has input, a number of inner states, and an output that is the outcome of the calculation—and predictive performance. There is no anticipatory dimension to account for. The "non-trivial machine" (von Foerster and Poerksen [50]) is essentially reactive: part of the calculation implies a dynamic dimension of the inner state connections. It is conceivable that along this line of an autonomic function associated with the inner state, anticipation could be defined as the result of the self-organization of such a machine. The DQN, like the professional human game testers, acquires a good understanding of the game algorithm, outperforming other reinforcement learning methods (including the training of living gamers).

2.3.1   The *a-Machine*

With the Turing machine, the real beginning of automated calculation was reached. Interestingly enough, behind his theoretic machine lies the same problem of automatic operations, in this case, the making and testing of mathematical statements. Hilbert was convinced that calculations were the basis for them. The meta-level of the enterprise is very relevant:

(a) objects in the reality of existence → representations → acts upon representations → new knowledge inferred from representations
(b) objects → numbers → counting → measurement → ideas about objects → ideas about ideas

The Turing saga was written so many times (and filmed with increased frequency) that it is hardly conceivable that the most important about it was not yet made public. Still, to understand the type of computation associated with his name—moreover, whether it is a possible path to anticipatory computing—a closer look is called for. Hilbert's conjecture that mathematical theories from propositional calculus could be decided—*Entscheidung* is the German for decision, as in proven true-or-false—by logical methods performed automatically was rejected. Indeed, Turing (after Gödel and Alonzo Church) disappoints Hilbert, the mathematician who challenged the community with quite a number of hard problems (some not yet elucidated).

First and foremost: Turing provided the mathematical proof that machines cannot do what mathematicians perform as a matter of routine: developing mathematical statements and validating them. This is the most important, and most neglected, contribution. Nevertheless, the insight into what machines can do, which we gain from Turing's analysis, is extremely important. Wittgenstein [51], recalling a conversation with Turing (in 1947) wrote: " 'Turing's machines': these machines are humans who calculate. And one might express what he says also in the form of games." Indeed, the idea behind digital computers is machines intended to execute any operation that could be done by a human computer. (Remember: initially, as of 1613, "computer" applies to a person employed to calculate, what Gregor Riesch meant by an algorist.) Turing [52] himself wrote, "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" [53]. Gödel would add, "mind, in its use, is not static, but constantly developing" [54]. "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, even a possible wrong decision. Therefore, non-algorithmic means not subject to pre-defined rules, but rather discovered as the process advances. For those who fail to take notice of Turing's own realization that not every computation is algorithmic, non-algorithmic computation does not exist.

Automatic machines (*a-machines* as Turing labeled them) can carry out any computation that is based on complete instructions; that is, they are algorithmic. One, and only one, problem remains: the machine's ability to recognize the end of the calculation, or that there is no end. This means that the halting problem turned out to be undecidable. This characterization comes from Gödel's work, where the undecidable names an entity that cannot be described completely and consistently. Turing's *a-machine* consists of an infinite tape on which symbols can be stored, a read/write tape head that can move left or right (along the tape), retrieve (read) symbols from the tape or store (write) to the tape. The machine has a transition tape and a control mechanism. The initial state (one from among many on the transition tape) is followed by what the control mechanism (checking on the transition tape) causes the machine to do. This machine takes the input values, obviously defined in advance; it operates on a finite amount of memory (from the infinite tape) during a limited interval. The machine's behavior is pre-determined; it also depends on the time context. Examining the design and functioning rules of the a-machine, one can conclude the following: 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. In some sense, the *a-machine* is no more than the embodiment of a physics-based view of all there is. This view ascertains that there are no fundamental differences between physical and living entities. This is a drastic epistemological reduction. 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.

2.3.2 Choice, Oracle, and Interactive Machines

Turing knew better than his followers. (Albeit, there is no benefit in making him the omniscient scientist that many proclaim him to be, reading into incidental notes ideas of a depth never reached.) In the same paper [53], Turing suggested different kinds of computation (without developing them). Choice machines, i.e., *c-machines*, involve the action of an external operator. The *a-machines* were his mathematical proof for the Hilbert challenge. Therefore, they are described in detail. The *c-machine* is rather a parenthesis. 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. Turing insisted that the oracle is not a machine; therefore 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. Predictive computation is attained; anticipation becomes possible.

The story continues. Actually, the theoretic construct known as the Turing machine—in it's *a-*, *c-*, and *o-* embodiments—will eventually become a machine proper within the ambitious Automatic Computing Engine (ACE) project. (In the USA, the EDVAC at the University of Pennsylvania and the IAS at Princeton University are its equivalents.) "When any particular problem has to be handled, appropriate instructions…are stored in the memory…and the machine is 'set up' for carrying out the computation," (Turing [55]). 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.

Von Neumann (who contributed not only to the architecture of the Turing machine-based computer, but also to the neural networks processing of data) asserted that, "…everything that can be described with a finite number of words, could be represented using a neural network" (Siegelmann and Sontag [56]). This is part of the longer subject of the Turing completeness or recurrent neural nets. Its relevance to the issue of anticipatory computing is indirect, via all processes pertinent to learning.

One more detail regarding Turing's attempt to define a test for making the distinction between computation-based intelligence and 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. It is well established by now that imitation, which was Turing's preferred game, is by no means indicative of intelligence.

Machines were programmed to answer questions in a manner that would make them indistinguishable from humans doing the same. This became the standard for winning in competitions meant to showcase progress in artificial intelligence (AI). To state that some entity—machine, person, simulation, or whatever else—can think is of low relevance, unless the thinking is about change, i.e., that it involves awareness of the future. The number of words necessary for describing such awareness is not finite; the number increases with each new self-realization of awareness. Creativity, in its broadest sense—to originate something (a thought, a melody, a theorem, a device, etc.)—is, of course, better suited to qualify as the outcome of thinking. However, at this level of the challenge, it should be clear that thinking alone is a necessary but not sufficient condition for creativity. Anticipation is the aggregate expression, in action, of all that makes up the living. Turing was not aware of this definitory condition of anticipation. It is difficult to speculate the extent to which he would have subscribed to it.

Not to be outdone by Google DeepMind, Facebook's AI research focused on understanding language (conversing with a human [6]. Learning algorithms in this domain are as efficient as those in playing games—provided that the activity is itself algorithmic. Unfortunately, the lack of understanding anticipation undermines the effort to the extent that the automated grammar deployed and the memory networks become subjects in themselves. The circularity of the perspective is its main weakness. One more observation: Imagine that you were to count the number of matchsticks dropped from a matchbox (large or small). It is a sequential effort: one stick after another. There are persons who know at once what the total number is. The label *savant syndrome* (from the French *idiot savant*, which would mean "learned idiot") is used to categorize those who are able to perform such counting (or other applications, such as multiplication in their head, remembering an entire telephone directory). Machines programmed to perform at this level are not necessarily different in both ability and degree of "autism"—impaired interaction, limited developmental dynamics.

But let us not lose sight of interactivity, of which he 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 (Eberbach et al. [57]) is not reducible to algorithmic computation. The most recent developments in the area of quantum computation, evolutionary computation, and even more so in terms of computational ubiquity, in mobile computing 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.

2.4   What Are the Necessary Conditions for Anticipatory Computing?

For a computation to qualify as anticipatory, it would have to be couched in the complexity corresponding to the domain of the living. Elsewhere [12], I argued that description of objects and phenomena, natural or artificial, that correspond to the intractable, make up the realm of G-complexity. Anything else corresponds to the physical.

2.4.1   Beyond Determinism

Anticipation comes to expression within G-complexity entities. 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 (I am personally aware only of the functioning of D-Wave, and there is some question whether it is a real quantum machine), one cannot expect them to output the same result all the time (Fig. 18).

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. 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. One case: Is the object in the frame analyzed a moving car? The answer is more like "It could be!" than "It is!" or "It's not!"

Predictive calculations are in some 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
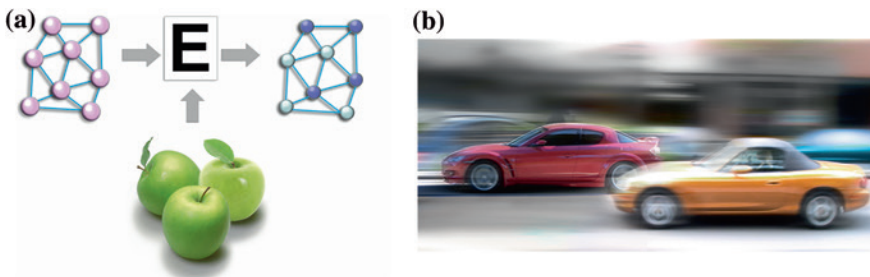

Fig. 18 (**a**) Quantum computation used in image recognition: apples, (**b**) a moving car

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. (For more details, the reader should consult Elsasser's Theory of Quantum Mechanical Description [58]. One more observation: the living is at least as sensitive to observation (representation, measurement) without necessarily qualifying as having a quantum nature.

Although very few scientists pursue this thought, it is significant to understand that Feynman argued for quantum computation in order to facilitate a better understanding of quantum mechanics, not for treating what are called "intractable problems." Factoring numbers, which are a frequent example of what quantum computation could provide, is important (for cryptography, for instance). However, it is much more relevant to better understand quantum phenomena. Paul Benioff and Richard Feynman (independently, in 1982) suggested that a quantum system can perform computations. Their focus was not on how long it takes to factor a 130-digit number (the subject of Shor's algorithm), not even the relation between time and the size of the input (the well-known P$\neq$NP problem of computer science).

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: the motoric expression (underlying the movement of humans and animals) might be regarded as an outcome of computation. Consider the classic example of touching the nose with the tip of the index finger (or any other finger, for that matter [59]). The physics of the movement (3 coordinates for the position of the nose) and kinematic redundancy (a wealth of choices given the 7 axes of joint rotation, 3 axes of shoulder rotation, elbow, joint, wrist rotation, etc.) lead to a situation in which we have three equations and seven unknowns. Of course, the outcome is indeterminate. The central nervous system, of extreme plasticity, can handle the richness of choices, since its own configuration changes as the action advances. However, those who perform computations in artificial muscles do not have the luxury of a computer endowed with plasticity. They usually describe the finite, and at most predict the way in which the physics of the artificial muscle works. There are, of course, many attempts to overcome such limitations. But similar to computer science, where computation is always Turing computation (i.e., embodied in an *a-machine*), biology-based computation, as practiced in our days, is more anchored in physics, despite the vocabulary. In reality, if we want to get closer to understanding the living, we need to generate a new language (Gelfand and Tsetlin [47, pp. 1–22]). Anticipation is probably the first word in this language.

2.4.2 An Unexpected Alternative

Mobile computing, which actually is the outgrowth of cellular telephony—i.e., not at all a computing discipline in virtue of its intrinsic hybrid nature of human-machine—offers an interesting alternative. From the initial computer-telephone integration (CTI) to its many current embodiments (tablets, note- and netbooks, smartphones, 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. Things get even more interesting as soon as we realize that 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 number of functions in addition to communication. Before focusing on the ubiquity of mobile computation, it is worth defining, in reference to the first part of this study, various forms of computation that make possible forecasting, prediction, planning, and even some anticipatory processes.

Regardless of the medium in which probability-based computing is attempted—any physical substratum (such as the artificial muscle mentioned above) can be used for computational purposes—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 (or, to be faithful to Peirce's terminology, *retroduction*), which is history dependent. Indeed, new ideas associated with hypotheses (yet another name for reduction) are not predictions, but an expression of anticipation (Fig. 19).

Alternatively, we can consider the interplay of probability and possibility. 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 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, [60] takes note of the fact that in Shannon's data-transmission theory (misleadingly called "information" theory), information is equated with a reduction in entropy—and not with form (not morphology). He understands this reduction to be the source of associating information with probability. But he also calls attention to possibilistic information, orthogonal to the probabilistic: one cannot be derived from the other. In his view (widely adopted in the scientific community), possibility refers to the distribution of meaning associated with a membership function. In more illustrative terms (suggested by Chin-Liang Chang), possibility corresponds to the answers to the question, "Can it happen?" (in respect to an event). Probability (here limited to frequency, which, as we have seen, is one
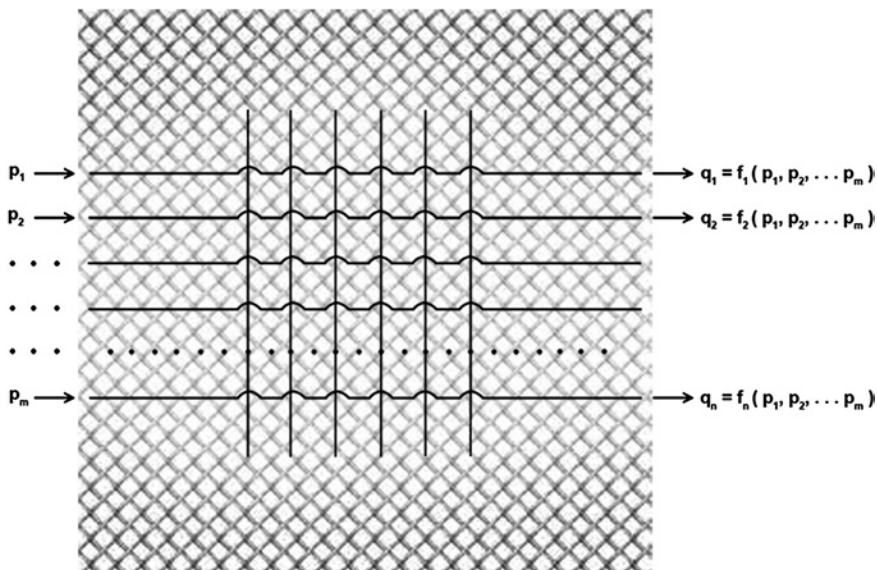
Fig. 19 Probability computer: the input values are probabilities of events. The integration of many probability streams makes possible dynamic modeling

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. 20).

One particular form of anticipative evaluation can be computing *perceptions* (Zadeh [61]). 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, as we suggested, an expression of anticipation (like/dislike, etc. expressed autonomously) based on perception. Soundscapes are yet another example (often of interest to virtual reality applications).

2.4.3 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, including integration of the human as a part of the interactive process.

In this sense we need to distinguish between actions initiated purposely by the person (let's say, taking a photo or capturing a video sequence) and actions triggered automatically by the behavior of the person carrying the device (sensing of
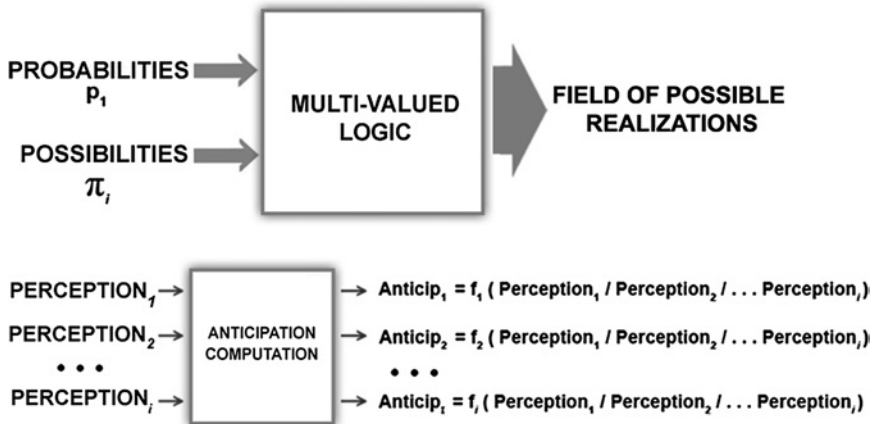
Fig. 20 Computing with probabilities and possibilities, computing with perceptions

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. The various sensors and the data they generate are of little, if any, significance to anticipation. If they could afford a holistic description, that would be conducive to anticipation (Nadin [62]). 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. 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. What emerge are characteristics reminiscent of cognitive processes traditionally associated with thinking. 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, 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, 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, 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.

These are Internet devices that share data, knowledge, experiences. In traffic, for instance, this sharing results in collision avoidance.

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 a 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. This is not to say that it is the only way to achieve anticipation performance.

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). A short list (incomplete, alas!) 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 (accelerometers, gyroscope, etc.), communication protocols (facilitating WiFi, Bluetooth, near-field communication), everything is in place to locate the user, the device, the interconnected subjects, 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 it), 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). Example: The attempts to deploy physical therapy 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 re-computed. Other such strategies are used in the use of the GPS facility (path tracking, but only as the device moves, i.e., the user is on a bike, on a car, train, etc.). Over all, the focus is on the minima (approximate representations). Instead of geo-location proper, infer location from data (as in the person's

calendar: restaurant, doctor, meeting, etc.). In some ways, the mobile device becomes an extension of the perception dimension of the living.

Although there is nothing that this kind of aggregated computation has in common with quantum computation, the focus on minima is relevant. As we have seen, 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. The metaphors of quantum computation, in particular the non-locality aspect, quite appropriately describe interactive processes, which no close algorithmic computation could perform. It is at this level where the once-upon-a-time classic texts of Bennett [63, 64], Bennett and Landauer [65], and others make evident the limits of an understanding of computation within the Turing machine model embodied in physical devices. Truth be told, no one has come up with a reassessment of the new context for open forms of computation.

I am inclined to doubt a statement such as "A computation, whether it is performed by electronic machinery, on an abacus, or in a biological systems such as the brain, is a physical process" [65, p. 58]. My position is that meaning is more important in the living than the outcome of any calculation is (should any take place). If there is computation within the living, chances are that it takes place differently from that on the abacus or in silicon. Moreover, I have doubts that the question "How much energy must be expended to perform a particular computation?" is very meaningful in respect to interactive computations. In an information processing environment, energy is not only that of the battery powering the device, but also of the interactions. Interactions in the living, as Niels Bohr suggested, continuously change the system. The neat distinctions of physics (often applied to living processes despite the fact that they are only partially relevant when it comes to life) are simply inadequate. Anticipation expressed in action has a specific energy condition corresponding to the fact that entropy decreases as a result of activity (Elsasser [58]).

Shannon's data transmission theory (improperly called "information theory") describes the cumulative effect of noise upon data. If a word or an image is transmitted over a channel, its initial order is subject to change, that is, it loses its integrity; or, in Shannon's view, its entropy increases. The Second Law of Thermodynamics (Bolzmann) contains a formalism (i.e., a mathematical description) similar to that of Shannon's law. But having the same description does not make the two the same, neither does it establish a causal relation. If we consider the genetic code, we'd better acknowledge that genetic messages do not deteriorate. The Laws of Thermodynamics apply, but not Shannon's law of data transmission. Information stability in processes of heredity, as Elsasser points out, makes the notion of information generation within the living necessary. This generated information guides predictive, as well as anticipatory, action.

2.4.4 No Awareness of the Future

Together with the statement that computation can be performed in any physical system comes the understanding that computers are, in the final analysis, subject to the laws of physics. This applies to energy aspects—how much energy it takes to perform a computation—as well as to computer dynamics. Stepney [66, p. 674] delivers a description of how machines "iteratively compute from the inputs to determine the outputs." Newton's physics and Lagrange's mathematics of the "principle of least action" are invoked and the outcome of the analysis is relatively straightforward: imperative languages (in Watt's sense [67], i.e., "based on commands that update variables held in storage") support Newton-based computations; logic languages (implementing relations) are Lagrangian. As such, the distinction does not really lead to significant knowledge from which computer science could benefit. But there is in the argument one aspect of relevance to anticipation: the time aspect. The underlying physical embodiment is, of course, described through physical laws. This applies to conventional computation as well as to quantum computation. To achieve even stationary condition, the computer would require awareness of the future (at least in terms of recognizing the end of execution time, if not the halting problem). Of course, no program is atemporal. For that matter, algorithms also introduce a time sequence (obviously different from that of programs).

When Turing modeled a calculation with pencil on paper on his abstract machine, his intention could not have been to ascertain a reductionist view. Rather, he focused on what it would take to transfer a limited human form of calculation—based on algorithms—to a machine. But outside that limited form remains a very large space of possibilities. Analog computation corresponds to another subset of calculations, with *ad hoc* rules reflecting a different heuristics. Neural networks, cellular automata, microfluidic processors, "wet computing," optical computing, etc. cover other aspects, and sometimes suggest calculations that might take place in the living (membrane computing, for instance) without being subject to self-control, or even being reflected in one's awareness. For all we know, neural networks dynamics, partially reflected in neural network computation, might even explain awareness and consciousness, but are not subject to introspective inquiry.

With all this in mind, we can, again making reference to our understanding of the difference between expectation, prediction, forecasting, etc., address the relation between computation in a physical substratum and that in a living substratum. A computer can predict its own outcome, or it can even forecast it. Everything driven by probability, i.e., generalizing from the past to the future, is physically computable. A physical machine can predict the functioning of another machine; it can simulate it, too. As a physical entity, such a machine is subject to the laws of physics (descriptions of how things change over time). A machine cannot anticipate the outcome of its functioning. If it could, it would choose the future state according to a dynamics characteristic of the living (evolution), not to that of physical phenomena (the *minima* principle). A machine, as opposed to a living medium of calculations, is infinitely reducible to its parts (the structure of matter down to its finest details, some of which are not yet fully described). 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. Within the living, there is no identity as we know it from physics. All electrons are the same, but no two cells are the same.

The Law of Non-Identity: The living is free of identity.

The living describes the world and itself in awareness of the act of describing. The living continuously remakes itself.

2.4.5 The Mobile Paradigm and Anticipatory Computing

But let's continue with more details of the mobile paradigm and the latter's relevance to anticipatory computing. The first aspect to consider is the integration of a variety of sensors from which data supporting rich interactions originate (Fig. 21).

Distinct levels of processing are dedicated to logical inferences (while driving, one is far from the desktop; or, while jogging, is not in the office, unless walking on a treadmill) with the purpose of minimizing processing. Technical details—the physics, so to say—are important, although for our concerns the embodied nature of interaction between user and device are much more relevant. 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 speech data. Still, a utility, such as StressSense, delivers useful information, which is further associated with blood pressure, heart rhythm, possibly EMG, and what results can assist the individual in mitigating danger. The spelling 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 we discuss.

El Kalioubi (whose work was mentioned previously) developed a similar facility for reading facial expression. In doing so, she facilitates the anticipatory dimension of emotions to a degree that this facility makes available information on attention— the most coveted currency in the world of computer-supported interactions. During a conversation we had (at SIGGRAPH 2010, Boston, when she was just starting her activity at MIT), she realized that MindReader—her program at the time—was merely making predictions under the guidance of a Bayesian model of probability

Fig. 21 Sensor integration with the purpose of facilitating rich interactions

inferences. Since that time, at Affidex, her focus is more and more on associating emotional states and future choices. It is easy to see her system integrated in mobile devices. Important is the realization that the description of physical processes (cause-and-effect sequence), and 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 determines the process (Landauer [68]). For instance, in machine learning, the structure of classifiers—simple, layered, complicated—is partially relevant. What counts is the training data, because once it is identified as information pertinent to a certain action, it will guide the performance. However, the curse of dimensionality does not spare mobile computing. Data sets scale exponentially with the expectation of more features. Many models excel in the number of features exactly because their designers never understood that the living, as opposed to the physical, is rather represented by sparse, not big, data. This is the result of the fact that living processes are holistic (Chaitin [69]).

At this time in the evolution of computation, the focus is changing from data processing to proving the thesis that all behavior, of physical entities and of organisms (the living) is either the outcome of calculations or can be described through calculations. This is no longer the age of human computers or of computers calculating the position of stars, or helping the military to hit targets with their missiles. Routine computation (ledger, databases, and the like) is complemented by intelligent control procedures. Self-driving cars or boats or airplanes come after the smart rockets (and everything else that the military commissioned to scientists). It is easy to imagine that the deep-Q network will soon give place to even higher performing means and methods that outperform not only the algorithms of games, but also of the spectacular intelligent weapons.

The Law of Outperforming Algorithms: For each algorithm, there is an alternative that will outperform it.

All it takes is more data, higher computer performance, and improved methods for extracting knowledge from data.

Thesis 3: Anticipatory computation implies the realization of necessary data (the *minima* principle).

Working only on necessary data (and no more) gives anticipatory computation an edge. It does not depend on technology (e.g., more memory, faster cycles) as does algorithmic computation.

Corollary: Predictive computation, as a hybrid of algorithmic and anticipatory computation, entails the integration of computation and learning.

Human level control is achieved not by outperforming humans playing algorithmic games, but by competing with humans in conceiving games driven by anticipation. Human-like understanding of questions in natural language is relevant to language at a certain moment in time (synchronic perspective) but lacks language dynamics (diachronic perspective).

2.4.6 Community Similarity Networks: How Does the Block Chain Model Scale up?

Without the intention of exhausting the subject, I will discuss a few issues pertinent to directions that pertain to the anticipation potential of interactive mobile devices. The tendency is 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. At this level, mobile computation (the smartphone in its many possible embodiments) becomes part of the ecology of billions of devices, some endowed with processing capabilities, others destined to help in the acquisition of significant data. Each device—phone, objects in the world, individuals, animals, etc.—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. Contracts are issued as necessary: deliver supplies (e.g., for house cleaning, or for a 3D print job). The smartphone can automatically post the bid ("Who has the better price?"), but so could any other device on the Internet-of-Everything. Peer-to-Peer in this universe allows for establishment of dynamic communities: they come into existence as necessary and cease to be when their reason for being no longer exists.

So-called community similarity networks (CSN) associate users—individuals or anything else—who share 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). A variety of similar contraptions (a chip in the shoe, a heart monitor, hearing- or seeing-aid devices) also generates data. The Apple Watch™, or any other integrating artifact, scales way further, as a health monitoring station. To quote a very descriptive idea from one of the scholars I invited to the upcoming conference on Anticipation and Medicine (Delmenhorst, Germany, September 2015):

> Real time physiological monitoring with reliable, long-term memory storage, via sophisticated "physiodiagnostics" devices could result in a future where diseases are diagnosed and treated before they even present detectable symptoms (McVittie and Katz [70]).

The sentence could be rewritten to apply to economic processes and transactions, to political life, to art, to education. The emphasis is on before, characterizing anticipation. Of course, understanding language would be a prerequisite. (As already mentioned, the AI group at Facebook [6] is trying to achieve exactly this goal.)

On this note I would argue with Pejovic and Musolesi [71] that neither MindMeld™ (enhancing online video conferencing), nor GoogleNow™, or

Microsoft's Cortana™ (providing functionality without the user asking for it) justifies qualifier anticipatory. Nevertheless, I am pretty encouraged by their project (anticipatory mobile dBCI, i.e., behavior change interventions), not because my dialog with them is reflected in the concept, but rather because they address current needs from an anticipatory perspective. Indeed, behavior change, informed by a "smart" device, is action, and anticipation is always expressed in action.

Just as a simple example: 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—not another app (there are too many already), but rather a dialog facility.

I would hope that similar projects could be started for the domains mentioned above (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.), we could test anticipatory ideas embodied in new forms of computation, such as those described so far. The progress in predictive computation (confusingly branded as anticipatory) is a promising start.

2.4.7 Robots Embody Predictive Computation (and Even Anticipatory Features)

Anticipatory computation conjures the realm of science fiction. However, 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 couched in complexity. 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.

Spectacular progress in the field of robotics comes close to what we can imagine when approaching the issue of anticipatory computation. 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* is the Russian word that inspired the Czech Karel Čapek to coin the term). Therefore, like the human being, 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 interest us here.

The predictive abilities of robots pose major computational challenges. In the living, the world, in its incessant change, appears as relatively stable. For the robot to adapt to a changing world, it needs a dynamic refresh of the environment in

which it operates. Motor control relies on rich sensor feedback and feed-forward processes. Guiding a robot (towards 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. 22).

Intelligent decisions require data from the environment also. Therefore, sensors of all kinds are deployed (to adaptively control the movement but also to make choices). 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 (Makin, Holmes & Ehrsson [72], Nadin [73]), the position of arms, legs, fingers, etc. corresponds to sensory information from skin, joints, muscles, tendons, eyes, ears, nostrils, tongue. Redundancy, which in other fields is considered a shortcoming (costly in terms of performance) helps eliminate errors due to inconsistencies or to sensor data loss,
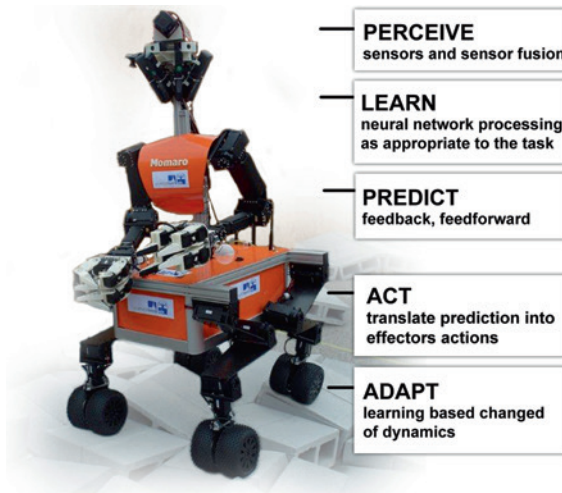

Fig. 22 Interaction is the main characteristic of robots. The robot displayed serves only as an illustration. It is a mobile manipulation robot, Momaro, designed to meet the requirements of the DARPA Robotics Challenge. It consists of an anthropomorphic upper body on a flexible hybrid mobile base. It was an entry from the Bonn University team NimbRo Rescue, qualified to participate in the DARPA Robotics Challenge taking place from June 5–6, 2015 at Fairplex, in Pomona, California

and to compensation of variances. The technology embodied in neuro-robots endowed with predictive and partial anticipatory properties (e.g., "Don't perform an action if the outcome will be harmful") integrates recurrent neural networks (RNN), multilayered networks, Kalman filters (for sensor fusion), and, most recently, deep learning architectures for distinguishing among images, sounds, etc., and for context awareness (Schilling and Cruse [74]). Robots require awareness of their state and of the surroundings in order to behave in a predictive manner. (The same holds for wearable computers.) Of course, robots can be integrated in the computational ecology of networks and thus made part of the Internet-of-Everything (IoE).

2.5 Computation as Utility

Based on 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
- 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 representations of somato-sensory relevance) (Fig. 23).

In view of 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 (the ever-expanding domain of apps is rather telling of their limitations), would be part of the complex process of forming, expressing, and acting in anticipation. Achieving an adaptive open system is most important. Outperforming humans in playing closed-system games is not a performance to be scorned. But it is only a first step. The same can be said of conceiving and implementing a so-called "intelligent dialog agent" as a prerequisite for understanding natural language. It is not the language that is alive, but those who constitute themselves in language (Nadin [41]). Memory Networks might deliver within a closed discourse universe, but not in an open pragmatic context. Understanding what anticipation is could spare us wasted energy and talent, as well as the embarrassment of claims that are more indicative of advancing deeper into a
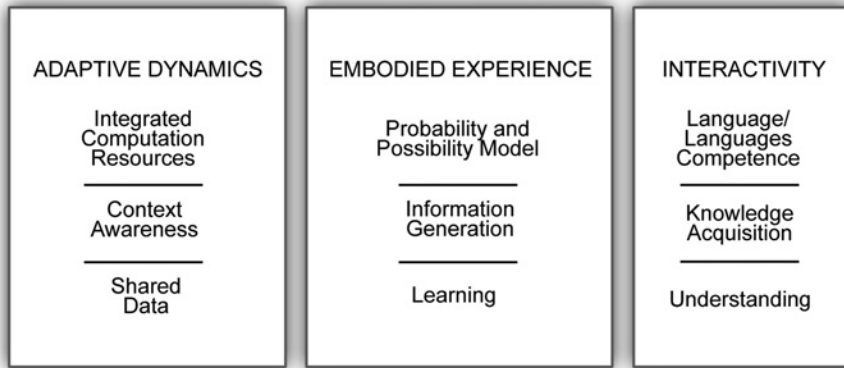
Fig. 23 Adaptive dynamics, embodied experience, and rich interactivity are premises for anticipatory performance

one-way street of false assumptions. Instead, we could make real progress in understanding where the journey should take us.

Cigna CompassSM is a trademark of the Cigna Corporation
MindMeld™ is a trademark of Expert Labs
Fitbit™ is a trademark of Fitbit, Inc.
AppleWatch™ is a trademark of Apple, Inc.

References

1. Mitchell, M.: Is the universe a universal computer? Science **298**(5591), 65–68 (2002)
2. Deutsch, D.: Quantum theory, the Church-Turing principle and the universal quantum computer. Proc. R. Soc. London A **400**, 97–117 (1985)
3. Hawking, S.: Stephen Hawking warns artificial intelligence could end mankind. Interview with Rory Cellan-Jones, BBC. http://www.bbc.com/news/technology-30290540. Accessed 2 Dec 2014
4. Kurzweil, R.: The Singularity is Near. Viking Press, New York (2005)
5. Shannon, C.: Interview: Father of the Electronic Information Age by Anthony Liversidge, Omni, August (1987)
6. Minh, V., Kavukculoglu, K., Silver, D., et al.: Human-level control through deep reinforcement learning. Nature **518**, 529–533 (2015). http://www.nature.com/search?journals=nature%2Cnews&q=volodymyr&shunter=1424698400067. Accessed 27 Feb 2015
7. Weston, J., Bordes, A., Chopra, S, Mikolov, T.: Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks (2015). http://arxiv.org/pdf/1502.05698v3.pdf. Accessed 3 March 2015
8. Monfort, N.: Twisty Little Passages: An Approach to Interactive Fiction. MIT Press, Cambridge (2005)
9. Wolfram, S.: A New Kind of Science. Wolfram Media, Champaign (2002)
10. Zuse, K.: Rechnender Raum. Friedrich Vieweg & Sohn, Braunschweig. English: Calculating space, An MIT Technical Translation Project MAC, MIT, 1970. AZTEC School of Languages, Inc., Cambridge (1969)

11. Gödel, K.: Über formal unentscheidbare Sätze der Principa Mathematica und verwandte Systeme, Monatshefte für Mathematik und Physik. **38**, 173–198 (1931). The first incompleteness theorem originally appeared as Theorem VI
12. Nadin, M.: G-Complexity quantum computation and anticipatory processes. Comput. Commun. Collab. **2**(1), 16–34 (2014)
13. England, J.L.: Statistical physics of self-replication. J. Chem. Phys. **139**, 12, 121293 (2013)
14. Carroll, L.: Through the Looking Glass. MacMillan, London (1871). http://www.brainyquote.com/quotes/quotes/l/lewiscarro389152.html
15. Newton, I.: Philosophiae Naturalis Principia Mathematica (Mathematical Principles of Natural Philosophy) (1687)
16. Windelband, W.: Aufsätze und Reden zur Philosophie und ihrer Geschichte, 5. Präludien, Tübingen (1915)
17. Montgomery, R.: Gauge theory of the falling cat. In: Enos, M.J. (ed.) Dynamics and Control of Mechanical Systems, American Mathematical Society, pp. 193–218 (1993). http://dictionary.sensagent.com/Falling%20cat%20problem/en-en/
18. Mehta, R.: Mathematics of the Falling Cat (2012). http://www.math.smith.edu/∼rmehta/slides/fallingcatnew.pdf
19. Kane, T.R., Scher, M.: A dynamical explanation of the falling of the cat phenomenon. Int. J. Solids Struct. **5**, 663–670 (Pergamon Press, London, 1969)
20. Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik [On the clear content of quantum-theoretic kinematics and mechanics]. Z. Angew. Phys. **43**(3–4), 172–198 (1927)
21. 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. **4**(10), 1043–1048 (2001)
22. Nadin, M.: Anticipation in Knowledge Media Design (2nd edn). Knowledge Media Design (Stefan, P., (ed.)), pp. 71–90. Oldenburg, Berlin (2006)
23. Kahneman, D. and Tversky A.: On the reality of cognitive illusions. Psychol. Rev. **103**(3), 582–591 (1996)
24. Delfabbro, P.: The stubborn logic of regular gamblers: obstacles and dilemmas in cognitive gambling research. J. Gambl. Stud. **20**(1), 1–21 (2004)
25. Gigerenzer, G., Muir Gray, J.A. (eds.): Better Doctors, Better Patients, Better Decisions: Envisioning Health Care 2020. MIT Press, Cambridge, (2011)
26. Sedlmeier, P., Gigerenzer, G.: Do studies of statistical power have an effect of the power of studies? Psychol. Bull. **105**(2), 309–316 (1989)
27. Tversky, A., Gilovich, T.: The "Hot Hand": Statistical reality or cognitive illusiion? Chance, vol. 2, no. 4., pp. 31–34. Taylor & Francis, London (1989, published on line 2012)
28. Miller, J.B., Sanjurjo, A.: A Cold Shower for the Hot Hand Fallacy. Social Science Research Network (Working Paper No. 518) (2014)
29. Hertwig, R., Ortmann, A.:The Cognitive Illusion Controversy: A Methodological Debate in Disguise that Matters to Economists. Experimental Business Research, pp. 113–130. Springer, New York (2005)
30. Tsetsos, K., Chater, N., Usher, M.: Salience driven value integration explains decision biases and preference reversal. Proc. Natl. Acad. Sci. U.S.A. **109**(24), 9659–9664 (2012)
31. Bernoulli, J.: Ars Conjectandi (The Art of Conjecturing, in Latin) (1713)
32. von Glasersfeld, E.: Radical Constructivism: A Way of Knowing and Learning. The Falmer Press, London/Washington (1995)
33. Salimpoor, V., Benovoy, M., Larcher, K., Dagher, A., Zatorre, R.J.: Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nat. Neurosci. **14**, 257–256 (2011)
34. Ekman, P., Rosenberg, E.L. (eds.): What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). Oxford University Press, New York (1997)
35. Ekman, P.: Emotions Revealed. Times Books, New York (2003)

36. Gladwell, M.: The Naked Face. Can experts really read your thoughts? The New Yorker, pp. 38–49 (2002)
37. Arsham, H.: Time Series Analysis and Forecasting Techniques (2002). http://ubmail.ubalt.edu/~harsham/stat-data/opre330.htm
38. Dayan, P., Kakade, S., Montague, P.R.: Learning and selective attention. Nat. Neurosci. **3**, 1218–1223 (2000)
39. Nicolelis, M.A.: Actions from thoughts. Nature **409**(6818), 403–407 (2001)
40. Nicolelis, M.A.L., Lebedev, M.A.: Principles of neural ensemble physiology underlying the operation of brain–machine interfaces. Nat. Rev. Neurosci. **10**(7), 530–540 (2009)
41. Nadin, M.: The Civilization of Illiteracy. Dresden University Press, Dresden (1997)
42. Nadin, M.: Antecapere ergo sum. What price knowledge? AI & Society 25th Anniversary Volume: A Faustian Exchange: What is To Be Human in the Era of Ubiquitous Technology, vol. 28, pp. 39–50. (Springer, London, 2013)
43. Nadin, M.: Anticipation and risk—from the inverse problem to reverse computation. In: Nadin, M. (ed.) Risk and Decision Analysis vol. 1, no. 2, pp. 113–139. IOS Press, Amsterdam (2009)
44. Nadin. M.: Anticipatory computing: from a high-level theory to hybrid computing implementations. Int. J. Appl. Res. **1**(1), 1–37 (2010) (Image based on earlier work: Hybrid Anticipatory Control Mechanisms (2002))
45. Taleb, N.N.: Learning to love volatility. Wall Street Journal (2012)
46. 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. **13**, 1–14 (1960)
47. Gelfand, I.M., Tsetlin, M.L.: Mathematical Modeling of Mechanisms of the Central Nervous System. Models of the Structural-Functional Organization of Certain Biological Systems (I.M. Gelfand, Editor; C.R. Beard, Trans.) MIT Press, Cambridge (1971)
48. Velupillai, K.V.: The Unreasonable *In*effectivenss of Mathematics in Economics. Working Paper No. 80 (2004). http://wwweocnomics.nugalway.ie/index/html [(Retrieved February 28, 2015) and in Camb. J. Econ. (Spec. Issue Econ. Future), **29**(6), 849–872 (2005)]
49. Hubel, D.H., Wiesel, T.N.: Receptive fields of single neurons in the cat's striate cortex. J. Physiol. **165**(3), 559–568 (1963)
50. von Foerster, H., Poerksen, B.: Understanding systems: conversations on epistemology and ethics. Kluwer, New York (2002)
51. Wittgenstein, L.: Remarks on the Philosophy of Psychology, Vol. 1, G. E. M. Anscombe and G. H. von Wright (Eds.), G. E. M. Anscombe (Trans.); Vol. 2, G. H. von Wright and H. Nyman (Eds.), C. G. Luckhardt and M. A. E. Aue (Trans.). Blackwell, Oxford. (1980) Cited in J. Floyd: Wittgenstein's Diagonal Argument: A Variation on Cantor and Turing. http://www.bu.edu/philo/files/2012/09/Wittgensteins-Diagonal-Argument.pdf
52. Turing, A.M.: Intelligent machinery [technical report]. Teddington: National Physical Laboratory (1948) cf. (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)
53. Turing, A.M.: Programmers' Handbook for Manchester electronic computer. Mark II (1951)
54. Gödel, K.: Some remarks on the undecidability results, Collected Works II, pp. 305–306. Oxford University Press, Oxford (1972)
55. Turing, A.M.: The ACE Report. A.M. In: Carpenter, B.E., Doran, R.W. (eds.) Turing's Ace Report of 1946 and Other Papers. MIT Press, Cambridge (1986)
56. Siegelmann, H.T., Sontag, E.D.: On the computational power of neural nets. J. Comput. Syst. Sci. **50**(1), 132–150 (1995)
57. Eberbach, E., Goldin, D., Wegner, P.: Turing's ideas and models of computation. In: Teuscher, C. (ed.) Alan Turing. Life and Legacy of a Great Thinker. Springer, Berlin/Heidelberg (2004)
58. Elsasser, W.M.: Theory of quantum-mechanical description. Proc. Natl. Acad. Sci. **59**(3), 738–744, (1968)

59. Latash, M.L.: Synergy. Oxford University Press, New York (2008)
60. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst. **1**, 3–28 (1978). (Elsevier, Amsterdam)
61. Zadeh, L.A.: Foreword. Anticipation—the end is where we start from (M. Nadin). Lars Müller Verlag (2003)
62. Nadin, M: Quantifying Anticipatory Characteristics. The AnticipationScope and the Anticipatory Profile. In: Iantovics, B., Kountchev, R. (eds.) Advanced Intelligent Computational Technologies and Decision Support Systems, Studies in Computational Intelligence, vol. 486, pp. 143–160. (Springer, New York/London/Heidelberg: Springer 2013)
63. Bennett, C.H.: Undecidable dynamics. Nature **346**, 606–607 (1990)
64. Bennett, C.H.: Universal computation and physical dynamics. Physica D Nonlinear Phenomena **86**(1–2), 268–273 (1995). (Elsevier, Amsterdam)
65. Bennett, C.H., Landauer, R.: The fundamental physical limits of computation. Sci. Am. **253** (1), 48–56 (1985)
66. Stepney, S.: Local and global models of physics and computation. Int. J. Gen Syst **43**(7), 673–681 (2014)
67. Watt, D.A.: Programming Language Design Concepts. Wiley, Hoboken (2004)
68. Landauer, R.: The physical nature of information. Phys. Lett. A **217**, 188–193 (1996)
69. Chaitin, G.: What is life? ACM SICACT News **4**(January), 12–18 (1970)
70. MacVittie, K., Katz, E.: Biochemical flip-flop memory systems: essential additions to autonomous biocomputing and biosensing systems. Int. J. Gen Syst **43**(7–8), 722–739 (2014)
71. Pejovic, V., Musolesi, M.: Anticipatory mobile computing: a survey of the state of the art and research challenges. ACM Comput. Surv. V(N) (Association for Computing Machinery, New York, 2015)
72. Makin, T.R., Holmes, N.P., Ehrsson, H.H.: On the other hand: dummy hands and peripersonal space. Behav. Brain Res. **191**, 1–10 (2008)
73. Nadin, M. Variability by another name: "repetition without repetition." In: Anticipation: Learning from the Past. The Russian/Soviet Contributions to the Science of Anticipation, pp. 329–336. Springer (Cognitive Systems Monographs), Chem, 2015
74. Schilling, M., Cruse, H.: What's next: recruitment of a grounded predictive body model for planning a robot's actions. Frontiers Psychol **3**, 1–19 (2013)