

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Quest for I (Intelligence) in AI (Artificial Intelligence): A Non-Elusive Attempt

Kumar S. Ray

Abstract

This chapter essentially makes a non-elusive attempt in quest of ‘I’ (Intelligence) in ‘AI’ (Artificial Intelligence). In the year 1950, Alan Turing proposed “the imitation game” which was a gaming problem to make a very fundamental question — “can a machine think?”. The said article of Turing did not provide any tool to measure intelligence but produced a philosophical argument on the issue of intelligence. In 1950, Claude Shannon published a landmark paper on computer chess and rang the bell of the computer era. Over the past decades, there have been huge attempts to define and measure intelligence across the fields of cognitive psychology and AI. We critically appreciate these definitions and evaluation approaches in quest of intelligence, which can mimic the cognitive abilities of human intelligence. We arrive at the Cattell-Horn-Carroll (C–H–C) concept, which is a three-stratum theory for intelligence. The C–H–C theory of intelligence can be crudely approximated by deep meta-learning approach to integrate the representation power of deep learning into meta-learning. Thus we can combine crystallized intelligence with fluid intelligence, as they complement each other for robust learning, reasoning, and problem-solving in a generalized setup which can be a benchmark for flexible AI and eventually general AI. In far-reaching future to search for human-like intelligence in general AI, we may explore neuromorphic computing which is essentially based on biological neurons.

Keywords: general AI, crystallized intelligence, fluid intelligence, deep learning, meta learning, deep-meta learning, neuromorphic computing

1. Introduction

In this chapter, we look for I (Intelligence) in AI (Artificial Intelligence). Still today the term “Intelligence” is not well-quantified for machine implementation. According to the definition of the Oxford dictionary, it is stated as:

“The ability to acquire and apply knowledge and skills.”

We assume rationality as Human Intelligence used in our day to day activity for planning, problem-solving, reasoning and others. With the tremendous growth and development of human civilization, different branches of science and technology are developed. Artificial Intelligence is one such branch which tries to mimic human intelligence through programs implemented into Human-made machines (computers).

In 2007, Legg and Hutter provided a survey of definitions of Artificial Intelligence/Intelligence with methods of evaluation. A decade later, in 2017, José Hernández-Orallo reported an extensive survey on evaluation methods. In this chapter, we describe AI as an attempt to imitate human intelligence in algorithmic form [1, 2].

Normally the rational behavior of an individual indicates his/her basic element of intelligence. Aristotle held the belief that the man is a rational animal. But a growing body of research suggests otherwise. From ancient times, philosophers have been proposing theories of human rationality. There are, however, many definitions of rationality and these change over time. For Plato and Aristotle, man has both a rational and an irrational soul in different proportions. According to Bertrand Russell, "Man is a rational animal. So at least we have been told. Throughout a long life, I have searched diligently for evidence in favour of this statement. So far, I have not had the good fortune to come across it." The term rationality has a handful of interpretations.

With the gradual growth of science and technology people try to adopt sophisticated computing facilities, which may be an attempt to substitute complex mental computation at any particular situation at hand. Thus life becomes smarter and faster to face different challenges of the universe. If we look back at the history of computing facilities for intelligent decision making we observe as follows:

In the year 1942, physicist John Mauchly proposed ENIAC (Electronic Numerical Integrator and Computer). ENIAC project was completed in 1945. It was the first operational computer in USA developed by Army ordinance to compute ballistic firing table during world war II.

In the year 1950, Alan Turing, a British Mathematician and Logician, who broke the German Enigma code during world war II, proposed "The Imitation Game" which was a gaming problem to make a very fundamental question "can machine think?"; which was an informal announcement of Artificial Intelligence. The question raised by Turing was not essentially concerned about an abstract activity like playing chess [3, 4].

In 1950, Claude Shannon published a landmark paper on computer chess and rang the bell of the computer era. At that instant ENIAC was a newborn baby. But visionary people like Shannon, Alan Turing could realize the tremendous potential for computer science and technology. During that period computers were mainly used for ballistic calculations for missiles whereas games appeared to be a natural application for a computer which average people could appreciate. The first working checkers' program was published in 1952. Chess programs followed shortly after that. Arthur Samuel published a strong checker-playing program based on machine learning concept. Samuel used a signature table together with an improved book learning procedure which was a superior approach compared to the earlier one. "alpha-beta" pruning and several forms of forward pruning were used to control the spread of the movement over search tree and allow the program to look ahead to a much deeper depth than it otherwise could do. Though it could not outplay checker masters, the program's playing capability was highly appreciated.

The early effort of Alan Turing, Claude Shannon, Arthur Samuel, Allan Newell, Herbert Simon and others generated tremendous impetus in researching computer-performance at games which could be a testbed for ultimate "intelligence" generated artificially (through computer program) to "exhibit" human-level "intelligence". In the year 1955 J. McCarthy, Marvin Minsky, N. Rochester and C.E. Shannon proposed to study "Artificial Intelligence" during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The basic objective of the study was to proceed based on the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can

be made to simulate it. Their basic ambition was to build a machine which can deal with the problems that are essentially reserved for humans. In those early days of "Artificial Intelligence" (AI), in 1958. Herbert A. Simon and Allen Newell published a paper titled, 'Heuristic problem solving: the next advance in operations research'. At the banquet of the Twelfth National Meeting of the Operations Research Society of America, Pittsburgh Pennsylvania, November 14, 1957, Simon presented the content of the paper as stated above. He brought the term "Heuristic" in practice. At that time it appeared to be an over-optimistic prediction, but its impact is still far-reaching. To establish the need for 'Heuristic' in real-life problem solving he precisely categorized two types of problems; well-structured problems and ill-structured problems. Well-structured problems can be solved explicitly by known existing computational techniques; whereas ill-structured problems are not well-structured. For instance, first, the variables are not numerical, but symbolic or verbal (linguistic), second, the truth status is vague multivalued, instead of precise two-valued, third, there are many practical problems where, in a time-critical situation, variables are not directly measurable (observable) and 'most practical problems' computational algorithms are not available under such circumstances. "Heuristic" can play a significant role to resolve some of the above mentioned ill-structured problems. The term "Heuristic" is essentially domain-specific information which can roughly quantify the perception and/or intelligence of an individual by estimating the intuition, experience and common sense in general for any judgemental decision process that cannot be reduced to systematic computational routine. The parameter 'Heuristic' is an added advantage for solving ill-structured practical problems associated with several environmental uncertainties. Under environmental uncertainty for any judgemental procedures, several 'hunches' and/or wild guess at random are considered as heuristic. The heuristic function may help find a feasible/reasonable (not necessarily optimal) solution of an ill-structured practical problem. Though the necessity for randomness is not proven, there is much evidence in its favor, as stated by Craik's model.

In May 1997 when the chess machine DEEP BLUE defeated world chess champion Garry Kasparov in an exhibition match, it was an indirect silent reply, "YES", to the very fundamental question "can machine think?" raised by Alan Turing in the year 1950. Of course, the thought process of DEEP BLUE machine is not comparable to that of a human being, but definitely, the DEEP BLUE machine very efficiently imitated the thoughtful mind of a world champion of chess. Thus, game-playing became a roseta stone of Artificial Intelligence (AI).

Programming computer to play games is definitely a step towards understanding the methods that may be employed for machine implementation of human intelligent behavior. We still have much to learn from the study of games, and these newer techniques may be applied in future to real-life situations to imitate human intelligence. But the basic question remains that how human becomes so intelligent.

In this chapter, however, we try to explore the cognitive abilities of human being through psychometric models of human intelligence. We observed that the present state of art of artificial intelligence can mimic human intelligence in a crude sense of approximation. At present AI cannot reach the top level of three stratum of Cattell-Horn-Carroll (C-H-C) theory of intelligence. AI can only model few lower level activities of fluid intelligence and crystallized intelligence. The present state of art of artificial intelligence is implemented through Von-Neumann computing. To breakout of Von-Neumann way of thinking, we also explore the possibility of neuromorphic computing. To develop new learning methods with the characteristic of biological brain it is necessary to learn from cutting edge research in neuroscience. As a part of this process there should be a theoretical understanding of "intelligence". Without the theoretical underpinning, we cannot implement

intelligence through neuromorphic computing. Under the present scenario of understanding “intelligence” and mimicking human intelligence in an artificial manner we should further move towards the understanding of native/natural intelligence (NI) which is organic/biological and which is essentially based on biological model of human brain.

2. Evaluation of human intelligence: a brief exposure

Research in the fields of psychology, cognitive science, anthropology, and biology cultivates a sophisticated study on how human intelligence evolved. Understanding about the brain of living humans and great apes and the intellectual abilities they support are enabling us to assess what is unique about human intelligence and what we share with our primate relatives. Examining the habitats and skeletons of our ancestors gives cues as to environmental, social and anatomical factors that both constrain and enable the evolution of human intelligence.

Many methods are used to assess human intelligence and its evolution. These includes (i) behavioral measures which may involve naturalistic observation or analyzing responses in laboratory experiments, (ii) artifactual measures which involve analysis of tools, art and so forth, (iii) anatomical/neurological measures which involve studies of the brain and cranium. Ideally, all three would converge upon a unified picture of how human intelligence evolved. However, this not always the case and indeed the assessment of human intelligence is still under several challenges.

3. Models of human intelligence

Basically there are four important models of human intelligence:

- i. Psychometric model
- ii. Cognitive model
- iii. Cognitive and contextual model
- iv. Biological model.

In this chapter, we consider the first three models which essentially deal with crystalized intelligence, fluid intelligence and combination of these two. We try to approximate or crudely approximate the above features of intelligence through deep learning, meta learning and deep meta learning approaches. We try to adopt in a very crude way the three stratum of Cattell-Horn-Carroll (CHC) theory of intelligence [5].

3.1 Psychometric model

Psychometric model is based on a composite abilities measured by mental tests. This model can be quantified.

One of the earliest of the psychometric model came from the British psychologist Charles E. Spearman (1863–1945), who published his first major article on intelligence in 1904: *The Abilities of Man: Their Nature and Measurement*.

Spearman did not know exactly what the general factor was, but he proposed in 1927 that it might be something like “mental energy.”

The debate between Spearman and Thurstone has remained unresolved.

The American psychologist John B. Carroll, in 1993, proposed a “three-stratum” psychometric model of intelligence that expanded upon existing theories of intelligence. The third stratum consisted solely of the general factor, g , as identified by Spearman. It might seem self-evident that the factor at the top would be the general factor, but it is not, since there is no guarantee that there is any general factor at all. Though there is long pending debates on g (general factor), in this chapter, we discuss this particular issue, based on some conjecture, in section 4.

3.2 Cognitive models

Underlying most cognitive approaches to intelligence is the assumption that intelligence comprises mental representations (such as propositions or images) of information and processes that can operate on such representations.

Other cognitive psychologists had studied human intelligence by constructing computer models of human cognition.

3.3 Cognitive-contextual models

Cognitive-contextual theories deal with the way that cognitive processes operate in various settings. Two of the major theories of this type are that of the American psychologist Howard Gardner and that of Sternberg.

4. Putative test of intelligence

The term putative is commonly used to describe an entity or a concept that is based on what is generally accepted or inferred even without direct proof of it. It means something like an inference or a supposition. There are several examples on putative test of intelligence, like picture completion, picture arrangements, block design, object assembly, etc.

4.1 Culture-fair test

A ‘culture-fair’ or culture-related test makes minimal use of language and not ask for specific facts. On culture-fair tests, Euro-American and African-American children differ because culture can influence a child’s familiarity with the entire testing situation.

Cattell argued that the observed variation among individuals in their scores on any intelligence can be regarded as depending on:

G: variation in the innate gene endowment.

dG: variations in environmentally-produced development of general ability.

C: variations in the closeness of the individual’s cultural training and experiences to the cultural medium in which tests are expressed.

t: variations in familiarity with tests and test situations generally.

f: fluctuations in the underlying capacity.

fr: fluctuations in the effective expression or application of the ability through strength and direction of volition.

s: specific abilities.

e: chance errors of measurement.

In describing the G term in this expression, Cattell had reference to a culture-fair concept of intelligence:

This being the case, *a combination of dG and C* would constitute a manifest general ability, *crystallized intelligence*, which might, if there was any validity to the notion of culture-fair tests, be distinguished from G, *fluid intelligence*.

Later Cattell made these ideas more explicit. He said general ability is of two kinds; (i) fluid ability which manifests perception in new situation, and (ii) crystallized ability which manifests itself in known situation.

He argued that the two abilities should show different development patterns of change.

4.2 Definitions of fluid and crystallized intelligence

Fluid Intelligence (G_f) involves concepts and can be obtained from experiences and opportunities that are afforded to the vast majority.

Thus, G_f involves learning and is a product of acculturation, *but it does not result primarily from differential opportunities in learning or from highly intensive acculturation, such as is promoted through educational programs, which, in one way or another, exclude substantial numbers of individuals.*

The mathematical model which would best represent the lawful combination of the above (and probably many other) factors might be highly complex, but in general form the theoretical terms can be represented as follows:

$$G_f = f(H, M, I, L1, T1, O1) \quad (1)$$

where G_f represents a performance involving fluid intelligence almost exclusively, f represents a function. H refers to a hereditary component. M to the maturation rate, I to injury, $L1$ to learning, $T1$ to the time over which these factors have operated, and $O1$ indicates the extent to which each of these factors has interacted optimally with each other and with environmental circumstances.

Crystallized intelligence (G_c) is an outgrowth of G_f . In the early years of development and under certain other conditions *the two may be so highly related and cooperative as to be virtually indistinguishable*. But over the course of development, when a properly broad view of this is taken, they may be seen to become separated by virtue of the fact that manifested intelligence is produced by a large number of factors which operated largely independently of those seen as accounting for basic intellectual potential. In general these can be classified as factors promoting intensification of acculturation.

$$G_c = f(G_f1, C, B, P, R, L, 2, T2, O2) \quad (2)$$

where G_c represents a performance involving crystallized Intelligence to a high degree, C refers to opportunities and encouragements (chances), E to ergs and sentiments (motive traits), P to non-intellectual personality traits (temperament), R to a factor of longterm memory, $L2$ to the degree of intensive learning distinct from that which is provided for most people. $T2$ to the time over which these factor have operated, $O2$ to the extent to which the combination of factors and development stages was optimal for development of G_c , and G_f refers to the level of Fluid intelligence that operated over this period.

Thus a performance which is said to characterize crystallized intelligence is also seen to contain at least a trace of fluid intelligence, so that to some extent this G_c measure can be said to be confounded with measure of G_f .

Practically, it must be recognized that the learning component in G_f is not completely devoid of exclusive and intensive acculturation, so that it too, is to some extent confounded with G_c . But the essential hypothesis is of this study is that the functions of equations of G_f and G_c can be separated as distinct linear components in performances on a wide sampling of putative tests of intelligence (see **Figure 1**).

4.3 The Cattell-horn-Carroll (CHC) theory of cognitive abilities

The Cattell-Horn-Carroll (CHC) theory of cognitive abilities is the most comprehensive and empirically supported psychometric theory of the structure of cognitive abilities to date. Simplified version of the Cattell-Horn and Carroll model of the structure of abilities is shown in **Figure 2**.



Figure 1.
Performance study between G_c and G_f .

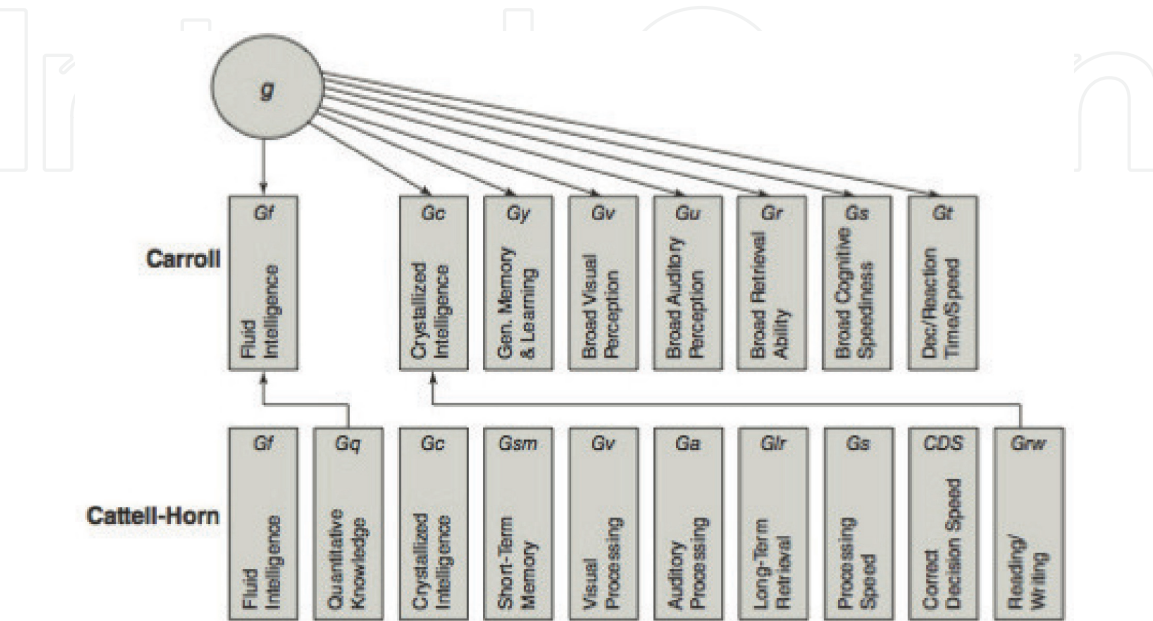


Figure 2.
Representation of Cattell-horn-Carroll three stratum theory.

5. Variant of putative test of intelligence: the factor structure of spearman

In section 3.1, we have already discussed Spearman's g-factor.

Spearman's theory postulated a general capacity, termed g — a kind of “mental energy/neural energy” ... from which all cognitive processes are derived. g-factor is a commonly accepted entity, but there is no evidence of how mental energy (neural energy) is generated. There are several pending debates on this particular issue. Setting aside the debates (section 3.1) we replace g-factor by a more intuitive term “Mood” which is also not measurable but which is more pragmatic to assume in the present context of a cognitive process. The mood is a favorable state of mind consists of the nervous system, to do something through a hierarchy of levels. The mood is placed at the top of the hierarchy and factors of varying degrees of generality further down. Thus, when a person is in a favorable state of mind to solve any problem, the person is in the right mood to solve it. That means, under a favorable state of mind (i.e. in the right mood) neural energy is charged at an absolute magnitude and initiate several levels of intelligence to solve a problem. If a person is in the off mood, then neural energy of the mind is not sufficiently charged to handle a problem.

5.1 Mood

The mood is a favorable state of mind of a person to do something with rationality.

5.2 Difference between mood and emotion

The mood can last for hours. It should not be confused with emotion which lasts, at most, anywhere from second to minutes. It is typically easier to identify emotional trigger but difficult to pinpoint the trigger for our mood. The mood does not have its unique facial expression, whereas emotions do.

5.3 Emotion intelligence

Emotional Intelligence (otherwise known as emotional quotient or EQ) is the abilities to understand, use and manage your own emotion in positive ways to relieve stress, communication effectively, empathize with others, overcome challenges and defense conflict. According to Goleman, emotion can be viewed as:

- Self-awareness; this is the ability to recognize and understand personal mood, emotion, drive and the effect of them on both self and others.
- Self-regulation.
- Internal motivation.
- Empathy (the ability to understand and share the feelings of another).
- Social skills.

In 1927, Spearman stated that g-factor might be something like “mental energy”. Alternatively, it might be viewed as neural energy of brain. But the question is how the energy is generated in brain.

5.4 Where the energy comes from?

Recent neuroscientific evidence suggests brain function is a product of the organization of energetic activity in the brain.

Treating brains as neural information processors does not help understand brain function (consciousness) as a physical process because information, according to the commonly accepted definitions, is not a physical property of brains at the neural level; there is no information in a neuron.

5.5 Brain energy and oxygen metabolism

Dynamic metabolic changes occurring in neurons are critically important in directing brain plasticity and cognitive function.

With dynamic changes in oxygen metabolism occurring during neuronal activity, dynamic changes are likely to be reflected in level of oxygen concentration, potentially having secondary effects on protein function and gene expression.

5.6 Link between mood states and creativity

Creativity is a multifaceted construct, in which different moods influence distinct component of creative thoughts.

Mood shifts are crucial in scaling creativity. The activating moods induce more creative fluency and originality (i.e. novelty) than deactivating moods.

Based on the above discussion in Section 4, we state that the long-pending debates on g-factor of Spearman can be replaced by the new conjecture as follows:

- **Kumar’s conjecture:** g-factor should be replaced by mood which, under favorable mental condition, can generate sufficient mental/neural energy inside the brain to activate different cognitive activities.

Thus, “g” of **Figure 2** is replaced by “mood” and generate a modified three stratum approach to Cattell-Horn-Carroll (CHC) theory of intelligence shown in **Figure 3**.

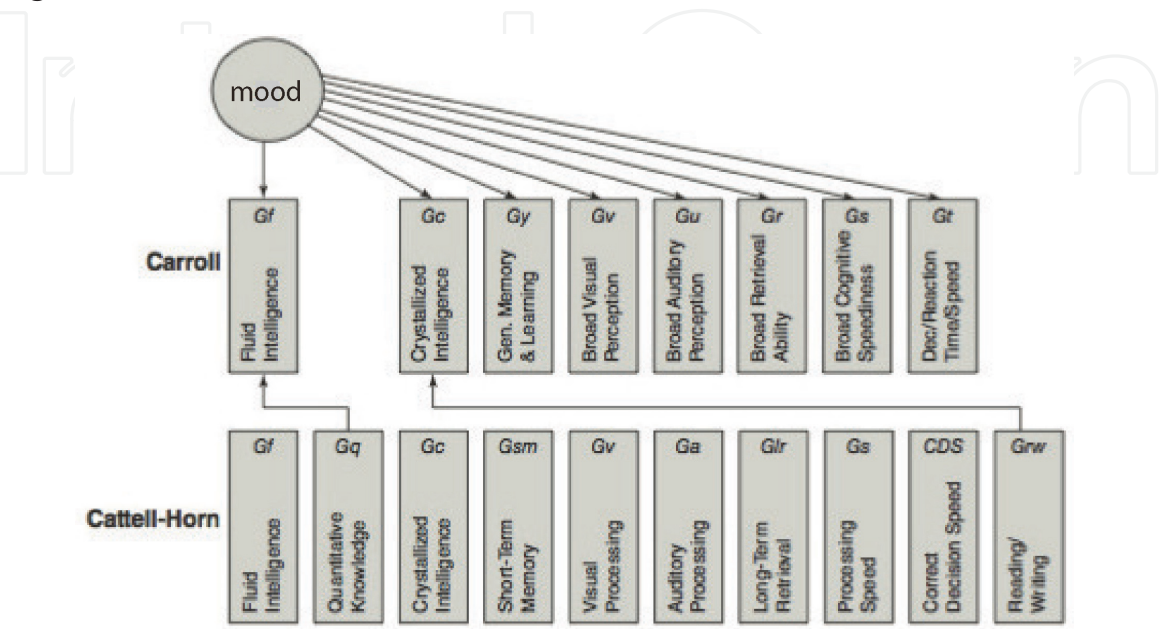


Figure 3.
Modified three stratum approach to C-H-C theory of intelligence.

6. Definitions of CHC abilities

These definitions are derived from an integration of the writings of Carroll (1993), Gustafsson and Undheim (1996), Horn (1991), McGrew (1997, 2005), and Schneider and McGrew (2012).

6.1 Fluid intelligence (G_f)

Figures 2, 3, Tables 1 and 2 provide the definition of CHC abilities.

6.2 Crystallized intelligence (G_c)

It reflects the knowledge and experience of a person.
 G_c includes both declarative (static) and procedural (dynamic) knowledge.

Table 2 shows the definition of narrow crystallized ability.
Definitions of broad crystallized and fluid abilities are available in Figures 2 and 3.

Narrow Stratum I Name (Code)	Definition
Fluid Intelligence (G_f) Induction (I)	Ability to discover underlying characteristic of a problem.
General Sequential Reasoning (RG)	Ability to discover rules to solve a novel problem.
Quantitative Reasoning (RQ)	Ability for inductively and deductively reasoning.

Note. Definitions were derived from Carroll (1993) and Schneider and McGrew (2012).

Table 1.
Narrow G_f stratum I ability definitions.

Narrow Stratum I Name (Code)	Definition
Crystalized Intelligence (G_c)	Depth and breadth of acquired knowledge.
General (verbal) Information (K0)	Domain of knowledge.
Language Development (LD)	Understanding of native language.
Lexical Knowledge (VL)	Content of vocabulary for oral communications.
Communication Ability (CM)	Speaking ability.
Grammatical Sensitivity (MY)	Knowledge of grammar of native language.
Oral Production and Fluency (OP)	Narrow oral communication skills.

Note. Definitions were derived from Carroll (1993) and Schneider and McGrew (2012).

Table 2.
Narrow G_c stratum I ability definitions.

7. Approximate model of crystalized intelligence: a non-elusive attempt

Crystalized intelligence (G_c) includes both declarative (static) and procedural (dynamic) knowledge (see Section 6.2). In the following section, we try to model G_c in an approximate sense through deep neural network model considering only declarative (static) knowledge. Convolutional neural network (CNN) is an implementation of deep neural network architecture. There are several variations of CNN architecture, e.g. Alexnet, Inception, Resnet, Demnet, etc. Input to the CNN is a static representation of knowledge represented by a matrix.

7.1 Why CNN?

Suppose, we have a 28×28 RGB image. So, the total number of inputs in a neural network will be $28 \times 28 \times 3 = 1872..$

Let we have a 1000×1000 RGB image. In this case the total number of inputs in a neural network will be 3 million, which is pretty large.

Since, the number of inputs have increased, the number of weight parameters, will also increase. If there are 1000 nodes in the first layer, the number of elements in the weight matrix of the first layer will be, 3 billion.

We see that with the increase in the dimension of the image, there is a huge increase in the number of parameters, in a feedforward neural network. Thus, it is pretty difficult to train a neural network with such a large number of parameters.

7.2 Computer vision problem

Suppose we have 6×6 grayscale image.

3	0	1	2	7	4
1	5	8	9	3	1
2	7	2	5	1	3
0	1	3	1	7	8
4	2	1	6	2	8
2	4	5	2	3	9

We wish to detect vertical edges in it.
So, the filter or kernel we use is as follows:

1	0	-1
1	0	-1
1	0	-1

After convolution, the resultant matrix, we get as:

-5	-4	0	8
-10	-2	2	3
0	-2	-4	-7
-3	-2	-3	-16

The filter can be learnt using neural networks, which will determine the 9 values of the filter.

We treat each element of the filter as parameters and learn these parameters using back-propagation, similar to the ordinary neural network.

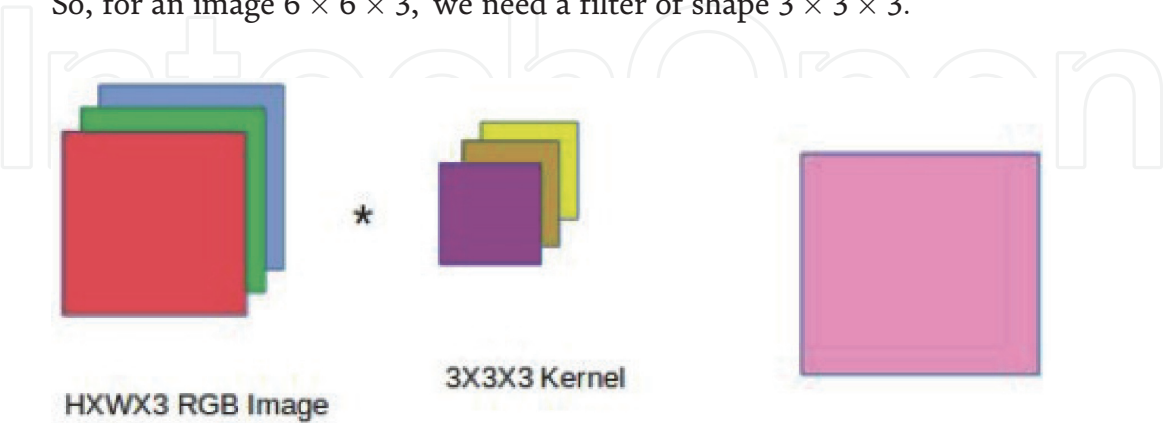
7.3 A short summary of convolutional operations

Summary of convolutions

$$\begin{matrix} n \times n \text{ image} & f \times f \text{ filter} \\ \text{padding } p & \text{stride } s \\ \left\lceil \frac{n + 2p - f}{s} + 1 \right\rceil \times \left\lceil \frac{n + 2p - f}{s} + 1 \right\rceil \end{matrix}$$

How to do convolutions on RGB Images?

Since an RGB image consists of 3 channels, we need to have 3 filters for each channel. So, for an image $6 \times 6 \times 3$, we need a filter of shape $3 \times 3 \times 3$.



How is this convolution computed?

As in 2D convolution, the first filter is convoluted with the Red channel, the second filter with the Green channel, and the third filter with the Blue channel. The values at each convolutional step are added over the channels to give the final result, which will output a single channel, or a 2D matrix.

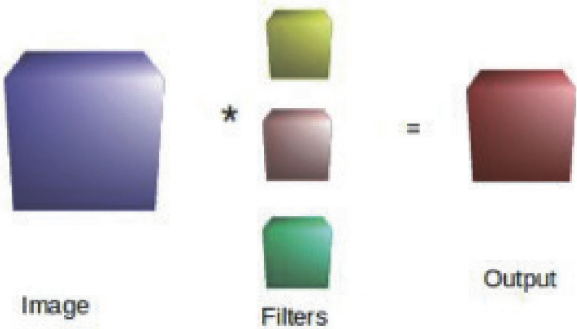
Suppose, the above $3 \times 3 \times 3$ filter used is for detecting vertical edges. Now, suppose that we also want to detect Horizontal edges. So, we need another $3 \times 3 \times 3$ filter for that purpose, which will again output a 2D matrix.

By stacking the output of these two filters, we get as follows:

$$[(H - f + 1) \times (W - f + 1) \times 2] \text{ output(considering no padding).}$$

The number of channels in the output is equal to the number of filters we are using. And, the number of channels in each filter = number of filters in the input. However, before stacking up the outputs, bias is added to the output and passed through the activation function, which is then used as input to the next layer.

Convolutional Layer.



Now, there are various types of layers in a CNN:

- 1. Convolutional
- 2. Pooling
- 3. Fully Connected

Pooling Layers:

Let us consider a 2D matrix for Max-Pooling:

1	3	2	1
2	9	1	1
1	3	2	3
5	6	1	2

Max pooling takes the max of the elements in a $f \times f$ region.
Suppose, we take a 2×2 filter, with strides 2, the output will be a 2×2 2D matrix.
Now, the elements in the output will be max of the elements in the 2×2 region, the filter is passed over.
Going by this way, the output will be.

9	2
6	3

If we have a 3D input, the max-pooling output will have the same number of channels as in input. If the number of channels in the input is n_c , then the number of channels in the output of max-pooling will also be n_c .
Average Pooling
Instead of taking the max of the elements, we take the average in this technique.

3.75	1.25
3.75	2

One important point to note about Pooling layers is that, there are no trainable parameters in Pooling layers.
The two important features of CNNs are:

- 1. **Parameter Sharing:** A filter learnt can be used to detect a feature over all of the input image.
- 2. **Sparsity of Connection?:** In each layer, each output value is dependent only on a small number of inputs.

Unfortunately, Deep learning models are often problematic. Though Deep learning models are robust under declarative (static) knowledge, it is not sufficient under procedural knowledge which refers to the process of reasoning with previously learned procedures to transform learning. Further, several abilities being assessed by psychometric intelligence tests are crystalized abilities which are acquired through experience and which are not distinguishable from skills (multipurpose skills). On the other hand, AI tests showed a focus on capabilities that enable new skill acquisition; hence crystalized abilities are not acceptable for intelligent decision making [6].

8. Approximate model of fluid intelligence: a nonelusive attempt

In this section, we model, in an approximate sense, the above said concept of fluid intelligence (G_f) for on-spot problem solving of previously unseen problems

through meta-learning (learning to learn) approach. Inductive and deductive reasoning are generally considered to be the hallmark narrow ability indicators of G_f . But in our study we do not consider such hallmark ability of G_f [7].

8.1 Meta-learning: learning to learn fast

Meta-learning, also known as “learning to learn”, intends to design models that can learn new skills or adapt to new environments rapidly with a few training examples. There are three common approaches: 1) learn an efficient distance metric (metric-based); 2) use (recurrent) network with external or internal memory (model-based); 3) optimize the model parameters explicitly for fast learning (optimization-based).

We expect a good meta-learning model capable of well adapting or generalizing to new tasks and new environments that have never been encountered during training time. The adaptation process, essentially a mini learning session, happens during test but with a limited exposure to the new task configurations. Eventually, the adapted model can complete new tasks. This is why meta-learning is also known as learning to learn.

8.2 Define the meta-learning problem

A good meta-learning model should be trained over a variety of learning tasks and optimized for the best performance on a distribution of tasks, including potentially unseen tasks. Each task is associated with a dataset $\mathcal{D}$, containing both feature vectors and true labels. The optimal model parameters are:

$$\theta^* = \arg \min_{\theta} \mathbb{E}_{\mathcal{D} \sim p(\mathcal{D})} [\mathcal{L}_{\theta}(\mathcal{D})] \quad (3)$$

It looks very similar to a normal learning task, but *one dataset* is considered as *one data sample*.

Few-shot classification is an instantiation of meta-learning in the field of supervised learning. The dataset $\mathcal{D}$ is often split into two parts, a support set S for learning and a prediction set B for training or testing, $\mathcal{D} = \langle S, B \rangle$. Often we consider a K -shot N -class classification task: the support set contains K labeled examples for each of N -classes.

Figure 4 shows an example of 4 shot 2-class image classification.

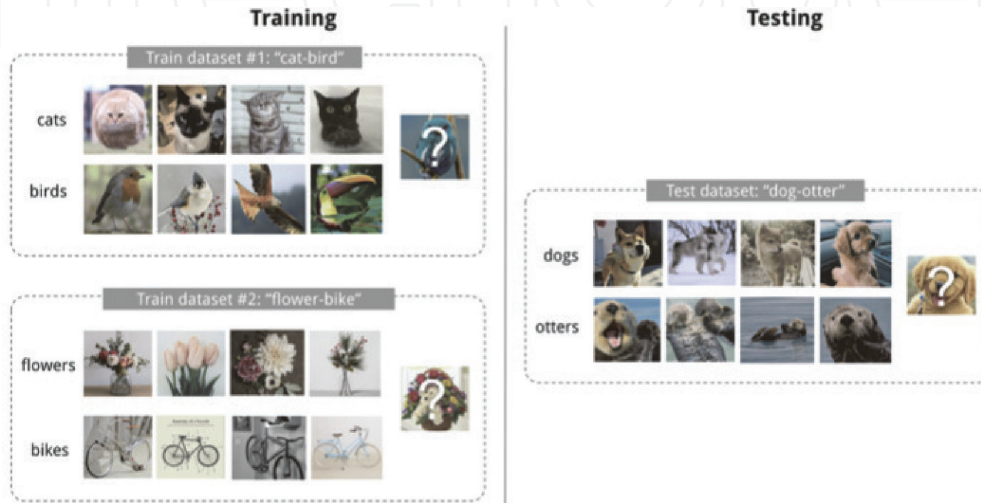


Figure 4.

An example of 4-shot 2-class image classification. (image thumbnails are from Pinterest).

From section 4.2, we understand that crystallized intelligence (G_c) is an out-growth of fluid intelligence (G_f). Thus the performance of crystallized intelligence is influenced by a trace of fluid intelligence, though (G_c) and (G_f) are two separate distinct components in putative test of intelligence. Also from **Figure 2**, we understand that the component reasons (G_f) and acquired knowledge (G_c) are derived from the top level mental (neutral) energy g . Hence to make a very ‘crude approximate’ of three stratum theory of C–H–C (see **Figure 3**), we adopt deep meta-learning approach where we integrate the power of deep learning approach into meta-learning. The (G_f) and (G_c) are both derived from the top level mental (neutral) energy as shown in **Figure 2** and try to follow the hierarchy of three layers to derive the broad and narrow abilities to perform the specific task of given job. Here we consider the term ‘crude approximation, because the top level of **Figure 2** or **Figure 3** can never be reached by the present state of art of artificial neural network. Specially “mood” at the top level of **Figure 3** is a biological phenomenon which generates sufficient mental energy (neural energy) inside the brain under favorable mental conditions. Hence, under such circumstances we assume sufficient neutral (mental) energy is generated for C–H–C theory to perform lower level of cognitive process like crystalized and fluid intelligence.

9. Concept space of deep meta-learning

Figure 5 shows the concept space of deep meta-learning. Eq. (4) represents the meta-learning process [8].

$$\min_{\theta_G, \theta_M, \theta_D} \mathbb{E}_{T \sim \mathcal{P}(T), (\mathbf{x}, \mathbf{y}) \sim \mathbb{D}} \left[J \left(\mathcal{L}_T(\theta_M, \theta_G), \mathcal{L}_{(\mathbf{x}, \mathbf{y})}(\theta_D, \theta_G) \right) \right], \tag{4}$$

where θ_G, θ_M and θ_D are the parameters of deep meta-learning. We assume that the top level mental (neural) energy is available for C–H–C theory of intelligence and a crude approximation of C–H–C theory to mimic human intelligence can be achieved through deep-meta-learning approach. In deep-metal learning approach we crudely approximate to integrate crystalized intelligence (G_c) into fluid intelligence (G_f).

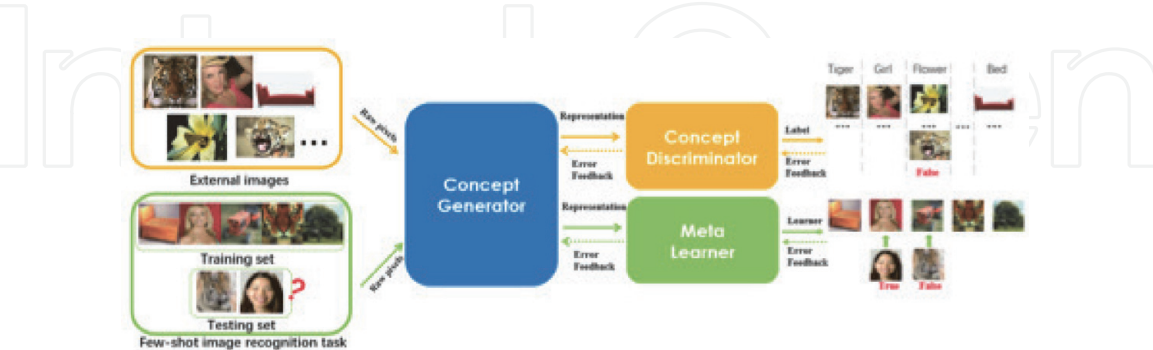


Figure 5.
Concept space of deep meta-learning.

10. Paradeigm shift from Von Neumann computing to neuromorphic computing

So far, we have approximately modeled the psychometric model of human intelligence and implemented in Von-Neumann computer system. Now we seek a

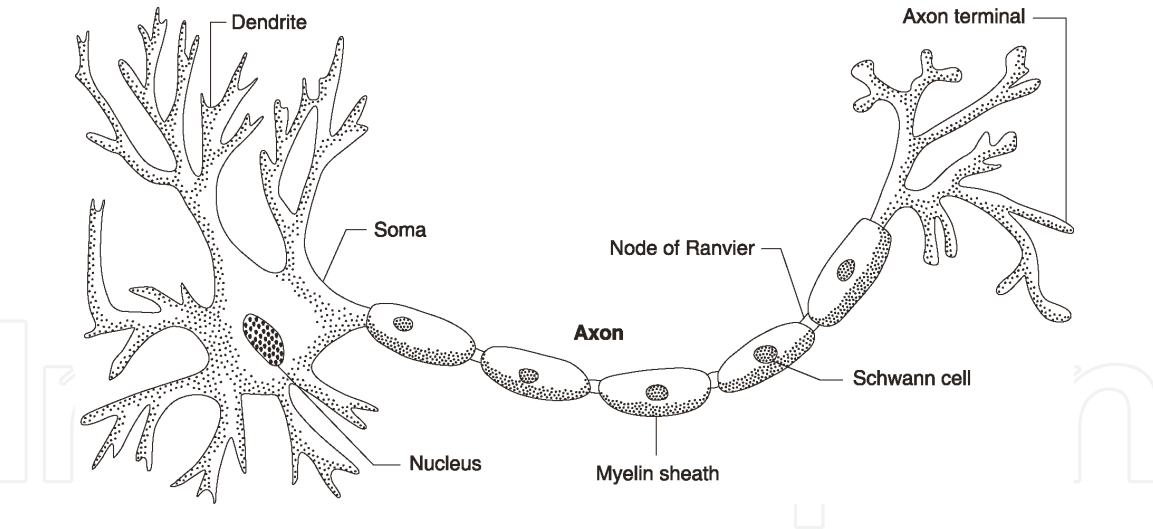


Figure 6.
Structure of simple neuron.

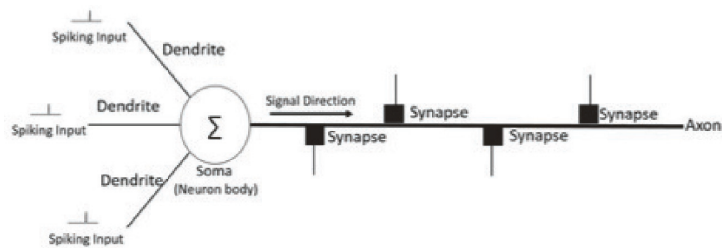


Figure 7.
Neuron design.

new type of computing device which is beyond Moore’s law and Von-Neumann architecture. This new type of computer can proactively interpret and learn from data, solve unfamiliar problems using what it has learned and separated with the energy efficiency of the human brain [9].

Inspired by the working mechanism of the nervous system, the performance development of the computing system has led to a novel non-traditional computing architecture, namely, the neuromorphic computing system. The neuromorphic computing system was proposed by Carver Mead in the 1980s to mimic the mammalian neurology using the very-large-scaled-integrated (VLSI) circuit. In order to physically realize the biological plasticity of a synapse, neuromorphic is combined with computing architecture memristors as electronic synapses.

Although fundamental functions of the brain are still under investigation, two main elements: neuron and synapse are well studied at the cellular level. The structure of a simple neuron is shown in **Figure 6**.

There are four main parts of each neuron, whose functionalities are summarized as shown in **Figure 7**.

Several well-known neuron models are investigated, such as integrate and fire IX model, Xitzhugh-Sitzhudh-Naguno (XN) model, Hodgain-Huxley (HH) model, Leaky integrate an fire (LIS) model, etc.

11. Concluding remarks

In our non-elusive attempt to search for I (intelligence) in AI (Artificial Intelligence), in the first part of this chapter, we crudely approximate the Cattle-Horn-

Carroll (C-H-C) theory of intelligence (see **Figure 3**) through Deep-meta learning approach where we integrate fluid intelligence (G_f) into crystallized intelligence (G_c). Thus problem-solving in unknown environment and robust task specific learning mechanism are combined. During this process of approximation of (C-H-C) theory of intelligence, we realize that with the present state of the art of Artificial Intelligence we can never reach the top-level g-factor/mood (see **Figures 2 and 3**). Hence the approximation process is crude. Though, g-factor as proposed by Spearman is not well defined, the mood which is an alternative conjecture to g-factor is basically a biological phenomenon which occurs inside the brain to generate sufficient mental (neural) energy under the favorable state of mind as stated above to perform lower level of cognitive activities. Thus, in the three stratum C-H-C theory of intelligence, we set aside all the debates on g-factor and inherently assume that such mental (neural) energy is already existing due to the above stated biological phenomenon, i.e. mood. Thus the lower level of cognitive activities can be performed. Hence we consider deep-meta learning approach to crudely approximate C-H-C theory of intelligence to mimic human intelligence in Artificial Intelligence (AI).

In the second part of this chapter, we consider a paradigm shift from Von-Neumann architecture to Neuromorphic computing. It is clear that an entirely new way of thinking about algorithm development is required for neuromorphic computing to break out of the Von-Neumann way of thinking. To develop new learning methods with the characteristics of biological brains, we need to learn from cutting edge research in neuroscience. As a part of this process, we need to build a theoretical understanding of “intelligence”. Without the theoretical underpinnings, we cannot implement true intelligent neuromorphic systems. One of the key features of biological brains that likely enables speedy learning from limited examples or trials is the structural features that are present in biological brains as a result of evolution which should be customized through the learning process. A neuromorphic system may include a long-term off-line training or learning component that may create gross network structures or modules which may be refined and tuned by shorter-term on-line training or learning component. The goal of a neuromorphic computer should not be to emulate the brain. We should instead take inspiration from biology but not limit ourselves to particular models or algorithms.

From the above study we understand that the present state of art of artificial intelligence algorithms which are implemented through Von-Neumann computing cannot model the top level factor (g-factor/mood) of three stratum C-H-C theory of intelligence. Instead, with some assumptions about the top level factor (g-factor/mood) present AI approach can realized some lower level cognitive activities of fluid intelligence and crystallized intelligence. Thus the attempt to mimic human intelligence by conventional AI algorithms is not that much successful as much we expect it to be through generalization of learning algorithm. On the other hand, alternative computing tool, i.e. neuromorphic computing device may attempt to adopt brain functioning for mimicing human intelligence provided the realization of plasticity of synaptic activity is achieved through electronic devices. Under the present scenario we should move towards native/natural intelligence (NI) which is organic/biological and which is essentially based on biological model of human brain. We should explore this new field and should no longer think of artificial intelligence as machines, robot and software code; rather we should think of biological artifacts. Thus in future we should welcome biological AI or BIO-AI.

IntechOpen

IntechOpen

Author details

Kumar S. Ray
Department of Computer Engineering and Applications, GLA University, Mathura,
India

*Address all correspondence to: ksray@isical.ac.in

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] A.M. Turing, Computing Machinery and Intelligence, MIND, vol LIX. No. 236. J, October 1950.
- [2] Claude Shannon, A Chess-Playing Machine, Scientific American, vol 182, NO 2, February 1950.
- [3] S. Legg, A Collection of Definitions of Intelligence, Proceedings of the 2007 Conference on Advances in Artificial General Intelligence.
- [4] José Hernández-Orallo, Evaluation in Artificial Intelligence: From Task-Oriented to Ability Oriented Measurement, Artificial Intelligence Review, 48(3) 397–447, 2017.
- [5] McGraw K.S., The Cattle-Horn-Carroll Theory of Cognitive Abilities: Past, Present and Future in D.P. Flanagan, J.L. Genshaft & P.L. Harrison (Eds.) Contemporary Intellectual Assessment Theories, tests and issue, (PP 136–182) New York, Guilford.
- [6] Schmidhuber, J., Deep Learning in Neural Networks: An Overview, Neural Networks, 61: 85–117, January 2015.
- [7] Chelsa Finn's BAIR blog on "Learning to Learn".
- [8] M. Huisman, J.N. Van Rojn, A. Pleat, A Survey of Deep-Meta learning, Arxiv: 2010. 03522VI[CS. LG], 7 Oct 2020.
- [9] Hongyu An, Kangjun Bai and Yang Yi, The Roadmap to Realize Memristive Three-Dimensional Neuromorphic Computing System, <http://dx.doi.org/10.5772/intechopen.78986>.