A formal definition of the phenomenon of Collective Intelligence and its IQ measure[1]

Tadeusz SZUBA

Department of Mathematics and Computer Science, Kuwait University,
E-mail: szuba@mcs.sci.kuniv.edu.kw

Abstract. This paper formalizes the concept of Collective Intelligence *(C-I).* Application of the Random PROLOG Processor (*RPP*) has allowed us to model the phenomenon of *C-I* in social structures, and to define a *C-I* measure (*IQS*). This approach works for various beings: bacterial ÷ insect colonies to human social structures. It gives formal justification to well-known patterns of behavior in social structures. Some other social phenomenon can be explained as optimization toward higher IQS. The definition of *C-I* is based on the assumption that it is a specific property of a social structure, initialized when individuals organize, acquiring the ability to solve more complex problems than individuals can. This property amplifies if the social structure improves its synergy. The definition covers both cases when *C-I* results in physical synergy or in logical cooperative problem-solving.

1. Introduction

Individual Intelligence (I-I) and *Artificial Intelligence* (*A-I*) have obtained research results, but for *Collective Intelligence (C-I),* the bibliography is poor. Only a few books can be found, e.g. [8], [12], but there is no formal approach. There are also books describing *C-I* biological behavior, e.g. of ant colonies [7], and papers on abstract ant colonies are available, e.g. [5]. The intelligence of humans is widely analyzed in psychology (psychometrics provides different IQ tests for *I-I*, e.g. [4]); sociology, and neuropsychology where there are several models of human intelligence, e.g. [3]. From the point of view of social structures, there is a library of experiments and observations of behavior and cooperation there, (group dynamics, e.g. [11]). The situation in the research on *C-I* might be explained by a widespread fear that *C-I* must be something much more complex than *I-I*. In addition, some top scientists support the conclusion of A. Newell, who objected to *C-I* [10], arguing: *"A social system, whether a small group or a large formal organization, ceases to act, even approximately, as a single rational agent. This failure is guaranteed by the very small communication bandwidth between humans compared with the large amount of knowledge available in each human's head. Modeling groups as if they had a group mind is too far from the truth to be useful scientific approximation very often."* This is a logical result of a priori assumed models of computations. The power and rapid growth of computers impress us; thus, we assume the determinism of computations, and that the information must have an

---

[1] This research is supported by Grant No. SM 174 funded by Kuwait University.

address, or at least must be rationally (in a deterministic way) moved around a computer system. However, alternative models of computations have been proposed, e.g. [1], and applied to *C-I* [13], [14]. Experiments with modeling of chaotic collective inferences in a social structure with use of such models demonstrate [14] that the group (under some restrictions) can complete an inference with an appropriate conclusion, even facing internal inconsistencies, i.e. it can work as a *group mind*.

2. The Basic Concepts of Modeling Collective Intelligence

It is a paradox that the evaluation of *C-I* seems easier than the evaluation of the *I-I* of a single being, which can only be evaluated on the external results of behavior during the problem-solving process. Neuropsychological processes for problem-solving are still far from being observable. Consequently, it is necessary to create abstract models of brain activity [3] based on neuropsychological hypotheses or computer-oriented models like *A-I*. In contrast, many more elements of collectively intelligent activity can be observed and evaluated in a social structure. We can observe displacements/actions of beings, as well as exchange of information between them (e.g. language or the pheromone communication system). *I-I* and behavior is scaled down as a factor – to accidental, local processes. *C-I* can be evaluated through chaotic models of computations like the RPP, and statistical evaluation of the behavior of beings in structured environments. *C-I* evokes some basic observations: • In a social structure, it is difficult to differentiate thinking and non-thinking beings. • The individuals usually cooperate in chaotic, often non-continuous ways. Beings move randomly because real life forces them to do so. Inference processes are random, starting when there is a need, when higher level needs are temporarily not disturbing a given being, or when there is a chance to make inference(s). Most inferences are not finished at all. This suggests using quasi-Brownian [15] movements for modeling social structure behavior [14]. • Individual inferences are also accidental and chaotic. • Resources for inference are distributed in space, time, and among beings. • The performance of a social structure is highly dependent on its organization. • Facts and rules can create inconsistent systems. Multiple copies are allowed. • Probability must be used as an IQ measure for a social structure.

3. Formal Definition of Collective Intelligence

The entry assumption is that *C-I* itself is a property of a group of beings and is expressed/observable and measurable. It is not assumed that beings are cooperating or are conscious or not; nothing is assumed about the communication; we don't even assume that these beings are alive. To better understand the above issues, let's look at some examples. Suppose that we observe a group of ants which have discovered a heavy prey that must be transported, and we also observe a group of humans who gather to transport heavy cargo. Ant intelligence is very low, and a simple perception/communication system is used – however, it is clear that ants display *C-I*. On the other hand, humans, after a lot of time, thought, and discussion will also move the cargo; this is also *C-I*. Because of such situations, the definition of *C-I* must be abstracted from possible methods of thinking and communication between individuals. The definition must be based on the results of group behavior. Let's look into another case. In medieval cities there were streets with shoemaker shops only. They gravitated there because the benefits gained overwhelmed the disadvantages, e.g. when some customers decided to buy shoes from a neighbor. Some shoemakers were sometimes in fact, even enemies. In this example, *C-I* emerges without any doubt; this is obvious just looking at the amount and quality of shoes produced on such a street. Thus we cannot assume willful cooperation for *C-I*, or the definition of cooperation would have to be very vague. Bacteria and viruses cooperate through exchange of (genetic) information; they use *C-I* against antibiotics, but the question is whether they are alive. Thus, the assumption about the existence of live agents must also be dropped. The definition we give now is based on these assumptions, and will formally cover any type of being, any communication system, and any form of synergy, virtual or real.

Let there be given a set *S* of individuals $indiv_1$, …, $indiv_n$ existing in any environment *Env*. No specific nature is assumed for the individuals nor for their environment. It is necessary only to assume the existence of a method to distinguish $indiv_i \;\; i = 1,\dots,n$ from the *Env*. Let there be also given a testing period $t_{start} - t_{end}$ to judge/evaluate the property of *C-I* of *S{…}* in *Env*. Let there now be given any universe *U* of possible problems $Probl_i$ proper for the environment *Env*, and be given the complexity evaluation for every problem $Probl_i$ denoted by $f_O^{Probl_i}(n)$. *C-I* deals with both *formal* and *physical* problems, thus we should write the following:

$$f_O^{Probl_i}(n) \stackrel{def}{=} \begin{cases} \textit{if } Probl_i \textit{ is a computational problem, then apply a standard} \\ \textit{complexity definition with n defining the size of the problem;} \\ \textit{if } Probl_i \textit{ is a physical problem, then use as a complexity} \\ \textit{measure any proper physical units, e.g. weight, size, time,} \\ \textit{etc to express n} \end{cases} \quad \textbf{(1)}$$

Let's also denote in the formula the ability to solve the problems of our set of individuals *S* over *U* when working/thinking without any mutual interaction (absolutely alone, far from each other, without exchange of information):

$$Abl_U^{all\ indiv} \stackrel{def}{=} \bigcup_{\boldsymbol{Probl_i} \in U} \max_S \left(\max_n f_O^{Probl_i}(n) \right) \quad \textbf{(2)}$$

This set defines the maximum possibilities of *S*, if individuals are asked e.g. one by one to display their abilities through all the problems. Observe that if any problem is beyond the abilities of any individual from *S*, this problem is not included in the set.

<u>DEFINITION 1. COLLECTIVE INTELLIGENCE AS A PROPERTY</u>

Now assume that individuals coexist together, and interact some way. We say that *C-I* emerges because of cooperation or coexistence in S, iff at least one problem $Probl'$ can be pointed to, such that it can be solved by a lone individual but supported by the group, or by some individuals working together, such that

$$f_O^{Probl_i}\left(n^{'}\right) \overset{significantly}{>} f_O^{Probl_i}\left(n\right) \in Abl_U^{all\ indiv} \quad \textbf{(3)}$$

$$or$$

$$\exists\, Probl^{'} \ such\ that \left(\forall n\ Probl^{'} \notin Abl_U^{all\ indiv}\right) \wedge \left(Probl^{'} \in U\right)$$

The basic concept of definition (3) is that the property of *C-I* emerges for a set of individuals S in an environment U iff there emerges a quite new problem $\in U$ which can be solved from that point, or similar but even more complex problems can be solved. Even a small modification in the structure of a social group, or in its communication system, or even in, e.g. the education of some individuals can result in *C-I* emergence, or increase. An example is when the shoemakers move their shops from remote villages into the City. This can be, e.g. the result of a king's order, or creation of a "free trade zone". The important thing is that the distance between them has been reduced so much that it triggers new communication channels of some nature (e.g. spying). Defining *C-I* seems simpler but measuring it is quite a different problem. The difficulty with measuring *C-I* lies in the necessity of using a specific model of computations, which is not based on the DTM Turing Machine.

4. The Random PROLOG Model of Computations

Upon observing the social structures of humans, it is striking that the inferences are parallel, chaotic, and at the same time different algorithms of inference are applied. The same agent can be the processor, the message carrier, and the message, from the point of view of different parallel-running processes. The inferring process can be observed as global, but the role of specific elements cannot be easily identified and interpreted. Thus, a mathematical logic-based representation of a social structure is necessary for analysis, separate from interpretation. We have to rely on only approximations of the inferences we observe, and we must be able to easily improve an approximation; thus declarative, loosely coupled inference systems are necessary. PROLOG fulfills the requirements after some modifications. PROLOG is declarative, i.e. clauses can be real-time added/retracted; independent variants of facts and procedures are allowed. This is basic for chaotic systems. Symbols and interpretation of clauses are independent because PROLOG is based on 1st order predicate calculus. Thus, clauses can describe roughly observable aspects of behavior (e.g. of ants) without understanding them very well. In PROLOG, it is easy to adopt a molecular model of computations [1]. The RPP model of computations includes these properties: • *Information_molecules* infer randomly and are independent which implies full parallelism. • The RPP allows us to implement more inference patterns than standard PROLOG, i.e. concurrent forward, backward, and rule-rule inferences, e.g. where an inference is backward only from the formula of the goal. • With simple stabilizing methods (e.g. self-elimination, self-destruction of inconsistent clauses) the RPP is resistant to inconsistent sets of clauses. • Different parallel threads of inferences can be run at the same time in the RPP. • Parallelism of multiple inference diagrams highly compensates for the low efficiency of the random inference process.

4.1. Clauses and Computational PROLOG Space in the RPP

The 1st level Computational Space (*CS)* with inside quasi-random traveling Clause Molecules (*CMs)* of facts, rules, and goals c_i is denoted as the multiset $CS^1 = \{c_1, \ldots, c_n\}$. Thus, clauses of facts, rules, and goals are themselves 0-level *CS*. For a given *CS*, we define a membrane similar to that of the Chemical Abstract Machine (*CHAM*) [2] denoted by $|\cdot|$ which encloses inherent facts, rules, and goals. It is obvious that $CS^1 = \{c_1, \ldots, c_n\} \equiv \{|c_1, \ldots, c_n|\}$. For a certain kind of membrane $|\cdot|$ its type p_i is given, which will be denoted $|\cdot|_{p_i}$ to define which *CMs* can pass through it. Such an act is considered Input/Output for the given *CS* with a given $|\cdot|$ It is also allowable in the RPP to define degenerated membranes marked with $\cdot|$ or $|\cdot$ i.e. a collision-free (with membrane) path can be found going from exterior to interior of an area enclosed by such a membrane, for all types of *CMs*. The simplest possible application of degenerated membranes in the *CS* simulating a given social structure is to make, e.g. streets or other boundaries. If the *CS* contains clauses as well as other *CSs*, then it is considered a higher order one, depending on the level of internal *CS*. Such internal *CS* will be also labeled with $\hat{v}_j$ e.g.

$$CS^2 = \{|c_1, \ldots CS^1_{\hat{v}_j}, \ldots c_n|\} \quad iff \quad CS^1_{\hat{v}_j} \equiv \{|b_1, \ldots, b_n|\}$$
$$where \quad b_i \quad i = 1\ldots m \quad and \quad c_j \quad j = 1\ldots n \quad are \quad clauses \tag{4}$$

Every c_i can be labeled with $\hat{v}_j$ to denote characteristics of its individual quasi-random displacements. The general practice will be that higher level *CSs* will take fixed positions, i.e. will create structures, and lower level *CSs* will perform displacements. For a given *CS* there is a defined position function *pos*:

$$pos: \ O_i \rightarrow \langle position\ description \rangle \cup undefined \quad where \quad O_i \in CS$$

If there are any two internal *CS* objects O_i, O_j in the given *CS*, then there is a defined distance function $D(pos(O_i), pos(O_j)) \rightarrow \Re$ and a rendezvous distance *d*. We say that during the computational process, at any time *t* or time period Δt, two objects O_i, O_j come to rendezvous iff $D(pos(O_i), pos(O_j)) \le d$. The rendezvous act will be denoted by the rendezvous relation ®, e.g. O_i ® O_j which is reflexive and symmetric, but not transitive. The computational PROLOG process for the given *CS* is defined as the sequence of frames *F* labeled by *t* or Δt, interpreted as the time (given in standard time units or simulation cycles) with a well-defined *start* and *end*, e.g. $F_{t_0}, \ldots, F_{t_e}$. For every frame its multiset $F_j \equiv (|c_1, \ldots, c_m|)$ is explicitly given, with all related specifications: *pos(.)*, membrane types *p*, and movement specifications *v* if available. The RPP is initialized for the start of inference process when the set of clauses, facts, rules, and goals (defined by the programmer) is injected into this *CS*. When modeling the *C-I* of certain closed social structures, interpretations in the structure will be given for all CS^m_n, i.e. “this *CS* is a message”; “this is a single human"; "this is a village, a city", etc. The importance of properly defining $\hat{v}_j$ for every CS_j^i should be emphasized. As has been mentioned, the higher level CS_j^i will take a fixed position to model substructures like villages or cities. If we model a single human as CS_j^1, then $\hat{v}_j$ will reflect displacement of the being. Characteristics of the given $\hat{v}_j$ can be purely Brownian or can be quasi-random, e.g. in lattice, but it is profitable to subject it to the present form of CS_j^i. Proper characteristics of $\hat{v}_j$ provide the following essential tools; • The goal clause, when it reaches the final form, can migrate toward the defined *Output* location (membrane of the main *CS* or even a specific *CS*). Thus, the appearance of a solution of a problem in the *CS* can be observable. • Temporarily, the density of some *CMs* can be increased in the given area of *CS*. After the given CS_j^i reaches the necessary form, it migrates to specific area(s) to increase the speed of selected inferences.

The right side of the rule and goal clause in the RPP is any set *S(...)* of unit clauses, contrary to standard PROLOG, where there is an ordered list of unit clauses. Such a set can even be considered a local, invariable *CS* of rules, containing unit clauses. This approach is compatible with the philosophy of the RPP and allows us to define traveling clusters of facts enclosed by membranes. It is allowed to introduce the *configuration*, to order, e.g. spatially, the unit clauses in the set *S(...)*.

4.2. The Inference Model in the RPP

The pattern of inference in Random PROLOG generalized for any CS has the form:

DEFINITION 2. GENERALIZED INFERENCE IN CS^N

Assuming that $CS = \{\dots CS_j^i \dots CS_l^k \dots\}$, we give these definitions:

CS_j^i ® CS_l^k and $U(CS_j^i, CS_l^k)$ and C(one or more CS_n^m of conclusions) $\vdash$

one or more CS_n^m of conclusions, $R(CS_j^i$ or $CS_l^k)$ **where:**

CS_j^i ® CS_l^k denotes rendezvous relation; $U(CS_j^i, CS_l^k)$ denotes unification of the necessary type successfully applied; C(*one or more* CS_n^m *of conclusions*) denotes that CS_n^m are satisfiable. Notice that the reaction → in (*CHAM* [2]) is equivalent to $\vdash$ inference here. $R(CS_j^i$ or $CS_l^k)$ denotes that any parent *CMs* are retracted if necessary. The standard PROLOG inferences are simple cases of the above definition. Later, when discussing N-element inference, we will only be interested in "constructive" inferences, i.e. when a full chain of inferences exists. Thus the above diagram will be abbreviated as $CS_j^i; CS_l^k \xrightarrow{RPP} \sum_n CS_n^m$ without mentioning the retracted *CMs* given by $R(CS_j^i$ or $CS_l^k)$. In general, successful rendezvous can result in the "birth" of one or more *CMs*. All of them must fulfill a *C(...)* condition;

otherwise, they are aborted. Because our RPP is designed to evaluate the *C-I* of closed social structures, simplifying assumptions based on real life can be made. It is difficult to find cases of direct rendezvous and inference between two CS_i^m *and* CS_j^n *if* $m,n \geq 1$ without an intermediary involved $CS_k^0 \quad k=1,2,\ldots$ (messages, pheromones, observation of behavior, e.g. the bee's dance, etc.). Even in *GA* [9], the crossover of genes can be considered the inference of the two genomes CS_i^0 *and* CS_j^0. Only if we consider CS^n on the level of nations, where exchange (migration) of humans takes place, can such a case be considered an approximation to such high level rendezvous and inferences. This is, however, just approximation, because finally, this exchange is implemented at the level of personal contact of humans, which are just rendezvous and inferences of two CS_i^0 *and* CS_j^0 with the help of $CS_k^0 \quad k=1,2\ldots$. Thus, rendezvous and direct inference between two CS_j^i *if* $i \geq 1$ are rare. We will only make use of a single CS_{main}^n *for* $n > 1$ as the main *CS*. Single beings (humans, ants) can be represented as $CS_{individual}^1$. Such beings perform internal brain-inferences, independently of higher, cooperative inferences inside CS_{main} and exchange of messages of the type CS^0. It will be allowable to have internal CS^k inside the main *CS*, as static ones (taking fixed positions) to define sub-structures such as streets, companies, villages, cities, etc. For simplicity, we will try to approximate beings as CS^0; otherwise, even statistical analysis would be too complicated. It is also important to assume that the results of inference are not allowed to infer recursively. They must immediately disperse after inference; however, after a certain time, inferences between them are allowed again (this is called *refraction* [6]).

5. Formal Definition of the Collective Intelligence Measure

The two basic definitions for *C-I* and its measure IQS have the form:

<u>DEFINITION 3: N-ELEMENT INFERENCE IN CS^N</u>

There is a given *CS* at any level $\mathrm{CS}^n = \left\{\mathrm{CS}_1^{a_1}, \ldots \mathrm{CS}_m^{a_m}\right\}$, and an allowed Set of Inferences SI of the form $\{\textit{set of premises CS}\} \xrightarrow{I_j} \{\textit{set of conclusions CS}\}$, and one or more CS_{goal} of a goal. We say that $\left\{I_{a0}, \ldots, I_{a_{N-1}}\right\} \subseteq SI$ is an N-element inference in CS^n, if for all $I \in \left\{I_{a0}, \ldots, I_{a_{N-1}}\right\}$ the premises $\in$ *the present state of* CS^n at the moment of firing this inference, all $\left\{I_{a0}, \ldots, I_{a_{N-1}}\right\}$ can be connected into one tree by common conclusions and premises, and $CS_{goal} \in \left\{\textit{set} \text{ of conclusions for } \mathrm{I}_{a_{N-1}}\right\}$.

<u>DEFINITION 4: *COLLECTIVE INTELLIGENCE* QUOTIENT (IQS)</u>

IQS is measured by the probability *P* that after time *t*, the conclusion CM_{goal} will be reached from the starting *state of* CS^n, as a result of the assumed N-element inference. This is denoted $IQS = P(t, N)$.

For evaluating *C-I* the last two definitions fulfill these requirements: • N-element inference must be allowed to be interpreted as any problem-solving process in a social structure or inside a single being, where N inferences are necessary to get a result; or any production process, where N-technologies/elements have to be found and unified into one final technology or product. Therefore in the same uniform way we model inferring processes or production processes within a social structure. This is very important because some inference processes can be observed only through resultant production processes or specific logical behavior (of ants, bees, bacteria). • Simulating N-element inference in the RPP allows us to model the distribution of inference resources between individuals, dissipation in space or time, or movements (or temporary concentration) in the *CS*. This reflects well the dissipated, moving, or concentrated resources in a social structure of any type. • Cases can be simulated where some elements of the inference chain are temporarily not available, but at a certain time *t*, another inference running in the background or in parallel will produce the missing components. This is well known in human social structures, e.g. when a given research or technological discovery is blocked until missing theorems or sub-technology is discovered. • Humans infer in all directions: forward, e.g. improvements of existing technology; backward, e.g. searching how to manufacture a given product going back from the known formula; and also through generalization, e.g. two or more technologies can be combined into one more general and powerful technology or algorithm. N-element inference simulated in the RPP reflects all these cases clearly.

EXAMPLE: HUMAN SOCIAL STRUCTURE FACING A SIMPLE PROBLEM

Let us consider a human social structure suffering from cold weather. Assume that a) there is one person with matches moving somewhere inside the social structure; b) another has a matchbox; c) a third person has an idea of how to use the matches and matchbox to make a fire, if both components are available; d) a fourth person is conscious that heat from fire is a solution for the cold weather problem in this social structure; e) a fifth person is conscious that fire is necessary. Other humans are logically void. Nobody has total knowledge of all the necessary elements and their present position; everybody can infer only locally. For this social structure the CS^1 (for the RPP) can be defined as set of CS^0: `CS¹ = {matches. matchbox. matches ∧ matchbox → fire. fire → heat. ?heat.}` Now it is necessary to make assignments: • The displacement abilities v for above defined CS^0. We can use pure or quasi-Brownian movements, e.g. beings run chaotically around on foot looking for the necessary components. Public transportation can change the characteristics of displacement. Virtual travel by telephone or Internet search can also be defined. Traveling may be on a day/week/year scale. • The internal structure of CS^1 which affect rendezvous probability should be defined (internal membranes reflecting boundaries, e.g. streets). • Artificial restrictions can be imposed on unification, e.g. different languages, casts, racial conflicts, etc. This system is a 4-element inference. The IQS can be computed here for this social structure on the basis of simulations as a curve showing how probability changes in the given time period, or as the general analytical function *P(4, t)*, if we can manage to derive it.

The proposed theory of *C-I* allows deriving formal definitions of some important properties of Social Structures obvious in real life: • LEVELS OF *C-I*; • SPECIALIZATION OF SOCIAL STRUCTURES; • EQUIVALENCE OF SOCIAL STRUCTURES OVER DOMAIN D.

6. Conclusion

The formal approach to the *C-I* presented here targets the future *C-I* Engineering of humans, and similar applications: • The Collective Intelligence of bacterial colonies against new drugs; • The Collective Intelligence of social animals and insects that create a threat, e.g. for farming. We should remember that the science of *C-I* is very sensitive, i.e. any definition or assumption can immediately be verified through numerous results of an observational or experimental nature, e.g. [7], [11], etc. Thus, every element from the formal approach has interpretation. So far this formalism has done very well when applied to human social structures [19], [14]. It has allowed us to confirm theoretically some behavior and properties of the social structures discussed in [11], as well as to discuss some new ones which until now had seemed “beyond” theoretical possibilities for evaluation, e.g. “why we build cities’.

Literature

1. Adleman L.: Molecular computations of solutions to combinatorial problems. *Science*. 11. 1994.
2. Berry G., Boudol G.: The chemical abstract machine. Theoretical Computer Science. 1992.
3. Das J. P. et al: Assessment of cognitive processes. Allyn and Bacon. 1994.
4. Goldstein A. P., Krasner L.: Handbook of Psychological Assessment. Pergamon. 1990.
5. Dorigo M., Gambardella L.M.: Ant Colonies... BioSystems, 1997.
6. Giarratano J., Riley G.: *Expert Systems*. PWS Pub. II ed. 1994.
7. Hölldobler E., Wilson O.: The ants. Harvard University. 1990.
8. Levy P.: Collective Intelligence. Plenum Trade 1997.
9. Michalewicz Z.: GA + data structures = evolution programs. Springer 1992.
10. Newell A.: Unified theories of cognition. Harvard University Press. 1990.
11. Shaw M. E.: Group Dynamics. IV ed. McGraw-Hill. 1971.
12. Smith J. B.: Collective Intelligence in Computer-Based Collaboration. Lawrence 1994.
13. Szuba T.: Evaluation measures for the collective intelligence of closed social structures. ISAS’97 Gaithersburg/Washington D.C., USA, 22-25 Sept., 1997
14. Szuba T.: A Molecular Quasi-random Model of Computations Applied to Evaluate Collective Intelligence. Future Generation Computing Journal. Volume/issue: 14/5-6. 1998.
15. Whitney C. A.: Random processes in physical systems. Wiley & Sons. 1990.