Towards a theory of epistemic information

Stefano Mizzaro

Department of Mathematics and Computer Science
University of Udine
Via delle Scienze, 206 — Loc. Rizzi, I 33100 Udine, Italy
mizzaro@dimi.uniud.it
http://www.dimi.uniud.it/~mizzaro

Abstract: The concept of information is still lacking a complete understanding, as witnessed by the large number of different meanings that overload the terms "data", "knowledge", and "information". In this paper, a definition of information (named epistemic information) based on the knowledge states of agents is proposed. It is used to analyze, understand, and explain some puzzling phenomena, like the subjectivity of information, and the inconsistency between amount of information contained in a datum and amount of information received. A "knowledge conservation principle" is also proposed and justified.

1 Introduction

The Association for Computing Machinery (ACM) defines computer science as the science of the algorithms that process information. Whereas, since Turing, Tarsky, and Church, the concept of algorithm is quite well understood, the concept of information is still lacking a complete understanding. The paramount manifestation of this impasse is perhaps the large number of different meanings that overload the terms "data", "knowledge", and "information" [15]. Many definitions of information have been proposed; we can classify them into two groups: "hard sciences" *vs.* "soft sciences" information theories. The first group (in which information is usually defined in an objective way) includes the well known Shannon's information theory [24], the Algorithmic Information Theory, independently developed by Chaitin, Kolmogorov, and Solomonoff [16], and the Semantic Information Theory introduced by Bar-Hillel, Carnap, and Popper [13] and further developed by Dretske [9], Barwise and Perry [2], and Devlin [7]. Besides these, there are others, perhaps less known, "soft sciences" approaches (in which information is usually defined in a subjective way): Bateson defined information as a difference [3], Brookes proposed that "information is a small bit of knowledge" [5], and, more recently, Clancey expanded Bateson's definition of information, proposing that information is the detection of a difference that is functionally important for an agent to adapt to a certain context [6].

In this paper I propose a new approach to this issue: I present definitions, examples, a partially formalized graphical notation, and an analysis that leads to a better understanding of some puzzling phenomena concerning information. This paper extends the work [19], where a similar, though narrower, analysis is used to understand some crucial concepts of information retrieval, with an emphasis on relevance [21]. Here I do not take into account the information retrieval field.

The paper is structured in the following way: Section 2 sketches some concepts that are the basis of *epistemic information*, defined in Section 2.7.3. In Section 2.9 the obtained scenario is analyzed and some of its limitations are singled out. The following of the paper extends the scenario for overcoming these limitations. Section 7 concludes the paper and sketches some future developments of this line of research.

2 Epistemic information: building blocks

2.1 Agents

I will assume that the world is populated by *agents*, that *act* in the *environment*, or *'External' World* (henceforth simply *world*), whatever it may be. I borrow the following definition of an agent from a book that uses this concept as a basis for giving a uniform description of the artificial intelligence field:

> An *agent* is anything that can be viewed as *perceiving* its environment through *sensors* and *acting* upon that environment through *effectors*. A human agent has eyes, ears, and other organs for sensors, and hands, legs, mouth, and other body parts for effectors. A robotic agent substitutes cameras and infrared range finders for the sensors and various motors for effectors. A software agent has encoded bit strings as its perceptions and actions. [23, page 31]

The agents can be biological (human, and more generally living, beings) or artificial (robots, softbots, computers). Complexity is not necessary: simple entities like a thermostat or an atom can be seen as an agent. There is an implicit but important assumption in this definition: that we can separate the agent and the world he (she, it) is in. This assumption is usually accepted in the artificial intelligence field, though there are some different viewpoints (*e.g.,* [17]). My position is: let us accept this assumption as a work hypothesis, not as an established truth, and let us see how far it will lead us.

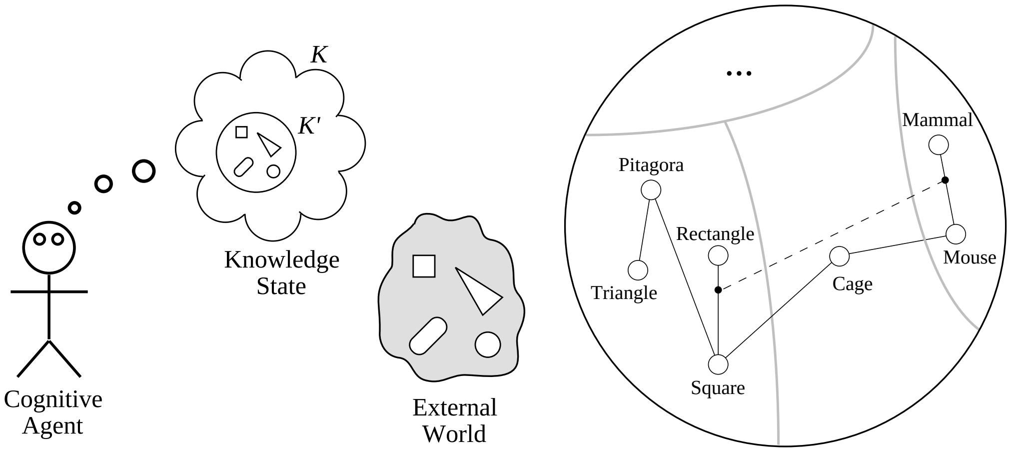
2.2 Knowledge states and items

Some agents (we might call them *cognitive agents*; for the sake of brevity, in the following I will omit the word "cognitive", since I will speak of this kind of agents only) have an internal *Knowledge State* (KS): through their sensors, the agents perceive (a portion of) the world and *represent* it into their KSs. The portion of a KS that corresponds to a portion of the world is said the *representation* of that portion of the world into the KS. The representation of the world can be more or less *correct* (*i.e.,* corresponding to the world) and *complete* (*i.e.,* taking into account every aspect of the world). The acting of the agent in the world takes place on the basis of his KS.

Each KS is a finite[1] collection of atomic components, that I call *Knowledge Items* (KI). Each KS is thus a set of KIs, and in the following I use some of the usual notation of set theory, as $\in$ (belong), $\subseteq$ (subset), $\setminus$ (set difference), $\emptyset$ (empty set), $\cup$ (union), $\cap$ (intersection), and so on, with the usual meaning extended to KSs and KIs. I will use the capital letter K (with some subscripts and/or superscripts) to denote the KSs, and lowercase k for the KIs. Figure 1(a) sketches the scenario presented so far: an agent perceives a portion of the world and represents it in K', a subKS of his whole KS K.

The reader can imagine many alternatives for having a more concrete picture of KSs (and KIs), for instance: logical theories (*i.e.,* sets of logical formulas) [12], semantic nets [23], sets

[1] Most of the paper would be unchanged if we relaxed this constraint. However, taking into account only finite KS avoids some problems, though it is a significant limitation. I will come back later on this issue.



(a) Agent, KS, world, representation. (b) Links and partitions of KSs.

Figure 1: KSs, KIs, and links.

of beliefs [12],[2] situations [1, 7, 9], recursive models [20], minds and ideas [3], and so on. I do not take position among these (and many others) alternatives in this paper: I am interested in proposing a higher level view, and I try to remain at a level of abstraction general enough to comprise all of them.

2.3 Links

Whatever KSs and KIs are, I suppose that some *links* exist among KIs, similarly to what happens in the well known Truth Maintenance Systems (TMS) or in the Recursive Models [20]. This means that a KS can be (fuzzyly) partitioned into subKS, each (fuzzy) partition containing the KIs more strongly linked. See for instance Figure 1(b): the KIs regarding, say, Euclidean Geometry (as the concepts of triangle, square, Pitagora's Theorem, and so on) belong to one partition, while the KIs regarding, say, mammals belong to another partition. Such partitions, besides being subjective, are neither absolute nor clear-cut: it is (almost) always possible to find a link chain among two KIs. It is a fuzzy, or perhaps fractal, situation. For instance, it is possible to link, say, a mammal with a square through the KSs about mice and cages (see the links chain in Figure 1(b)). Moreover, the links themselves are a kind of KIs, and this allows to have links among links (and so on), as the one, represented by the dashed line in Figure 1(b), that links two links of the same kind, *i.e.*, two "is-a" links.

The above assumptions are widely spread in many fields, for instance: artificial intelligence (under the label "logicism" [12, 22]), situation semantics [7], cognitive science [11], and human-computer interaction [8]. They can be criticized from many points of view (*e.g.*, [4, 17, 18, 6]), but they will be useful in the following of this paper. Thus, again, I do not take them as established truths, but as useful work hypotheses: for the sake of brevity, I avoid to analyze the (many) philosophical implications of these issues.

[2] In this paper I will not distinguish between belief and knowledge (defined as a true belief, see [12]): an agent can not distinguish if his KIs (or beliefs) are true or not.

2.4 Transitions

The KS of an agent may change as time goes on: when this happens, I say that a *transition* from an *initial* KS K_I to a *final* KS K_F takes place. A transition can take place for two different reasons:

- by (internal) inference: the agent reasons, reflects, and modifies his KS without any input from the world. This will be called *inferential transition*; it is the only kind of transition that can take place in an agent without sensors;
- by perceiving a datum: through his sensors, the agent perceives something (a datum) from the world and this leads to the modification of his KS (a transition into another KS). This will be called *noninferential transition*. If the datum leads to a change of the KS then the datum is said to *carry* information. Note that everything can be a datum, even *nothing* (*i.e.*, no reception from the world), because nothing is different from something, and so it can carry information [3]. Thus an inference and a null datum are two different things.

2.5 K_A and K_R

Let us analyze in more detail a single transition. The modification that takes place in the KS can be described by the differences between the initial KS K_I and the final KS K_F. What is added to K_I is a subKS here indicated by K_A and what is removed from K_I is indicated by K_R. In Figure 2(a) the two KSs K_I and K_F are represented by circles, the subKS K_A added to the KS by the little white semi-circle on the border of the KS, and the removed subKS K_R by the little black semi-circle. The transition between the two KSs is represented by an arrow labeled by the corresponding datum (or by 'infer' if it is an inferential transition). Finally, a time instant can be associated to each KS (in the figure, t_I and t_F are the time instants of K_I and K_F, respectively).

K_A and K_R can be defined in a formal way, using set difference:

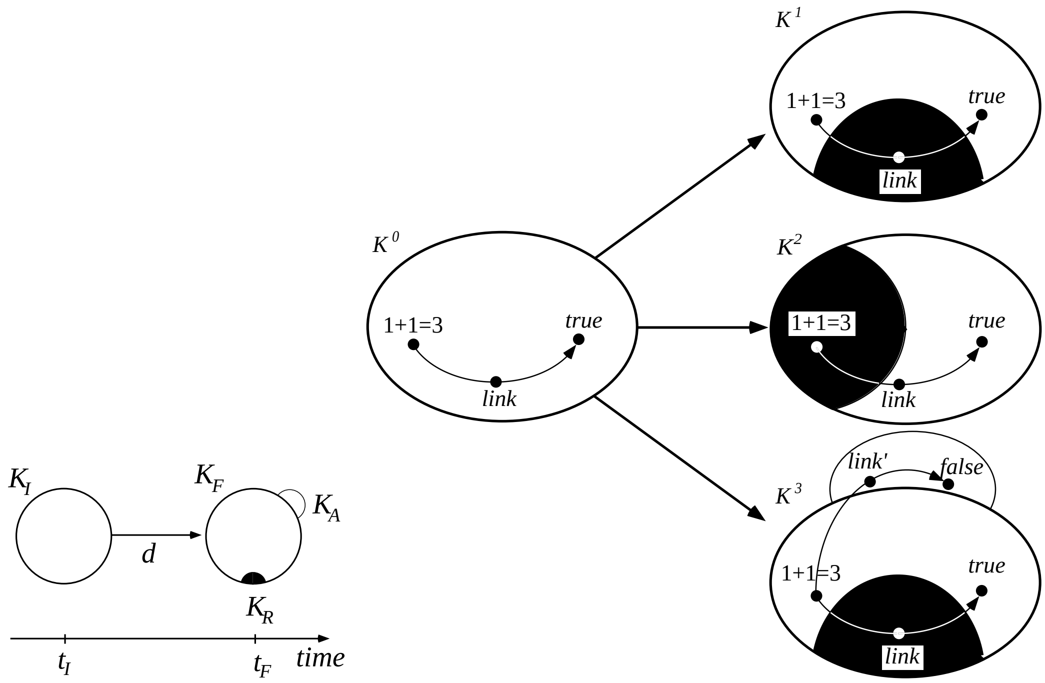
Definition 1 (K_A and K_R) *Given an initial KS K_I and a final KS K_F,*

$$K_A \triangleq K_F \setminus K_I; \; K_R \triangleq K_I \setminus K_F. \qquad \square$$

Besides adding new KIs (K_A), a transition can also lead to the removal of some subKS (K_R). This happens, for instance, when the agent forgets something, or when the agent changes his knowledge about a fact, believed true in the KS before a transition and false later. As an example, consider an agent knowing that $1 + 1 = 3$; in his KS (K^0 in Figure 2(b)) he has a KI which represents the expression (labeled with "$1 + 1 = 3$" in figure), a KI to represent the truthness (labeled with *true*), and a link, that is a KI too (the arrow), for linking the two previous KIs.[3] Now, many transitions can take place, for instance the following three (see the figure): (i) the agent forgets that "$1 + 1 = 3$" was true (K^1): $K_R = \{$*link*$\}$; (ii) the agent forgets what it was that was true (K^2): $K_R = \{$"$1+1=3$"$\}$; and (iii) the agent discovers that "$1 + 1 = 3$" is false, and not true (K^3): $K_R = \{$*link*$\}$, $K_A = \{$*link',false*$\}$.

Finally, note that the addition and removal of KIs can be complex operations, since the KIs linked to the added or removed KIs are affected too, in a recursive way. This will be discussed in Section 2.8.

[3] Of course, in the KS there will be KIs for "1", "+", "3", "=", and perhaps further ones. Let us disregard them.



(a) K_A and K_R.

(b) Forgetting and changing knowledge.

Figure 2: Differences among the initial and final KS of a transition.

2.6 Networks of possible KSs

It is possible to imagine a *network* (*à la* Kripke [14]) of the possible KSs of an agent: the nodes of the network are KSs, among which some are 'real' KSs (*i.e.*, KSs actually possessed by the agent, sometime) while other ones are possible KSs that do not become 'real' ones (*i.e.*, the agent does not possess them, though he could); the arcs of this network are the transitions from one KS to another one. We have already seen an example of a network of KSs in Figure 2(b); in Figure 3 a more complex situation is presented, and some transitions among possible KS are represented. The KSs and transitions in the figure are the possible ones, but only one path from K^1 to K^7 is followed in the reality, for instance the one with the thickest lines, while the other KSs remain only possible ones.

2.7 Data, knowledge, and information

The concepts presented so far (agents, knowledge states, knowledge items, links, transitions, K_A, K_R, and networks) allow us to propose some definitions of three terms (and concepts) "data", "knowledge", and "information", clearly distinguishing among them.

2.7.1 Data

A *datum* is a *difference* in the world: it is something physical, that can be observed. Given a particular datum, an agent may be capable of perceiving it or not (depending on his sensors). Sometimes, even if the sensors would allow to perceive the datum, the actual perception does not take place, for instance because the agent does not note the datum, or because he is not

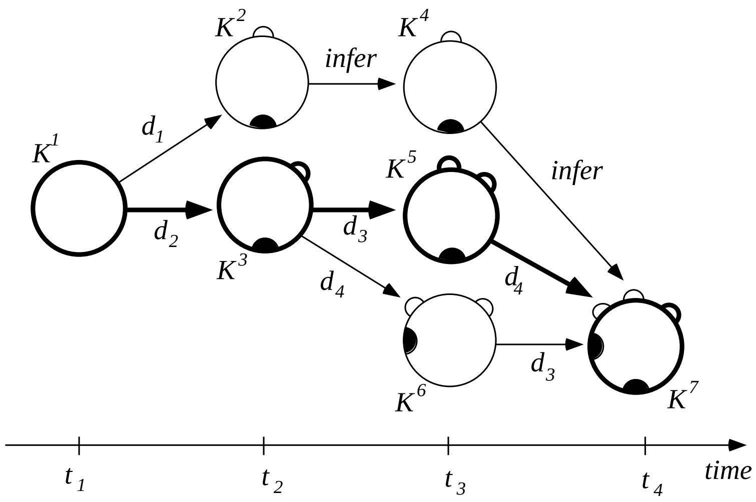


Figure 3: KSs and transitions among them.

interested in it. For instance, an (ultra)sound produced by an ultrasonic dog whistle is a datum that cannot be perceived by the human sensors; a dog is able to perceive it, though the perception might not happen at all if the dog is strongly focused on something more important (eating a tasty bone). Thus a datum can be, for a given agent, *potential* (if the datum is perceivable) or *actual* (if the datum is perceived): this dichotomy is quite important, and I will extensively discuss it later, in Section 4.

There are some well known definitions of the "amount of information" in a datum (*e.g.*, Shannon, Chaitin, Solomonoff, Kolmogorov). I will use the term "information" for something different, and I will speak of the *amount of a datum* d, denoted by $|d|$. And I will not choose a specific definition among them. Note that it is often very difficult, if not impossible, to measure the amount of a datum using one of the above well known definitions. For instance, if the datum is a book, or a painting, or a natural language utterance, and so on. Sometimes, the measurement is feasible.

2.7.2 Knowledge

We can assume that *knowledge* exists inside a KS only: knowledge is what the KIs are made of, like "matter" is what the atoms are made of. In this way, either a (printed) book does not contain knowledge, or it is an agent. Both positions seem maintainable: on the one side, we all should agree that if a book is an agent, it is at least a very peculiar kind of agent (no input from the world, and no action), but on the other side we often say, for instance, "there is a lot of knowledge in that book". I will come back later on this issue.

2.7.3 Information

When a datum is perceived by an agent, it can lead to a modification of the agent's KS, that changes, say, from K_I to K_F. When this happens, the datum has carried some *information* to the agent. We have two choices here: either the information can be intrinsic in the datum (and an agent can perceive all or a part of it), or the information is subjective, contextual and depends also on the agent's KS. Here I follow the latter choice: the information carried by a datum is the difference between the two (final and initial) KSs. I call the information defined here *epistemic information*, since it is based on the KS of an agent. More in detail:

Definition 2 (Epistemic information) *Given an agent in an initial KS K_I, receiving a datum d, and ending up in a final KS K_F, the information carried by d is defined as the ordered pair*

$$I(d, K_I, K_F) \triangleq \langle K_A, K_R \rangle$$

(see Figure 2(a); K_A and K_R are defined in Definition 1). □

This is a way of expressing the difference among the two (final and initial) KSs. We could also define $\Delta K \triangleq \langle K_A, K_R \rangle$ and write $I(d, K_I, K_F) = \Delta K$.

Let's give some definitions that will be useful in the following.

Definition 3 (Quantity of knowledge, $|K|$) *The* quantity of knowledge $|K|$ *in a KS K is the number of KIs in K.* □

This definition is perhaps too naive, but it is adequate in the following. Let's remark that if we had infinite KSs, there would be a problem here: a KS K with an infinite numerable number of KIs would have the same quantity of knowledge of another KS K' with the same KIs plus one. But an agent with the second KS (K') "knows more" than an agent with the first KS (K), and this seems not satisfactory.

Definition 4 (Variation of the quantity of knowledge, $\Delta|K|$)

$$\Delta|K| = |K_F| - |K_I| = |K_A| - |K_R|. \quad \square$$

If $\Delta|K|$ is positive (negative), there is more (less) knowledge after the transition.

Definition 5 (Quantity of information, $|I|$)

$$|I| = |\Delta K| = |\langle K_A, K_R \rangle| \triangleq |K_A| + |K_R|. \quad \square$$

The quantity of an ordered pair $|\langle A, B \rangle|$ is defined here as $|\langle A, B \rangle| \triangleq |A| + |B|$, where A and B are disjunct, and $|\Delta K|$ is different from $\Delta|K|$ (the variation of the quantity of knowledge, Definition 4). K_A and K_R must be disjoint: if a KI were in both K_A and K_R, it would have been added and then removed (or vice-versa) in the same transition.

2.8 Perception and restructuring

A noninferential transition between two KSs is not a mere accumulation of knowledge, but involves a restructuring of the KS. The research in the field of belief revision [10] is on this topic; three kinds of transitions are defined: *expansion* (just adding something to a KS), *contraction* (removing something), and *revision* (modifying something, *i.e.*, a contraction followed by an expansion). All three of these have to preserve two conditions on the KS: *consistency* (an agent cannot believe both one thing and its negation) and *logical omniscience* (an agent believes all the logical consequences of anything he believes). Many considerations could be done on this issue (*e.g.*, are logical omniscience and consistency too strong requirements?), but they would lead us too far. What is relevant here is the need of restructuring the KS after receiving a datum. On the basis of the above sketched framework, this is explained by means of the links: the KIs linked to the added or removed ones are affected too, in a recursive way.

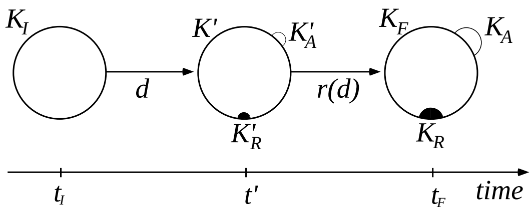


Figure 4: Perception and restructuring transitions.

Therefore, the noninferential transition caused by a datum d can be divided into two parts, as shown in Figure 4: a first *perception* transition (labeled with the datum d in figure) in which the datum is perceived and something is immediately added to or removed from K_I, obtaining K'; and a second *restructuring* transition (labeled with $r(d)$ for "restructuring because of datum d") in which the restructuring operation takes place, and the KIs linked to the added (${K_A}'$) or removed (${K_R}'$) KIs are affected. The resulting modification of the KS is anyway fully represented by K_A and K_R, thus in the following I will sometimes treat a noninferential transition as an atomic one.

Of course, there are two problems: (i) understanding when a perceptual transition ends and a restructuring transition starts, and (ii) understanding the difference between restructuring and inferential transitions. I do not investigate the first point and I will assume that restructuring and inferential transitions are the of same kind.

Summarizing, the KS plays a fundamental role in an agent receiving data: the information carried depends on the KS of the agent, and it should be said that a datum is 'interpreted' (not 'received') by an agent on the basis of his KS. It is possible to define an *interpretation* function

$$\textit{int: Data} \times \textit{KS} \rightarrow \textit{KS}$$

that, given as argument a datum and a KS, assumes as value the KS resulting from the transition: in Figures 2(a) and 4, *int*$(d, K_I) = K_F$. On the basis of what above said (see Figure 4), this function can be divided into two components (perception and restructuring)

$$\textit{perc: Data} \times \textit{KS} \rightarrow \textit{KS};\ \textit{restr: Data} \times \textit{KS} \rightarrow \textit{KS},$$

such that *int(d,K) = restr(d,perc(d,K))*.

Let's note that the *meaning* of a datum can be defined in a similar way to information: in Figure 4, the meaning of the datum d is *Meaning*$(d, K_I) \triangleq \langle {K_A}', {K_R}' \rangle = \langle \textit{perc}(d, K_I) \setminus K_I, K_I \setminus \textit{perc}(d, K_I) \rangle$.

2.9 Some remarks

Many concepts have been defined: agents, KSs, KIs, links, transitions, K_A, K_R, network of possible KSs, data, knowledge, and, finally, information.

Also Bateson [3] and Brookes [5] define information as a difference, but in ways that are different from the above proposed one. Bateson defines an item of information as a difference (in the world) that makes another difference: on the basis of the above definitions, Bateson's difference is a (perceived) datum. Brookes proposes in his "fundamental equation" $K[S] + \Delta I = K[S + \Delta S]$, that "information is a small bit of knowledge": a "knowledge structure" $K[S]$ is changed to a new knowledge structure $K[S + \Delta S]$ by the information ΔI. Brookes's view is more similar to the one proposed here than Bateson's one, but it is anyway different: Brookes's knowledge and information are measured in the same units, while this is not true for epistemic information.

With the above proposed epistemic information, we can remark that the same datum can carry different information, both to two different agents (if the KSs of the two agents are different) and to the same agent in different time instants (if the KSs of the agent in the two time instants are different). For instance, if the datum is an utterance in some language, an agent understanding such a language can obtain information, while an agent not understanding the language cannot.

Moreover, two different data can carry the same information. For instance, an utterance uttered in two different languages carries the same information to an agent understanding both languages (and already knowing that the speaker knows both languages). A number expressed through different 'formats' (8, "eight", VIII, 1000 in binary notation, 10 in octal base) carries the same information to an agent not 'sensible' to the difference of the base.

There are anyway some issues deserving further explanation:

1. The same datum sometimes carries the same information to different (human) agents. Why? It is not maintainable that there are human beings with identical KSs: we need further explanation.

2. The amount of received information seems not always related to the amount of data:

   (a) A huge datum d (*i.e.*, a datum with a high $|d|$) can give a very small amount of information to an agent (for instance, if he cannot understand d, or if he already knows d, and so on).

   (b) On the other side, a single 'bit' (*i.e.*, an atomic datum, as yes/no, 0/1, true/false, on/off) can carry an huge amount of information to an agent in a proper KS, *i.e.*, a KS with a high knowledge "pressure" (borrowing the term from physics), in which a single bit triggers some transitions with a high difference between the initial KS and the final KS.

   This data-information inconsistency is quite disappointing. If we used everyday speaking, it should be named information-information inconsistency. In everyday use of the language, there are (at least) two meanings of "information". On the one side we often say "there is a lot of information in that book (that figure, those words, and so on)", thus supporting an objective view of information, a view that considers information as intrinsic in its support. On the other side we often also say "this book (figure, words, ...) gives no information to me". A longer Japanese text *contains* more information than a shorter one, but a Japanese text *gives* no information at all to me (because I do not understand Japanese). With the terminology proposed here, the inconsistency is less worrying (a longer Japanese text is a large data, and its interpretation gives no information to me) but still deserving further analysis.

3. There seems to be no "knowledge conservation principle". This is not necessarily a problem, but it is at least strange that information and knowledge, sometimes, seem to be created from nothing, or to disappear. Moreover, such a principle could help us in better understanding the concepts of information and knowledge.

These remarks can not be explained in the above framework. Some new concepts are needed: they are introduced in the following sections.

3 Prerequisite KS

As above said, the definition of information proposed here implies a subjective concept of information, accordingly to what happens in the real world. However, in everyday life, the same datum sometimes brings the same information to different agents (remark 1 in Section 2.9). Why? This may be explained assuming that the KSs of the agents that populate the real world (human beings) are similar for genetic and social factors. But this is less true if we consider people from different cultures, *e.g.*, European *vs.* Asiatic, or different kinds of agent, *e.g.*, human beings *vs.* computers,[4] and it is surely true that there are no human beings with identical KSs: some further explanation is needed.

The explanation proposed here is that the information received through a particular datum does not depend on the whole initial KS, but only on a part of it, that can be named *prerequisite* subKS. In this way, the subjectiveness of information is less evident. This prerequisite subKS, indicated by K_P, must be such that the information received by the agent would not change if the initial KS of the agent were just K_P instead of the whole KS.

Let us see an example. An agent believes a wrong version of Pitagora's Theorem (for instance "the sum of the squares on the two catheti is greater than the square on the hypotenuse", or $a^2 + b^2 > c^2$, instead of the well known correct version $a^2 + b^2 = c^2$). When the agent receives the proof of the right version of the theorem (a datum), his KS changes accordingly. Referring to Figure 5(a) (in which the noninferential transition is taken as an atomic one), we have: K_I is the initial KS of the agent; d is the proof of the right version of the theorem; K_P is the prerequisite KS and represents the notions of triangle, square, and so on, necessary for understanding the theorem; K_F is the final KS of the agent; K_A is the subKS representing the correct version of the theorem; and K_R is the subKS representing the truthness of the wrong version of the theorem. Obviously, the KS of the agent may contain something more than K_A, K_R, and K_P, but this is not relevant in this example.

Intuitively speaking, it is possible to characterize K_P referring to what the agent has to do with the received information, with his motivations, aims, interests. A more formal definition is the following.

Definition 6 (Prerequisite KS, K_P) *Given an initial KS K_I, a final KS K_F, and a transition labeled with a datum d, K_P is defined as a KS such that:*

(i) *It is a subKS of the initial KS: $K_P \subseteq K_I$.*

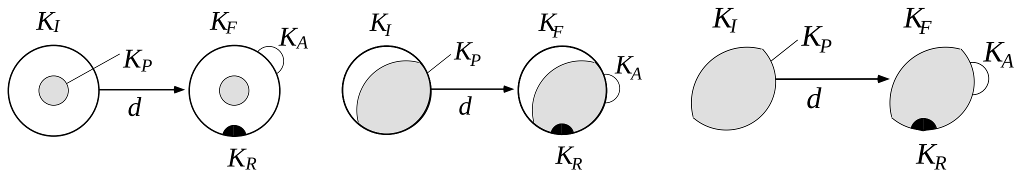
(ii) *The information carried by d to the agent in a KS K_I would not be different if the initial KS were just K_P. Using the int function defined at the end of Section 2.8: $I(d, K_I, \mathit{int}(d, K_I)) = I(d, K_P, \mathit{int}(d, K_P))$.*[5]

(iii) *K_P is* minimal*: $\nexists K' \subseteq K_P$ such that the previous condition (ii) holds.* □

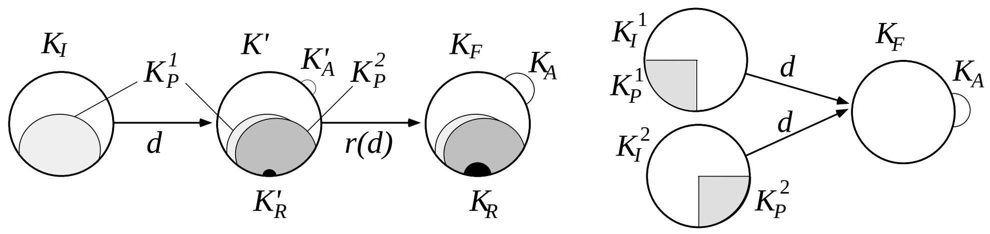
Note that on the basis of this definition we obtain a restriction on K_R and K_P: $K_R \subseteq K_P$ (a particular case being $K_R = \emptyset$). This means that the removed subKS must be a part of the prerequisite KS (otherwise (ii) does not hold, as it is easy to see), and this is quite reasonable. Therefore, Figure 5(a) should be modified as Figure 5(b) (where $K_R \subseteq K_P$), and Figure 5(c) shows how the information carried by a datum depends on K_P only: K_A and K_R are the same in Figures 5(b) and 5(c).

[4] This might be an explanation of all the difficulties encountered in computer science, especially in artificial intelligence: a high difference between the KSs of the two kinds of agents, human beings and computers.

[5] Or: $\langle \mathit{int}(d, K_I) \setminus K_I, K_I \setminus \mathit{int}(d, K_I)\rangle = \langle \mathit{int}(d, K_P) \setminus K_P, K_P \setminus \mathit{int}(d, K_P)\rangle$.



(a) K_P: a first attempt. (b) K_P contains K_R. (c) Information depends on K_P only.



(d) K_P and perception and restructuring. (e) Non univocal K_P.

Figure 5: K_P: the interpretation of a datum does not depend on the whole KS.

There are some problems with this definition. If we had infinite KSs, there are some situations in which a prerequisite KS does not exists, because of the minimality condition: in some infinite sets, like in the open interval $(0..1]$, the minimum does not exist, and the infimum does not belong to the set. This is another reason to restrict ourselves to the finite case. However, even in the finite case, the prerequisite KS might be not univocal, since the information carried by a datum could be the same also with disjoint K_P (as in Figure 5(e)). These issues are not studied in more detail in this paper.

If we split the noninferential transition into the perception and restructuring ones, the situation becomes slightly more complex. The prerequisite KS changes after receiving d (see Figure 5(d)): K_P^1 is the prerequisite KS for the perception transition, K_P^2 for the restructuring one, and the prerequisite KS for the whole noninferential transition is $K_P = K_P^1 \cup K_P^2$. The above conditions (i)–(iii) hold for such a K_P, and analogous ones can be defined (using the *perc* and *restr* functions) for each of the two transitions:

(i_1) $K_P^1 \subseteq K_I$;

(ii_1) $\langle \mathit{perc}(d, K_I) \setminus K_I, K_I \setminus \mathit{perc}(d, K_I)\rangle = \langle \mathit{perc}(d, K_P^1) \setminus K_P^1, K_P^1 \setminus \mathit{perc}(d, K_P^1)\rangle$;[6]

(iii_1) K_P^1 is minimal, that is $\nexists K \subseteq K_P^1$ such that the previous condition (ii_1) holds.

(i_2) $K_P^2 \subseteq K_I$;

(ii_2) $\langle \mathit{restr}(d, K') \setminus K', K' \setminus \mathit{restr}(d, K')\rangle = \langle \mathit{restr}(d, K_P^2) \setminus K_P^2, K_P^2 \setminus \mathit{restr}(d, K_P^2)\rangle$;

(iii_2) K_P^2 is minimal, that is $\nexists K \subseteq K_P^2$ such that the previous condition (ii_2) holds.

[6] Or: $\mathit{Meaning}(d, K_I) = \mathit{Meaning}(d, K_P)$.

4 The actual-potential dichotomy

The data-information inconsistency (Section 2.9, remark 2) is not completely explained by K_P. We can understand it by making explicit the dichotomy between actual and potential data, knowledge, and information, as proposed in the following subsections.

4.1 Data

We have already seen in Section 2.7.1 that a datum can be actual or potential: remember the whistle that is a potential but not an actual datum for the busy dog (and it is not even potential for a human being, because his sensors do not allow him to perceive it), or think of a detail not seen by a detective. Both of these data are differences in the world (potential data), but they are actual data only if perceived. This might be an explanation of the data-information inconsistency, but it is only partial: if it is true that an agent might not perceive a datum, and thus obtain no information from it (remark $2(a)$ in Section 2.9), the symmetric remark $2(b)$ cannot be explained on these basis.

4.2 Knowledge

We can divide the KS of an agent into two parts, and speak of *actual knowledge* (*Kact*) and *potential knowledge* (*Kpot*). I do not give formal definitions of *Kact* and *Kpot*. Instead, I propose some examples. Let us consider a theorem prover. It is a particular kind of agent: once axioms and inference rules are fixed, the set of the derivable theorems is determined, and thus one could say that the theorem prover has the knowledge of all the derivable theorems before deriving them. Now, a better distinction can be obtained: the theorem prover has the potential knowledge of the derivable, and not yet derived, theorems, and has the actual knowledge of the derived theorems only. Another example is presented in Figure 6: an agent has a KS with the KI "$x + 1248 = 2726$"; if the agent knows elementary arithmetic, he "knows" in some sense that $x = 1478$. But this knowledge can be only potential, and become never actual if the agent is not interested in knowing the value of x. As a last example, let us explain the belief that "mathematicians are doing nothing", meaning that they are producing no new knowledge. After the actual-potential distinction, it is correct to say that mathematicians are not producing new *Kpot*, while they are producing new *Kact*.

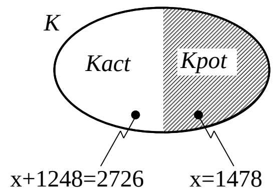


Figure 6: $K =$ *Kact* $\cup$ *Kpot*.

Therefore, we can say that $K =$ *Kact* $\cup$ *Kpot*, and we can divide each KS in two parts, as in Figure 6: the white part represents the *Kact*, the grey one the *Kpot* of the KS K.

It is important not to mistake the *Kact*-*Kpot* dichotomy with the manifest-latent knowledge dichotomy: they are similar but different.

Definition 7 (*Kman* and *Klat*) *The* manifest knowledge *(abbreviated with Kman in the following) of an agent is the knowledge that is revealed by his behavior, while the* latent knowledge *(Klat) is the knowledge that the agent anyway has, but cannot be ascertained by observing the agent's behavior only.*

In other terms, an agent can use his knowledge for acting, and on the basis of his actions it is possible to understand what knowledge that agent has. Here are some examples of this distinction:

- A book can be regarded as an agent that has no knowledge, since it has no perception of the world, and does not act. But I prefer to say that a book does possess knowledge, even if it is only latent and not manifest.

- I did not know that a friend of mine knows Latin, until when I heard him translating a Latin sentence into English. His knowledge of Latin was latent, but actual.

- A theorem prover that prints the theorems it is deriving on the screen is an agent that has the actual knowledge of the printed theorems, and this knowledge is manifest. It is also possible to imagine a computer program that prints only *some* of the derived theorems; in that case, the *Kman* is only a portion of its *Kact*: there is also the *Klat* of all the not printed but derived theorems (that can be used for deriving other theorems).

The total actual knowledge of an agent is obtained adding *Kman* and *Klat*: $\mathit{Kact} = \mathit{Kman} \cup \mathit{Klat}$; let us also note that the inferential and restructuring transitions are transformations of *Kpot* into *Kact*: they do not add any new knowledge.

4.3 Information

Also information can be actual or potential. As we have seen in the previous section, the equation $x + 1248 = 2726$ is a datum that can carry the information that $x = 1478$. But this information can be only potential, and not become actual, for an agent not interested in knowing x's value. In the following, I will use *Iact* and *Ipot* for actual and potential information, respectively.

Iact and *Ipot* can be formally defined on the basis of *Kact*-*Kpot* distinction (see Figure 7(a), that refines Figure 2(a)). First of all, let us note that K_A and K_R are made up by an actual part and a potential one: $K_A = \mathit{Kact}_A \cup \mathit{Kpot}_A$ and $K_R = \mathit{Kact}_R \cup \mathit{Kpot}_R$, (we also have $\mathit{Kact}_A = \mathit{Kact}_F \setminus \mathit{Kact}_I$, $\mathit{Kpot}_A = \mathit{Kpot}_F \setminus \mathit{Kpot}_I$, $\mathit{Kact}_R = \mathit{Kact}_I \setminus \mathit{Kact}_F$, and $\mathit{Kpot}_R = \mathit{Kpot}_I \setminus \mathit{Kpot}_F$). Now we can define:

Definition 8 (*Iact* and *Ipot*)

$$\mathit{Iact}(d, K_I, K_F) \triangleq \langle \mathit{Kact}_A, \mathit{Kact}_R \rangle; \ \mathit{Ipot}(d, K_I, K_F) \triangleq \langle \mathit{Kpot}_A, \mathit{Kpot}_R \rangle. \qquad \Box$$

We can also write $\mathit{Iact}(d, K_I, K_F) = \Delta \mathit{Kact} = \langle \mathit{Kact}_F \setminus \mathit{Kact}_I, \mathit{Kact}_I \setminus \mathit{Kact}_F \rangle$ and $\mathit{Ipot}(d, K_I, K_F) = \Delta \mathit{Kpot} = \langle \mathit{Kpot}_F \setminus \mathit{Kpot}_I, \mathit{Kpot}_I \setminus \mathit{Kpot}_F \rangle$. So far, I have spoken of one kind of information only, without distinguishing between *Iact* and *Ipot*.

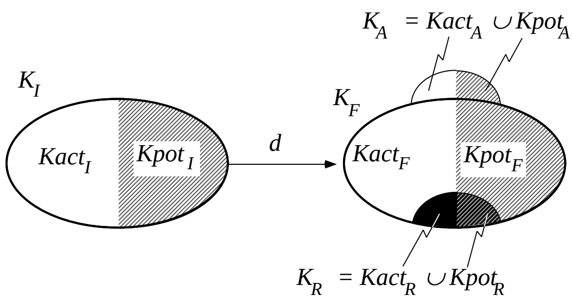
To relate I, *Iact*, and *Ipot*, I define the sum of two information items as the sum of two ordered pairs, taking a special care for keeping disjoint the added and removed parts K_A and K_R:

Definition 9 (Information sum, $\oplus$)

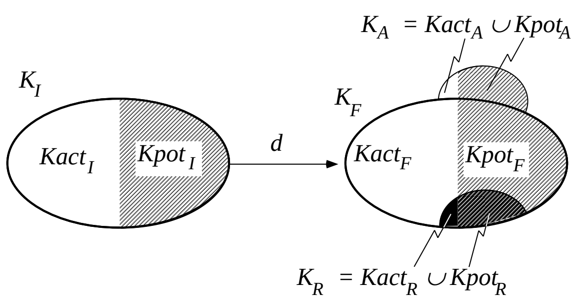
$$\langle A, B \rangle \oplus \langle C, D \rangle \triangleq \langle (A \cup C) \setminus (B \cup D), (B \cup D) \setminus (A \cup C) \rangle. \qquad \Box$$

On the basis of this definition, $I = \mathit{Iact} \oplus \mathit{Ipot}$ (*i.e.,* $I(d, K_I, K_F) = \mathit{Iact}(d, K_I, K_F) \oplus \mathit{Ipot}(d, K_I, K_F)$). In the definition, it is necessary to subtract $B \cup D$ and $A \cup C$ to avoid that a potential KI that becomes actual (or viceversa) increases $|K|$. As a matter of fact, even if we know that $A \cap B = C \cap D = \emptyset$, we do not know if $A \cap D$ and $B \cap C$ are empty or not, *i.e.,* if there is some potential KI that becomes actual (or viceversa), and these transformations should not increase $|K|$.

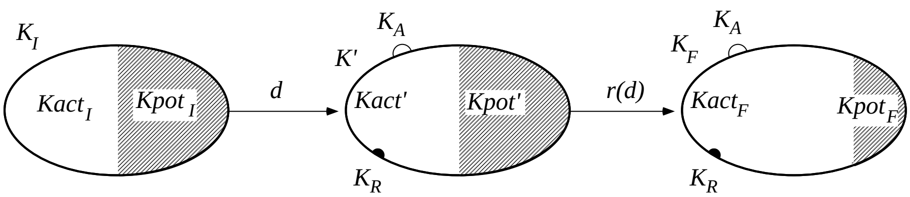
We can also define:



(a) $I = Iact \oplus Ipot$.



(b) A transition with a huge change in *Kpot*.



(c) *Kpot* and *Kact* in perception and restructuring transitions.

Figure 7: *Kact*, *Kpot*, *Iact*, and *Ipot*.

Definition 10 (Amount of *Iact* and *Ipot*) *The amount of Iact (or, alternatively, Ipot) is the number of added actual (potential) KIs plus the number of removed actual (potential) KIs:*

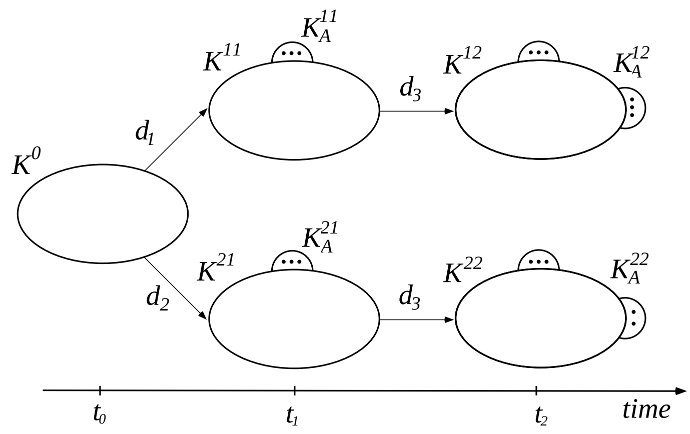
$$|Iact| = |\Delta Kact| \triangleq |Kact_A| + |Kact_R|;\ \ |Ipot| = |\Delta Kpot| \triangleq |Kpot_A| + |Kpot_R|. \qquad \Box$$

Now, it follows that: $|I| = |Iact \oplus Ipot| \leq |Iact| + |Ipot|$. The equality holds if and only if there are no transformations of *Kpot* in *Kact*, or viceversa.

4.4 Discussion

The actual-potential knowledge dichotomy permits to better understand the data-information inconsistency. Concerning remark $2(a)$ in Section 2.9, a huge datum d (a lot of bits) can lead to a small amount of changes in the *Kact* part of K, but to a huge amount of changes in the *Kpot* part (see Figure 7(b)). Concerning remark $2(b)$, if an agent has a high *Kpot*, a small datum can lead to a very different KS (by a heavy restructuring operation): a lot of *Kpot* becomes *Kact*. See Figure 7(c): the initial KS K_I is partitioned into $Kact_I$ and $Kpot_I$. The perceiving of a small datum d leads, after the perception transition, to adding a very small K_A and removing a very small K_R, without changing in a significant way the partition between actual and potential KSs: $Kact_I \cong Kact'$ and $Kpot_I \cong Kpot'$. After the restructuring transition, the actual and potential KSs change, since actual knowledge is derived from potential knowledge: $Kact' \ncong Kact_F$ and $Kpot' \ncong Kpot_F$ (and $Kact' \subsetneq Kact_F$ and $Kpot' \supsetneq Kpot_F$).

Now the data-information inconsistency appears less worrying: when an agent perceives a huge datum, a part of it can lead to a change in *Kpot*. When the agent receives a small datum (a few bits), the *Iact* received before the restructuring operation is a small amount, that can trigger a huge restructuration transforming a lot of *Kpot* into *Kact*. The human agents are conscious of their *Kact* only; therefore they perceive the modification in their *Kact* while



(a) Network of KSs.

$$d_1 \equiv \begin{cases} a+b=1 \\ b+c=2 \\ e+f=3 \end{cases}$$
$$d_2 \equiv \begin{cases} a+b=1 \\ c+d=2 \\ e+f=3 \end{cases}$$

(b) Two equations systems as data.

$$d_1 \equiv \begin{cases} a_1+a_2=1 \\ a_2+a_3=2 \\ \dots \\ a_i+a_j=i \\ \dots \\ a_{100}+a_{101}=100 \end{cases} \qquad d_2 \equiv \begin{cases} a_1+a_2=1 \\ a_3+a_4=2 \\ \dots \\ a_{2i-1}+a_{2i}=i \\ \dots \\ a_{199}+a_{200}=100 \end{cases}$$

(c) Two longer equations systems.

Figure 8: Knowledge pressure.

they do not perceive (or at least perceive in a more confused way) the modification in their *Kpot*. The data-information inconsistency is thus perceived in an amplified way, while it is relieved if *Kpot* is taken into account.

5 Knowledge pressure

The *Kact-Kpot* dichotomy helps to explain the data-information inconsistency described in Section 2.9. But there are still certain information phenomena observed (usually in an introspective way) by human beings that can not be explained on this basis and require further analysis. For instance, consider a simple agent, knowing elementary arithmetic. It has (Figure 8(a)) a KS K^0 at time t_0, and can receive either one of the two data d_1 and d_2 (equations systems) in Figure 8(b). d_1 and d_2 can lead to two KSs K^{11} and K^{21} at time t_1. The ${K_A}^{11}$ and ${K_A}^{21}$ are similar, and contain the three equations in d_1 and d_2, respectively. There seems to be no significant difference between K^{11} and K^{21}, and if the datum $d_3 \equiv \{e = 1\}$ were received at time t_1, the two KSs K^{12} and K^{22} at time t_2 would be obtained, with two new KIs in each of ${K_A}^{12}$ and ${K_A}^{22}$ for the values of e and f. However, if $d_3 \equiv \{a = 1\}$ is received at time t_1, the difference between K^{11} and K^{21} becomes important: ${K_A}^{22}$ contains only two KIs for $a = 1$ and $b = 0$, while ${K_A}^{12}$ contains three KIs for $a = 1$, $b = 0$, and $c = 2$. The difference between K^{11} and K^{21} is more evident if d_1 and d_2 are the equation systems in Figure 8(c) and $d_3 \equiv \{a_1 = 1\}$, since in this case ${K_A}^{12}$ has 101 KIs and ${K_A}^{22}$ has just 2 KIs.

The difference between the two KSs K^{11} and K^{21} in Figure 8(a) cannot be explained on the basis of the *Kact-Kpot* distinction introduced in the previous section: K^{11} and K^{21} have the same *Kpot* (that we can assume empty). The explanation is that the knowledge in a subKS of K^{11} is, in some sense, more concentrated, since a single datum can trigger a lot of restructuring or inferential operations: there is a subKS of K^{11} that has a high *knowledge pressure* ("tension" or "density" might be two alternative terms). On the other side, the knowledge in K^{21} is more spread, since there are many data that can trigger some restructuring or inferential operations. More precisely, in K^{21} there are 200 of these data, the values of $a_1 \cdots a_{200}$, while the corresponding data in K^{11} are 101, the values of $a_1 \cdots a_{101}$.

6 Knowledge conservation principle

It is commonly accepted that conservation principles for information and for knowledge do not exist. After a first analysis, this seems not questionable. Human beings reason, think, and find (discover, invent) something "new". When an agent "tells" something to another agent, the knowledge seems duplicated. When a teacher gives a lesson, the knowledge in his mind is multiplied, since each of the students has (more or less) the same knowledge after the lesson. Finally, people do forget: where has the knowledge that was in their mind gone? In this section I instead show that, on the basis of the concepts presented in the previous sections, it is possible to define a *conservation principle for knowledge*, analogous to conservation principles of physics (energy, momentum, and so on). The aim is to show that for each isolated system the total amount of knowledge is constant. I will present the knowledge conservation principle in the next subsections, with increasing generality.

6.1 Agent types

It is possible to imagine agents of different kinds, depending on their capabilities. The capabilities that an agent could have can be classified as follows:

Perception. Depending on how much an agent perceives of the world. There is a range of values (a continuum), with two extremes: a *complete perception* agent has a complete perception of the world, whereas a *no perception* one is completely isolated from his environment and thus his KS can change by means of inferential transitions only. In between we have *perceiving* agents with various perception levels. Human beings are not complete perception agents: we cannot perceive some kinds of data (infrared light, ultrasounds, etc.) and our perception of a datum depends on our previous KS.

Reasoning. Depending on how much the agent is capable of deriving new KIs from his KS, expliciting the potential KIs. Also in this case there is a range of values, with two extremes: an *omniscient* agent, capable of making actual all his potential knowledge (while a non-omniscient agent cannot derive something), and a *nonreasoning* agent, completely unable of deriving new KIs. In between there are many different agents, with different reasoning kinds: monotonic or non monotonic, deductive or inductive, and so on.

Memory. Depending on how much the agent is capable of keeping, conserving his KS. Again, we have a range of values. A *permanent memory* agent does never lose portions of his KS, while a *no memory* agent immediately loses his KS. In between we have various kinds of *volatile memory* agents. For instance, human beings sometime forget, while computers can be regarded as permanent memory agents (apart from failures).

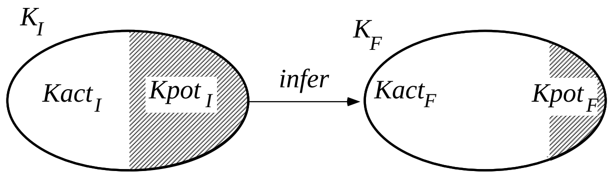


Figure 9: The transformation of *Kpot* in *Kact* during an inferential transition.

Therefore, a no perception, nonreasoning, and with permanent memory agent (*e.g.,* a book) will have a *static* KS (a KS that does not change as time goes on), while an agent with a *dynamic* KS must have some degree of either perception, inference, or volatile memory.

These agent kinds are important also because, in my opinion, a science of information can be found only if the right simplifications and abstractions are made. In physics, one speaks of ideal conditions: no friction, point mass, ideal environment with no external forces, and so on. I believe that for finding some good results in the information world we need to make similar simplifications and thus not to start from *human* information processing, but from *simple-agent* information processing. Some examples of the results that one can obtain are presented in the following.

6.2 A no perception, permanent memory agent

Let us take a no perception, permanent memory agent with a KS $K_I = \mathit{Kact}_I \cup \mathit{Kpot}_I$. The total amount of knowledge of the agent is $|K_I| = |\mathit{Kact}_I| + |\mathit{Kpot}_I|$. If there is an inferential transition transforming *Kpot* into *Kact* (see Figure 9) and leading to a KS $K_F = \mathit{Kact}_F \cup \mathit{Kpot}_F$, then we can postulate that $|\mathit{Kact}_F| > |\mathit{Kact}_I|$ (the actual knowledge increases) and $|\mathit{Kpot}_F| < |\mathit{Kpot}_I|$ (the potential knowledge decreases). In a more restrictive way, no knowledge is actually *created*, there is just a *transformation* process, and the amount of (actual and potential) knowledge in the KS of a no perception agent cannot increase.

Postulate 1 (Knowledge conservation principle) *The knowledge of a no perception, permanent memory agent agent remains constant as time goes on:*

$$\mathit{Kpot}_F \cup \mathit{Kact}_F = \mathit{Kpot}_I \cup \mathit{Kact}_I. \qquad \Box$$

This postulate implies that $|K(t)| = \mathit{const}$ (or $\frac{d|K(t)|}{dt} = 0$) and $\Delta|K| = \Delta|\mathit{Kpot}| + \Delta|\mathit{Kact}| = 0$ (explicitly resembling the “action and reaction” law of dynamics).

6.3 A complete perception, permanent memory agent

Let us consider an agent with a KS $K_I = \mathit{Kpot}_I \cup \mathit{Kact}_I$ (see Figure 7(a)). If the agent receives a datum d that leads to a transition to another KS $K_F = \mathit{Kpot}_F \cup \mathit{Kact}_F$, the information received is $I = \Delta K = \langle K_A, K_R \rangle = \langle K_F \setminus K_I, K_I \setminus K_F \rangle$, and we can not postulate that $|K_F| = |K_I| + |d|$ (whatever definition of $|d|$ we use), because, for instance, the agent might already know what d carries.

On the basis of the distinction between *Kpot* and *Kact*, we can only postulate that $|K_F| \leq |K_I| + |d|$. The equality can hold only if $|d|$ is, besides completely perceived by the agent, completely novel to him. If we take into account actual information only (as human beings tend, introspectively, to do), the equality does not hold, since some of the information received might be potential only. Thus, a huge datum d that leads to a little amount of *Iact* if one considers *Kact* only, actually might lead to a huge amount of I (*Ipot*) as soon as the *Kpot* is taken into account.

Such an agent is not an isolated system, thus his amount of knowledge can be not constant. Nevertheless, if we take into account a wider and closed system (containing the agent receiving the datum and the agent sending it), we can maintain that a conservation principle holds, as shown below.

6.4 An agency

Sometimes a creation of knowledge can seem to take place, for instance if an agent A_0 is communicating to a set of n agents $\{A_1, \ldots, A_n\}$ something that only he knows (some KIs in the KS of A_0 and not in the KSs of $A_1, \ldots, A_n$). After the communication act, *Kact* and *Kpot* of A_0 are constant, whereas *Kact* and *Kpot* of the other agents have changed. The higher n is, the higher the proliferation of knowledge.

Again, this is a short-sighted view. Let us consider the *agency* (the set of all the agents) $\mathcal{A} = \{A_0, \ldots, A_n\}$. Now $\mathcal{A}$ is an agent itself: it acts socially, interacting with the world. But if $\mathcal{A}$ is an agent, it has a KS, with actual and potential parts, and when A_0 communicates to $A_1, \ldots, A_n$, the *Kact* of $\mathcal{A}$ increases, while the *Kpot* of $\mathcal{A}$ decreases. Postulate 1 holds in this case too.

Let's take another example: if I receive a Japanese text on paper, my KS does not change (apart from details). If I also have a Japanese-English dictionary, the agency made up by myself, the Japanese text, and the dictionary has the potential knowledge of what the translated text means to me.

If an agency is made up of permanent memory agents and completely isolated, its *Kpot* tends to become actual. When all the agents have the same KS, nothing can change. It is something like the second principle of thermodynamics: the entropy of a closed system tends to increase, and such a system will end up in a "thermic dead". This might mean that we are going towards an "information dead", but the situation is not so simple, since new agents are continuously created in the real world.

6.5 A forgetting agent

The conservation principle just presented does not hold in some situations, since an agent may *forget* some of his subKSs. If the agent is isolated, then knowledge is lost, but this is an ideal case (as soon as something examines the agent, it is no more isolated). If we have an agent A_0 in an agency $\mathcal{A} = \{A_0, A_1\}$, as soon as A_0 forgets a KI k, the *Kpot* of $\mathcal{A}$ increases, since A_1 can tell k to A_0: there is no knowledge destruction, but only transformation from *Kact* (of A_0) into *Kpot* (of $\mathcal{A}$).

7 Conclusions and future work

This research is still at a preliminary stage, and there are many future developments.

Infinite KSs should perhaps be taken into account. One might wonder if there exists some agent in the real world with an infinite number of KIs in the actual part of his KS, or even in the potential part of his KS. A theorem prover seems to be such an agent, with an infinite amount of potential knowledge, though it needs an infinite amount of time for having an infinite number of actual KIs. Besides discussing if in the real world there exist infinite KSs, an analysis of the cases that are left out by restricting ourselves to the finite case should be done, and the effectiveness of the approximation obtained with the finite case should be evaluated. Also the problems about existence and univocity of the prerequisite KS K_P (Section 3) deserve further attention.

The scenario presented in this paper should be enriched in order to include into the description some new items: a more detailed analysis of what is inside the KSs (the dichotomies knowledge *vs.* metaknowledge [12], implicit *vs.* explicit knowledge, and other ones should be taken into account; also the links should be analyzed more in depth, as they seem to play an important role in the transitions between KSs); a more dynamic vision of KSs (in this paper, I have preferred to define static KSs in order to avoid the problems related to logical omniscience [12], and this is the reason for putting the inferences outside the KSs; the alternative way of including the inferences inside the KSs—and thus take into account the area of belief revision [10]— should be considered); and the intention of an agent [7] (the aims and goals play a crucial role in the interpretation of a datum; they might shed light on the above recalled existence and univocity problems of K_P; and they are dynamic, *i.e.*, they can change when a transition between KSs takes place).

The subjectivity of epistemic information might lead to important changes to the data transmission field. Source encoders might add semantic redundancy for safer transmissions; the data processing theorem might no longer hold; if a sender agent has to send some huge data to a receiver, and if the sender knows something about the KS of the receiver, the transmission can be done more effectively (what is already known by the received has not to be sent).

Many questions are still open. Data themselves might be considered as agents, thus having their own KSs: does this answer the problem of measuring the amount of a datum $|d|$ (it could be defined as the quantity of the KS of the datum)? The links between the theory sketched here and physics should be clarified: is there something analogous to heat or friction? What happens in the “agency of all the agencies”? Is its knowledge constant? What happens when new agents are created? How do these ideas relate to the field of agent computing? How do they affect the human computer interaction research? Do they support the search for more personalized user interfaces and systems? Is it possible to fully formalize this work, looking for an axiomatic theory? Is it possible to experimentally verify these ideas? Since the KS of an agent observing another agent may change (thus leading to a change in the agency made of by the two agents), do we have some kind of “knowledge indetermination principle”?

As a final remark, let’s note that if the ideas presented in this paper are correct, the two worlds of information and matter seem to be described by similar laws. What does this mean with respect to the Cartesian mind-body dualism? Do they support the view that mathematical descriptions of the nature are consistent with our perceptions only because we invented mathematics using our mental categories, and the same categories are used for perceiving the nature with our senses?

Acknowledgments

Many thanks to Antonello Conti, my brother Marco, and three anonymous referees (especially one of them!) for many useful comments.

References

[1] J. Barwise. *The Situation in Logic*. CSLI Lecture Notes 17, Stanford, CA, 1989.

[2] J. Barwise and J. Perry. *Situations and attitudes*. MIT Press, Cambridge, MA, 1983.

[3] G. Bateson. *Mind and Nature — A Necessary Unity*. Dutton, E. P., 1979.

[4] L. Birnbaum. Rigor mortis: a response to Nilsson's "Logic and artificial intelligence". *Artificial Intelligence*, 47:57–77, 1991.

[5] B. C. Brookes. The foundations of information science. Part I. Philosophical aspects. *Journal of Information Science*, 2:125–133, 1980.

[6] W. J. Clancey. *Situated Cognition—On Human Knowledge and Computer Representations*. Cambridge University Press, Cambridge, UK, 1997.

[7] K. Devlin. *Logic and Information*. Cambridge University Press, Cambridge, UK, 1991.

[8] A. Dix, J. Finlay, G. Abowd, and R. Beale. *Human-computer interaction*. Prentice-Hall, New York, 1993.

[9] F. Dretske. *Knowledge and the Flow of Information*. Bradford Books, MIT Press, 1981.

[10] P. Gärdenfors. *Knowledge in Flux – Modeling the Dynamics of Epistemic States*. MIT Press, London, 1988.

[11] H. Gardner. *The mind's new science: A history of the cognitive revolution. With a new epilogue by the author: Cognitive science after 1984*. Basic Books, New York, 1987.

[12] M. R. Genesereth and N. J. Nilsson. *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann Publishers, Los Altos, California, 1987.

[13] J. Hintikka. On semantic information. In J. Hintikka and P. Suppers, editors, *Information and Inference*, pages 3–27. D. Reidel publishing company, Dordrect-Holland, 1970.

[14] G. E. Hughes and M. J. Cresswell. *An Introduction to Modal Logic*. Methuen and Co. Ltd., London, 1968.

[15] F. Lehner and R. Maier. Can information modelling be successful without a common perception of the term "information"? In H. Kangassalo, J. F. Nilsson, H. Jaakkola, and S. Ohsuga, editors, *Information Modelling and Knowledge Bases VIII*, pages 181–198. IOS Press, Amsterdam, 1997.

[16] M. Li and P. Vitanyi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer Verlag, New York, 1997.

[17] H. R. Maturana and F. J. Varela. *The tree of knowledge*. Shambala, Boston and London, 1992. ISBN 0-87773-642-1.

[18] S. Mizzaro. La conoscenza in intelligenza artificiale. In *Proceedings of the annual conference AICA'95*, volume II, pages 1066–1073, 1995. In Italian. Translation of the title: "Knowledge in artificial intelligence".

[19] S. Mizzaro. A cognitive analisys of information retrieval. In P. Ingwersen and N. O. Pors, editors, *Information Science: Integration in Perspective — Proceedings of CoLIS2*, pages 233–250, Copenhagen, Denmark, Oct. 1996. The Royal School of Librarianship. Paper awarded with the "CoLIS2 Young Scientist Award".

[20] S. Mizzaro. Towards recursive models—A computational tool for the semantics of temporal presuppositions and counterfactuals in natural language. *Informatica - An International Journal of Computing and Informatics (ISSN 0359-5596)*, 21(1):59–77, Mar. 1997. The Slovene Society Informatika, Ljubljana, Slovenia.

[21] S. Mizzaro. How many relevances in information retrieval? *Interacting With Computers*, 10(3):305–322, June 1998. ISSN: 0953-5438. Elsevier, The Netherlands. Paper awarded with the Informer (British Computer Society IR Group newsletter) 'Best Student Paper in IR', a prize for the best paper by an European student in the period 1st November 1996 - 1st November 1997.

[22] N. J. Nilsson. Logic and artificial intelligence. *Artificial Intelligence*, 47:31–56, 1991.

[23] S. Russel and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River, NJ, 1995.

[24] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. Journal*, (27):379–423, 623–656, 1948. http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html.