The measurement of Artificial Intelligence – An IQ for machines?

JOVELINO FALQUETO[1], WALTER C. LIMA[2], PAULO S. S. BORGES[1], JORGE M. BARRETO[1]
[1]Departamento de Informática e Estatística, Universidade Federal de Santa Catarina,
[2]Universidade do Estado de Santa Catarina,
Florianópolis, SC, Brazil
falqueto@inf.ufsc.br, d2wcl@pobox.udesc.br, pssb@inf.ufsc.br, barreto@inf.ufsc.br

Abstract

Once talking about the possibilities of creating 'thinking machines', Turing properly said that this task should start by defining of what is understood as 'machines' and of what is 'think'. Regarding the pursuit of modeling intelligence, two large avenues were opened by researchers in almost the same epoch: Symbolic Artificial Intelligence – SAI – and Connectionist Artificial Intelligence – CAI – based respectively on symbols and rules and in artificial neurons. It seems that the time is come to start thinking (and acting) to establish a standard of comparison, that could objectively tell how far we have gone along the road of constructing ever better AI systems. Devising an Intelligence Quotient IQ – for machines or any intelligent system would be, perhaps, an advancement but unfortunately, the history of the development of techniques to measure human IQ, the first source checked to find applications to AI, points to a very fuzzy zone. Admitting that possibility, we present some conjectures. For example, introducing some metric to evaluate the redundancy of the rules of an Expert System, or the efficiency of a given topology in a Neural Net could bring new insights on ranking AI paradigms and indicate which are the most promising ones.

Keywords: Modeling, machine intelligence, symbolic AI, connectionist AI.

1. Introduction

Since man got consciousness he has been seeking the power to control life and death. He wants to imitate God and struggles hard to create beings similar to himself. But many philosophical and practical difficulties have appeared. The great obstacles concerning the problems of reconstructing the human body are now apparently being progressively worked out by important advances in biological and chemical sciences, aided by modern technologies. This view is easily corroborated by the daily news in the popular media, which routinely announces new achievements, e.g. the increasing progress in the description of human genome, the ability of cloning animals and artificially generating fragments of natural tissues, among many other similar feats.

Another field where modern scientific research has allocated a large amount of resources is the construction of intelligent machines. Should this deed be accomplished along with the correlated problem of creating a human body, two ancient problems would be elegantly solved and it could be said that these newly created beings would have AI – Artificial Intelligence. Until now, we cannot even precise what "natural intelligence" is, which means that we are still uncertain if the attribute of *intelligence* occurs only in humans or, according to some authors, also in animals [1]. But one straightforward objective of AI, the "study and pursuit of mental faculties to be implemented with the use of computers" [2] seems clear to anyone involved in this broad area of Computer Science. Then, in a first stage, we endeavor to find models that may be abble to mirror what we expect by 'intelligence' and as a second step we try to use these models in computers systems to solve problems.

Turing[3], in his considerations about the possibilities of building thinking machines properly said that this task should start by the definitions of what could be accepted as 'machines' and of what is 'think'. In connection with the first concept, Turing himself supplies an answer: "This special property of digital computers, that they can mimic any discrete state machine, is described by saying that they are universal machines". In this way they are able to execute any computable process, including the simulation of analog computers. But, even if the strong Turing assumption is true, the other side, that is, the concept of "thinking" appears far more complicated. Accordingly, Turing, in an scheme largely known as having the inspiration of a genius, proposed an escape solution. It is his famous test that, in all, admits as intelligent a machine that could act "as intelligently as" a human being facing the same circumstances. In fact, we are confronting here with consequences of the mind-body problem, and "a number of philosophers consider it as the most difficult of all human problems, that is, the relation between our minds and the universe ... and its modern version generally has the form of the question: *how does our mind relate to our brain?*" [4].

A set of different strategies are adopted to shed some light on this issue, and Psychology has knowledge areas — like Psychometrics — that, by supposing that the mental abilities are measurable (a matter not universally settled), works with tests to quantify them. From those studies come some insights regarding possible measures of AI, which will be discussed in the sequence. Cognitive

Psychology represents another approach and deals with the processes by mean of which the mind hypothetically functions. From this approach also derives another, concerned with the inter-influences of the environment on the intelligence. Still another conception looks the problem of studying intelligence by means of the biological aspects supposedly involved. This direction is mainly composed of three areas: one interested in specific regions of the brain responsible for individualized characteristics of intelligence – for example, language ability is mainly situated in the region of the temporal lobe of the brain [5]. The two others refer to the electromagnetic waves originated from the brain activities and to the blood flow involved.

In all we can only state that intelligence is little known and the whole field remains a large area of speculations. A lengthy debate has also been taking place on the conjecture of the existence of multiple types of intelligence, and H. Gardner [6] is frequently cited by the specialized texts suggesting at least seven types, among them linguistics, music, mathematics, spacial ability.

2. IQ and others Q's

The intelligence quotient – IQ – was proposed in 1905 by the French Psychologists Alfred Binet and his assistant Theodore Simon, with the objective of measuring the "beautiful pure intelligence", that is, without intervention of external factors [7]. Their proposition was a consequence of a government commission assigned to study forms of ensuring adequate education to mentally handicapped children. Binet observed that these children worked out problems in a similar manner used by younger "normal" children. Therefore, he tested the possibility that the intelligence level might be directly related to age. The two scientists experimentally designed lists of questions for each age and abandoned those that had been wrongly answered by more than 25% of the "normal children". With this method they managed to build a set of questions well adapted to each age, every one expected to mirror the average knowledge dominated by the specific age stages. The supposition is that if one child properly answers the questions belonging to the eight year old set and fails with the nine year old set, then that child has a mental age of eight years. The quotient of the age thus obtained and the chronological age gives the decimal IQ of that child. In a percent form:

$IQ = 100 \times (\text{Mental age} / \text{Chronological age})$

The experience has demonstrated a statistical distribution of the measured IQ near the normal curve and also that two-thirds of the people – known as "normal persons" – have this index between the extremes of $IQ = 85$ and $IQ = 115$. Unfortunately, persons soon were labeled as "normal", "geniuses", "idiots" and so on, having as basis only the weak evidence provided by their IQ scores. This index was initially enthusiastically adopted by different psychology streams, but soon reticent or totally discordant voices were heard, mainly pointing out the strong dependency on the kind of education of the individuals being measured, their social and familiar status, etc. Then, the IQ index could, at most, indicate the minute performance, similar to a static snapshot, but would not stand for a good indication referring to a long period of life or to the future capacity of the individual, consequently entailing its limited use. Also, being obtained under artificial conditions, this index could point out both the "beautiful pure intelligence" and the knowledge acquired in specific situations experimented by the individual in his interaction with the environment to which he casually belonged. The Russian Psychologist L. S. Vygotsky [8] separates two very distinct sets: one called "really developed functions" and the other "potentially developed functions", corresponding to abilities already dominated and others in a dormant stage, respectively. Clearly the IQ as defined by Binet gives only a socially and culturally biased measure of Vygotsky's "developed abilities" and a small hint, if any, to the "potential abilities".

3. IA Paradigms

Knowing the controversies about the difficulties with "natural" IQ, one should be quite cautious in proposing a similar measure to machines and systems. Additionally, as very superficially reviewed, the history of development of techniques to measure human IQ, certainly the first source to be examined to check the real possibilities of finding applications to AI, points to a very fuzzy zone. But we can admit the possibility of that challenge and offer some conjectures.

In their efforts to create intelligence, or at least to model it, two large avenues were opened by researches and they took place at almost the same epoch. One, now known as Symbolic Artificial Intelligence – SAI – tried to explain and simultaneously model intelligence with basis on the Physical Symbol System hypothesis [9]. Those who adopt this statement as a corner stone work with the theoretic assumption that to wit the emergence of intelligence nothing more is needed than two sets of objects: one with suitable symbols and another formed with rules dictating how these symbols must be manipulated. There is also a group that advocates a more extreme position: symbols and rules are necessary and also sufficient, that is, they must be present as the basis of any intelligent system, and nothing else is necessary to convey intelligence.

On the other side of the AI land – some would call it a quagmire – stand those who support the thesis that intelligence emerges from the neurons. They set up their belief on the following hypothesis: "If a sufficiently accurate model of the brain is built, then this model will show intelligent behavior. If only one a small part of the brain is fabricated, then the brain function exerted by that part will appear in the model" [10]. Accordingly, an extensive field of study has been developed – Artificial Neural Nets and the whole area is denominated Connectionist Artificial Intelligence – CAI.

A broad research area, globally called Evolutionary Computation (EC) [11], also deals with development of methods to solve difficult problems, thus having a strong relation with the process of simulating intelligence. Observing the different manners that Nature has employed to create, sustain and develop life — an amazing task — the field of Evolutionary Computation uses the inspiration in those natural processes to design algorithms to be executed in a computer that may be able solve complex problems, where neither analytic nor heuristic methods exist or are efficient, and where no data is available. Being aware that the theory of evolution of species enunciated by Charles Darwin in 1859 [12] is not "proven" in the mathematical sense, the known evolutionary factors: recombination, selection and mutation are used, guided by random laws, to emulate Nature's evolutionary process.

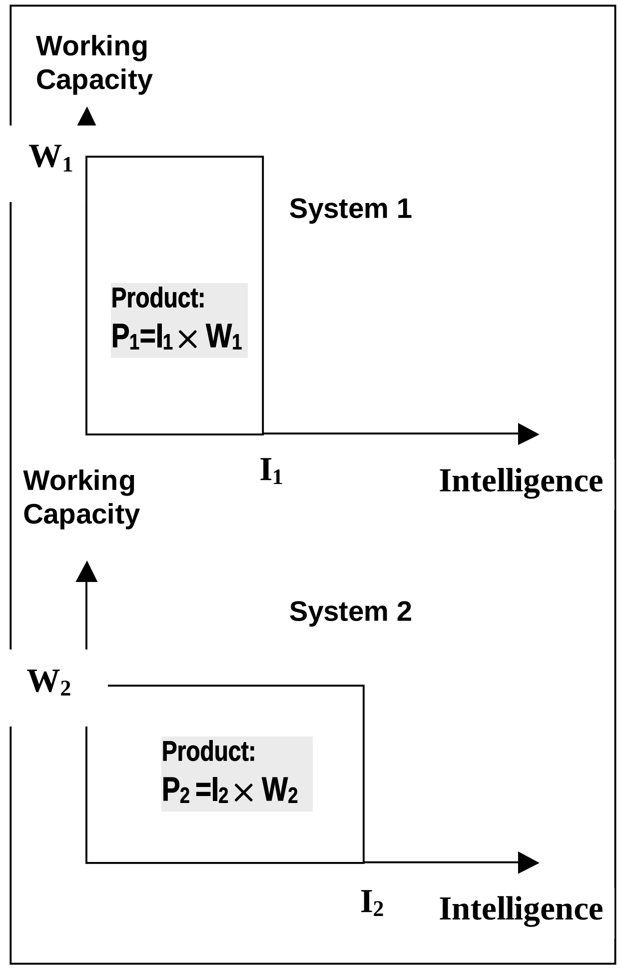
Although the method cannot ensure that a global optimum solution to the problem under experimentation will eventually be found, at least near-optimum answers can be provided. EC is suitable to be applied to complex problems, where the term 'complex' does not refer to the difficulty in understanding it, but to the the vast solution space to be searched with little or no clues that can be used to direct the search. This characteristic makes the brute force of the most powerful computer useless to screen the whole space. Satisfability problems are an example where this kind of complexity arises.

Moreover, EC relies on a very large amount of calculations, thereby being a technique that owes its feasability to the existence of computers. No human, no matter how 'intelligent' he or she is, would be capable of either solving such complex problems or performing, in a lifetime, the operations required by EC methods. How come, then, that a "dumb machine" — a computer — can perform so nicely towards the attainment of an objective, thereby defying human abilities to do the same? If we scape from the definition of intelligence as a pure and abstract concept and concentrate in seeing it as the power to solve problems, then we must take into account another factor: the "working capacity".

We suggest that intelligence, to be effective in reaching the goals that demand its application, must be put "in motion". In other words, we believe that it is of little use to count on a 'dormant' or 'lazy' great intelligence to solve our problems. Maybe we would be better off with a not-so-brilliant but very industrious intelligence. When we see a computer performing complicated calculations, though knowing it, we are hardly aware that the elementary atomic operations underneath that job are quite simple, each of them requiring little intelligence to be understood. What is at play here is the incredible speed with which such elementary operations are being carried out, which, altogether, may provide, in a relatively short time, a fine solution to a problem, often unreachable by even the most intelligent human.

The figure below illustrates this point. Although System 1 is less intelligent than System 2 ($I_1 < I_2$), it has a greater "working capacity" ($W_1 > W_2$), thereby yielding the same final product as System 2 ($P_1 = P_2$).

4. A methodology for IA measurement: Some ideas



The availability of a simple and reliable instrument to measure AI would be of great value since it would furnish an important comparative instrument to evaluate intelligent software and hardware systems, or in a more general expression, intelligent machines. The development of an index that yields the intelligence degree of such a system would permit, for instance, to make decisions about which are more adequate, efficient, or cheaper among a number of machines under appraisal.

Investigation on how to determine this index might bring new ideas and new directions to the development of applications in the AI area and others. But the examples regarding the Psychology field offer no reason to optimism, even showing it is not a simple task. With the difficulties already pointed out for the "human IQ", speaking about an IQ for machines should be thought in a loose sense, having in mind that attributing intelligence to machines is still a high controversial issue. Is that measurement possible? Denying it beforehand is embarrassing, on account of the small amount of knowledge currently available on the theme. Moreover, a formal, classic inductive demonstration that a method of measurement is appropriate is not easy, if even possible. Perhaps one possibility would be to develop not just one, but several instruments, combining them accordingly in regard to the different AI paradigms.

5. Algebraic considerations

It seems difficult in the present incipient stage of research in intelligence to define and rank the set of IQ measures of machine intelligence. Hence, this would only be possible if all the machines were working with the same problem and the same AI methodology and the measures would necessarily reflect that methodology. We could have, maybe, a complete ranking for a particular methodology, but not for all sets of IQ measures of diverse intelligent machines operating under other methodologies. In fact, the materialization of intelligence employing a certain paradigm may be rather different when using another, even if both materializations are aimed to the solution of the same problem. The set of those IQ measures would have scientific and practical interest only if their elements could be quantitatively compared to each other. These conditions seem too simple; yet one should be aware that extant systems admitted as intelligent and solving the same problem, are not easily compared [13] [14]. It appears that the same flaws that affect the measurement of "human IQ" are also present when we deal with the "machine IQ". In other words, as we do not know exactly what is to be measured, other practical predicaments arise, as for example, how to measure, with what, where, when, etc.

A problem can be modeled as the interaction of three sets: *I*, *T* and *O*, where *I* indicates the input elements, *O* the output elements and *T* a transformation that describes how *O* can be obtained from *I*. In the case of measuring the intelligence of machines, the definitions of the variables constitute the set *I*. The yielded measures form the set *O* and *T* would be the measurement methodology function — the mapping $I \rightarrow O$. As mentioned before, we generally stumble in big problems when defining *what is* an intelligent machine and still bigger in regard to the determination of an adequate IQ. Consequently, finding *T* is not a simple task. We are looking for a function that generates IQ measures having as input the set of intelligent machines and as output one element of the *O* set with the *order relation property*. It would be very nice if the set of measures *O* had the *complete order property* that is defined below [15].

Given the arbitrary sets *E* and *F*, a subset of the Cartesian product $E \times F$ is defined as a relation *R*. In the particular case where $E = F$, it is said that "*R* is a relation in *E*". In the case of the existence of a set *M* of intelligent machines and assuming that the IQ of each machine of this set has been obtained, then we have the set *O* of all intelligence measures for those machines. *O* is said to have complete order if its elements obey the axioms:

i. *Reflexive property:* $\forall i \in O,\ i\ R\ i$
ii. *Antisymmetric property:* $\forall\ i, j \in O$, *if* $i\ R\ j$ *and* $j\ R\ i$, *then* $i = j$
iii. *Transitive property:* $\forall i,j,k \in O$, *if* $i\ R\ j$ *and* $j\ R\ k$, *then* $i\ R\ k$
iv. *Comparability property:* $\forall\ i, j \in O,\ i\ R\ j$ *or* $j\ R\ i$

A special case is when *R* is the regular common order, generally denoted by the ≤ sign. In this case the relation *R* is said to be of "total order", or alternatively, the set *O* is of "total order". If only the first three axioms are verified it is said that *R* has "partial order" because not all elements of *O* can be compared using the relation *R*.

6. An IQ for "symbolic intelligent machines"

One of the best-known approaches of the SAI paradigm is the construction of Expert Systems based on Post's production rules [16]. Suppose that one of such a system is made with *n* rules, but *x* of these are not necessary, because of redundancy: it could be said, without mathematical rigor, that only (*n*-*x*) rules would be "linearly independent". We could also admit that among the *x* rules some do not convey real "intelligence" to the system. With basis on these assumptions, the quotient $IQs = 100\ (n\text{-}x)/n$ could be a measure of the amount of intelligence pertaining to the system. An actual case of the proposed IQ measurement would be an Expert System to deal with, say, the genealogical tree of a person (or to construct family relationships of a person). Suppose that we have built the set of Post's rules in PROLOG code, including the five lines shown below:

(i) *father (joseph, jacob);*
(ii) *grandfather (joseph, isaac);*
(iii) *father (jacob, isaac);*
(iv) *grandfather (jacob, abraham);*
(v) *father (isaac, abraham);*

It can be clearly seen that the rules (ii) and (iv), for instance, convey no additional information to the system, since the statements "*jacob is father of joseph*" and "*isaac is father of jacob*" logically leads to the consequence that "*isaac is grandfather of joseph*". Consequently, the rules (ii) and (iv) could be eliminated from the set; they add no more "intelligence" to the set, they are "*linearly dependent*" of others already embodying the system. In the CAI case one could admit that a neural net with *n* neurons solves a problem. But with the utilization of another topology, another type of neurons, such as dynamic ones, it would be enough if only *n*-*x* neurons were adopted.

7. Conclusion

This paper presents some ideas of how the actual frontiers of the ever-evolving area of research known by the generic name of Intelligent Computation could be enhanced by finding ways to measure and therefore compare the efficiency of its techniques — an IQ for machines. The area of AI is growing both in the effort dedicated to research and in operational products already commercialized. Perhaps some kind of objective measures of how far we have walked in the roads that presumably lead to intelligence could help. As recalled in the beginning, Turing's ideas about intelligent machines must be seriously taken. The concepts of *intelligence*, *machine*, *working capacity*, as well as many others in Computer Science, are neither completely nor adequately defined. Until they are, admitting the real possibility of this achievement, the set formed by the union of measures of the IQs of machines, could not be accepted as having complete order, because we have not yet established a way to compare the basic elements of the paradigms involved. Nevertheless, we believe that the concept of intelligence, taken alone, is too abstract to be objectively measured. To make sense, it must be associated with some task. Nowadays, there is a noticeable trend in the popular literature that tries to show that there is more than one type of intelligence, and the different kinds regard specific and distinct abilities of understanding something. If we compare a computer running a program with a human being, maybe we could say that that the hardwired part of that machine would correspond to the genetic code of the human, while the software would stand for the acquired knowledge (training, education) — the *memetic* component. So, like humans, perhaps computers also have different intelligences.

References

[1] D. Zohar and I.N. Marshall, *SQ: Connecting with our Spiritual Intelligence* (USA: Bloomsbury Pub Plc, 2000).

[2] E. Cherniac and D. McDermott: *Introduction to Artificial Intelligence* (Massachusetts: Addison-Wesley, 1985).

[3] A . M. Turing: *Computing Machinery and Intelligence* (Mind: 59(236):433-460, 1950).

[4] J. Searle: *Mente, cérebro e ciência* (Lisboa: Ed. 70 Ltda.,1984).

[5] Churchland, P. S., *Neurophilosophy: Toward a Unified Science of the Mind/Brain* (Cambridge, MA: MIT Press/ A Bradford Book, 1986).

[6] H. Gardner: *Estruturas da mente: A teoria das inteligências múltiplas* (P. Alegre: Ed. Artes Médicas Sul Ltda.,1994).

[7] D. Huisman, A . Vergez: *Curso Moderno de Filosofia* (R. Janeiro: Freitas Bastos, 3ª ed. 1964, p. 20).

[8] L.S. Vygotsky: *A formação social da mente* (S. Paulo: Martins Fontes, 1994).

[9] A .Newell: *Physical Symbol Systems* (Cognitive Science, 18:87-127, 1982).

[10] J.M. Barreto: *Inteligência Artificial no limiar do século XXI* (Florianópolis: Duplic Prest. Serviços, 2000, p. 5).

[11] J. Falqueto, J. M. Barreto, P.S.S. Borges: *Amplification of perspectives in the use of Evolutionary Computation* (IEEE Symposium on Bio-Informatics and Bio-Engineering, Arlington, VA, 2000)

[12] C. R. Darwin: *On the origin of species by means of natural selection* (Connecticut: Grolier Enterprises Corp. Danbury, 1859).

[13] J. Tanomaru: *Motivação, fundamentos e aplicações de algoritmos genéticos* (Curitiba: III Congresso Brasileiro de Redes Neurais e III Escola de Redes Neurais, 1995).

[14] D. R. Tveter: *The pattern recognition basis of Artificial Intelligence* (Los Alamitos, CA: IEEE Computer Society Press, 1998).

[15] L. H. Jacy Monteiro: *Iniciação às Estruturas Algébricas* (São Paulo: Liv. Nobel S.A ., 1971).

[16] E. Post: *Formal reductions of the general combinatorial problem* (American Journal of Mathematics, 65:197-268, 1943).