Epistemological Issues in Human Computation

Helmut Nechansky

Defining Epistemology

Traditional Epistemology is the branch of philosophy concerned with individual human knowledge, its base, its content and its validity. The focus on the individual results from the fact that there is no knowledge without an individual carrier.

There is no unequivocal definition of knowledge yet, but there is some shared understanding that knowledge is determined by the following four aspects: (1) The human senses and the human mind form its structural base; This base determines (2) what can become its possible content; (3) This content may or may not amount to a representational model corresponding to the world external to the individual carrier; (4) If the content does amount to a valid representational model can be confirmed by repeated observation of a correspondence with the external world, by observation of predicted states, and by goal-orientated actions leading towards predicted goal-states.

Social epistemology adds that the knowledge of any single individual depends on and is interrelated with the knowledge of other individuals, since any human is born, brought up and mostly lives in a social world. Therefore individual knowledge cannot be studied alone.

There are many other branches of epistemology we cannot mention here due to limitations of scope and space. The Internet Encyclopedia of Philosophy (2013) and the Stanford Encyclopedia of Philosophy (2013) offer the easiest access to this wide field, while Goldman (1999) offers an interesting discussion. However, this chapter provides a context sufficient for understanding the role of epistemology in human computation.

H. Nechansky (✉)
Nechansky—Engineering Efficiency, Vienna, Austria
e-mail: hn@nechansky.co.at

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_8,

The Role of Epistemology

Epistemology as a branch of philosophy may seem outdated in a time of ‘knowledge society’, of cloud computing, and human computation. But this is not the case, since we do not yet have an unequivocal, agreed on, scientific definition what actually constitutes ‘knowledge’. So any dealing with knowledge is ultimately still a philosophic endeavor.

And knowledge is the base of all our actions. Questions about this base arise often: How do we know? What can we know? Is this knowledge valid? Is it complete, i.e. sufficient to reach a goal? These are *epistemological* questions at the core of all human endeavors. Usually they do not get the attention they would deserve. And the more complex the systems, on which we rely, become, the more important become *answers* to these questions.

A Cybernetic Approach to Main Aspects of Epistemology

Cybernetics is the general theory of control in technical, biological and sociological systems. Control is pursuing and maintaining a goal-value, i.e. a certain physical state, against a changing environment, i.e. against physical influences disturbing that state. The process of control consists of (a) observing the environment with sensors, (b) comparing the sensor data with a goal-value and (c) deciding for an action to achieve that goal. Standard example for that process is a temperature controller, which aims at a desired room temperature as goal-value; to achieve that it (a) observes the current temperature, (b) compares it with the desired room temperature and (c) decides between the actions “heating” or “cooling” to achieve that.

In the following we will consider humans as complex controller structures. Here the brain has *in principle* controller functions similar to a temperature controller, but in much larger numbers and much more complex forms. Primarily the brain has to enable survival by maintaining some existential goal-values (necessary air, water and food supply; the body temperature). To achieve that it has (a) to observe the state of the environment, (b) to compare if that state serves the existential goal-vales, and (c) to decide for actions to enable that. Secondarily the brain has additional controller functions, which enable making a model of the environment, and, based on that, making predictions, concepts and setting long-term and short-term goal-values. To realize these future goal-values the brain again carries out the controller functions of (a) observing the environment, (b) checking if it corresponds to the goals and (c) deciding for actions to make it so.

Of course, the preceding description of brain functions is a crude simplification (for some important underlying complexities see Nechansky 2012a, b, 2013a, b), but we do maintain that the brain has primarily controller functions. For reasons of brevity we consider here just a few of these controller functions, each illustrating an epistemological problem. We will first describe these controller functions for individuals (illustrating the core problems of traditional epistemology) and then analyze how they work when two individuals interact (illustrating the core problems of social epistemology). Then we equip these two individuals with connected computers and discuss the resulting options. And this will be the base to finally place human computation within epistemology.

Epistemological Aspects of the Individual

To illustrate basic epistemological aspects of human reasoning we present here a complex controller structure (see Fig. 1): Here sensors provide input data used to model aspects of the environment, which are relevant to the given goal-values. Then these models are used for two purposes: internally, occasionally, to modify the goal-values and externally, continuously, to make decisions for goal-orientated actions. In more detail this structure carries out the following functions:

Sensor Inputs: Sensors allow the observation of certain physical aspects of the environment, and turn these into internal sensor data, which somehow represent and map them.

Modeling Decisions: Under this heading we summarize all decisions that have to do with sensor data. This includes what is usually called ‘learning’, but goes beyond that.

Primarily modeling decisions are about what to *ignore*. Humans are permanently confronted with more stimuli than they can observe. So they have to decide to pay

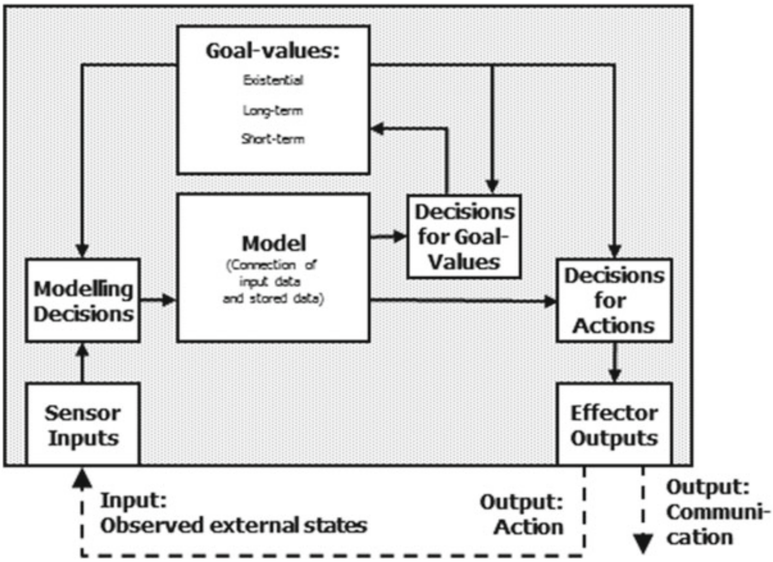

Fig. 1 Epistemological aspects of the individual: the external loop with modeling decisions and decisions for actions and the internal loop with decisions for goal-values

attention to some, and to ignore others considered irrelevant in relation to their *goal-values* (see below). And from the stimuli humans pay attention to they produce permanently more sensor data than they can use for active data processing. So they have to ignore sensor data considered irrelevant.

Secondarily modeling decisions are about what do to with the data considered to be relevant. These decisions are about (a) storing actual sensor data; (b) retrieving stored data for comparison with actual sensor data for pattern recognition; (c) connecting actual and stored sensor data to patterns, sequences and more complex *models* (see below) representing aspects of the environment considered relevant in relation to goal-values; (d) replacing partial or whole models that did not serve the realization of goal-values (the last two points form the core of 'learning'); (e) using models to make predictions of possible future states and events; (f) starting to make new models in relation to newly set long term or short term goal-values.

The decisive point here is that sensor inputs alone are of no value for the individual. Active decisions are required to use them. In these decisions the sensor inputs are evaluated in relation to *goal-values* (see below), i.e. how valuable the data are to pursue certain goals.

Goal-values: All decision processes we discuss here aim at what we call summarizing 'goal-values'. These are all the objectives, which humans partly have to maintain and partly want to achieve.

We distinguish the following three individual goal-values: We are born with just a few fixed (a) existential goal-values (necessary air, water and food supply; the body temperature) plus basic emotions about what is good or bad in relation to these goal-values. While we grow up we learn external states that serve these existential goal-values, for better or worse. Based on that we make *models* (see below) and predictions, which lead to *decisions for goal-values* (see below). These decisions set (b) long-term goal-values (e.g. learning a profession, participating in a human computation project) and (c) short-term goal-values (e.g. how to be successful now in that profession, or project).

Models: Models are the result of previous *modeling decisions* (see above) about the use of actual sensor data and stored data in relation to certain *goal-values* (existential, long-term or short-term).

Models consist primarily of stored relevant sensor data, which represent previous observations in the form of (a) patterns mapping single external states; (b) sequences of patterns representing external events; (c) interrelated sequences of patterns interconnecting events to whole stories experienced in the past.

Models are secondarily organized as plans and concepts consisting of chains of causes and effect leading towards certain goal-values. So there are large numbers of models standing side by side, representing chains of cause and effect to serve existential goal-values (like eating), long-term (like professional conduct), and short-term goal-values (like the necessary steps of a project). There is generally a hierarchy with hierarchically higher goal-values requiring higher priority of models (e.g. eating has to occasionally interrupt professional conduct, which in turn determines the necessary steps of a project).

Models allow (a) identifying current sensor inputs as corresponding to certain patterns or as being part of a previously observed sequence of patterns; (b) deriving predictions of possible future states and events from known sequences or interrelated sequences of patterns.

Models can be confirmed by repeated observations, by observing predicted states, and by goal-orientated actions leading towards predicted goal-states.

Model based predictions are used for two different types of further decisions, leading to two different kinds of feedback loops—one internal and the other external:

Decisions for goal-values: Predictions may be occasionally used to set long-term or short term goal-values (see above). This is an *internal* feedback loop. Here primarily existential goal-values are applied to make decisions for long-term goal-values (e.g. trying to make a living within a certain profession). Then secondarily these long-term goals are used to make decisions for short-term goals. So decisions for goal-values are primarily related to the existential goal-values and create secondarily a hierarchy of subordinated goal-values, by adding, changing or deleting long-term and short term goals.

This is the most important and least understood process of individual epistemology. It determines the entire further behavior of the individual: The previously set goal-values determine directly what is considered important in *modeling decisions* (see above), i.e. which models are made, and indirectly which predictions become possible and which *decisions for actions* (see below) are made.

Decisions for actions: Normally predictions derived from models are just used to trigger one of the effectors (muscles generally, but mainly arms, hands, legs, feet or mouth) to take an appropriate physical action or to start a communication. This is the usual *external* feedback loop, trying to change the external world in some way towards a goal-value.

So the goal-value (whether existential, long-term, or short-term) currently applied determines which action to choose (e.g. to eat, work or communicate, etc.).

Effector Outputs: Decisions for actions trigger the effectors to cause external effects, either physical actions or communication, i.e. primarily words, addressing other individuals.

In summary, these two feedback loops work as follows: Humans make observations of their environment. Based on that, they make primarily models that serve their existential goals. From learning what serves these needs best they secondarily derive models to serve long-term and short-term wants. The sum of these goal-values for needs and wants determines their *modeling decisions* and their *decisions for actions*, i.e. their entire further individual behavior.

Aspects of Social Epistemology

Now let us apply this controller model of a human to the interaction of two individuals (see Fig. 2). This illustrates the problems of social epistemology:

An interaction starts when individual A acts towards B. B observes these actions and evaluates the corresponding sensor inputs in relation to currently important

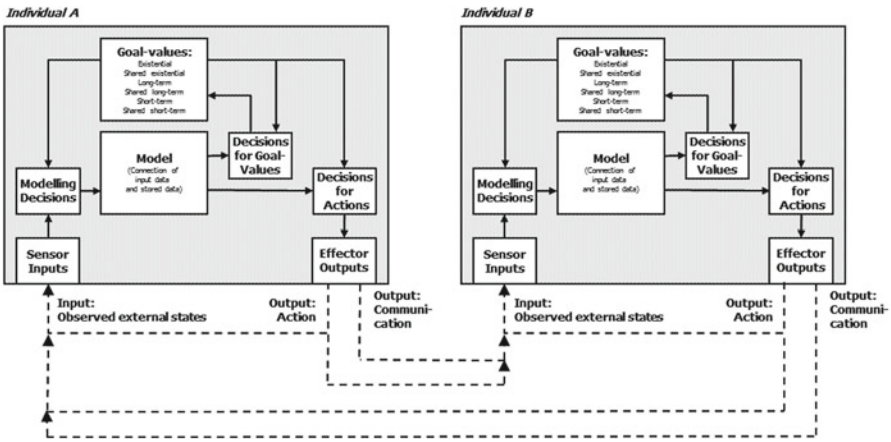

Fig. 2 Aspects of social epistemology: interactions can lead to individual decisions for shared goal-values

goal-values (existential, long-term, short-term). Making a *modeling decision* (see above) B develops a model of A's behavior, predicts its usefulness or danger, and decides for an appropriate action. Then A runs the same process in relation to B's response.

Repetition of this basic exchange may cause at some point in time a *decision for goal-values* (see above) in A and/or in B: Repeated usefulness of A's behavior will cause B to consider A as predictably 'good' or 'interesting'. Then B may decide to add goal-values referring to A to the list of B's already given individual goal-values. Now A may, but need not do the same.

Ideally, of course, this process leads to the development of *shared goal-values* (existential, long-term or short-term), which all interacting parties agree on and add to their individual goal-values. The basic form of this process is realized, of course, in the upbringing of a child. Here the parents serve the needs of the child. So the child will develop shared goal-values with the parents.

Let us mention that the development of shared goal-values may happen spontaneously (e.g. when people face the same problem or threat).

Or this process may be skipped, because interacting people already came independently to shared goal-values (e.g. the same interests or profession).

But mostly, shared goal-values result just from stipulating reciprocally advantageous exchanges of goods, or services, or money and labor, etc.

On the other hand shared goal-values may be propagated by manipulation (A may control the data available to B, using e.g. advertising, censored news, political propaganda, etc.; thus A can limit what may enter into B's models); or they may be enforced to a certain degree (A may have power to control B's access to important resources, like income, etc., or may even be able to apply force; thus A can make B subordinate to and serve his or her goal-values).

The general constraint on developing shared goal-values is the scarcity of goods or societal positions (A and B cannot eat the same bread, or fill the same position in a hierarchy, etc.). Therefore individual goal-values do remain important.

Once parties do share goal-values this will lead to similarities in *modeling decisions* (see above). So they will consider similar data as relevant, will remember and store similar data, and will be interested in making similar models containing certain sequences of cause and effect and enabling particular predictions. Shared goal-values will lead, too, to similar *decisions for actions* (see above).

Shared goal-values will only lead to similar, but not equal, models, as long as the parties rely just on their individual modeling decisions and model making. Only if they cooperate to make externalized mutual verbal concepts, plans, computer programs or mathematical models, they can get to increasingly equal or even unequivocal models.

In summary social epistemology is about human interactions, which make individuals activate their *internal feedback loop for decisions for goal-values*, the process least understood in individual epistemology. The best result is that A and B end up with *individual as well as some shared some goal-values*. And whenever they apply shared goal-values in their current decisions for actions, they will cooperate.

Individuals and Computers: Structures and Interactions

Now let us introduce computers into the relationship of the individuals A and B (see Fig. 3). We characterize computers as controller structures, too, which, of course, differ from humans:

The main differences are: (a) Computers work usually with *fixed* goal-values set by the programmer (we show that in Fig. 3 with the bold arrows directly setting goal-values). So (b) computers lack the *internal feedback loop for making decisions for goal-values*. (In machine learning we occasionally allow computers to make decisions for short-term goals. But we definitely do not want a computer to change its long-term goal-values by itself, so that e.g. a computer programmed to analyze climate data decides on its own to analyze some other data.)

The *external feedback loop* of humans and computers is widely similar: Computers also have sensor inputs (via a keyboard, sensors or data lines). They apply *modeling programs* (matching human *modeling decisions*, but with fixed goal-values) and derive *models* from them (containing here mainly data and mathematical functions, which represent external patterns and sequences), which are used to make predictions. Based on these predictions *programs deciding for actions* are applied (again matching human *decisions for actions* with fixed goal-values).

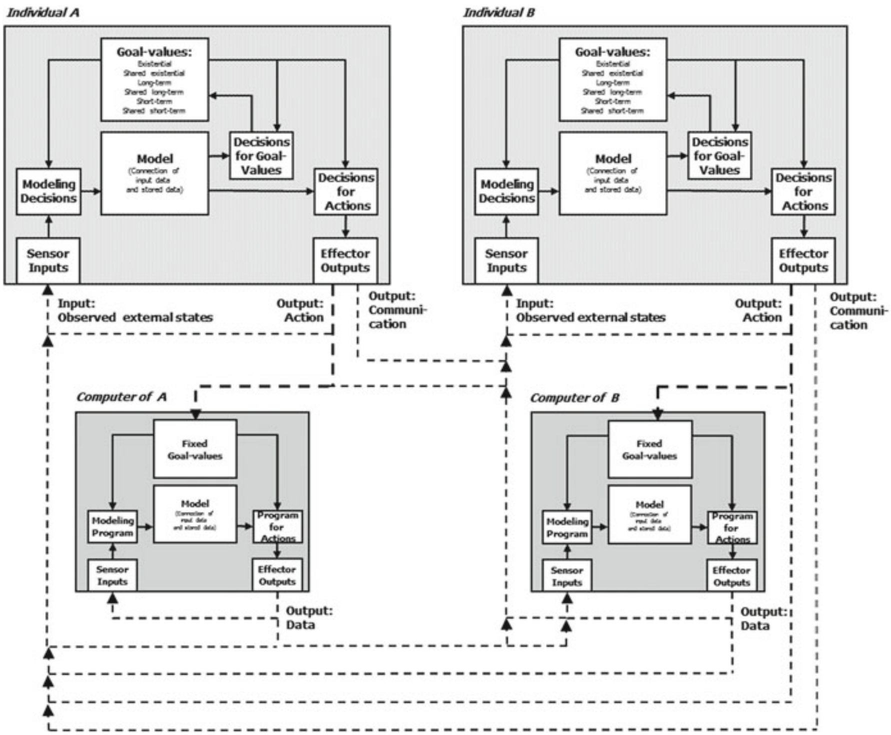

Fig. 3 Individuals and computers: interaction channels added to the context of social epistemology

The *effector outputs* are actions like sending data to other computers, controlling some technical device, or making printouts for human users, etc.

If we take this basic scenario with two individuals and two computers n—times, to match a network, we will get hierarchies (Nechansky 2008) of individuals and computers. We cannot detail that here. We can only assert that this does not change the involved basic epistemological processes.

The Epistemic Processes of Human Computation

Now let us apply all we developed above to human computation:

A human computation project starts with an *initiation phase*, when an initiator defines the long-term goals and short-term tasks. Since human computation is generally applied to problems that require some human contribution, reaching these goals includes tasks that computers cannot yet perform. So the usual advantage of computers we emphasized in section "Individuals and Computers: Structures and Interactions", that they can be directly programmed to work towards a goal, is not available here.

Therefore collaborators have to be sought. The long-term goals and short-term tasks have to be communicated to them. They have to agree on them. And then they have to make *individual decisions for values*, accepting them as *shared goal-values*. We cannot overemphasize that:

1. The success of a human computation projects depends widely on the precise descriptions of long-term goals and short-term tasks, so that the collaborators can understand them and can later make the appropriate individual *modeling decisions* (see above).
2. So the decisive step of a human computation project is a successful finalization of the basic process of social epistemology, as discussed in section "Aspects of Social Epistemology", leading to the acceptance of shared goal-values. Persons focused too much on computation may easily overlook that.

   Dividing a project into subprojects may, but need not weaken that requirement: Now shared goal-values are just needed for the subprojects. But some people might deny contributing, because they do not share the goal-values of the whole project (e.g. a pacifist might deny to contribute to a subproject of a military project).

Once collaborators are found the *computation phase* of the project can start. It may take various forms (see e.g. Quinn and Bederson 2011), which may use any of the possible interconnections between individuals and their computers shown in Fig. 3.

After data acquisition and data distribution, the decisive step is, of course, the human evaluation of the data. Here the short term task of the human contributors is to make *modeling decisions*, judging if data meet the goal-values of the project (e.g. if pictures contain certain patterns, or data sets belong to a certain category, etc.). As emphasized above, the quality of this step depends primarily on clarity of goal-values, i.e. the preceding initiation phase. But it is important, too, that the collaborators do not have any conflicts of interest, i.e. that no other competing long-term and short-term goal-values influence them in their *modeling decisions*. So the success of the project depends to a large degree on the precise consideration and crafting of goal-values the collaborators can fully agree on. More on the importance of goal setting, and its interrelation with motivation and task performance, can be found in Locke's and Lathan's (1990, 2002) classic works on organizational psychology.

After the collection of the results from the collaborators questions of quality control arise. Some evaluation of the results by the initiator must be performed to check if the contributors acted as expected. Since computation is not directly available for obvious reasons, this can only be done indirectly, with approximate use of computers, applying statistics, employing experts, or another round of human computation. Anyway the understanding of the decisive *individual modeling decisions* of the contributors remains vague. So the validity of results obtained using this method remains in question.

However the results may be aggregated in computer models. These models can be confirmed with the usual options of repeated observations, observation of predicted states, and goal-orientated actions leading towards predicted goal-states.

Conclusion

Human computation projects are aiming at goal-values formulated by an initiator. Their critical phase is the initiation, when collaborators have to be found to subscribe to these goals. At that point clearness, specificity, and completeness are the base for the future success. Unfortunately these are difficult to ensure and hardly to measure.

So human computation projects are firmly intertwined in the loops which form the core of individual and social epistemology, i.e. how to interact and communicate with other individuals to make them decide for shared goal-values and cooperate towards them. The results achieved in human computation depend on the success of these processes.

If the results produced by human computation projects contribute to creating further shared goal-values is still another question. That may happen, if these results impress individuals because they show an important relation to their previous goal-values (existential, long-term or short-term), so that they make new decisions for goal-values, set new shared goal-values and start cooperating towards them. Of course, that requires again running through the core processes of individual and social epistemology. We can never escape these loops.

Given the experiences with the precise computer models of climate change, we should not be overly optimistic that human computation projects will lead towards new shared goal-values. All these climatological models predict a threat to the *existential goal-values of all humans*. But not even these threatening results have led to widespread decisions for new goal-values among the endangered people. These decisions always remain individual ones. We can try to influence them, as discussed above, but we cannot directly activate the loops of individual and social epistemology to achieve shared goal-values.

References

Goldman A (1999) Knowledge in a social world. Oxford University Press, Oxford

Internet Encyclopedia of Philosophy (2013) http://www.iep.utm.edu. Accessed 12 Apr 2013

Locke EA, Latham GP (1990) A theory of goal setting and task performance. Prentice Hall, Englewood Cliffs

Locke E, Lathan G (2002) Building a practically useful theory of goal setting and task motivation. Am Psychol 57(9):705–717. doi:10.1037//0003-066X.57.9.705

Nechansky H (2008) Elements of a cybernetic epistemology: decisions, control and principles of societal organization. Kybernetes 37(1):83–93. doi:10.1108/03684920810851005

Nechansky H (2012a) Elements of a cybernetic epistemology: sequence learning systems. Kybernetes 41(1/2):157–176. doi:10.1108/03684921211213007

Nechansky H (2012b) Elements of a cybernetic epistemology: pattern recognition, learning and the base of individual psychology. Kybernetes 41(3/4):444–464. doi:10.1108/03684921211229514

Nechansky H (2013a) Elements of a cybernetic epistemology: elementary anticipatory systems. Kybernetes 42(2):185–206. doi:10.1108/03684921311310567

Nechansky H (2013b) Elements of a cybernetic epistemology: complex anticipatory systems. Kybernetes 42(2):207–225. doi:10.1108/03684921311310576

Quinn AJ, Bederson BB (2011) Human computation: charting the growth of a burgeoning field. In: CHI'11, proceedings of the SIGCHI conference on human factors in computing systems, pp 1403–1412. doi:10.1145/1978942.1979148

Stanford Encyclopedia of Philosophy (2013) http://plato.stanford.edu. Accessed 12 Apr 2013