About the Necessary Move from Cognitics to Ethics; Additional Definitions, and Contributions to Metrics in MCS

Jean-Daniel Dessimoz

HES-SR, West Switzerland University of Applied Sciences (formerly HESSO),
HEIG, Colleges of Engineering and Business Administration (formerly EIVD),
Yverdon-les-Bains, Switzerland, Jean-Daniel.Dessimoz@eivd.ch

Abstract

Progresses in microelectronics and software engineering have since long pushed information processing operations up to abstraction levels well above transmission and coding marks, up to heights better characterized by cognition. Technical activities and automation in this context, along with the associated scientific developments, have lead to the notion and practice of cognitics. Metrics have been introduced, including specific units, in particular for knowledge, expertise or learning.

In an evident manner, metrics show that reality cannot be perceived and known in any degree of completeness, but infinitesimal. So much for the cognitive path.

A fundamental change of paradigm is therefore to be made. Instead of the traditional deductive pattern which makes us hope to forecast and control the future as a result of past and current conditions, we must in priority start by freely projecting our goal(s) in some convenient future; then induction and backtracking dictates our intermediary actions and indeed the proper selection and/or elaboration of (ancillary) models. Ethics is what ultimately characterizes the process of choosing good goals (and conditions).

The paper presents the fundamental definitions and units of our MSC theory, presents illustrative examples, and elaborates on the necessity of addressing ethical issues. In this process new notions are rigorously added to the theory, relating to sapience and wisdom.

Quantitative assessment of cognitive properties will be made, in representative cases of autonomous, mobile robots. Interaction of a robot with humans, or with a group of other, cooperating robots will be discussed.

1. INTRODUCTION

In the history of mankind, tools and machines have progressively been invented, and have to a large extent replaced human contributions in most working tasks relating to physical world (e.g. wrought stones, then hammers, knifes, much later on power, and technological processes). During last century, machine-based artefacts have started replacing humans for non-physical, information related tasks as well, such as for control or complex data processing. The concept of information has been technically defined and made measurable already half a century ago [1].

Progresses in microelectronics and software engineering have since 2 decades pushed information processes up to abstraction levels well above the ones required for usual transmission and coding protocols. At these heights, information processing is better characterized as cognition. Technical activities and automation in this context, along with the associated scientific developments, have lead to the notion and practice of cognitics.

Complexity and intelligence have been addressed in the past. Some of the best-accepted references include [2-5],

Then some metrics have been introduced, including specific units, in particular for knowledge, expertise or learning since the early 90's ([6-7]), and have already been applied to several application areas (in particular [8-10]), along with general concerns such as for concrete domains and robotics (e. g. [11-15]) or knowledge management (e. g. [16]). Evidence of progress made can be seen today on the net (e.g. [17]) and in new companies, curricula or institutions (e.g. [18]).

Now new developments are possible. In evident manner (this point will be developed below), metrics show that reality cannot be perceived and known in any degree of completeness, but infinitesimal. A fundamental change of paradigm is therefore to be made. Instead of the traditional deductive pattern which makes us hope to forecast and control the future as a result of past and current conditions, we must in priority start by freely projecting our goal(s) in some convenient future; and then let induction and backtracking dictate the requires intermediary actions and indeed the proper selection and/or elaboration of (ancillary) models. Ethics is what characterizes the process of choosing good goals (and conditions).

The paper presents in section 2 the fundamental definitions and units of our MSC theory, along with illustrative examples for autonomous, mobile robots. In section 3, we elaborate on the necessity of addressing ethical issues. Therefore new notions are rigorously added to the theory, in part 4, relating to sapience and wisdom. Then quantitative assessments of cognitive properties are made, in representative. Finally, interaction of a robot with humans, or with a group of other, cooperating robots is discussed.

2. FUNDAMENTAL DEFINITIONS AND UNITS OF OUR MSC THEORY

The MCS theory has already rigorously introduced most of the main cognitive concepts, which are briefly reminded in Appendix A.

The two fundamental notions, "Model", and "Information", however have turned out to be very surprising. In principle they had been considered as obvious, classical concepts, on which it would be safe to build right away. Time and experience indeed confirm that those two notions provide a safe and strong basis on which to build MCS. But time and experience have also proven that "Model" and "Information" still require more discussions, and repeated presentations to be widely and deeply understood.

In general terms, a model is typically a very simple representation of reality. By nature, it is very much incomplete. Nevertheless such a partial view may prove useful for many specific applications and contexts.

Information is what allows a receiver to create or update his/her/its model of some specific object or reality. This is totally compatible with Shannon's theory, definitions and units. But in the context of cognitics, we are no longer necessarily limited to technical communications channels, and sources are no longer required to be much considered. What remains and really becomes an essential component is the receiver. In particular, for a cognitive agent (e.g. human, or situated automaton) the sunlight and a tree may combine to become an adequate source of information about time.

Among the key concepts defined and ***made measurable*** in MCS theory, let's briefly comment here the most prominent ones: complexity, knowledge, expertise and intelligence. Complexity characterizes the quantity of information necessary for an exhaustive description. For example an open question calls for more complex answers (e.g. 100 bit) than a tick-box (1 bit). Knowledge is the property of an agent capable of generating correct information ("doing right"). For example a standard TTL NAND gate (2 input, logic levels) has about 2 lin of knowledge. And a contradictor, 1 lin. Expertise is the most important property of a cognitive agent; it denotes the combined ability to elaborate correct information (knowledge) and to do it fast (fluency). For example Dude, one of our 2005 mobile robots, normally can visually detect the position of a bridge, out of 4 possible configurations, within 0.1 s. Dude's level of expertise in this domain is therefore of 30lin/s. In MCS, intelligence is the property of a cognitive agent to increase its expertise through time and experience (learning). By this definition, in particular, situated automata without any kind of memory and real-time modelling have 0 (lin/s)/bit of intelligence.

The reader can refer to Appendix-A for extended definitions.

3. NECESSITY OF ADDRESSING ETHICAL ISSUES

When we directly consider reality, it appears that complexity is infinite, no matter how constrained and focused the domain to assess is. Unfortunately, we are used to restrict our awareness to common models rather than to the reality behind them. For example, the painter Magritte regularly makes the observer surprised when he proposes the sentence « Ceci n'est pas une pipe – this is not a pipe » below his faithful representation … of a pipe. Similarly, in AI, it is common to state that « the map is not the territory ».

Consider for example the trajectory of planet Earth around the Sun. It is hard to convince people that this trajectory is circular, rather than elliptical. Considering the trajectory as circular yields a complexity of 7 bit (at 1% accuracy on the circle). Considering it as elliptical increases the complexity to 14 bit (at 1% accuracy on the ellipse). In fact, real world observation shows that the peak-to-peak difference between best-fitting circle and ellipse is of about 4%. And the peak-to-peak difference between real Earth trajectory and best-fitting ellipse is of 11%! Due to Moon effect. We start now (?) to realize that, far from being a simple ellipse, in reality the actual trajectory of planet Earth around the Sun is a strange distorted curve which cannot be fully described, no matter how many bits of information we are ready to throw into that description… complexity is infinite.

When one considers reality directly, one is usually overwhelmed by information. Therefore the first step to do in cognition and cognitics is to elaborate (or select) a (predefined) model. The choice of an appropriate model (in terms of type and complexity) is not dictated the intrinsic complexity of any reality domain to be represented, but rather by the goal one has elected to aim at. Referring again to the above example, what is critical is the selection of goal. If one wishes to explain how seasons pass, on Earth, modelling a trajectory as circular will be the best approach: it is the least complex and yet sufficient. If one is concerned by our distance to the next galaxy, the model could even be simpler: Earth and Sun are at the same location.

Even though reality is factual and therefore seems to prevent any choice, what really matters in cognition and cognitics is the goal one defines and aims at. From some specific requirements relating to that goal, it may then become possible to elaborate an appropriate model, of reasonable complexity and to work out a successful plan or sequence of actions.

In conclusion, the overall cognitive process is bound to start by an adequate choice of goals. In the long term if an agent wants to survive, it must make those choices in compliance with the requirements of its environment. By definition, ethics is what ultimately dictates his/her/its choices.

4. THEORETICAL DEFINITIONS FOR SAPIENCE AND WISDOM.

The current paper adds to MCS model two rigorous definitions, one for sapience and the other one for wisdom. Before addressing these notions however, it is useful to distinguish clearly between three different concepts closely related to each other: right, true, and good.

4.1 SAPIENCE

In other times and elsewhere, various views have been communicated about sapience ([19] could be a good starting point for readers wishing to explore some of such typical, past or current views). Yet, as far as the author knows, no one has ever proposed a way to estimate it quantitatively or to measure it.

Here and now we define as part of MCS theory that sapience is the essential property of a cognitive agent, i.e. of an active structure capable of cognition. Sapience can appear under a number of signs, such as knowledge, expertise, or intelligence, which for all the main ones have already been defined and made measurable in MCS. If one wishes to quantify sapience, a worthwhile approach is to assign it an index, rating an agent performance by reference to the one of typical humans ("homo sapiens”). A machine passing the test of Turing could be assigned the value of 1. If an agent is ten times more expert than typical humans, in a given domain, it should be characterized by a sapience value of 10 in this domain. Being a ratio, sapience has no specific unit in MCS.

4.2 RIGHT, TRUE OR GOOD; AND THEIR COMPLEMENTS

“Right” is usually considered as a logic, Boolean value, complementary to “wrong”. Let us define “right” as the qualifier of an entity that complies with a given law (assertion). For example if the law is “to move ahead”, a step forward is “right”.

In MCS, true and good can readily be defined on the basis of “right”: “True” is “right”, when the law to comply with is “correspondence to reality”. For example it is true that braking reduces speed. “Good” is “right” when the law to comply with is “to progress towards a chosen goal”. For example, if a robot is required to move, it is good for it to switch on some power circuits.

False and bad are the respective complementary qualities of true and good.

4.3 WISDOM

In MCS, we define wisdom as a specific property of cognitive agents, which refers to their ability to take good decisions, i.e. to be expert in delivering the messages that make an agent reach a given goal.

In 2005 Eurobot competition about bowling, for example, it is “wise” for the robot Dude to start moving, when the referee gives his “go”, rather than halting. Similarly, it is wise to cross a bridge towards skittles to hit, rather than to rush into the holes in the middle of the playground.

To make it simple and easy, we propose here to estimate in Boolean terms the quantity of wisdom for an agent, on a given domain: true or false, reflecting the fact that the goal is being reached or not by the agent.

Without being essential, a usual feature of wisdom is to relate to complex situations and major or “meta”-goals: to survive, to win the game, to gain a place in the Hall of Fame. Or, in agreement with “The Stranger” in “The Big Lebowski” movie [20]: … to eat the bar, rather than being eaten by the bar.

Reference [21] could be a useful starting point for readers wishing to explore some other, typical, past or current views on wisdom.

5. INTERACTION OF A ROBOT WITH HUMANS, OR WITH A GROUP OF OTHER, COOPERATING ROBOTS

Trends in robotic developments as well as progress in MCS theory bring us to the point of considering collective entities. We shall discuss this new situation in three contexts: collective cognitive agents, groups of cooperating robots, and robot-human interactions.

5.1 COLLECTIVE COGNITIVE AGENTS

In principle, considering a collective cognitive agent rather than an individual one does not change anything in terms of definitions and metric equations for the MCS theory. The behavioural model adopted in MCS, can be applied in sub-systems as well as on the global level of an entity. What may be new is that at a “meta-level” a new entity is formed, which can be referred to as a “group” (e.g. team, society, nation, and mankind).

If we distinguish each of the members of a group, it becomes in general necessary to assume that some kind of communication exist between members. Classically, such a collective entity has been viewed under three different angles: the whole, i.e., as mentioned, the “group”; the typical individual member of group, e.g. the “lambda” member; and some set of intangible underlying factors which ensure the coordination of individuals so as to achieve a specific collective identity and behaviour i.e.

the “spirit” or culture. The “spirit” consists in a system of common, shared references, values and objectives, which may dynamically evolve, and yet does not exist per se, i.e. out of the members.

As an example, consider an orchestra playing without conductor: the group is the orchestra, a lambda member is a musician, and the “spirit” is what makes it possible for the musicians to play together in a coherent way, even when there is no additional conductor nor outside regulating factor.

5.2 INTERACTION BETWEEN MUTIPLE, COOPERATING ROBOTS

Robots become more common and some of them start to be able to cooperate with each other, i.e. to behave and act as a (self-) coordinated group.

Numerous examples already exist, where robots exchange information according to predefined schemes, with more or less adaptation to environment. For example, actions can be triggered by light signals (e.g. Alf 2001, with its 5 autonomous mobile sub-robots); or even between opponents in Eurobot competitions, some cooperation is achieved between robots in order to avoid mutual collisions (e.g. Diego[3] ’98, uses vision for this purpose; Lomu ’04 uses ultrasound sensors).

It seems that much development is still to be achieved before groups of robots are able to collectively elaborate and dynamically adapt complex behaviours, i.e. to demonstrate instances of significant robotic cultures or “spirits” (Even though like fractals or systemics tend to demonstrate, what globally looks like inherently complex constructions may often result from a rather straightforward sequence and replication of trivial basic behaviours.)

5.3 INTERACTION BETWEEN ROBOTS AND HUMANS

The case of robots interacting with humans is discussed below in three points. While some science-fiction authors already propose guidelines for some hypothetical future worlds, we shall consider below more pragmatically the current state and the challenges of close future.

I. Asimov is well known for having expressed his three laws for robots meant to interact with humans, and various authors have suggested recent extensions [22]. They all converge to require robots to be helpful and non-threatening.

Today already, safety concerns are dominant. And robots do not really differ from other machines in the way they are treated. In particular, robots are mostly present in industrial factories; they are typically fixed on the floor and are therefore easy to confine in areas forbidden to humans. For mobile vehicles, safe solutions include separate travel tracks, or – consider especially planes- preventive maintenance, monitoring by humans, and extensive control and checking procedures. In the case of Eurobot, the usual restrictions for air pressure and voltage levels apply; pyrotechnic and chemical products are forbidden.

For interaction, nowadays, the best practices with robots reflect what is considered good in associated domains, such as micro technology devices, software engineering, or generally machines and commercial products: clear man-machine interface (e.g. Eurobot rules for “remote” start by players, and emergency power stop by referee); design for ease of test and maintainability, extensive and efficient visual messages and graphic displays, restrictions imposed by laws, rules and standards. When starting in Eurobot competitions, we found critical to design a new, specific language (Piaget), which was later adapted to include a maximal compatibility with the excellent VAL language for (industrial) robots (re. Unimation and Stäubli/Adept). The multi-agent Piaget language and operating environment makes it simple and efficient to communicate to robots the tasks to be done and the strategies to follow.

In future, we can imagine many applications of robots cooperating with humans. Many of them will gain to be humanoids, for at least two major reasons: 1. it is more convenient for layman users to communicate with an agent similar to himself/herself; 2. it would be most beneficial to be able to substitute case by case man and robot, and to be able to share common resources: access to shelves and fridge for food loading and unloading; capability to walk upstairs and downstairs, to walk through garden, to drive a car, etc.

An interesting case is the one of pet robots. Technically, it is surely easier to design animats than humanoids. It remains however to be explored to what extent humans will care for such non-biological artefacts, especially if those systems do not bring some services of practical utility, such as, notably, transportation, monitoring or guidance. Nevertheless new possibilities will emerge, such as novel playing abilities or network-based behaviour.

6. CONCLUSION

Technical definitions and associated metrics for "information" have been introduced half a century ago, in order to tackle telecommunication issues. It has been proposed to extend the domain of applicability of this concept to cognition (re. cognitics). In this effort, the very old notion of modelling must be revisited, as well as the one of information: modelling is meant to be primarily simple and good, rather than complete and true; here information mostly relates to receiver's capabilities.

In cognitics, a particular theory as been built, MCS; metrics have been introduced, including specific units, in particular for the quantitative assessment of complexity, knowledge, expertise or learning. One of the results is the following: in an evident manner, reality cannot be perceived and known in any degree of completeness, but infinitesimal.

A fundamental change of paradigm is therefore to be made. Instead of the traditional deductive pattern which makes an agent hope to forecast and control the future as a result of past and current conditions in reality, the agent must in priority start by freely projecting his/her/its goal(s) in some convenient future; then induction and backtracking dictates his/her/its intermediary actions and indeed the proper selection and/or elaboration of (ancillary) models. Ethics is what characterizes the process of choosing good goals (and conditions).

The paper rigorously adds new definitions and units of our MSC theory: sapience denotes the cognitive capabilities of agents, in quantities expressed as ratios to corresponding, typical, human performance levels. Distinction is made between "right" (conforming to some given law), "true" (conforming to reality), and "good" (leading to corresponding goal) qualifiers. Wisdom characterizes an agent capable of reaching (most significant) goals.

The paper finishes with a discussion of robots cooperating with each other, or with humans. Humanoids are found especially useful in the future for best communication possibilities with humans, and for best possibilities to assist them with shared resources, while cooperating.

For future research, it might be worth to introduce, for wisdom, some quantitative index finer than the one proposed, which consists in only two Boolean values; it might for example be a function of (inverse) distance between goal and reached states in model space.

APPENDIX – SUMMARY OF MCS ESSENTIAL DEFINITIONS

The appendix briefly presents, by alphabetical order, the core definitions in MCS - our model for cognitive sciences. This can validly be viewed as a glossary, ontology or an axiomatic declaration. The number behind each concept denotes the logical order in which definitions are introduced (re. above the novel definitions, for sapience, wisdom, true, right, and good).

Abstraction (3b). Property of a system that generates less information than it receives. The abstraction index, iabs, is the ratio of incoming information quantity (ni [bit]). over the out coming information quantity (no [bit]). Inverse of concretization. Equ.: iabs=ni/no [without unit]

Complexity (3a). Property of a model that requires a lot of information in order to be exhaustively described. Quantitatively, complexity is the amount of required information. Unit: same as information, i.e. [bit]

Concretization (3c): Property of a system that generates more information than it receives. The concretization index, ic, is the ratio of out coming information quantity (no [bit]) over the incoming information quantity (ni [bit]). Inverse of abstraction. Equ.: ic=no/n1 [without unit]

Experience: (4b) Property of a system that has been exposed to a cognitive domain. Quantitatively, it is usually evaluated in terms of time (duration) [s]. An alternate (better?) view is to assess experience, R, in terms of number Na of witnessed input-output associations. Equ.: R=Na*(ni+no) [bit]

Expertise (5a). Property of a cognitive system which delivers fast the pertinent output. Quantitatively, it is the product of knowledge, K, and fluency, f. Equ.: E=K*f . The unit is [lin/s]. In general terms, synonyms for expertise include know-how, skill, competence and excellence.

Fluency(4c): Property of a system which delivers information fast. It can be viewed as a processing speed. Fluency, f, is the inverse of the time duration , Δt, necessary to deliver output information. Equ. : f=1/Δt [1/s]

Information (2). Information is what allows a receiver to update his model. Quantitatively, it is the difference of model size in terms of information content, between the states “before” and “after” message arrival. Computation is made on the basis of message probabilities, which are essential elements in the model Consider that the incoming message is one among N possible variants. If the probabilities of those various occurrences of the message are pi, where pi is the probability of the ith message, then the average quantity Qa is given by the following equation: Qa:= Sum $_{\text{for } i:=1 \text{ to } N}$ of (pi log 1/pi). The log is usually taken in base 2, thereby yielding [bit].

Intelligence (7). Intelligence is the property of a system capable of learning. In quantitative terms, intelligence can be assessed as an index, i, which is the ratio of learning with respect to experience. Depending on the intuitive or more rigorous choice of formulations introduced for experience, we have two equations. Equ.: i=L/Δt [lin/s^2] (or i=L/ΔR [lin/s/bit])

Knowledge (4a). Knowledge is the property of a system which delivers the pertinent output, either proactively or in response to incoming messages. Quantitatively it is given by the following equation: K=log(no*2^{ni}+1). The log is in base 2, and the unit is the [lin].

Learning (6). Learning is the property of a system capable of increasing its expertise level as time goes (or better: as experience goes). Equ: L=E(t1)-E(t0) . Alternate view: L=E(r1)-E(r0). In both cases the unit is [lin/s]

Model (1). In general terms, a model is a simplified (that is, incomplete by essence) representation of reality, which is found useful in order to reach some specific goal. In MCS the basic reference model is behavioral. It can be viewed as a kind of (virtual) table, which contains as many states as possible incoming message types; each state contains the instant probability of occurrence for the corresponding input message, and also contains the corresponding output message. The goal of this model is to allow the quantitative assessment of key cognitive properties, such as knowledge, expertise, or learning.

Reductibility (5b): Property of a system which can be implemented by subsystems of integral complexity smaller than the complexity of the system itself.

Simplicity (4d): Property of a model which requires little information in order to be exhaustively described. Quantitatively, simplicity, exactly like complexity is the amount of required information. Unit: inverse of information unit, [1/bit].

REFERENCES

[1] Claude Elwood Shannon, coll. papers, ed. by N.J.A. Sloane, A. D. Wyner, Piscataway, NJ, IEEE Press, 1993

[2] G.J.Chaitin,Scientific American 232(1975)47-52.

[3] A. N. Kolmogorov, Information theory and the theory of algorithms , ed. by A. N. Shiryayev, transl. from the Russian by A. B. Sossinsky Dordrecht (etc.) (Kluwer Academic Publishers, 1993)

[4] D.Michie, On Machine Intelligence, 2nd Ed., (Ellis Horwood Ltd, Chichester, W Sussex, England, 1986).

[5] S. J. Rosenschein, Formal Theories of Knowledge in AI and Robotics, New Generation Computing, Ohmsha Ltd and Springer Verlag, 3 (4)(1985)345-357.

[6] J.-D. Dessimoz, "Knowledge in formulas", SGAICO Newsletter, SI Information, Soc. Suisse des Informaticiens, Zürich, Aug. 1991.

[7] J.-D. Dessimoz and Giovanni Mele, “Performance assessment of cognitive systems; case of elementary mobile robots”, Proc. ECAI 94, 11th European Conference on Artificial Intelligence, Amsterdam, 7-12 Aug., A. Cohn. ed., John Wiley & Sons, New York, pp. 689-693, 1994

[8] J.-D. Dessimoz, "Is a Robot that can Autonomously Play Soccer Intelligent?", Proc. IAS-6, Intern. Conf. on Intelligent Autonomous Systems, co-sponsored by IEEE, Venice, Italy, July 2000, pp. 8
[9] J.-D. Dessimoz, "Beyond information era : cognition and cognitics for managing complexity; the case of "enterprise", from a holistic perspective", Proc. ICIMS-NOE (Intelligent Control and Integrated Manufacturing Systems- Network of Excellence), ASI - 2000 (Advanced Summer Institute) and Annual Conference, CNRS-ENSERB-Université de Bordeaux, Bordeaux, France , 18-20 Sept. 2000, pp.164-170
[10] J.-D. Dessimoz, « Knowledge Management and Cognitics; Some Fundamental Aspects", in Proceedings « Workshop IPLnet 2002 ; From Research to Application », IPLnet National network of excellence in Integral Production Engineering and Logistics, c/o ZPA, FHA, Northwestern University of Applied Sciences, Switzerland, Febr 2003.
[11] http://fab.media.mit.edu/fab/
[12] http://Eurobot.org
[13] http://Robots05.ch
[14] http://Robocup.org
[15] http://Robot-CH.org
[16] Knowledge management case book : Siemens best practices / ed. by Thomas H. Davenport ... [et al.] 2nd ed. Erlangen : Publicis Corporate Publishing, 2002 pp. 336, ISBN 3-89578-181-9
[17] http://cognitics.org
[18] http://www.scico.u-bordeaux2.fr/index.html Institut de Cognitique
[19] http://en.wikipedia.org/wiki/Sapience
[20] The Big Lebowski, Movie by Ethan Coen, Joel Coen (1998) (The Stranger: One a those days, huh. Wal, a wiser fella than m'self once said, sometimes you eat the bar and sometimes the bar, wal, he eats you.)
[21] http://www.cop.com/info/wisdompg.html
[22] http://perso.wwanadoo.fr/monot.jc/index.htm#/monot.jc/themes/robotici.htm