ON KNOWLEDGE EVOLUTION: ACQUISITION, REVISION, CONTRACTION

Eliezer L. Lozinskii
Institute of Computer Science
The Hebrew University, Jerusalem 91904, Israel.
e-mail: lozinski@cs.huji.ac.il

Abstract. We consider evolution of Knowledge Bases caused by a sequence of basic steps of acquisition of a new information, either consistent or inconsistent with the original system. To make this process comply with the Principle of Minimal Change, a special *evidence metric* is introduced for measuring distance between states of knowledge. Then a novel semantics of knowledge bases is developed suggested by the *heuristics of weighted maximally consistent subsets* . The latter is efficiently applied to the processes of consistent and inconsistent acquisition of knowledge, belief revision and contraction.

1. Introduction

Consider a Knowledge Base S describing a world W . To be specific, let S be expressed in a First Order language, and regarded as a set of well-formed consistent formulae or, interchangeably, as a conjunction of all formulae of the set. We adopt the assumption that the real world is definite and non-contradictory. So, ontologically, in the reality any sentence is definitely either true or false, but never both true and false. On the other hand, a logic system S is an epistemic entity reflecting our current knowledge about the world. So, a main characteristic of S corresponding to natural limitations of human cognition is its incompleteness which means that there exist formulae ϕ such that neither ϕ nor $\neg\phi$ follows from S . On the other hand, a system may even be inconsistent, since humans (perhaps, as well as robots) may hold contradicting beliefs, e.g., ψ and $\neg\psi$ at the same time.

A knowledge base S grows by acquiring new information δ about the world. The process of growth of a knowledge base S raises the problem of determining its state S' resulting from accommodation of δ , such that many important questions should be asked and answered in this course. In particular, Is δ consistent with S or contradicts it? In the latter case, is the contradiction detected at the time of first acquisition of δ or turns out later during a subsequent usage of the knowledge base? What is the credibility of δ relative to the contradicting part of S ? Should δ be accepted or rejected? Does δ assert the knowledge stored in S or deny a part of it? In the latter case, which fragment of the knowledge should be removed from S to accommodate δ while preserving as much of the prior knowledge as possible? Should consistency of the system be restored or the contradiction of δ can (at least temporarily) be tolerated by the reasoning process?

The following special cases are characteristic of the process of knowledge evolution (they are displayed in Figure 1.1; $SAT(S)$ means that S is satisfiable).

-

BOXA: box wid 1 ht .5 "$SAT(S)$" box at center of BOXA wid .9 ht .4 move right 3 BOXB: box wid 1 ht .5 " $\neg SAT(S)$" box at center of BOXB wid .9 ht .4 # loop 1 line from BOXA.t up .6 left .4 spline up .15 left .1 then up .15 arc cw line right "$SAT(S \cup \{\delta\})$" above "1" below arc cw spline down .15 then down .15 left .1 arrow down .6 left .4 # loop 2 line from BOXA.ne up .6 right .4 spline up .15 right .1 then up .15 arc cw line right 1.5 " $\neg SAT(S \cup \{\delta\})$" above .cp 0 "2" below arc cw spline down .15 then down .15 right .1 arrow down .6 right .4 #loop 5 line from BOXB.t up .6 left .4 spline up .15 left .1 then up .15 arc cw line right " $\neg SAT(S \cup \{\delta\})$" above .cp 0 "5" below arc cw spline down .15 then down .15 left .1 arrow down .6 left .4 # loop 4 line from BOXA.b down .6 right .4 spline down .15 right .1 then down .15 arc cw line left "4" above "$CONTRACT(S, \delta)$" below arc cw spline up .15 then up .15 right .1 arrow up .6 right .4 # loop 3 line from BOXB.sw down .6 left .4 spline down .15 left .1 then down .15 arc cw line left 1.5 "3" above "$MEND(S, \delta)$" below arc cw spline up .15 then up .15 left .1 arrow up .6 left .4 # loop 6 line from BOXB.b down .6 right .4 spline down .15 right .1 then down .15 arc cw line left "6" above "$CONTRACT(S, \delta)$" below arc cw spline up .15 then up .15 right .1 arrow up .6 right .4

Figure 1.1. Knowledge base evolution.

(1) The original system S is satisfiable, and δ is consistent with S. In this case the new state of the system $S' = S \cup \{\delta\}$ is also satisfiable.

(2) S is satisfiable as well as δ, but δ is inconsistent with S, so, the new state $S' = S \cup \{\delta\}$ becomes unsatisfiable. Although inconsistency of a knowledge base is a most undesirable state, it should for many reasons be considered as a possible case. First, despite its inconsistency with S, δ may present an information that is more credible than the contradicting part of the system, and so, should be accommodated by S as a new reliable knowledge. Second, checking satisfiability of a large system is an extremely hard computational task, so, inconsistency of δ with S may not be detected, at least, at the moment of its acquisition. Third, the system may deliberately be "polluted" with an inconsistent data by some hostile activity. Besides, an inconsistent state of $S \cup \{\delta\}$ can be regarded as an intermediate stage in the process of restoring consistency of the system.

(3) S has been polluted by δ such that $S \cup \{\delta\}$ is unsatisfiable, but now consistency of the system must be restored by transforming $S \cup \{\delta\}$ to a satisfiable system T that preserves as "much" of the knowledge of S as possible. The quotation marks around *much* reflect the need to define properly the quantity of knowledge in S, and the difference between S and T in order to determine T.

(4) Suppose that the recent consistent state of knowledge asserted a sentence δ such that $S \models \delta$, but it has been newly discovered that truth of δ in the world is rather uncertain, and should not be believed. So, S should be transformed to a new state T such that $T \not\models \delta$. Let us say that δ should be *retracted* from S.

(5) Another special case of growing knowledge is that of an unsatisfiable system S acquiring a sentence δ. The resulting state $S \cup \{\delta\}$ remains unsatisfiable as well.

(6) A consistent subset S' of an unsatisfiable system S may assert a sentence δ that has to be retracted from S, so, S' has to be transformed to T such that $T \not\models \delta$, similarly to the case considered in (4). Generally, $(S - S') \cup T$ remains unsatisfiable, however, in certain cases retraction of δ may restore consistency of S. In these cases the process is similar to that of (3).

2. Principle of Minimal Change

It is commonly accepted that a proper process of knowledge growth should be governed by the *Principle of Minimal Change* in the sense that given a state S of knowledge, and a new information δ, the knowledge state S' resulting from accommodation of δ by S should differ from S as little as possible. So, an implementation of this principle requires a clear definition of the notion of difference between knowledge bases. The value of this difference is usually quantified by placing knowledge bases in a metric space, and characterizing the difference between the systems by the distance between their representations in the space.

Many existing approaches [7, 20, 39, 43, 44] determine the distance between knowledge bases by estimating a semantic difference between them. Assume that S is presented as a set of first-order formulae, and is logically equivalent to a conjunction of all these formulae. Then the meaning of S is determined by the set of its models $MOD(S)$. So, *semantically*, the difference between two states of knowledge, S and S', is determined by the distance in a certain space between the sets of models $MOD(S)$ and $MOD(S')$.

On the other hand, every separate formula of S presents in its own right a piece of knowledge about the world acquired from some source. The set of formulae constituting S determines a specific syntactic form of this state of knowledge, while each formula has its own meaning and importance. So, *syntactically* the difference between S and S' is determined by the distance between the corresponding sets of formulae measured in a proper way.

In the sequel we develop an approach to the estimation of difference between knowledge bases taking into account both aspects: the semantic and the syntactic one.

3. Model space

Given a first-order system S , let the *ground base* of S be a set of all ground atomic formulae of S defined over a given domain D :

$$GBase(S) = \{ A_1, A_2, \ldots \} .$$

Let us first consider only systems with finite ground bases, and then look for a way of relaxing this limitation (see Section 12).

Let N be the cardinality of $GBase(S)$. Consider a N-dimensional unit cube, the *model space* of S , and let every model $m_i \in MOD(S)$ be represented by a vertex v_i of the cube with coordinates $(x_{i1}, \ldots, x_{iN})$ such that for all $j = 1, \ldots, N$:

$$x_{ij} = \begin{cases} 1 & if \quad m_i \models A_j \\ 0 & otherwise . \end{cases} \tag{3.1}$$

Let $V(S)$ denote a set of all nodes representing the models of S .

Trying to express a semantic difference between two logic systems, we intend to define the distance between the two sets of their models represented by sets of nodes in the corresponding model space. As the meaning of a system S is determined by the set of its models, let each node of $V(S)$ representing a model $m_i \in MOD(S)$ be assigned a certain "mass" μ_i according to the semantic significance of m_i . Then it appears reasonable to choose the *mass center C(S)* of the set $V(S)$ as the single representative of the set of models of S . So, for coordinates $y_1, \ldots, y_N$ of $C(S)$:

$$y_j = \frac{\sum_{v_i \in V(S)} x_{ij} \cdot \mu_i}{\sum_{m_i \in MOD(S)} \mu_i} , \quad for \quad all \quad j = 1, \ldots, N . \tag{3.2}$$

Then the semantic difference between two logic systems is determined by the distance between the mass centers of their sets of models (measured due to a special metric presented in Section 5).

To simplify the analysis, in the next sections we assume that all models of a given system have equal semantic significance, and so,

$$y_j = \frac{\sum_{v_i \in V(S)} x_{ij}}{|MOD(S)|} , \quad for \quad all \quad j = 1, \ldots, N . \tag{3.3}$$

However, in Section 4 we shall relax this restriction, and show how different masses can be accommodated in the reasoning process.

We shall assume temporarily that the set $MOD(S)$ is finite, but in Section 12 will consider systems with infinite sets of models as well.

It turns out that $C(S)$ carries a deep semantic meaning. To reveal the latter, let us consider a system S describing a world W, and a sentence about W expressed by a formula ϕ. If ϕ is true in all models of S, then ϕ is a logical consequence of S, $S \models \phi$, and so, we conclude that ϕ is true in the reality. If ϕ is true in no model of S (i.e., $S \models \neg\phi$) then we say that ϕ is false in W. But if neither ϕ nor $\neg\phi$ follows from S, what can we conclude about the truth value of ϕ in W? Indeed, S does not give a definite answer, but it contains some incomplete information concerning ϕ that can be interpreted as an evidence of truth of ϕ. So, the problem is how to evaluate the *evidence* $E(S, \phi)$ of ϕ provided by S. The next section presents an approach that evaluates the evidence and shows its relationship with the notion of mass center.

4. Evidence and beliefs

Consider a system S describing a real world W that can exist in different states. The ground atomic formulae of $GBase(S)$ correspond to certain basic features of W. If W is rich enough, then S contains usually only a partial information of the world. This incompleteness of knowledge is twofold. On one hand, some relationships among facts and events of W are not known in S, and so, the set $MOD(S)$ of all models of S is a proper superset of the set of those combinations of the world features (described by $GBase(S)$) that appear in the reality in different states of W. If this containment holds indeed, we say that S describes W faithfully. In this case every formula ϕ that follows from S (and so, is true in all models of S) is true also in W, that is, in all its real states. On the other hand, some features of W may not be presented in S at all, and hence, a single model of S may represent several real states of the world.

So, a Knowledge Base S presents to its users an *image* of the real world W created by the set $MOD(S)$ in such a way that each model of S represents one state w of the world which is *possible* according to the knowledge embodied in S. Let $\{w\}$ denote the set of these *possible worlds*. Then there is a bijection between $MOD(S)$ and $\{w\}$.

Suppose now that we wish to find out the truth of a sentence ϕ in the present state of W. If $S \models \phi$ then ϕ is true in every state of W. But if the information in S about ϕ is incomplete, then all we can get from S is an answer to the question "What is *more likely*, that ϕ is true or false in W?". Should a probability distribution of states w be known, the question could be answered using the probability theory. Unfortunately, an exhaustive statistics of possible worlds is usually either unavailable or unreliable.

As we would not like to give up the task, we begin with a simplifying assumption that all possible worlds are equiprobable. The idea to assume a uniform probability distribution in the absence of a reliable statistics goes back to works of Keynes [21] and Carnap [5, 6]. It had later been adopted by many researchers [2, 19]. In practice, however, different possible worlds may have different probabilities. With this in mind, we show at the end of this section how the uniformity assumption can be relaxed while preserving its main results. Under this assumption, the more models of S assert ϕ (and so, the more numerous are possible worlds in which ϕ is true), the stronger should be the evidence that ϕ is likely to be true in an arbitrary state w. These observations suggest the following definition of evidence provided by a satisfiable system S for a formula ϕ.

Definition 4.1. [27] *Evidence.*

$$E(S, \phi) = \frac{|MOD(S \cup \{\phi\})|}{|MOD(S)|} \ . \tag{4.1}$$

The notion of evidence is very close to that of degree of confirmation introduced by Carnap [5].

By expressions (3.1), (3.3), (4.1), for $GBase(S) = \{A_1, ..., A_N\}$, and for all coordinates $y_1, ..., y_n$ of $C(S)$:

$$y_i = E(S, A_i) \ . \tag{4.2}$$

This reveals a strong semantic significance of the notion of mass center of a system S as a representation of the set of evidences provided by S for all atomic formulae of its ground base.

Further, if $E(S, \phi) > 0.5$ then by (4.1), ϕ is true in a majority of models of S, and hence, in a majority of possible worlds. In this case, under the equiprobability assumption, ϕ is more likely to be true than false in an arbitrary state of the world. However, in the reality possible worlds may not be distributed uniformly, and this, of course, must be taken into consideration in computation of evidence. Although an exhaustive statistics required by the existing methods of probabilistic reasoning (e.g., [31, 34]) is usually either not available or not satisfactorily reliable, there often is some restricted statistical information about rather a small subset of objects and events of the world. Suppose that prior probabilities of certain states of the world are known (as all the states are mutually exclusive, their mutual conditional probabilities equal 0). Let M denote a set of models of S representing these states such that $\rho(m)$ is the prior probability of $m \in M$, and $p(M) = \sum_{m \in M} \rho(m)$. Then assuming that all the states of the world with unknown probabilities are equiprobable, we get the following expression for evidence of ϕ in S:

$$E(S, \phi) = (1 - p(M)) \frac{|MOD(S \cup \{\phi\}) - M|}{|MOD(S) - M|} + \sum_{m \in (M \cap MOD(S \cup \{\phi\}))} \rho(m) \ . \tag{4.3}$$

If no prior probabilities of possible worlds are known such that $M = \emptyset$, then expression (4.3) becomes identical to that of Definition 4.1. In other special case, if prior probabilities are given not for all possible worlds, but for all of them asserting ϕ, such that $MOD(S \cup \{\phi\}) \subset M \subset MOD(S)$, then the evidence $E(S, \phi)$ amounts to the probability of ϕ. If $M = MOD(S)$ then reasoning can be performed by purely probabilistic methods.

Given a probability distribution of possible worlds, we may "load" every node $v_i \in V(S)$ with a mass μ_i proportional to the probability of the possible world represented by v_i. In particular, given a partial statistics discussed above,

$$\mu_i = \begin{cases} \rho(m_i) & \text{if} \quad m_i \in M \\ \dfrac{1 - p(M)}{|MOD(S) - M|} & otherwise\ . \end{cases} \tag{4.4}$$

In this case expression (4.2) holds for any distribution of possible worlds, namely: the evidences provided by a system S for all its ground atomic formulae are coordinates of its mass center.

The abovementioned considerations yield the following principle of producing plausible beliefs.

Principle of reasoning by evidence. Believe that ϕ is true if $0.5 < E(S, \phi) \leq 1.0$, but false if $0 \leq E(S, \phi) < 0.5$. If $E(S, \phi) = 0.5$ then both beliefs are equally credible.

5. Evidence metric

An image of the world induced by a knowledge base S is determined by the beliefs sanctioned by S, and the latter are strongly related to the evidence provided by S. Given two systems, S and T (over the same ground base), we intend to figure out the semantic difference between them by estimating differences between the evidences provided by S and T for the same formulae. As shown in Section 4, a set of evidences of all ground atoms of a system determines its mass center in the model space. So, we have to evaluate the distance between $C(S)$ and $C(T)$ due to a proper metric defined over the model space.

Definition 5.1. A *displacement* $\Delta(S, \phi)$ of a formula ϕ in S:

$$\Delta(S, \phi) = \begin{cases} 1 & \text{if} \quad E(S, \phi) > 0.5 \\ 0 & \text{if} \quad E(S, \phi) = 0.5 \\ -1 & \text{if} \quad E(S, \phi) < 0.5\ . \end{cases} \quad \blacksquare$$

Consider the model space of S and T, and let us define the distance $d_A(S, T)$ between $C(S)$ and $C(T)$ along an axis $A \in GBase$, where $GBase$ is the common ground base of S and T. The distance function should possess the following reasonable properties:

(i) $d_A(S, T) = d_A(T, S)$;

(ii) If $E(S, A) = E(T, A)$ then $d_A(S, T) = 0$;

(iii) If $E(S, A) = 1$ and $E(T, A) = 0$ then the distance reaches its maximum, $d_A(S, T) = 1$ (as the model space is a unit cube);

(iv) If A has the same displacement in both S and T, this means, due to the Principle of reasoning by evidence (Section 4), that both $E(S, A)$ and $E(T, A)$ suggest the same belief of A. So, the evidences differ only quantitatively. In this case

$$d_A(S, T) = c_1 \cdot |E(S, A) - E(T, A)|\ ,$$

where c_1 is a constant;

(v) If A has opposite displacements in S and T, e.g., $\Delta(S, A) = 1$, but $\Delta(T, A) = -1$, this means that S suggests that A is true, while T supports just the opposite belief. In this case $E(S, A)$ and $E(T, A)$ are dissimilar not only quantitatively (as in paragraph (iv)), but qualitatively as well. As S and T sanction opposite beliefs of A, the distance between S and T with respect to A has to be larger than a distance between them with respect to any atom $B \in GBase$ for which $\Delta(S, B) = \Delta(T, B)$, since the latter means that both S and T suggest the same belief of B. So, let

$$d_A(S, T) = c_1 \cdot |E(S, A) - E(T, A)| + c_2 .$$

Then for all A with opposite displacements, $E(S, A) \neq E(T, A)$, hence,

$$d_A(S, T) > c_2 .$$

For all B with identical displacements, $|E(S, B) - E(T, B)| < 0.5$, hence by (iv),

$$d_B(S, T) < \frac{c_1}{2} .$$

So, if $c_2 = \frac{c_1}{2}$ then for all $A, B \in GBase$ such that $\Delta(S, A) \cdot \Delta(T, A) = -1$ and $\Delta(S, B) = \Delta(T, B)$,

$$d_A(S, T) > d_B(S, T) ,$$

in accord with the abovementioned intuition.

(vi) If $\Delta(S, A) \neq 0$, but $\Delta(T, A) = 0$, then S suggests a certain belief of A, while T is most uncertain regarding A. So, the dissimilarity determined by $E(S, A)$ and $E(T, A)$ is greater than in the case considered in (iv), but less than in that of (v):

$$d_A(S, T) = c_1 \cdot |E(S, A) - E(T, A)| + c_3 ,$$

where $0 < c_3 < c_2$. Besides, it is reasonable to make the distance determined by evidences of 1 and 0 to be twice as large as that determined by 1 and 0.5 (or by 0.5 and 0). This is satisfied by assigning $c_3 = \frac{c_2}{2}$.

Finally, from (iii) and (v) we get $c_1 = \frac{2}{3}$, and so, $c_2 = \frac{1}{3}$, $c_3 = \frac{1}{6}$, yielding the following definition of $d_A(S, T)$.

Definition 5.2. *Evidence distance* between systems S and T *with respect to* a formula A:

$$d_A(S, T) = \frac{2}{3} |E(S, A) - E(T, A)| + c ,$$

where

$$c = \begin{cases} 0 & \text{if} \quad \Delta(S, A) = \Delta(T, A) \\ 1/3 & \text{if} \quad \Delta(S, A) \cdot \Delta(T, A) = -1 \\ 1/6 & \textit{otherwise} . \quad \blacksquare \end{cases}$$

Now we can complete a definition of *evidence metric* in the model space by defining the *evidence distance* between two systems, and showing that it satisfies the triangle inequality. The evidence distance is normalized in the interval [0, 1].

Definition 5.3. *Evidence distance* $D(S, T)$ between systems S and T over the same ground base $GBase$:

$$D(S, T) = \frac{\sum_{A \in GBase} d_A(S, T)}{|GBase|} . \qquad \blacksquare$$

Theorem 5.1. Distance function D (Definition 5.3) satisfies the triangle inequality, that is, for any systems S_1, S_2, S_3 over the same ground base,

$$D(S_1, S_3) \le D(S_1, S_2) + D(S_2, S_3) .$$

Proof. First, we show that for all $A \in GBase$, d_A satisfies the triangle inequality. With no loss of generality (by property (i)), let $E(S_1, A) \ge E(S_3, A)$. Then Definition 5.2 implies four different cases of displacement.

Case 1: $\Delta(S_1, A) = \Delta(S_3, A)$. Then $d_A(S_1, S_3) = \frac{2}{3}(E(S_1, A) - E(S_3, A));$

$$d_A(S_1, S_2) \ge \frac{2}{3}(E(S_1, A) - E(S_2, A)); \qquad d_A(S_2, S_3) \ge \frac{2}{3}(E(S_2, A) - E(S_3, A)).$$

Case 2: $\Delta(S_1, A) = 1$, $\Delta(S_3, A) = -1$. Then $d_A(S_1, S_3) = \frac{2}{3}(E(S_1, A) - E(S_3, A)) + \frac{1}{3}$.

If $\Delta(S_2, A) = 1$ then

$$d_A(S_1, S_2) \ge \frac{2}{3}(E(S_1, A) - E(S_2, A)); \qquad d_A(S_2, S_3) = \frac{2}{3}(E(S_2, A) - E(S_3, A)) + \frac{1}{3} .$$

If $\Delta(S_2, A) = 0$ then

$$d_A(S_1, S_2) = \frac{2}{3}(E(S_1, A) - 0.5) + \frac{1}{6} ; \qquad d_A(S_2, S_3) = \frac{2}{3}(0.5 - E(S_3, A)) + \frac{1}{6} .$$

If $\Delta(S_2, A) = -1$ then

$$d_A(S_1, S_2) = \frac{2}{3}(E(S_1, A) - E(S_2, A)) + \frac{1}{3} ; \qquad d_A(S_2, S_3) \ge \frac{2}{3}(E(S_2, A) - E(S_3, A)).$$

Case 3: $\Delta(S_1, A) = 1$, $\Delta(S_3, A) = 0$. Then $d_A(S_1, S_3) = \frac{2}{3}(E(S_1, A) - 0.5) + \frac{1}{6}$.

If $\Delta(S_2, A) = 1$ then

$$d_A(S_1, S_2) \ge \frac{2}{3}(E(S_1, A) - E(S_2, A)); \qquad d_A(S_2, S_3) = \frac{2}{3}(E(S_2, A) - 0.5) + \frac{1}{6} .$$

If $\Delta(S_2, A) = 0$ then

$$d_A(S_1, S_2) = \frac{2}{3}(E(S_1, A) - 0.5) + \frac{1}{6} ; \qquad d_A(S_2, S_3) = 0 .$$

If $\Delta(S_2, A) = -1$ then

$$d_A(S_1, S_2) = \frac{2}{3}(E(S_1, A) - E(S_2, A)) + \frac{1}{3} ; \qquad d_A(S_2, S_3) > \frac{2}{3}(E(S_2, A) - 0.5) + \frac{1}{6} .$$

Case 4: $\Delta_A(S_1, A) = 0$, $\Delta_A(S_3, A) = -1$ is similar to case 3.

Hence, in all cases

$$d_A(S_1, S_3) \le d_A(S_1, S_2) + d_A(S_2, S_3) .$$

This implies by Definition 5.3 that for all $A \in GBase$, D (as well as d_A) satisfies the triangle inequality. ■

It is worth noting that for complete logic systems having unique models, the evidence metric coincides with that of Hamming. Indeed, let S_1, S_2 be complete systems (over the same $GBase$) with unique models m_1, m_2 , respectively. The *Hamming distance* between S_1 and S_2 , $H(S_1, S_2)$, equals the number of ground atoms on which m_1 and m_2 disagree. Let $h(m_1, m_2)$ denote a set of all such atoms. Then for all $A \in GBase$,

$$d_A(S_1, S_2) = \begin{cases} 1 & \text{if} \quad A \in h(m_1, m_2) \\ 0 & \textit{otherwise} \end{cases}$$

and so,

$$D(S_1, S_2) = \frac{\sum_{A \in GBase} d_A(S_1, S_2)}{|GBase|} = \frac{|h(m_1, m_2)|}{|GBase|} = \frac{H(S_1, S_2)}{|GBase|} .$$

Hence, the evidence distance between two complete systems equals the normalized Hamming distance between them.

6. Consistent acquisition

Let us now study in turn different transitions of a growing Knowledge Base considered in the Introduction (cf. Figure 1.1).

If a satisfiable system S is augmented by a new data expressed by a formula δ consistent with S , then the resulting state of knowledge $S' = S \cup \{\delta\}$ is also satisfiable such that $MOD(S') \subseteq MOD(S)$. For any formula ϕ ,

$$E(S', \phi) = \frac{|MOD(S \cup \{\delta\} \cup \{\phi\})|}{|MOD(S \cup \{\delta\})|} . \tag{6.1}$$

If ϕ (or $\neg\phi$) is a logical consequence of S , then ϕ is true (false, respectively) in all models of S and S' . So, its evidence is not changed by acquisition of δ , such that $E(S', \phi) = E(S, \phi) = 1$ (0, respectively). However, if neither ϕ nor $\neg\phi$ follows from S , then the evidence of ϕ may be changed significantly, as shows the following example.

Example 6.1. Consider $S = \{ p_i \to q$ for $i = 1, 2, ..., n \}$, $\delta = \neg q$, $S' = S \cup \{\delta\}$. S has $2^n + 1$ models such that q is true in 2^n of them. So, for large n , $E(S, q)$ approaches 1, while $E(S', q) = 0$. ■

If δ is itself a logical consequence of S then $MOD(S') = MOD(S)$, and so, for any ϕ , $E(S', \phi) = E(S, \phi)$.

7. Inconsistent acquisition

If S is satisfiable, but δ is inconsistent with S, then $S' = S \cup \{\delta\}$ becomes unsatisfiable. As it is argued in Introduction, despite the unsatisfiability of S', it should for many reasons be still regarded as a source of knowledge of the world, and so, a way has to be found for extracting some reliable information from an inconsistent logic system.

Any knowledge base is designed and maintained to describe faithfully the ontologically consistent reality, so it is reasonable to assume that the inconsistency of S' indicates that S has been "polluted" with some wrong data, however, most of the system content still reflects the world truthfully. So, despite the pollution, there must be a meaningful subset of S', and in order not to lose correct data, we would like such a subset to be as large as possible.

Definition 7.1. *Mc-subset.* A subset σ of S is a *maximally consistent subset* (briefly, *mc-subset*) if σ is consistent, but for all formulae $\phi \in (S - \sigma)$, $\sigma \cup \{\phi\}$ is inconsistent. $MC(S)$ denotes a set of all mc-subsets of S. If S is consistent then it is its unique mc-subset. ■

Since only consistent subsets of the system can conform to the world, it appears reasonable to assume that *the semantics of a logic system is determined by that of its mc-subsets*. This principle goes back to the work of Rescher and Manor [38], and is widely adopted by researchers in the field [1, 3, 4, 10, 14, 23, 25, 30, 32, 40]. However, to be implemented, the principle requires a specific treatment, since different mc-subsets of an inconsistent system may suggest contradicting beliefs. The common approaches to semantics of mc-subsets reflect either a credulous or a skeptical way of reasoning. A *credulous* reasoner is ready to accept a belief that is true in *some* mc-subsets, losing the information contained in the rest of the system. On the other extreme, a *skeptical* reasoner believes only statements that are true in *all* mc-subsets of a given system, so, he/she can make no decision regarding numerous statements on which different mc-subsets disagree. Thus, both types of reasoners make no use of a significant part of the system's information, and, in particular, have no means for resolving contradictions between different parts of the same system.

We would like to develop an approach to semantics of an inconsistent system S that makes a *full* utilization of the knowledge contained in the system by taking into account *all* mc-subsets of S and evaluating the relative contribution of each of them to the system semantics. And we should also find a way to estimate plausibility of this approach, say, by comparing its decisions with beliefs suggested by a common sense intuition. Such a comparison is possible since certain cases are suitable for an intuitive judgment, like the following rather simplistic example.

Example 7.1. Consider the following consistent set of 5 clauses:

$$S = \{\ p_1,\ p_1 \to q,\ p_2,\ p_2 \to q,\ q \to r\ \}\ .$$

The unique model of S is

$$\mu = \{\ p_1,\ p_2,\ q,\ r\ \}\ .$$

Now, suppose that a new message claims that r is false:

$$\delta = \neg r \ .$$

A direct addition of δ makes the system inconsistent:

$$S' = \{ \ p_1, \ p_1 \rightarrow q, \ p_2, \ p_2 \rightarrow q, \ q \rightarrow r, \ \neg r \ \} \ .$$

Given S', let us try to justify intuitively a plausible belief concerning the truth of q. First, in S', there are two independent ways of proving q (namely, $\{ \ p_1, \ p_1 \rightarrow q \ \}$ and $\{ \ p_2, \ p_2 \rightarrow q \ \}$), but only one proof of $\neg q$ ($\{ \ q \rightarrow r, \ \neg r \ \}$). Second, if the new clause δ is more reliable than any old clause of S (for instance, because it is the most recent message, or has been received from a most credible source), even in this case q appears very likely to be true in S', since to prove $\neg q$ unambiguously, at least two clauses of S must be erroneous, but if only $q \rightarrow r$ is wrong then the truth of q is supported in S' as strongly as in S. So, due to this intuition, despite δ, S' still sanctions a belief that q is more likely to be true than false.

7.1. Heuristics of weighted mc-subsets

As the real world is supposed to be unambiguous in its properties, only consistent subsets of a knowledge base have a chance to conform to the reality, and so, are relevant to the process of producing plausible beliefs. This presumption has led to the idea (cf. [33]) that the meaning $< S >$ of an inconsistent system S should be determined by a disjunction of all its mc-subsets (each one of the latter considered as a conjunction of all its formulae):

$$< S > = \bigvee_{\sigma \in MC(S)} \ \bigwedge_{\psi \in \sigma} \psi \ . \tag{7.1}$$

Then

$$MOD(< S >) = \bigcup_{\sigma \in MC(S)} MOD(\sigma) \ ,$$

and hence, the evidence of a formula ϕ in S suggested by this approach is

$$E(< S > \ , \phi) = \frac{\sum_{\sigma \in MC(S)} |MOD(\sigma \cup \{\phi\})|}{\sum_{\sigma \in MC(S)} |MOD(\sigma)|} \ . \tag{7.2}$$

If this *disjunctive* approach is applied to a consistent system S augmented by a formula δ that is inconsistent with S, but more reliable than the formulae of S, then only those mc-subsets of $S \cup \{\delta\}$ should be considered which contain δ:

$$< S \cup \{\delta\} > = \bigvee_{\substack{\delta \in \sigma \\ \sigma \in MC(S)}} \ \bigwedge_{\psi \in \sigma} \psi \ . \tag{7.3}$$

The disjunctive approach indeed takes into consideration all relevant mc-subsets of a given system, but it grants all the subsets with the same role in determining the system's semantics. However, different mc-subsets may make different contribution to the image of the world inspired by the system. In particular, if this image should comply with the Principle of Minimal Change, then the closer to S is a mc-subset σ of $S \cup \{\delta\}$, the stronger should be its influence on plausible beliefs. This can be achieved by assigning a *semantic weight* $u(\sigma)$ to every mc-subset σ such that the closer to S is σ , the larger is $u(\sigma)$. If $D(\sigma, S) = 0$ then let us assign to $u(\sigma)$ its maximal normalized value of 1. If $D(\sigma, S) = 1$ then $u(\sigma)$ gets its minimal value. To normalize $u(\sigma)$ in the interval [0, 1], this minimal value could be 0, but since $u(\sigma)$ is deemed as a multiplicative factor, such an assignment would cause a total neglection of the information contained in σ . However, as any consistent subset of a system is supposed to contain some meaningful data, it appears reasonable to allow any mc-subset to make a non-zero contribution to the system semantics, even if this mc-subset is maximally remote from S at distance 1. If $D(\sigma, S) = 1$, this means that σ disagrees with S on all atomic formulae of its ground base, so, the larger is GBase, the larger is the disagreement, and so, the smaller should be the minimal value of $u(\sigma)$. A simple way to meet this condition is to make $u(\sigma)$ reciprocal to $|GBase|$ if $D(\sigma, S) = 1$. These observations lead to the following definition.

Definition 7.2. *Semantic weight* $u(\sigma)$ of a mc-subset σ of $S \cup \{\delta\}$:

$$u(\sigma) = 1 - \left(1 - \frac{1}{|GBase|}\right) \cdot D(\sigma, S) . \quad \blacksquare$$

Now we can compute evidence E_U sanctioned by mc-subsets weighted by $u(\sigma)$:

$$E_U(S \cup \{\delta\}, \phi) = \frac{\sum_{\sigma \in MC(S)} |MOD(\sigma \cup \{\phi\})| \cdot u(\sigma)}{\sum_{\sigma \in MC(S)} |MOD(\sigma)| \cdot u(\sigma)} . \tag{7.4}$$

Example 7.2. Consider again the system of Example 7.1:

$$S = \{ p_1, p_1 \rightarrow q, p_2, p_2 \rightarrow q, q \rightarrow r \} .$$

If $\delta = \neg r$ is more credible than every clause of S , then we should consider only the following mc-subsets of $S \cup \{ \neg r\}$ containing $\neg r$:

$$\sigma_1 = \{ p_1, p_2, q \rightarrow r, \neg r \} ,$$
$$\sigma_2 = \{ p_1, p_2 \rightarrow q, q \rightarrow r, \neg r \} ,$$
$$\sigma_3 = \{ p_1 \rightarrow q, p_2, q \rightarrow r, \neg r \} ,$$
$$\sigma_4 = \{ p_1 \rightarrow q, p_2 \rightarrow q, q \rightarrow r, \neg r \} ,$$
$$\sigma_5 = \{ p_1, p_1 \rightarrow q, p_2, p_2 \rightarrow q, \neg r \} .$$

Table 7.1 presents (in columns 1-5) evidences of p_1, p_2, q, r provided by $S, \sigma_1, ..., \sigma_5$. Columns 6,7 show for $\sigma_1, ..., \sigma_5$ their distances $D(\sigma, S)$ from S in the evidence space, and the corresponding semantic weights $u(\sigma)$. The content of columns 8-10 will be explained in the sequel.

-

Table 7.1. Mc-subsets of $S \cup \{\neg r\}$ (Examples 7.2, 7.3).

	Evidence p_1	p_2	q	r	$D(\sigma, S)$	$u(\sigma)$	$\lvert S - \sigma \rvert$	$w(\sigma)$	$contrib(\sigma)$
1	2	3	4	5	6	7	8	9	10
S	1	1	1	1					
σ_1	1	1	0	0	0.50	0.625	2	0.08	0.050
σ_2	1	0	0	0	0.75	0.438	2	0.08	0.035
σ_3	0	1	0	0	0.75	0.438	2	0.08	0.035
σ_4	0	0	0	0	1.00	0.750	2	0.08	0.020
σ_5	1	1	1	0	0.25	0.813	1	0.20	**0.163**

Table 7.2. Sets of weighted mc-subsets (Examples 7.2, 7.3).

	Evidence p_1	p_2	q	r	D
S	1.000	1.000	1.000	1.000	
$< S \cup \{\neg r\} >$	0.600	0.600	0.200	0	0.600
$(S \cup \{\neg r\})_U$	0.732	0.732	0.317	0	0.537
$(S \cup \{\neg r\})_{UW}$	0.818	0.818	**0.537**	0	**0.388**
$(S \cup \{\neg r\})_W$	0.692	0.692	0.385	0	0.538

Let us compute the evidence of p_1, p_2, q, r provided by $S \cup \{\neg r\}$ due to the disjunctive approach (expression (7.2)), and the heuristics of semantic weights (expression (7.4)). Table 7.2 shows these evidences in lines denoted by $< S \cup \{\neg r\} >$ and $(S \cup \{\neg r\})_U$, respectively. As a set of evidences represents a node in the evidence space, the last column of Table 7.2 gives a distance D of $< S \cup \{\neg r\} >$ and $(S \cup \{\neg r\})_U$ from S . It shows that $(S \cup \{\neg r\})_U$ is closer to S than $< S \cup \{\neg r\} >$, and even every particular evidence provided by $(S \cup \{\neg r\})_U$ is closer to S than that of $< S \cup \{\neg r\} >$. However, due to the Principle of reasoning by evidence (Section 4), both approaches suggest the same set of beliefs, namely, that p_1 and p_2 are true, but q and r are false. In the original system S the propositions p_1, p_2, q, r were all true. The change of the belief regarding r is easily acceptable, as the new message $\delta = \neg r$ is more credible than the knowledge in S . However, the belief that q is false contradicts the intuition developed in Example 7.1. So, the application of semantic weight $u(\sigma)$ has improved quantitatively the set of evidences relatively to the disjunctive approach, but does not yet fully satisfy the common sense intuition. The next section presents a further improvement of the heuristics of weighted subsets.

7.2. Syntactic weight

Given a system S , and looking for a proper representation of its meaning, many researchers choose the set of models of the system $MOD(S)$ [7, 23, 25, 43] or its theory $Th(S)$ [1, 13-15, 29, 30] (the latter is a set of all logical consequences of S , denoted also by $Cn(S)$). On the other hand, S is usually considered as a set of formulae such that each formula represents a separate piece of knowledge that may have come from a specific source, may have a distinct importance to a reasoner, and may be either credible, or uncertain, or unreliable, or even incorrect. Therefore the specific form in which the knowledge of S is expressed, is an important aspect of the system representation. As different sets of formulae may have the same set of models and the same theory, the syntactic form of a specific system must be taken into consideration when the system is changed. This attitude has been adopted by many other researchers [17, 32, 33].

Willard and Li [42] developed an approach presenting a compromise between semantic and syntactic considerations. First, they consider only a proper subset of $Cn(S)$ consisting of all formulae minimal under subsumption. Second, they assume that formulae explicitly given in S have priority over any derived formula. Among the formulae of S , any non-atomic formula may optionally be given a priority over any atomic formula. However, there still exist several different candidate subsets of the system that cannot be compared and discriminated under these assumptions.

This section develops a measure of a relative syntactic value of every mc-subset making all of them comparable.

Consider now an inconsistent set of formulae S , and its mc-subset σ . The more formulae of S are contained in σ , the fuller σ represents the content of the system, and so, the greater must be the role of σ in determining the meaning of S . Let this role be evaluated by a *syntactic weight* of σ , $w(\sigma)$. The largest mc-subset of S may contain as many as $|S| - 1$ formulae, and there can be at most $|S|$ such mc-subsets. S may contain at most $\binom{|S|}{k}$ mc-subsets of size $|S| - k$. Due to the abovementioned observation regarding the relative role of subsets, any mc-subset of size $|S| - k$ is more significant for reasoning about S than any mc-subset of size $|S| - k - 1$, and it appears reasonable to assume that the set of all possible mc-subsets of size $|S| - k$ outweighs the set of all possible mc-subsets of size $|S| - k - 1$. As shown in [28], this leads to the following definition.

Definition 7.3. *Syntactic weight* $w(\sigma)$ of a mc-subset σ of S :

$$w(\sigma) = \frac{|S - \sigma|!}{|S|^{|S - \sigma|}} . \quad \blacksquare$$

It should be noted that since $w(\sigma)$ reflects the given syntactic form of S , its value may change if the syntactic representation of S is altered. For instance, the weight $w(\sigma)$ may change if some part of S is duplicated or if a logical consequence of S is added to the system. Such alterations do not affect semantic weights, as the system remains logically unchanged.

Taking into account both semantic and syntactic weights yields the following expression for evidence:

$$E_{UW}(S \cup \{\delta\}, \phi) = \frac{\sum_{\sigma \in MC(S \cup \{\delta\})} |MOD(\sigma \cup \{\phi\})| \cdot u(\sigma) \cdot w(\sigma)}{\sum_{\sigma \in MC(S \cup \{\delta\})} |MOD(\sigma)| \cdot u(\sigma) \cdot w(\sigma)} , \tag{7.5}$$

where S and δ are separately consistent, but $S \cup \{\delta\}$ is not.

Example 7.3. Let us now apply both $u(\sigma)$ and $w(\sigma)$ to the system $S \cup \{\neg r\}$ of Examples 7.1, 7.2, and to its mc-subsets $\sigma_1, ..., \sigma_5$. Table 7.1 (column 9) gives syntactic weights of the mc-subsets. Table 7.2 presents the set $(S \cup \{\neg r\})_{UW}$ of evidences computed due to expression (7.5), and the distance D between $(S \cup \{\neg r\})_{UW}$ and S . We notice that all evidences of $(S \cup \{\neg r\})_{UW}$ are closer to S than that of $(S \cup \{\neg r\})_U$. But most significant is that unlike $(S \cup \{\neg r\})_U$ and $< S \cup \{\neg r\} >$, the application of both weights yields a belief that q is true (as the evidence of q in $(S \cup \{\neg r\})_{UW}$ is greater than 0.5), which accords with S . It is worth noting that applying the syntactic weight alone would not be enough to satisfy the Principle of Minimal Change, as follows from the last line of Table 7.2, where the evidence of q in $(S \cup \{\neg r\})_W$ is 0.385 suggesting a belief that q is false.

To summarize the heuristics of weighted mc-subsets, let us say that this is merely one of possible heuristics, and so, it may be improved and possibly adapted to a specific domain of knowledge. As a general approach, the heuristics provides a sound way to recovering plausible beliefs from polluted inconsistent systems, as suggested by the following procedure.

Procedure Inconsistent_acquisition(S, δ, φ) ;
(* Given a consistent system S augmented by a consistent formula δ such that $S \models \neg\delta$, returns a plausible belief of a given formula ϕ . *)

(i) Determine a set $\{\sigma\}$ of *relevant* mc-subsets of $S \cup \{\delta\}$. If there exist certain *integrity constraints* C or an importance ordering PI (see Section 9) that must be satisfied in every state of knowledge, then each relevant mc-subset has to satisfy C and PI . In addition, if δ is more credible than any formula of S , then every relevant mc-subset must contain δ ;

(ii) Given a formula ϕ , compute its evidence provided by $S \cup \{\delta\}$ due to expression (7.5), where MC is a set of all relevant mc-subsets of $S \cup \{\delta\}$;

(iii) To produce a plausible image of the world, believe that a formula ϕ is true iff its evidence provided by $S \cup \{\delta\}$ is greater than 0.5.

8. Belief revision

The process of accommodating a contradictory message by a knowledge base has been extensively studied in many works [1, 7-15, 17, 18, 23-25, 33, 35, 36, 39-44]. In [1, 13, 14] it has been presented as a problem of *Belief Revision* or *Theory Change* . Namely,

given a consistent system S and a consistent formula δ such that $S \models \neg\delta$, transform S into a consistent system T containing δ , and preserving as much as possible of the knowledge stored in S .

This transformation is supposed to be performed by applying an operator of *revision* (denoted $\dot{+}$) to S and δ . As δ is inconsistent with S , the revised system T may retain only a proper subset of the original one. So, some of the formulae of S have to be given up once for all. As there usually are different ways of doing so, revision should be performed carefully in order not to lose important knowledge.

To guide the revision process, Alchourron, Gardenfors and Makinson [1] have formulated a set of postulates that should be obeyed by any *rational reasoner* . The following 6 postulates constitute the *basic set* of these *AGM postulates* as presented in [14] (K is a set of all logical consequences of a given system S , $K = Cn(S)$).

$(K\dot{+}1)$ $K\dot{+}\delta$ is a theory;
$(K\dot{+}2)$ $\delta \in K\dot{+}\delta$;
$(K\dot{+}3)$ $K\dot{+}\delta \subseteq Cn(S \cup \{\delta\})$;
$(K\dot{+}4)$ If $\neg\delta \notin K$ then $Cn(S \cup \{\delta\}) \subseteq K\dot{+}\delta$;
$(K\dot{+}5)$ $K\dot{+}\delta$ is consistent iff δ is so;
$(K\dot{+}6)$ If $\delta \leftrightarrow \psi$ then $K\dot{+}\delta = K\dot{+}\psi$.

Some supplementary postulates have been also formulated in [1, 13, 14].

Recently different revision methods have been proposed and studied [1, 8, 13-15, 17, 23-25, 29, 32, 33, 35, 36, 39-44].

Maxichoice . This method presents the result of revising a knowledge base as one of its consistent subsets chosen by a special *selection function* . Specifically, due to [1, 14],

$$K\dot{+}\delta = Cn[\gamma(K\bot\neg\delta) \cup \{\delta\}] , \tag{8.1}$$

where $K\bot\neg\delta$ denotes a set of all mc-subsets of K failing to imply $\neg\delta$, and γ is a certain selection function. As many different selection functions satisfy the AGM postulates, a right choice of γ for a proper revision still remains a problem.

Full meet . This method builds the revised system on the knowledge that is common to all elements of $K\bot\neg\delta$ [1, 14]:

$$K \dot{+} \delta = Cn[(\bigcap_{\sigma \in K \perp \neg\delta} \sigma) \cup \{\delta\}] \ . \tag{8.2}$$

If maxichoice appears to be too "liberal" an operation, full meet is by far too restrictive, since a full meet revision retains rather a small part of the original system. Indeed, due to Theorem 6 in [14], if $\neg\delta \in K$ then $K \dot{+} \delta$ computed by full meet contains only all logical consequences of δ itself regardless of the content of K .

Partial meet . To relax the vice of full meet, and to preserve more of the original knowledge, one may build revision only on some of the elements of $K \perp \neg\delta$. This can be achieved using a special selection function, β , that picks out a proper subset of $K \perp \neg\delta$ consisting of most "significant" elements. Then a partial meet revision is

$$K \dot{+} \delta = Cn[(\bigcap_{\sigma \in \beta(K \perp \neg\delta)} \sigma) \cup \{\delta\}] \ . \tag{8.3}$$

So far we have reviewed revision methods defined on closed theories, or *belief sets* , that are sets of all logical consequences of a given knowledge base. This approach is appropriate if the Principle of Minimal Change is imposed on the states of the world described by the original and the revised systems. On the other hand, a knowledge base consists of distinctive pieces of knowledge (like laws of the nature), so, it is reasonable to apply the principle directly to the set of sentences constituting the knowledge base, usually called a *belief base* . This view has been adopted in numerous studies of system revision and update [12, 17, 32, 33].

As restoration of consistency of a polluted unsatisfiable system requires abandoning of some of its data, the guiding principle of revision should be that the revised system must retain the most valuable elements of the original one. This process could be performed efficiently, if all elements of a given system were ordered according to their importance or relevance to revision.

The idea of considering priorities among elements of knowledge goes back at least to the work of Rescher [37]. This idea has been employed in Circumscription [26], Autoepistemic logic [22], and Default reasoning [4, 16]. Several types of ordering have been proposed quite recently.

Due to the approach developed by Gardenfors and Makinson [14, 15], every sentence of the adopted language has a degree of *epistemic entrenchment* (relative to other sentences of the language) that reflects its importance in a given knowledge domain. When a belief set K is revised, the sentences with the lowest degree of epistemic entrenchment must be given up first. Fagin, Ullman and Vardi [12] introduced a notion of *database priorities* that is similar to the epistemic entrenchment.

Nebel [33] developed revision methods for belief bases. He imposed a complete preorder, called *epistemic relevance ordering* , on the elements of a given belief base, and defined an operation of *prioritized base revision* as a logical closure of a disjunction of its maximally consistent subsets that contain elements of the base with the highest degree of epistemic relevance.

Katsuno and Mendelzon [20] consider a total preorder over a set of all interpretations of a given language. This ordering is sanctioned by a given system, and determines which of the interpretations become models of the system after its revision.

Rott [40] compares the epistemic entrenchment with an ordering of mc-subsets of the system.

Roos [39] defines a reliability relation that reflects the relative reliability of premises. Then he defines a preference order on the system models such that the more premises with a higher reliability are satisfied by a given model, the higher is its preference.

Benferhat, Cayrol, Dubois, Lang and Prade [3] consider different preference relations on subsets of a given system S based on a complete priority preordering over the formulae of S. This preordering divides S into a collection $\{S_1, S_2, ..., S_n\}$ of disjoint subsets such that for all $1 \le i < j \le n$, formulae of S_i have a priority higher than those of S_j. Now any subset $A \subseteq S$ can be presented as a collection $\{A_1, A_2, ..., A_n\}$, where $A_i = A \cap S_i$. In particular, the authors define a *lexicographic ordering* $<^{lex}$ in such a way that for any two subsets A, B of S, $A <^{lex} B$ iff exists i such that $|A_i| < |B_i|$ and for any $j < i$, $|A_j| = |B_j|$.

Should one of the abovementioned orderings be attached to a knowledge base, this would facilitate its revision with a minimal change (relative to the ordering). Unfortunately, one can hardly expect such an ordering to be unambiguously defined in a real situation. One of the main reasons is that a definition of a total order (unless it is a trivial one) would require a very large amount of information due to a huge number of elements to be ordered (such as a number of sentences of a language, or even the number of their interpretations). In particular, it is rather unlikely that an expert in a knowledge domain can possess such an amount of data, and convey it to the system. At the same time, a rather limited information regarding relative importance of some elements may very well be available, determining a partial order. So, it appears worthwhile to find a way of accommodating a limited information about relative priorities among several elements of a given knowledge base.

9. Accommodating partial order

As mentioned in Section 8, in practice the ordering relation does not usually involve all elements of a given system, since relative significance or credibility of many elements may not be known or assessed at all, and so, elements mutually related by a preference may constitute rather a small subset of S.

Let $\le_i$ denote a transitive and reflexive relation of a *partial preorder of importance*, PI, over a subset S' of S such that for $\phi, \psi \in S'$, $\phi \le_i \psi$ means that ψ is at least as important (or as credible) as ϕ, and $\phi <_i \psi$ means that $\phi \le_i \psi$, but $\psi \not\le_i \phi$. Relative importance of formulae is intended to control transformations of S in the following way.

Observation 9.1. Given $(\phi \le_i \psi) \in PI$, if in the process of constructing a consistent version of S the formulae ϕ and ψ cannot be kept together, and there are two alternatives: one, to keep ϕ but abandon ψ, the other, to keep ψ but give up ϕ, then the latter one should be preferred. ■

Given an importance preorder PI, a set of formulae S can be partitioned into disjoint subsets, called *ranks* and numbered by natural numbers, in the following way. Let $r(\phi)$ denote the number of the *rank* containing a formula ϕ. If ϕ is not mentioned in PI then $r(\phi) = 0$. If ϕ is mentioned in PI, but there is no formula $\psi \in S$ such that $(\phi \le_i \psi) \in PI$, then $r(\phi) = 1$, otherwise, $r(\phi) = k + 1$, where k is the maximal rank among all formulae ψ such that $(\phi \leq_i \psi) \in PI$. So, the higher is the importance of ϕ, the smaller is the value of $r(\phi)$.

Definition 9.1. *Rank of formula.*

$$r(\phi) = \begin{cases} 0 & \text{if} \quad \neg \exists\, \psi[(\phi \leq_i \psi) \in PI \vee (\psi \leq_i \phi) \in PI\,] \\ 1 & \text{if} \quad \neg \exists\, \psi[(\phi \leq_i \psi) \in PI\,] \wedge \exists\, \psi[(\psi \leq_i \phi) \in PI\,] \\ k+1 & \text{if} \quad \exists\, \psi[(\phi \leq_i \psi) \in PI\,] \quad \textit{and} \quad k = \max\limits_{\psi\, \in\, S}(r(\psi) \mid (\phi \leq_i \psi) \in PI\,)\,. \end{cases} \quad \blacksquare$$

Thus, the rank of a formula is an indication of its importance or credibility. If a mc-subset σ contains a formula ϕ but not ψ, while $(\phi \leq_i \psi) \in PI$, and $r(\psi) = k$, then we say that σ *violates PI* at rank k. So, if an importance preorder PI is defined for a system S, then in a process of the system transformation we should consider only those of its mc-subsets which *comply with PI* in the sense that if Observation 9.1 cannot be satisfied for all ordered pairs $(\phi \leq_i \psi) \in PI$, then the highest violated rank should be pushed down as low as possible.

Definition 9.2. *Compliance with PI*. Given S and PI, denote by $P[MC(S)]$ a set of all elements of $MC(S)$ complying with PI. For all $\sigma \in MC(S)$ denote by $\sigma(k)$ a set of all formulae $\phi \in \sigma$ with a rank $r(\phi) = k$. Then $P[MC(S)]$ is produced by the following function *Comply* such that $P[MC(S)] := Comply(MC(S), PI)$.

Function Comply(X, Y);
(* Given a set X of sets σ of formulae, and an importance preorder Y, returns a subset of X containing only those sets $\sigma \in X$ that comply with Y. *)
$Z := X$;
For all elements $\sigma \in Z$ do
    For all elements $\sigma' \in Z$ such that $\sigma(i) = \sigma'(i)$ for all $0 \leq i < k$, but $\sigma(k) \subset \sigma'(k)$ do
        For all formulae $\phi \in (\sigma'(k) - \sigma(k))$ do
            If there is a formula $\psi \in \sigma$ such that $(\psi \leq_i \phi) \in PI$, then delete σ from Z, $Z := Z - \{\sigma\}$;
$Comply(X, Y) := Z$
end (* Comply *). ■

Actually, Definition 9.1 completes the partial preorder PI, and thus, allows us to produce $P[MC(S)]$ by applying any efficient method based on a total preorder, referred to in Section 8. So, what we would like to stress is not how to distinguish mc-subsets complying with PI from those violating it, but the point that although the heuristics of weighted mc-subsets does not need any ordering information, if such an information of any extent is provided, there is a natural way to make the suggestions of the heuristics comply with the given ordering by considering $P[MC(S)]$ instead of $MC(S)$.

It is worth noting that if PI is a total preorder like the *epistemic relevance ordering* defined in [33], then $P[MC(S)]$ is similar to the set of *prioritized removal* introduced in [33, p.55].

10. Restoring consistency

Let us reconsider the problem of restoring consistency of a knowledge base (cf. arc 3 in Figure 1.1). Namely,

given (a) a consistent system S,
(b) a preorder of importance PI,
(c) a set of constraints C (C may contain, for instance, formulae that are undoubtedly true in every possible world), and
(d) a consistent formula δ which is inconsistent with S,

transfer S to a system T that
(i) is consistent,
(ii) contains δ (assuming that δ is more reliable than the formulae of S),
(iii) satisfies PI and C, and
(iv) complies with the Principle of Minimal Change relative to S.

Assume that a restoration of consistency is performed by an operator $MEND$ such that

$$T = MEND(S, \delta) . \tag{10.1}$$

The approach developed in the foregoing sections suggests the following solution to the problem.

Definition 10.1. *Subsets relevant to revision.* A mc-subset σ of $S \cup \{\delta\}$ is *relevant to revision* if it (i) contains δ, (ii) satisfies all constraints C, and (iii) complies with PI, that is, $\sigma \in P[MC(S \cup \{\delta\})]$. R_{REV} denotes a set of all elements of $MC(S \cup \{\delta\})$ relevant to revision. ■

Now T is supposed to replace $S \cup \{\delta\}$, while the semantics of the latter based on the Principle of reasoning by evidence is designed to satisfy the Principle of Minimal Change relative to S. So, to comply with this principle, MEND has to choose an element σ of R_{REV} that makes the largest contribution to the semantics of $S \cup \{\delta\}$. This *relative contribution*, $contrib(\sigma)$, is proportional to the number of models of σ, and to its semantical and syntactical weights:

$$contrib(\sigma) = \frac{|MOD(\sigma)| \cdot u(\sigma) \cdot w(\sigma)}{\sum\limits_{x \in R_{REV}} |MOD(x)| \cdot u(x) \cdot w(x)} . \tag{10.2}$$

Definition 10.2. *Restoring consistency.*

$$MEND(S \cup \{\delta\}) = \sigma \mid \sigma \in R_{REV} \wedge contrib(\sigma) = \max_{x \in R_{REV}} (contrib(x)) . \quad ■$$

Example 10.1. Consider the system of Example 7.2 for which both PI and C are empty. As q has two independent proofs, and $\neg r$ is most credible, this suggests that q does not imply r, and so, the clause $q \rightarrow r$ is most doubtful. Trying to restore consistency of $S \cup \{\neg r\}$, we notice (see Table 7.1, column 10) that σ_5 makes the largest contribution to the semantics of the system. Hence, by Definition 10.2,

$$MEND(S \cup \{\neg r\}) = \sigma_5 = (S \cup \{\neg r\}) - \{q \rightarrow r\} . \quad \blacksquare$$

To relate MEND to belief revision, let us examine the operator from the point of view of AGM postulates (Section 8), where $K \dot{+} \delta$ is replaced by $Cn(MEND(S, \delta))$, and K by $Cn(S)$.

Theorem 10.1. $MEND(S, \delta)$ complies with the basic AGM postulates for revision (Section 8).

Proof. $(K \dot{+} 1)$ $MEND(S, \delta)$ is a set of formulae, so, $Cn(MEND(S, \delta))$ is a theory.

$(K \dot{+} 2)$ $\delta \in Cn(MEND(S, \delta))$ since, by Definitions 10.1, 10.2, $\delta \in MEND(S, \delta)$.

$(K \dot{+} 3, K \dot{+} 4)$ If $\neg\delta \notin Cn(S)$ then $S \cup \{\delta\}$ is consistent, hence, $MEND(S, \delta) = S \cup \{\delta\}$ and $Cn(MEND(S, \delta)) = Cn(S \cup \{\delta\})$. Otherwise, $S \cup \{\delta\}$ is inconsistent, so, $Cn(S \cup \{\delta\})$ contains all sentences of the language. In this case $Cn(MEND(S, \delta)) \subset Cn(S \cup \{\delta\})$, since $MEND(S, \delta)$ is consistent. Hence, in any case $Cn(MEND(S, \delta)) \subseteq Cn(S \cup \{\delta\})$.

$(K \dot{+} 5)$ By Definitions 10.1, 10.2, $MEND(S, \delta)$ is consistent and contains δ , that is possible iff δ is consistent.

$(K \dot{+} 6)$ If $\delta \leftrightarrow \psi$ then $MEND(S, \delta)$ and $MEND(S, \psi)$ differ only in that δ of $MEND(S, \delta)$ is replaced in $MEND(S, \psi)$ by ψ . Hence, $Cn(MEND(S, \delta)) = Cn(MEND(S, \psi))$. $\blacksquare$

In determining the revised state of knowledge the MEND operator takes into consideration both the semantic and syntactic difference between knowledge bases. As logically equivalent systems may have different syntactic representation, any syntax-dependent operator applied to such systems may produce results that differ not only syntactically, but semantically as well. This phenomenon (although well justified in many cases) has been considered by some researchers as an undesirable property of a revision process. So, recently several approaches have been developed that perform belief revision on the model-theoretic level [7, 23, 25, 43]. Namely, a logic system S is presented by the set of its models, and then the revision is performed by selecting those models of δ which are the closest in a certain sense to the models of S .

Due to the approach developed by Dalal [7], the distance $d(m, m')$ between two models m, m' is the Hamming distance measured by the number of ground atomic formulae on which the models disagree. Let $g^{(i)}(m)$ denote a set of interpretations $\{\mu\}$ such that $d(\mu, m) \leq i$. Then the revision of S by δ , denoted $S \circ \delta$, is represented by a set of models

$$MOD(S \circ \delta) = \Big(\bigcup_{m \in MOD(S)} g^{(k)}(m) \Big) \cap MOD(\delta) , \tag{10.3}$$

such that k is the least value for which $MOD(S \circ \delta)$ is not empty.

$S \circ \delta$ is very close semantically to the original system, however, this is achieved at the expense of ignoring the syntactic structure of S . Indeed, this structure is totally lost in the revision process, as the latter produces only the set of models of the revised system without any suggestion regarding its syntactic representation. This loss is responsible for counterintuitive results that may be produced by this kind of non-syntactic revision, as illustrated by the following example.

-

Table 10.1 Comparing $S \circ \delta$ with $MEND(S, \delta)$ (Example 10.2).

	Evidence p	Evidence q	Evidence r	$D(\sigma, S)$	$u(\sigma)$	$\lvert S - \sigma\rvert$	$w(\sigma)$	$\lvert MOD(\sigma)\rvert$	$contrib(\sigma)$
1	2	3	4	5	6	7	8	9	10
S	1	1	1/2						
σ_1	1	0	0	0.500	0.667	1	0.333	1	**0.600**
σ_2	0	0	0	0.833	0.444	1	0.333	1	0.400
$S \circ \delta$	1	0	1/2	0.333					

Example 10.2. $S = \{p,\ p \rightarrow q,\ r \rightarrow q\}$, $\delta = \neg q$, with the following sets of models:
$MOD(S) = \{\{p, q, r\}, \{p, q, \neg r\}\}$;

$MOD(\delta) = \{\{p, \neg q, r\}, \{p, \neg q, \neg r\}, \{\neg p, \neg q, r\}, \{\neg p, \neg q, \neg r\}\}$.

Due to the approach of [7], already the closest neighbors of $MOD(S)$, for $k = 1$, provide the revision:

$$MOD(S \circ \delta) = \{\{p, \neg q, r\}, \{p, \neg q, \neg r\}\} .$$

Let us compare $S \circ \delta$ with $MEND(S, \delta)$. There are two relevant mc-subsets of $S \cup \{\neg q\}$:

$$\sigma_1 = \{p,\ r \rightarrow q,\ \neg q\} , \quad \sigma_2 = \{p \rightarrow q,\ r \rightarrow q,\ \neg q\} .$$

The evidences provided by σ_1, σ_2 together with their weights are shown in Table 10.1. It follows from the data (see column 10) that

$$MEND(S, \neg q) = \sigma_1 .$$

The last line of Table 10.1 presents $S \circ \delta$. It shows (column 5) that $S \circ \delta$ is closer semantically to S than $MEND(S, \delta)$. But what knowledge does $S \circ \delta$ preserve? Given its models, the simplest form of $S \circ \delta$ is $\{p, \neg q\}$. The absence of $r \rightarrow q$ is unjustified. Indeed, $r \rightarrow q$ is consistent with $S \circ \delta$, and since $\neg q$ is regarded as most credible, $r \rightarrow q$ yields that r is false, just as it follows from $MEND(S, \delta)$. ■

11. Knowledge contraction

A specific transformation of a knowledge base that has been studied extensively [1, 14, 15, 29, 30] is *contraction* . Namely,

> if a consistent system S implies a formula δ , but a new message informs that there is not enough grounds for believing that δ is true, then δ should be *retracted* from S by transforming S to a state T that does not imply δ , and is as close to S as possible.

Then T is said to represent S *contracted* by δ .

Alchourron, Gardenfors and Makinson [1, 14] have formulated the following *basic AGM postulates* that should be obeyed by any rational contraction operator "$\dot{-}$" ($K \dot{-} \delta$ denotes a theory K contracted by δ , where $K = Cn(S)$):

$(K \dot{-} 1)$ $K \dot{-} \delta$ is a theory;
$(K \dot{-} 2)$ $(K \dot{-} \delta) \subseteq K$;
$(K \dot{-} 3)$ If $\delta \notin K$ then $K \dot{-} \delta = K$;
$(K \dot{-} 4)$ If δ is not a tautology then $\delta \notin (K \dot{-} \delta)$;
$(K \dot{-} 5)$ If $\delta \in K$ then $K \subseteq Cn((K \dot{-} \delta) \cup \{\delta\})$;
$(K \dot{-} 6)$ If $\delta \leftrightarrow \psi$ then $K \dot{-} \delta = K \dot{-} \psi$.

The AGM postulates of contraction are similar to those of revision (Section 8). Due to this similarity, the solution methods for revision described in Section 8 are applicable to contraction as well.

Maxichoice [1, 14]. Given a selection function γ that chooses one element of a set $K \perp \delta$ of all maximal (by set inclusion) subsets of K not implying δ,

$$K \dot{-} \delta = Cn(\gamma(K \perp \delta)) . \tag{11.1}$$

Full meet [1, 14].

$$K \dot{-} \delta = Cn(\bigcap_{\sigma \in (K \perp \delta)} \sigma) . \tag{11.2}$$

Partial meet [1, 14].

$$K \dot{-} \delta = Cn(\bigcap_{\sigma \in \beta(K \perp \delta)} \sigma) , \tag{11.3}$$

where β selects a proper subset of $K \perp \delta$.

Epistemic entrenchment [9, 14, 15]. $K \dot{-} \delta$ is an element of $K \perp \delta$ produced from K in such a way that formulae with the lowest degree of epistemic entrenchment are given up first.

These methods applied to contraction have the same advantages and, unfortunately, the same shortcomings as in the case of revision (see Section 8). As an alternative to the existing methods that is based on the heuristics of weighted mc-subsets and allows accommodation of a partial priority ordering of elements of a given knowledge base, we define a contraction operator $CONTRACT(S, \delta)$ similar to $MEND$ of Section 10.

Given a consistent system S (together with a set of constraints C and a preorder of importance PI), and a non-tautological formula δ that should be retracted from S , we have to transform S to a system $T = CONTRACT(S, \delta)$ which (i) does not imply δ , (ii) satisfies C , (iii) complies with PI , and (iv) obeys the Principle of Minimal Change relative to S (cf. arc 4, Figure 1.1).

Definition 11.1. *Subsets relevant to contraction.* A subset σ of S is *relevant to contraction* if it (i) is a maximal subset not implying δ , that is, $\sigma \in S \perp \delta$, (ii) satisfies all constraints C , and (iii) complies with PI, that is, $\sigma \in Comply(S \perp \delta, PI)$. R_{CON} denotes a set of all elements of $S \perp \delta$ relevant to contraction. ■

By the argument presented in Section 10, S should be contracted to one of its relevant subsets making the largest contribution to the system semantics.

Definition 11.2. *Contraction.*

$$CONTRACT(S, \delta) = \sigma \mid \sigma \in R_{CON} \wedge contrib(\sigma) = \max_{x \in R_{CON}} (contrib(x)) . \quad \blacksquare$$

-

Table 11.1. Relevant subsets of S (Example 11.1).

	Evidence				$D(\sigma, S)$	$u(\sigma)$	$\lvert S - \sigma \rvert$	$w(\sigma)$	$\lvert MOD(\sigma) \rvert$	$contrib(\sigma)$
	p_1	p_2	q	r						
1	2	3	4	5	6	7	8	9	10	11
S	1	1	1	1						
τ_1	1	1	1/3	2/3	0.250	0.813	2	0.08	3	0.147
τ_2	1	1/4	1/2	3/4	0.375	0.719	2	0.08	4	0.173
τ_3	1/4	1	1/2	3/4	0.375	0.719	2	0.08	4	0.173
τ_4	1/3	1/3	2/3	5/6	0.472	0.646	2	0.08	6	0.234
τ_5	1	1	1	1/2	0.125	0.906	1	0.20	2	**0.273**

Example 11.1. Consider the system of Example 7.2,

$$S = \{p_1,\ p_1 \rightarrow q,\ p_2,\ p_2 \rightarrow q,\ q \rightarrow r\}\ ,$$

and suppose that r has to be retracted from S. There are 5 subsets of S relevant to this contraction:

$$\begin{aligned}
\tau_1 &= \{\ p_1,\ p_2,\ q \rightarrow r\}\ ,\\
\tau_2 &= \{\ p_1,\ p_2 \rightarrow q,\ q \rightarrow r\}\ ,\\
\tau_3 &= \{\ p_1 \rightarrow q,\ p_2,\ q \rightarrow r\}\ ,\\
\tau_4 &= \{\ p_1 \rightarrow q,\ p_2 \rightarrow q,\ q \rightarrow r\}\ ,\\
\tau_5 &= \{\ p_1,\ p_1 \rightarrow q,\ p_2,\ p_2 \rightarrow q\}\ .
\end{aligned}$$

Table 11.1 presents data characterizing these subsets relative to S. Due to column 11, by Definition 11.2, $CONTRACT(S, r) = \tau_5$ preserving the beliefs suggested by S for all atoms except r (see columns 2-5). ■

Theorem 11.1. $CONTRACT(S, \delta)$ complies with the AGM postulates $(K \dot{-} 1) - (K \dot{-} 4)$ and $(K \dot{-} 6)$, where $K \dot{-} \delta$ is replaced by $Cn(CONTRACT(S, \delta))$, and K by $Cn(S)$.

Proof. $(K \dot{-} 1)$ $Cn(CONTRACT(S, \delta))$ is a theory, since $CONTRACT(S, \delta)$ is a set of formulae.

$(K \dot{-} 2)$ $CONTRACT(S, \delta)$ is a subset of S, hence $Cn(CONTRACT(S, \delta)) \subseteq Cn(S)$.

$(K \dot{-} 3)$ If $\delta \notin Cn(S)$ then $S \not\models \delta$, hence S is its only maximal subset not implying δ. So, $CONTRACT(S, \delta) = S$ and $Cn(CONTRACT(S, \delta)) = Cn(S)$.

$(K \dot{-} 4)$ If δ is not a tautology then by Definitions 11.1, 11.2, $CONTRACT(S, \delta) \not\models \delta$. Hence, $\delta \notin Cn(CONTRACT(S, \delta))$.

$(K \dot{-} 6)$ If $\delta \leftrightarrow \psi$ then $S \perp \delta$ and $S \perp \psi$ are identical. Hence, by Definitions 11.1, 11.2, $CONTRACT(S, \delta) = CONTRACT(S, \psi)$. ■

AGM postulate $(K \dot{-} 5)$ deserves a separate discussion. The postulate is intended to express a property of *recovery*, namely, that a system S contracted by δ should retain enough information to recover all the beliefs of S when δ is added back. However, $(K \dot{-} 5)$ allows the recovered system to contain some beliefs that did not belong to S before the contraction. As stated in [14, p.11], $(K \dot{-} 5)$ "has turned out to be the most controversial among the AGM postulates for contraction". Indeed, Makinson [30] presents arguments, both intuitive and formal, showing that the recovery postulate should not necessarily govern contraction of knowledge. Hansson [17] gives a nice example showing that a reasonable contraction of a set of formulae may not satisfy $(K \dot{-} 5)$.

Example 11.2. [17] Consider $S = \{p, q, r\}$. S contracted by $p \vee q$ becomes $T = \{r\}$ such that neither p nor q belongs to $Cn(T \cup \{p \vee q\})$. Hence,

$$Cn(S) \not\subseteq Cn(T \cup \{p \vee q\})\ ,$$

and even

$$Cn(T \cup \{p \vee q\}) \subset Cn(S)\ . \quad ■$$

It appears reasonable that if a formula that follows from a system is retracted from the system, and then is added back to the contracted one, this process (that can be viewed as a reflection of doubts and hesitation) is not supposed to create any new knowledge relative to the original system. The following theorem shows that $CONTRACT(S, \delta)$ possesses this desirable property.

Theorem 11.2. Let CON be any operator that contracts a set of formulae S to its subset $CON(S, \delta)$. If $\delta \in Cn(S)$ then $Cn(CON(S, \delta) \cup \{\delta\}) \subseteq Cn(S)$.

Proof. Since $CON(S, \delta)$ is a subset of S, $S \models CON(S, \delta)$. If $\delta \in Cn(S)$ then $S \models \delta$, and so, $S \models (CON(S, \delta) \cup \{\delta\})$. Hence, $Cn(CON(S, \delta) \cup \{\delta\}) \subseteq Cn(S)$. ■

The theory of a system S can be , however, fully recovered after a contraction of S by δ if the part of S, lost because of the contraction, is implied by δ.

Theorem 11.3. Let CON be any operator that contracts a set of formulae S to its subset $CON(S, \delta)$. Then for all formulae δ such that $S \models \delta \models (S - CON(S, \delta))$,

$$Cn(CON(S, \delta) \cup \{\delta\}) = Cn(S) .$$

Proof. Since $S \models \delta$ and $CON(S, \delta) \subseteq S$, we have $S \models (CON(S, \delta) \cup \{\delta\})$. Since $\delta \models (S - CON(S, \delta))$ and $CON(S, \delta) \cup (S - CON(S, \delta)) = S$, we have $(CON(S, \delta) \cup \{\delta\}) \models S$. Hence, $Cn(CON(S, \delta) \cup \{\delta\}) = Cn(S)$. ■

The following example shows that the converse of Theorem 11.3 does not hold, which means that, fortunately, the theory of S can be recovered after a contraction in more cases than those determined by Theorem 11.3.

Example 11.3. $S = \{p, q\}$, but it turns out that there is no ground to believe that p and q always have the same truth value, so, $\delta = p \wedge q \vee \neg p \wedge \neg q$ should be retracted from S. There are two possibilities for contracting S by δ:

(1) $CON(S, \delta) = \{p\}$. In this case $\delta \not\models (S - CON(S, \delta))$, however,

$$Cn(CON(S, \delta) \cup \{\delta\}) = Cn(\{p,\ p \wedge q \vee \neg p \wedge \neg q\}) = Cn(\{p, q\}) .$$

(2) $CON(S, \delta) = \{q\}$. In this case $\delta \not\models (S - CON(S, \delta))$ as well, however,

$$Cn(CON(S, \delta) \cup \{\delta\}) = Cn(\{q,\ p \wedge q \vee \neg p \wedge \neg q\}) = Cn(\{p, q\}) .$$

Thus, in both cases $Cn(CON(S, \delta) \cup \{\delta\}) = Cn(S)$ despite the fact that δ does not imply $S - CON(S, \delta)$. ■

Many researchers consider as a desirable property of revision and contraction operators that they should satisfy the following expressions.

Levi identity [24]:

$$K \dot{+} \delta = Cn((K \dot{-} \neg\delta) \cup \{\delta\}) . \tag{11.4}$$

Harper identity [18]:

$$K \dot{-} \delta = (K \dot{+} \neg\delta) \cap K . \tag{11.5}$$

A relationship between these identities and the AGM postulates is determined by the following propositions.

Theorem 11.4. [14, Theorem 1] If a contraction operator $\dot{-}$ satisfies $(K \dot{-} 1) - (K \dot{-} 4)$ and $(K \dot{-} 6)$, then the revision operator $\dot{+}$ obtained from Levi identity (11.4) satisfies $(K \dot{+} 1) - (K \dot{+} 6)$. ■

Theorem 11.5. [14, Theorem 2] If a revision operator $\dot{+}$ satisfies $(K \dot{+} 1) - (K \dot{+} 6)$, then the contraction operator $\dot{-}$ obtained from Harper identity (11.5) satisfies $(K \dot{-} 1) - (K \dot{-} 6)$. ■

Theorems 11.4, 11.5 indicate that there may exist revision and contraction operators that satisfy the AGM postulates, but do not satisfy the identities. To justify rationality of such a pair of operators, let us consider a revision operator $REV(S, \delta)$ that returns a mc-subset of $S \cup \{\delta\}$ containing δ, and a contraction operator $CON(S, \delta)$ returning a maximal subset of S not implying δ. For these operators, Levi and Harper identities take, respectively, the following form:

$$Cn(REV(S, \delta)) = Cn(CON(S, \neg\delta) \cup \{\delta\}) ; \tag{11.6}$$

$$Cn(CON(S, \neg\delta)) = Cn(REV(S, \delta)) \cap Cn(S) . \tag{11.7}$$

Lemma 11.1. To satisfy the identities it must hold that

$$REV(S, \delta) = CON(S, \neg\delta) \cup \{\delta\} . \tag{11.8}$$

Proof. If $S \not\models \neg\delta$, then $CON(S, \neg\delta) = S$ and $REV(S, \delta) = S \cup \{\delta\}$, satisfying expressions (11.6)-(11.8).

If $S \models \neg\delta$, then consider a set σ such that

$$REV(S, \delta) = \sigma \cup \{\delta\} .$$

Both σ and $CON(S, \neg\delta)$ are proper subsets of S. It follows from (11.6) that $REV(S, \delta)$ is logically equivalent to $CON(S, \neg\delta) \cup \{\delta\}$. Hence, $REV(S, \delta) \cup CON(S, \neg\delta) \cup \{\delta\}$ is consistent, and so, $\sigma \cup CON(S, \neg\delta) \cup \{\delta\}$ is consistent as well. This means that $(\sigma \cup CON(S, \neg\delta)) \not\models \neg\delta$. So, $\sigma \subseteq CON(S, \neg\delta)$, since $CON(S, \neg\delta)$ is a maximal subset of S not implying $\neg\delta$. But if σ is a proper subset of $CON(S, \neg\delta)$, then $REV(S, \delta) \subset (CON(S, \neg\delta) \cup \{\delta\})$ contradicting the premise that $REV(S, \delta)$ is a mc-subset of $S \cup \{\delta\}$. Hence, $\sigma = CON(S, \neg\delta)$, and $REV(S, \delta) = \sigma \cup \{\delta\} = CON(S, \neg\delta) \cup \{\delta\}$. ■

Imagine now two reasoners, A and B, each in possession of a copy of a system S implying that a formula δ is false. A and B have received from reliable sources messages α and β, respectively:

α : There is not enough ground for being sure that δ is false.

β : δ is true.

Upon receiving the messages, the reasoners transform their systems to S_α and S_β, respectively, such that $S_\alpha = CON(S, \neg\delta)$, $S_\beta = REV(S, \delta) = \sigma \cup \{\delta\}$, where $CON(S, \neg\delta) \subseteq S$ and $\sigma \subseteq S$. To satisfy the identities (by Lemma 11.1) σ must be identical to $CON(S, \neg\delta)$. However, it appears rather unlikely that to accommodate such different messages as α and β, both reasoners will *in all cases* retain the same subset of their system to represent the new state of knowledge.

Regarding the operators *MEND* and *CONTRACT* (Definitions 10.2, 11.2), we note that their application to the system S of Examples 7.2, 10.1, 11.1 satisfies Levi and Harper identities, since $MEND(S, \neg r) = CONTRACT(S, r) \cup \{ \neg r \}$. However, in other cases $MEND$ and $CONTRACT$ may not satisfy the identities, which is quite acceptable due to the abovementioned observations.

12. Infinite ground bases

So far we have considered systems with finite ground bases (cf. Section 3). Now we extend this study to systems with infinite ground bases, and so, with infinite models and possibly with infinite sets of models.

Since a set of predicate symbols occurring in a first order system S is finite, the reason for an infinite ground base of S is infiniteness of the domain of terms of the system. Let DOM be an infinite enumerable set of terms of S, $rdom$ denote a reduced domain that is a finite subset of DOM, and S^{rdom} stand for the original system S defined over $rdom$. Consider an infinite sequence $rdom_1, rdom_2, \ldots$ of finite subsets of DOM, such that for all $1 \leq i < j$, $rdom_i \subset rdom_j \subset DOM$, and $\lim_{i \to \infty} rdom_i = DOM$. To extend the approach developed in the preceding sections to systems with infinite ground bases, let us view a system S as a limit of growing S^{rdom} while $rdom$ tends to DOM by assuming consecutive values of $rdom_i$ for $i = 1, 2, \ldots$ The ground base of S^{rdom} is finite, thus allowing the following extended definition.

Definition 12.1. Given systems S, T over the same ground base with an infinite domain of terms DOM, let $rdom_1$, $rdom_2$, . . . be a sequence of finite subsets of DOM such that

$$\lim_{i \to \infty} rdom_i = DOM \ . \tag{12.1}$$

Then

$$E(S, \phi) = \lim_{i \to \infty} E(S^{rdom_i}, \phi) \ , \tag{12.2}$$

$$D(S, T) = \lim_{i \to \infty} D(S^{rdom_i}, T^{rdom_i}) \tag{12.3}$$

and other characteristics of S and T are determined by the corresponding limits if these limits exist. ■

-

Table 12.1. Finite base mc-subsets of $S \cup \{\delta\}$ (Example 12.1).

	Evidence for all $x \in rdom_i$ $P(x)$	$Q(x)$	$u(\sigma)$	$w(\sigma)$	$\lvert MOD(\sigma)\rvert$	$contrib(\sigma)$
1	2	3	4	5	6	7
S^{rdom_i}	1	1				
$\sigma_1^{rdom_i}$	1	$\frac{2^{i-1}-1}{2^i-1}$	$\frac{2}{3}+O\left(\frac{1}{i}\right)$	$\frac{1}{2}$	2^i-1	$\frac{1}{1+O\left(\left(\frac{3}{2}\right)^i\right)}$
$\sigma_2^{rdom_i}$	$\frac{3^{i-1}-2^{i-1}}{3^i-2^i}$	$2\,\frac{3^{i-1}-2^{i-1}}{3^i-2^i}$	$\frac{1}{2}+O\left(\frac{1}{i}\right)$	$\frac{1}{2}$	3^i-2^i	$\frac{1}{1+O\left(\left(\frac{2}{3}\right)^i\right)}$

Example 12.1. Consider a system

$$S = \{(\forall x)[P(x)], (\forall x)[P(x) \rightarrow Q(x)]\}$$

with x ranging over all natural numbers. Suppose that $\delta = (\exists x)[\neg Q(x)]$, which is inconsistent with S .

Let us restore consistency of $S \cup \{\delta\}$ by computing $MEND(S, \delta)$. There are two relevant mc-subsets of $S \cup \{\delta\}$:

$$\sigma_1 = \{(\forall x)[P(x)], (\exists x)[\neg Q(x)]\} ; \quad \sigma_2 = \{(\forall x)[P(x) \rightarrow Q(x)], (\exists x)[\neg Q(x)]\} .$$

Let $rdom_i$ consist of i natural numbers. Table 12.1 characterizes S^{rdom_i} , $\sigma_1^{rdom_i}$, $\sigma_2^{rdom_i}$. It suggests that $MEND(S, \delta) = \sigma_2$, since (see column 7)

$$contrib(\sigma_1) = \lim_{i \rightarrow \infty} contrib(\sigma_1^{rdom_i}) = 0 ,$$

while

$$contrib(\sigma_2) = \lim_{i \rightarrow \infty} contrib(\sigma_2^{rdom_i}) = 1 .$$

13. Summary

We study Knowledge Bases described in a First Order language as sets S of well-formed consistent formulae. A system grows by acquiring new information δ that may be either consistent or inconsistent with S . In the latter case the system becomes unsatisfiable, and its users have to cope with "polluted" data, otherwise consistency of the system should be restored. Besides, a new information may require retraction of certain data from the system.

A general rule guiding all system transformations is the Principle of Minimal Change requiring that a new state of knowledge S' resulting from accommodation of δ by S should differ from S as little as possible. Implementation of this principle requires a precise definition of the notion of difference between knowledge bases.

To express a semantic difference between two logic systems we present their models as nodes in a *model space* , and then determine the semantic distance between the systems by the distance between the *mass centers* of the sets of their models measured due to a special metric (Section 3).

To define this metric we present the notion of *evidence* $E(S, \phi)$ provided by a system S to a formula ϕ (Section 4) that is a measure of credibility of a belief in truth of ϕ . Then Section 5 introduces *evidence metric* in which the *evidence distance* $D(S, T)$ between two systems S and T is determined by the evidence provided by the systems to the elements of their common ground base. For complete logic systems the evidence metric coincides with that of Hamming.

The next sections investigate typical cases of a knowledge base transformation (see Figure 1.1) such as consistent acquisition of new knowledge (arc 1), inconsistent acquisition (arc 2), restoring consistency of a system (arc 3), knowledge contraction (arc 4).

The simplest case of consistent acquisition is briefly presented in Section 6, while the complex case of inconsistent acquisition is discussed in Section 7 leading to definition of *semantic* and *syntactic weights* of maximally consistent subsets of a system, and to introduction of a *heuristics of weighted mc-subsets* .

The processes of restoring consistency of a knowledge base, and contracting its data have been extensively studied in the framework of *Belief Revision* . This well known theory is briefly surveyed in Sections 8, 11. We develop two operators based on the heuristics of weighted mc-subsets: MEND for restoring consistency (Section 10), and CONTRACT for knowledge contraction (Section 11). It is shown that MEND and CONTRACT satisfy the AGM postulates.

The case of an unsatisfiable system S augmented by a new knowledge δ (arc 5 in Figure 1.1) has not received a special analysis in this study. The reason is that in this case $S \cup \{\delta\}$ remains unsatisfiable, and so, can be treated by the method developed for the case of inconsistent acquisition (Section 7). The semantics of both the original system S and its transformed state $S \cup \{\delta\}$ is determined by the sets of their mc-subsets due to the heuristics of weighted mc-subsets (Section 7). Although the old inconsistencies of S persist in the transformed state, two special cases (pointed out by one of the anonymous referees of this paper) should be distinguished as their complexities differ significantly: in one case δ introduces a new inconsistency to S , while in the other δ is consistent with all mc-subsets of S , thus, causing no additional "pollution".

If S is unsatisfiable then there is a statement δ and mc-subsets σ', σ'' of S such that $\sigma' \models \delta$, $\sigma'' \models \neg\delta$. We say that δ is *ambiguous* in S . If δ has to be retracted from S (arc 6 in Figure 1.1), this means that all mc-subsets of S asserting δ have to be contracted by δ (Section 11). If δ is the only ambiguous statement in S , then this process will restore consistency of the knowledge base, since in this case $(S - \bigcup_{\sigma' \models \delta} \sigma') \cup \bigcup_{\sigma' \models \delta} CONTRACT(\sigma', \delta)$ is satisfiable.

Finally, Section 12 extends the approach developed in this paper to systems with infinite domains and ground bases.

Acknowledgment

I have greatly benefited from numerous fruitful discussions with Norman Foo. I am grateful to the anonymous referees for their most apt and helpful comments. This research was supported by the Israel Science Foundation administered by the Israel Academy of Sciences and Humanities.

References

1. C. E. Alchourron, P. Gardenfors and D. Makinson, On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic,* v. 50 (1985) 510-530.

2. F. Bacchus et al., From statistics to beliefs. In: *Proceedings of the 9th National Conference of the American Association for Artificial Intelligence,* (1992) 602-608.

3. S. Benferhat et al., Inconsistency management and prioritized syntax-based entailment. In: *Proceedings of the 13th International Joint Conference on Artificial Intelligence, IJCAI-93,* Chambery, France (1993) 640-645.
4. G. Brewka, Preferred subtheories: an extended logical framework for default reasoning. In: *Proceedings of the 11th International Joint Conference on Artificial Intelligence, IJCAI-89,* Detroit, Michigan (1989) 1043-1048.
5. R. Carnap, On inductive logic. *Philosophy of Science,* v. 12, no. 2 (1945) 72-97.
6. R. Carnap, *The Continuum of Inductive Methods.* The University of Chicago Press (1952).
7. M. Dalal, Investigations into a theory of knowledge base revision: preliminary report. In: *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence,* Saint Paul, Minn. (1988) 475-479.
8. J. Doyle, Reason maintenance and belief revision: foundations vs. coherence theories. In: P. Gardenfors, ed., *Belief Revision,* Cambridge University Press (1992) 29-51.
9. D. Dubois and H. Prade, Epistemic entrenchment and possibility logic. *Artificial Intelligence,* v. 50 (1991) 223-239.
10. D. Dubois and H. Prade, A survey of belief revision and updating rules in various uncertainty models. *International Journal of Intelligent Systems,* v. 9 (1994) 61-100.
11. T. Eiter and G. Gottlob, On the complexity of propositional knowledge base revision, updates, and counterfactuals. *Artificial Intelligence,* v. 57 (1992) 227-270.
12. R. Fagin, J. D. Ullman and M. Y. Vardi, On the semantics of updates in databases. In: *Proceedings of 2nd ACM SIGACT-SIGMOD Conference,* Atlanta (1983) 352-365.
13. P. Gardenfors, *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* The MIT Press, Cambridge, Mass. (1988).
14. P. Gardenfors, Belief revision: an introduction. In: P. Gardenfors, ed., *Belief Revision,* Cambridge University Press (1992) 1-28.
15. P. Gardenfors and D. Makinson, Revision of knowledge systems using epistemic entrenchment. In: *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge,* Los Altos, Ca. (1988).
16. H. Geffner, *Default Reasoning: Causal and Conditional Theories.* MIT Press (1992).
17. S. O. Hansson, A dyadic representation of belief. In: P. Gardenfors, ed., *Belief Revision,* Cambridge University Press (1992) 89-121.
18. W. L. Harper, Rational conceptual change. In: *Bulletin of the Philosophy of Science Association,* v. 2 (1977) 462-494.
19. E. T. Jaynes, Where do we stand on maximum entropy? In: R. D. Levine and M. Tribus (eds.) *The Maximum Entropy Formalism,* MIT Press, Cambridge, Mass. (1979) 15-118.
20. H. Katsuno and A. O. Mendelzon, On the difference between updating a knowledge base and revising it. In: P. Gardenfors, ed., *Belief Revision,* Cambridge University Press (1992) 183-203.

21. J. M. Keynes, *A Treatise on Probability,* Macmillan, London (1921).

22. K. Konolige, Hierarchic autoepistemic theories for nonmonotonic reasoning. In: *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence,* Saint Paul, Minn. (1988).

23. Lea Sombe, A glance at revision and updating in knowledge bases. *International Journal of Intelligent Systems,* v. 9 (1994) 1-27.

24. I. Levi, Subjunctives, dispositions and chances. *Synthese,* v.34 (1977) 423-455.

25. F. Levy, A survey of belief revision and updating in classical logic. *International Journal of Intelligent Systems,* v. 9 (1994) 29-59.

26 V. Lifschitz, Computing circumscription. In: *Proceedings of the 9th International Joint Conference on Artificial Intelligence,* Los Angeles (1985) 121-127.

27. E. Lozinskii, Information and evidence in logic systems. *Journal of Experimental and Theoretical Artificial Intelligence,* v. 6 (1994) 163-193.

28. E. Lozinskii, Resolving contradictions: a plausible semantics for inconsistent systems. *Journal of Automated Reasoning,* v. 12 (1994) 1-31.

29. D. Makinson, How to give it up: a survey of some formal aspects of the logic of theory change. *Synthese,* v. 62 (1985) 347-363.

30. D. Makinson, On the status of the postulate of recovery in the logic of theory change. *Journal of Philosophical Logic,* v. 16 (1987) 383-394.

31. R. Neapolitan, *Probabilistic Reasoning in Expert Systems.* John Wiley & Sons (1990).

32. B. Nebel, Reasoning and Revision in Hybrid Representation Systems. *Lecture Notes in Artificial Intelligence,* v. 422, Springer-Verlag (1990).

33. B. Nebel, Syntax-based approaches to belief revision. In: P. Gardenfors, ed., *Belief Revision,* Cambridge University Press (1992) 52-88.

34. N. Nilsson, Probabilistic logic. *Artificial Intelligence,* v. 28 (1986) 71-87.

35. A. S. Rao and N. Y. Foo, Formal theories of belief revision. In: *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning,* Toronto, Ontario (1989) 369-380.

36. A. S. Rao and N. Y. Foo, Minimal change and maximal coherence: a basis for belief revision and reasoning about actions. In: *Proceedings of the 11th IJCAI,* Detroit, Mich. (1989) 966-971.

37. N. Rescher, *Hypothetical Reasoning.* North-Holland (1964).

38. N. Rescher and R. Manor, On inference from inconsistent premises. *Theory and Decision,* (1970) 179-219.

39. N. Roos, A logic for reasoning with inconsistent knowledge. *Artificial Intelligence,* v. 57 (1992) 69-103.

40. H. Rott, Two methods of constructing contractions and revisions of knowledge systems. *Journal of Philosophical Logic,* v. 20 (1991) 149-173.

41. C. B. Schwind, Revision and action. *International Journal of Intelligent Systems,* v. 9 (1994) 143-152.

42. L. Willard and Y. Y. Li, The revised Gardenfors postulates and update semantics. In: *Proceedings of the 3rd International Conference on Database Theory,* Paris (1990). Lecture Notes in Computer Science, v. 470, Springer-Verlag, Berlin.

43. M. S. Winslett, Reasoning about action using a possible models approach. In: *Proceedings of the 7th National Conference of American Association for Artificial Intelligence,* Saint Paul, Minn. (1988) 89-93.

44. M. S. Winslett, *Updating Logical Databases.* Cambridge University Press (1990).