# Resolving Contradictions: A Plausible Semantics for Inconsistent Systems

Eliezer L. Lozinskii
Institute of Mathematics and Computer Science
The Hebrew University, Jerusalem 91904, Israel

**Abstract.** The purpose of a Knowledge System $S$ is to represent the world $W$ faithfully. If $S$ turns out to be inconsistent containing contradictory data, its present state can be viewed as a result of information pollution with some wrong data. However, we may reasonably assume that most of the system content still reflects the world truthfully, and therefore it would be a great loss to allow a small contradiction to depreciate or even destroy a large amount of correct knowledge. So, despite the pollution, $S$ must contain a meaningful subset, and so it is reasonable to assume (as adopted by many researchers) that the semantics of a logic system is determined by that of its maximally consistent subsets, *mc-subsets*. The information contained in $S$ allows deriving certain conclusions regarding the truth of a formula $F$ in $W$. In this sense we say that $S$ contains a certain amount of *semantic information*, and provides an *evidence of* $F$. A close relationship is revealed between the evidence, the quantity of semantic information of the system, and the set of models of its mc-subsets. Based on these notions, we introduce the *semantics of weighted mc-subsets* as a way of reasoning in inconsistent systems. To show that this semantics indeed enables reconciling contradictions and deriving plausible beliefs about any statement including ambiguous ones, it is successfully applied to a series of justifying examples, such as chain proofs, rules with exceptions, and paradoxes.

## 1. Introduction

Consider a Knowledge System $S$ describing (a part of) the real world $W$. $S$ is viewed and used as a source of information about $W$ in the sense that in order to find out whether (and to what extent) a statement $\alpha$ is true in $W$, one has to check the truth value of $\alpha$ in $S$. So, the purpose of the Knowledge System is to represent $W$ faithfully, and the meaning of the system must conform to the world in such a way that each model of $S$ corresponds to a possible state of $W$. Although $S$ is intended for an adequate description of $W$, the very nature of designing and maintaining of a Knowledge System may affect the degree of its adequacy. In particular, the system may contain contradictory data implying, for instance, that a statement $\alpha$ is true and at the same time false (for some other reason). There are numerous possible causes of contradiction in Knowledge Systems (as it has been pointed out repeatedly, e.g., [2, 3, 9, 10]). Indeed, different experts that provide the system with their professional opinion may not agree on some issues (such as a medical diagnosis), different witnesses may give conflicting testimonies in a court room, a control system may collect contradictory measurements from imprecise or faulty instruments, there may take place erroneous updates, wrong data may penetrate the system intentionally by some hostile or subversive activity, etc.

The real possibility that a Knowledge System contains contradictory data has stimulated an extensive search for methods of reasoning in inconsistent logic systems [4, 6, 8, 9, 14, 16, 23-25, 28, 29, 34].

The classical logic, unfortunately, can be of little use for reasoning with inconsistency. The problem is that for any formula $F$, a conjunction $F \wedge \neg F$ is unsatisfiable in the classical logic, and so contradictory data make a system $S$ trivially meaningless, since *any* formula $\phi$ becomes its logical consequence. The need of special tools for reasoning with inconsistent information led to developing of *paraconsistent logics* [4, 8, 9, 14, 16, 25, 28, 29, 34], in which a conjunction of contradictory statements can be assigned a certain meaning, or at least does not depreciate the content of the rest of the system. To achieve this goal, paraconsistent methods introduce special axioms, and use multiple-valued logics containing in addition to the conventional truth values *'true', 'false'* also a value *'both true and false'* . If $S$ contains $\phi \wedge \neg\phi$ , then assigning of the latter truth value to $\phi$ provides it with a certain meaning, and allows "recovering" of some information not affected by the contradiction. For instance, considering $S = \{p, \neg p, q\}$ , one can view as a model of $S$ a structure in which $p$ is *both true and false* , but $q$ is just *true* .

Use of a multiple-valued logic provides an inconsistent system $S$ with an *epistemological* meaning in the sense that $S$ can reflect beliefs of a reasoner (human or machine) who may happen to hold conflicting beliefs [4, 16, 28]. However, the system remains *ontologically* inconsistent as far as it is agreed that the reality complies with the laws of *Excluded Middle* and *Non-Contradiction* , which means that in every possible state of the real world every statement is either true or false exclusively.

Thus, given a logic system presenting some knowledge about the real world, one tries to figure out what is the truth value of a given statement in the reality. In the case that the system contains uncertain or even contradictory information, the goal is to decide whether the statement is more likely to be true than false in the world. The recent study of the problem of recovering the meaning of contradictory data [2, 16, 34] produced results for some limited cases of inconsistency. The following sections present a general approach to resolving contradictions in First Order logic systems. Plausibility of the proposed solution is illustrated by examples appealing to the common sense intuition. The solution takes into account a semantics based on consistent subsets of a given logic system.

## 2. Semantics of consistent subsets

Consider a First Order logic system $S$ as a set of well-formed consistent formulas or, interchangeably, as a conjunction of all formulas of the set. We assume that the domain of $S$ is specified. However, all considerations in the sequel hold for the Herbrand Universe of $S$ .

**Definition 2.1.** *Mc-subset.* A subset $\sigma$ of $S$ is a *maximally consistent subset* (briefly, *mc-subset* ) if $\sigma$ is consistent, but for all formulas $\phi \in (S - \sigma)$ , $\sigma \cup \{\phi\}$ is inconsistent. If $S$ is consistent then it is its only mc-subset. ■

**Definition 2.2.** *Classification of formulas.* If a formula $\psi$ is a logical consequence of all mc-subsets $\sigma$ of $S$ , $\sigma \models \psi$ , then $\psi$ is a *fact* in $S$ . If $\psi$ is consistent with every mc-subset of $S$ , and a logical consequence of some but not all of them, then $\psi$ is *supported* by $S$ . If for all mc-subsets $\sigma$ of $S$ , $\psi$ is consistent with $\sigma$ , but not a logical consequence of $\sigma$ , then $\psi$ is *vague* in $S$ . If $S$ contains consistent subsets $\alpha, \beta$ such that $\alpha \models \psi$ and $\beta \models \neg\psi$ , then $\psi$ is *ambiguous* in $S$ . ■

**Example 2.1.** $S = \{ \alpha, \beta, \neg\beta, \neg\beta \vee \gamma, \delta \vee \neg\eta \}$ .
The system has two mc-subsets:

$$\sigma_1 = \{\ \alpha,\ \beta,\ \neg\beta \vee \gamma,\ \delta \vee\ \neg\eta\ \}\ ,$$
$$\sigma_2 = \{\ \alpha,\ \neg\beta,\ \neg\beta \vee \gamma,\ \delta \vee\ \neg\eta\ \}\ .$$

So, by Definition 2.2, $\alpha$ is a fact in $S$, $\beta$ is an ambiguous formula, $\gamma$ is supported by $S$, while $\eta$ is vague. ■

Suppose that a Knowledge System $S$ turns out to be inconsistent containing contradictory data, in particular, some ambiguous formulas. Since the system has been designed to contain data adequate to the real world, we may view its present inconsistency as a result of *information pollution* with some wrong data. Unfortunately (cf. Section 1), there exist numerous causes of information pollution. However, we may reasonably assume that (say, due to a proper maintenance) most of the system content still reflects the world truthfully, and therefore it would be a great loss to allow a small contradiction to depreciate or even destroy a large amount of correct knowledge. So, despite the pollution, $S$ must contain a meaningful subset, and we want the latter to be maximally consistent in order not to lose any correct data. On the other hand, since the world is supposed to be consistent, only consistent subsets of the system have a chance to conform to it.

Thus, it appears reasonable to assume that *the semantics of a logic system is determined by that of its mc-subsets*. This principle, adopted by many researchers in the field (e.g., [2, 31]), is trivially correct for consistent systems, but is not easy to implement in a general case, since different mc-subsets of an inconsistent system may contain different and even conflicting data.

Two approaches are commonly applied to the semantics of mc-subsets: a credulous or a skeptical reasoning [31, 32]. A *credulous* reasoner is ready to accept a belief that is true in *some* consistent theory of the world corresponding to one of its possible states. On the other extreme, a *skeptical* reasoner believes only statements that are true in *all* mc-subsets of a given system. If a credulous reasoner wishes to avoid contradictions in his/her beliefs, he/she should not make decisions regarding ambiguous statements. The reasoner may prefer a certain mc-subset as a faithful representation of the world, but in this case he loses the information contained in other mc-subsets of the system, and so his beliefs may become ill-founded. On the other hand, a skeptical reasoner looks for a very strong justification of his beliefs, and so he usually ends up with a very few conclusions, never making any decision regarding ambiguous statements.

Thus, both types of reasoners make no use of a large part of the system's information, and, in particular, have no means for resolving contradictions.

Our goal is to develop tools for deriving plausible beliefs about any statement, including ambiguous ones. So, given any system $S$ (either consistent or not), and a consistent formula $\phi$, we intend to answer the question of whether $\phi$ is more likely to be true in the reality than false. A successful answering of this question would allow to reconcile contradictions in Knowledge Systems. And we wish to build our conclusions on a full utilization of the information contained in all mc-subsets of the system.

To see an intuitive way of reconciling contradictions, let us consider the following motivating example.

**Example 2.2.** Let $S$ be a consistent and correct knowledge system:

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

such that all the atomic formulas, $p_1, p_2, \ldots, p_k, q, r$ are true in $S$. Suppose that $\neg r$ is now added to $S$ (possibly by mistake) producing

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

such that all the atoms become at once ambiguous in $S'$, since now for each atom $x$ there is a consistent subset of $S'$ proving $x$ (e.g., $S$ itself), and another one proving $\neg x$ (indeed, $\{q \rightarrow r, \neg r\} \models \neg q$, and for all $1 \leq i \leq k$, $\{p_i \rightarrow q, q \rightarrow r, \neg r\} \models \neg p_i$). However, this does not mean that there remains no way of judgement about plausible truth values of the atoms. Consider, for instance, $q$. The modified system $S'$ contains $k$ independent proofs of its truthfulness, namely, $\{p_1, p_1 \rightarrow q\}, \ldots, \{p_k, p_k \rightarrow q\}$, while there is only one reason for $q$ to be false: $\{q \rightarrow r, \neg r\}$. So, if $k > 1$, and in the absence of more information about the credibility of the clauses of $S'$, one may reasonably assume that $q$ is more likely to be true than false. And the larger is $k$, the more justified appears this conclusion. We shall analyze this example in more details in the sequel. First, the next section presents an estimation of the quantity of semantic information.

## 3. Semantic information

Let us denote by $MC(S)$ the set of all mc-subsets of $S$, and by $MOD(\sigma)$ the set of all models of a consistent subset $\sigma$ of $S$. We assume that every maximally consistent subset of $S$ contributes to the semantics of the system. With this in mind we call each model of every mc-subset of $S$ a *quasi-model* of $S$, and then $QM(S)$ denotes the set of all quasi-models of $S$:

$$QM(S) = \bigcup_{\sigma \in MC(S)} MOD(\sigma).$$

If $S$ is consistent, then it is the only mc-subset of itself, and so, quasi-models of $S$ are just its models: $QM(S) = MOD(S)$.

As a Knowledge System allows assessment of the truth of various statements, it must contain *semantic information*, which we wish to utilize as fully as possible for making plausible decisions. So, we have to estimate the *quantity of semantic information*, $I(S)$, contained in a logic system $S$.

In [21] quantity of information has been defined for consistent logic systems. Here this definition is extended for any First Order logic system, including inconsistent ones. First, let us examine desirable properties that semantic information should possess to conform to human common sense intuition.

(i) An empty system contains no information: $I(\emptyset) = 0$.

(ii) Let the *base* of a system $S$ (denoted $Base(S)$) be the set of all ground atomic formulas of $S$ (that is, all ground instances of all predicates appearing in $S$). If $S$ consists of pairs of complement literals $A, \neg A$ for all atoms $A \in Base(S)$, then all the atoms of $Base(S)$ are ambiguous, so $S$ contains no semantic information as well.

Consider a system $S$, and a formula $F$ such that $Base(F) \subseteq Base(S)$.

(iii) If $S$ is consistent, and $F$ is its logical consequence, then addition of $F$ to $S$ should not change the information of $S$. Let $I(\phi \mid Cond)$ denote the quantity of information of a formula $\phi$ provided that a condition $Cond$ is satisfied. Then $I(S \cup \{F\} \mid S \models F) = I(S)$. As a generalization of this property, if $S$ is inconsistent, and $F$ is a fact in $S$, that is a logical consequence of every mc-subset of $S$ (cf. Definition 2.2), then addition of $F$ to $S$ should not change $I(S)$.

(iv) If $S$ is consistent, and $F$ is consistent with $S$, but not a fact in $S$, then the truth of $F$ in $S$ is uncertain. Addition of $F$ to $S$ eliminates this uncertainty, and so increases $I(S)$. By this intuition, addition of $F$ to an inconsistent system $S$ should increase $I(S)$ if $F$ is not a fact in $S$, but consistent with every mc-subset of $S$ (i.e., $F$ is either supported or vague in $S$, by Definition 2.2).

(v) If $S$ is consistent, and a consistent formula $F$ contradicts it (that is, $S \models \neg F$), then addition of $F$ to $S$ must decrease $I(S)$. This property implies that an inconsistent system contains less semantic information than any of its mc-subset $\sigma$, since $S - \sigma$ contradicts $\sigma$ (cf. Definition 2.1).

(vi) A system $S$ contains a maximum possible quantity of semantic information if it assigns a unique truth value to every atom $A \in Base(S)$, i.e., for all $A \in Base(S)$ exactly one of $A$ or $\neg A$ is a logical consequence of $S$. In this case $S$ has a unique model, and is consistent. Therefore no inconsistent system can possess a maximum quantity of information possible for a consistent system with the same base.

Properties (i)-(vi) indicate a close connection between the quantity of semantic information of a system and the set of its quasi-models. Indeed, to contain the maximum possible quantity of information, a system must be consistent and have only one model, that is the minimum number of models of a consistent system. At the other extreme, if a base $B$ is associated with an empty system containing, by property (i), the minimum possible (zero) quantity of information, then any interpretation of $B$ satisfies the system, so it has $2^{|B|}$ models, that is the maximum possible number of models in the given domain. In between, e.g., by property (iii), if a formula $F$ is a logical consequence of every mc-subset of $S$, then its addition to $S$ does not change $I(S)$. But in this case $F$ is true in every quasi-model of $S$, so its addition to $S$ causes no changes in $QM(S)$ either.

Thus, the quantity of semantic information contained in a logic system $S$ is a decreasing function of the size $|QM(S)|$ of the set of its quasi-models, and according to the widely adopted general definition of information quantity [1, 5], $I(S)$ must be a logarithmic function of $|QM(S)|$. (Here we consider systems with finite bases and sets of models, however, the approach and main results of this study can be applied to infinite systems. An example is given in Section 8.3).

These observations suggest the following definition of quantity of information of a logic system (in the sequel all logarithms are to the base 2).

**Definition 3.1.** *Quantity of information.*

$$I(S) = \log \frac{2^{|Base(S)|}}{|QM(S)|} = |Base(S)| - \log |QM(S)|. \quad \blacksquare$$

**Theorem 3.1.** Definition 3.1 is in accord with properties (i)-(vi).

**Proof.** (i) If $S = \varnothing$ then $S$ is its only mc-subset, and so, $|QM(S)| = |MOD(S)| = 2^{|Base(S)|}$. Hence, by Definition 3.1, $I(S) = 0$.

(ii) If $S = \bigcup_{A \in Base(S)} \{A, \neg A\}$ , then there are $2^{|Base(S)|}$ mc-subsets such that each mc-subset includes exactly one of $A$ or $\neg A$ for all $A \in Base(S)$ . Every mc-subset has only one model, so $|QM(S)| = 2^{|Base(S)|}$ , and again $I(S) = 0$ .

Consider a system $S$ and a formula $F$ such that $Base(F) \subseteq Base(S)$ .

(iii) Let $F$ be a logical consequence of every mc-subset of $S$ , $Base(F) \subseteq Base(S)$ , and $S' = S \cup \{F\}$ . If $F$ is added to $S$ then every mc-subset $\sigma$ of $S$ absorbs $F$ in the sense that if $\sigma' = \sigma \cup \{F\}$ , then (a) $MOD(\sigma) = MOD(\sigma')$ , and (b) $\sigma \in MC(S)$ iff $\sigma' \in MC(S')$ . Hence, $QM(S') = QM(S)$ , $Base(S') = Base(S)$ , and so, $I(S') = I(S)$ .

(iv) If $F$ is not a fact in $S$ , but consistent with every mc-subset of $S$ , then there is a mc-subset $\sigma$ of $S$ such that $\sigma \not\models F$ . Let $T_1$ denote the set of all such mc-subsets, while $T_2 = MC(S) - T_1$ . Then $T_2$ is either empty or consists of mc-subsets $\omega$ such that $\omega \models F$ (since there is no mc-subset proving $\neg F$ ). Denote $S' = S \cup \{F\}$ , $\sigma' = \sigma \cup \{F\}$ , $\omega' = \omega \cup \{F\}$ . If $F$ is added to $S$ , then every mc-subset of $S$ absorbs $F$ in such a way that (a) for all $\sigma \in T_1$ , $MOD(\sigma') \subset MOD(\sigma)$ , since the set $MOD(\sigma')$ contains only those models of $\sigma$ in which $F$ is true, and (b) for all $\omega \in T_2$ , $MOD(\omega') = MOD(\omega)$ , since $\omega \models F$ . Hence, $QM(S') \subset QM(S)$ , and so, $I(S') > I(S)$ .

(v) Let $S$ be a consistent set of formulas, $F$ a consistent formula such that $S \models \neg F$ , and $S' = S \cup \{F\}$ . $S$ is an mc-subset of the inconsistent $S'$ , and since $F \notin S$ , there must be another mc-subset $\sigma$ of $S'$ such that $F \in \sigma$ . $S$ and $\sigma$ have no model in common, since should such a model exist, it would be a model of $S \cup \{F\}$ , while the latter is inconsistent. Hence, $MOD(S) \subset QM(S')$ , and so, $I(S') < I(S)$ .

(vi) If a system $S$ has only one model then, by Definition 3.1, the quantity of its information reaches the maximum of $|Base(S)|$ . If $S$ is inconsistent then it contains more than one mc-subset. Hence, $|QM(S)| > 1$ , and so, $I(S) < |Base(S)|$ . ■

Theorem 3.1 can be extended to provide a characterization of different types of formulas classified by Definition 2.2.

**Theorem 3.2.** Let $\phi$ be a consistent formula, and $Base(\phi) \subseteq Base(S)$ .
(a) If $\phi$ is a fact in $S$ , then $I(S \cup \{\phi\}) = I(S)$ , but $I(S \cup \{\neg\phi\}) < I(S)$ ;
(b) If $\phi$ is supported by $S$ , then $I(S \cup \{\phi\}) > I(S)$ ;
(c) If $\phi$ is vague in $S$ , then both $I(S \cup \{\phi\})$ and $I(S \cup \{\neg\phi\})$ are greater than $I(S)$ .

**Proof.** (a) Implied by Theorem 3.1(iii, v).

(b) Implied by Theorem 3.1(iv).

(c) If $\phi$ is vague in $S$ , then every mc-subset of $S$ is consistent with each one of $\phi$ or $\neg\phi$ . So, being added to $S$ , each one of $\phi$ or $\neg\phi$ becomes absorbed by every mc-subset of $S$ decreasing the number of its models. Hence, $QM(S') \subset QM(S)$ , and $I(S') > I(S)$ . ■

A few remarks on Theorem 3.2 are in order.

**Remark 3.1** (a) If $\phi$ is supported by $S$ , then $I(S \cup \{\neg\phi\})$ can be either greater than, or equal to, or less than $I(S)$ (cf. Theorem 3.2(b)). Indeed, consider $S' = S \cup \{\neg\phi\}$ . Let $MC(S) = T_1 \cup T_2$ such that for all $\sigma \in T_1$ , $\theta \in T_2$ , it holds that $\sigma \models \phi$ , $\theta \not\models \phi$ , $\theta \not\models \neg\phi$ . Addition of $\neg\phi$ to $S$ affects mc-subsets of $S$ and sets of their models in three different ways. First, mc-subsets of $T_1$ and their models remain unchanged. Second, $\neg\phi$ becomes absorbed by every $\theta \in T_2$, and decreases the number of models of $\theta$ (by cancelling all models of $\theta$ that assert $\phi$ ). So, some quasi-models of $QM(S)$ do not belong to $QM(S')$. Third, $\neg\phi$ makes up new mc-subsets of $S'$ by joining proper subsets of the mc-subsets of $T_1$. Models of the new mc-subsets (in contrast to the second way) belong to $QM(S')$, but not to $QM(S)$. So, the total difference $|QM(S')| - |QM(S)|$ depends on the specific structure of $S$, as illustrated by Example 3.1.

(b) If $\phi$ is ambiguous in $S$, then addition of $\phi$ or $\neg\phi$ to $S$ can either increase $I(S)$, or decrease it, or leave it unchanged. Indeed, if $\phi$ is ambiguous in $S$, then $\neg\phi$ is ambiguous as well. Let $MC(S) = T_1 \cup T_2 \cup T_3$ such that for all $\sigma \in T_1$, $\theta \in T_2$, $\omega \in T_3$, the following holds: $\sigma \models \phi$, $\theta \not\models \phi$, $\theta \not\models \neg\phi$, $\omega \models \neg\phi$. Then by an argument similar to that of (a), the total change $I(S \cup \{\phi\}) - I(S)$ or $I(S \cup \{\neg\phi\}) - I(S)$ can be either positive, or zero, or negative depending on the specific structure of $S$. ■

**Example 3.1.** This example presents 3 instances of an inconsistent system, in which $p$ is a supported atom (cf. Remark 3.1(a) ). It is shown that depending on the content of the system, addition of $\neg p$ may either change the quantity of its information in any direction, or leave it unaffected. The following notation is common to all 3 instances: $\sigma, \theta$ denote mc-subsets of $S$ such that $\sigma \models p$, $\theta \not\models p, \theta \not\models \neg p$; $\omega_1, \omega_2, \omega_3$ are mc-subsets of $S'$; $m_i, \mu_j$ denote quasi-models of $S, S'$, respectively.

$$\boldsymbol{S_1 = \{\, a, \neg a, a \rightarrow p \,\}},$$

$\sigma = \{\, a, a \rightarrow p \,\}$ $\quad$ $m_1 = \{\, a, p \,\}$
$\theta = \{\, \neg a, a \rightarrow p \,\}$ $\quad$ $m_2 = \{\, \neg a, p \,\}$
$m_3 = \{\, \neg a, \neg p \,\}$.

$$S_1' = S_1 \cup \{\, \neg p\} = \{\, a, \neg a, a \rightarrow p, \neg p \,\},$$

$\omega_1 = \{\, a, a \rightarrow p \,\} = \sigma$ $\quad$ $\mu_1 = \{\, a, p \,\}$
$\omega_2 = \{\, \neg a, a \rightarrow p, \neg p \,\} = \theta \cup \{\, \neg p\}$ $\quad$ $\mu_2 = \{\, \neg a, \neg p \,\}$
$\omega_3 = \{\, a, \neg p \,\}$, *this is a new mc-subset* $\quad$ $\mu_3 = \{\, a, \neg p \,\}$

$$|QM(S)| = |QM(S')| = 3, \qquad \boldsymbol{I(S') = I(S)} = 0.42 .$$

$$\boldsymbol{S_2 = \{\, a, \neg a, a \rightarrow p, a \rightarrow b \,\}},$$

$\sigma = \{\, a, a \rightarrow p, a \rightarrow b \,\}$ $\quad m_1 = \{\, a, b, p \,\}$

$\theta = \{\, \neg a, a \rightarrow p, a \rightarrow b \,\}$ $\quad m_2 = \{\, \neg a, b, p \,\}$

$m_3 = \{\, \neg a, \neg b, p \,\}$

$m_4 = \{\, \neg a, b, \neg p \,\}$

$m_5 = \{\, \neg a, \neg b, \neg p \,\}$

$$S_2' = S_2 \cup \{\, \neg p\} = \{\, a, \neg a, a \rightarrow p, a \rightarrow b, \neg p \,\},$$

$\omega_1 = \{\, a, a \rightarrow p, a \rightarrow b \,\} = \sigma$ $\quad \mu_1 = \{\, a, b, p \,\}$

$\omega_2 = \{\, \neg a, a \rightarrow p, a \rightarrow b, \neg p \,\} = \theta \cup \{\, \neg p\}$ $\quad \mu_2 = \{\, \neg a, b, \neg p \,\}$

$\mu_3 = \{\, \neg a, \neg b, \neg p \,\}$

$\omega_3 = \{\, a, a \rightarrow b, \neg p \,\}$, *this is a new mc-subset* $\quad \mu_4 = \{\, a, b, \neg p \,\}$

$$|QM(S)| = 5, \quad |QM(S')| = 4, \quad I(S) = 0.68, \quad I(S') = 1.00, \quad \boldsymbol{I(S') > I(S)}.$$

$$\boldsymbol{S_3 = \{\, a, \neg a, a \rightarrow p, \neg a \rightarrow b \,\}},$$

$\sigma = \{\, a, a \rightarrow p, \neg a \rightarrow b \,\}$ $\quad m_1 = \{\, a, b, p \,\}$

$m_2 = \{\, a, \neg b, p \,\}$

$\theta = \{\, \neg a, a \rightarrow p, \neg a \rightarrow b \,\}$ $\quad m_3 = \{\, \neg a, b, p \,\}$

$m_4 = \{\, \neg a, b, \neg p \,\}$

$$S_3' = S_3 \cup \{\, \neg p\} = \{\, a, \neg a, a \rightarrow p, \neg a \rightarrow b, \neg p \,\},$$

$\omega_1 = \{\, a, a \rightarrow p, \neg a \rightarrow b \,\} = \sigma$ $\quad \mu_1 = \{\, a, b, p \,\}$

$\mu_2 = \{\, a, \neg b, p \,\}$

$\omega_2 = \{\, \neg a, a \rightarrow p, \neg a \rightarrow b, \neg p \,\} = \theta \cup \{\, \neg p\}$ $\quad \mu_3 = \{\, \neg a, b, \neg p \,\}$

$\omega_3 = \{\, a, \neg a \rightarrow b, \neg p \,\}$, *this is a new mc-subset* $\quad \mu_4 = \{\, a, b, \neg p \,\}$

$\mu_5 = \{\, a, \neg b, \neg p \,\}$

$$|QM(S)| = 4, \quad |QM(S')| = 5, \quad I(S) = 1.00, \quad I(S') = 0.68, \quad \boldsymbol{I(S') < I(S)}. \quad \blacksquare$$

Given a description $S$ of a world $W$, a reasoner has to derive beliefs of various events and situations in $W$ from the semantic information presented in $S$. With this in mind, we try in the following section to asses the credibility of such beliefs on the bases of quantity of semantic information of $S$, and its connection to sets of the system models.

## 4. Evidence

Consider a system $S$, and a formula $F$. Given the information about a world $W$ contained in $S$, we may derive certain conclusions regarding the truth of $F$ in $W$. In this sense we say that $S$ provides a certain *evidence of* $F$, denoted $E(S, F)$. We wish the evidence to provide a measure of credibility of a belief in truth of $F$. This section presents a way of estimation of $E(S, F)$, and shows its connection with the semantic information of $S$.

**4.1. Postulates of evidence**

To present a measure of credibility of a belief in truth of $F$, the evidence should comply with the following common sense postulates.

(1) If $F$ is a logical consequence of all mc-subsets of $S$, then the truth of $F$ is most credible, so $E(S, F) = 1$ ( let the range of evidence be normalized in the interval [0, 1] ). In this case a belief in falsehood of $F$ has no credibility, so $E(S, \neg F) = 0$ .

(2) In particular, if $\phi$ is a tautology, and $\psi$ is an unsatisfiable formula, then for all $S$, $E(S, \phi) = 1$ and $E(S, \psi) = 0$ .

(3) For all formulas $\phi, \psi$,
$E(S, \phi \vee \psi) \geq max( E(S, \phi) , E(S, \psi) )$ and $E(S, \phi \wedge \psi) \leq min( E(S, \phi) , E(S, \psi) )$ .
In particular, $E(S, \phi \vee \psi) = E(S, \phi) + E(S, \psi) - E(S, \phi \wedge \psi)$ .

(4) Postulates (2) and (3) imply for all $\phi$, $E(S, \phi) + E(S, \neg\phi) = 1$ .

(5) Consider a system $S$ (as in (ii), Section 3) consisting of pairs of complementary ground literals for all ground atoms of its base, $S = \bigcup_{A \in Base(S)} \{A, \neg A\}$ . This system must provide the same evidence for every ground atom and its negation, such that for all $A \in Base(S)$, $E(S, A) = E(S, \neg A) = 0.5$ .

(6) The more likely is that $F$ is true in $W$, the larger must be the value of $E(S, F)$ . And the most important: if $F$ is more likely to be true than false, then $E(S, F) > E(S, \neg F)$ .

## 4.2. Evidence vs. information

If a formula $F$ is considered on a background of some prior knowledge, then the quantity of *useful information* carried by $F$ should be estimated relative to this knowledge.

To reveal a connection between evidence and semantic information, let us first consider a consistent set of formulas $\sigma$, and a formula $F$ consistent with $\sigma$. Let $I(F/\sigma)$ denote the quantity of information contained in $F$ *relative* to $\sigma$. With regard to $\sigma$, $F$ is either a fact or a vague formula, so, by Theorem 3.2(a, c), $I(\sigma \cup \{F\}) \geq I(\sigma)$. If an addition of $F$ to $\sigma$ increments $I(\sigma)$, this means that $F$ indeed carries some amount of new information relative to $\sigma$, and the larger is this increment, the larger is $I(F/\sigma)$. This observation implies the following definition. Note that for a consistent $\sigma$, Definition 3.1 becomes $I(\sigma) = |Base(\sigma)| - \log |MOD(\sigma)|$.

**Definition 4.1.** *Relative information.*

$$I(F/\sigma) = I(\sigma \cup \{F\}) - I(\sigma) = \log \frac{|MOD(\sigma)|}{|MOD(\sigma \cup \{F\})|} . \quad \blacksquare$$

Adding $F$ to $\sigma$ means claiming that $F$ is true. So, the larger is the evidence of $F$ provided by $\sigma$, the smaller must be the amount of information contributed to $\sigma$ by $F$. Indeed, if, for instance, $\sigma \models F$ then by postulate (1), $E(\sigma, F) = 1$. But in this case, by Theorem 3.2(a), $I(\sigma \cup \{F\}) = I(\sigma)$, so $I(F/\sigma) = 0$. Hence, $E(S, F)$ must be a decreasing function, $g$, of $I(F/S)$. Then, for formulas $\phi$, $\psi$ consistent with $\sigma$, postulate 3 requires

$$g(I(\phi \vee \psi/\sigma)) = g(I(\phi/\sigma)) + g(I(\psi/\sigma)) - g(I(\phi \wedge \psi/\sigma)) ,$$

that is

$$\begin{aligned} & g(log \frac{|MOD(\sigma)|}{|MOD(\sigma \cup \{\phi \vee \psi\})|}) \\ & = g(log \frac{|MOD(\sigma)|}{|MOD(\sigma \cup \{\phi\})|}) + g(log \frac{|MOD(\sigma)|}{|MOD(\sigma \cup \{\psi\})|}) - g(log \frac{|MOD(\sigma)|}{|MOD(\sigma \cup \{\phi \wedge \psi\})|}). \end{aligned} \tag{4.1}$$

Note that

$$\begin{aligned} & |MOD(\sigma \cup \{\phi \vee \psi\})| \\ & = |MOD(\sigma \cup \{\phi\})| + |MOD(\sigma \cup \{\psi\})| - |MOD(\sigma \cup \{\phi \wedge \psi\})| . \end{aligned} \tag{4.2}$$

So, (4.1) and (4.2) yield for $g$:

$$g(x) = c2^{-x} ,$$

so that

$$E(\sigma, F) = c2^{-I(F/\sigma)} .$$

To satisfy postulate 1,

$$E(\sigma, F \mid \sigma \models F) = c2^{-I(F/\sigma \mid \sigma \models F)} = c = 1 .$$

Hence, finally,

$$E(\sigma, F) = 2^{-I(F/\sigma)} = \frac{|MOD(\sigma \cup \{F\})|}{|MOD(\sigma)|}, \tag{4.3}$$

that complies with the rest of postulate 1, since if $S \models F$ then $|MOD(S \cup \{\neg F\})| = 0$ .

### 4.3. Evidence vs. sets of models

The postulates of evidence, and the expression (4.3) show a close relationship of evidence to the semantic information of the system, and through the information — to the set of quasi-models. Indeed, if a system $S$ contains mostly truthful information about the world $W$ (along with a minor portion of wrong data), then most quasi-models of $S$ correspond to possible states of $W$ . The more states of $W$ affirm a statement $\alpha$ , the more quasi-models assert its truth, and vice versa. But the more states of $W$ affirm $\alpha$ , the more likely is that $\alpha$ is true in a given state of the world. So, in general, the more quasi-models of $S$ assert $\alpha$ , the larger must be the evidence $E(S, \alpha)$ (cf. postulate (6)). Expression (4.3) is in full accord with this intuition with respect to consistent systems. However, applying this intuition to inconsistent systems, we must take into consideration the following difference.

Every model of a consistent system $S$ satisfies all constituent formulas of $S$ , so all the models of $S$ have the same semantic status, and so they are just counted up for computing $E(S, \alpha)$ . However, a quasi-model of an inconsistent system $S'$ satisfies only a proper subset of $S'$ , so different quasi-models may differ in their contribution to the evidence, since the corresponding mc-subsets represent different portions of $S'$ . Let us illustrate this observation by the following example.

**Example 4.1.** $S = \{ a, b, \neg b, a \vee c, \neg a \vee c, \neg b \vee \neg c \}$ .
There are 4 mc-subsets, and 4 quasi-models:

$\sigma_1 = \{ a, b, a \vee c, \neg a \vee c \}$ , $\quad m_1 = \{ a, b, c \}$ ;
$\sigma_2 = \{ a, b, a \vee c, \neg b \vee \neg c \}$ , $\quad m_2 = \{ a, b, \neg c \}$ ;
$\sigma_3 = \{ a, \neg b, a \vee c, \neg a \vee c, \neg b \vee \neg c \}$ , $\quad m_3 = \{ a, \neg b, c \}$ ;
$\sigma_4 = \{ b, \neg a \vee c, \neg b \vee \neg c \}$ , $\quad m_4 = \{ \neg a, b, \neg c \}$ . ■

All atomic propositions of Example 4.1 are ambiguous in $S$ , since for each one of $a, b, c$ there is a mc-subset of $S$ proving its truth, and another one proving its falsehood. However, the propositions are defined in $S$ in different ways, so $S$ provides them with different evidence. One may try to figure out this evidence by applying a common sense intuition, like this. The evidence of $a$ seems to be larger than that of $\neg a$ , since $a$ is stated explicitly and unconditionally, while to prove $\neg a$ (from $\neg a \vee c$ ) a proviso is needed, namely that $c$ is false, but the latter implies $a$ as well (due to $a \vee c$ ). It appears that $\neg b$ is provided with a stronger evidence than $b$ (although both are presented explicitly), since there is another way to prove $\neg b$ , if $c$ is true (from $\neg b \vee \neg c$ ). Finally, the evidence of $c$ looks stronger than that of $\neg c$ , since any truth value of $a$ implies $c$ , while to prove $\neg c$ (from $\neg b \vee \neg c$ ), $b$ must be true, which is less evident than $\neg b$ , as suggested above.

Looking for a way of computing evidence, we note that a simple counting of quasi-models would not be in accord with the abovementioned intuition. For instance, $b$ is true in 3 quasi-models out of 4, but we have just concluded that $b$ is more likely to be false than true. On the other hand, $\neg b$ is asserted by the largest mc-subset $\sigma_3$ containing 5 clauses of $S$ out of 6 (while $|\sigma_1| = |\sigma_2| = 4$ , and $|\sigma_4| = 3$ ). This might be decisive with respect to the evidence of $b$ . The next section shows a way to resolving this problem.

## 5. Heuristics of weighted mc-subsets

To make our beliefs most adequate to the reality, the information presented in a given system $S$ should be utilized as fully as possible. So, the larger part of $S$ is contained in a mc-subset $\sigma$ , and the smaller is the subset $S - \sigma$ , the more significant must be the contribution of $\sigma$ to our judgment about the semantics of $S$ . With this in mind, let us assign to every mc-subset $\sigma$ of $S$ a *weight* $w(\sigma)$ as a function of the size of $\sigma$ , $w(\sigma) = f(\,|S - \sigma|\,)$ , such that

$$\textit{for all integers} \quad 0 < k < |S| - 1, \quad f(k) > f(k+1)\ . \tag{5.1}$$

If $S$ is inconsistent, then its largest mc-subset may contain $|S| - 1$ formulas, and there can be at most $|S|$ such mc-subsets. Similarly, $S$ can contain at most $\binom{|S|}{2}$ mc-subsets of size $|S| - 2$ , etc. Any mc-subset of size $|S| - k$ is more significant than any one of size $|S| - k - 1$ , and it appears reasonable to assume that the set of all possible mc-subsets of size $|S| - k$ should outweigh the set of all possible mc-subsets of size $|S| - k - 1$ . So,

$$f(k) \cdot \binom{|S|}{k} \;>\; f(k+1) \cdot \binom{|S|}{k+1} ,$$

$$\frac{f(k+1)}{f(k)} \;<\; \frac{k+1}{|S| - k}\ . \tag{5.2}$$

Due to (5.1), $f(k)$ must be a monotonously decreasing function, so we assume

$$\textit{for all integers} \quad 0 < k < |S| - 1, \quad \frac{f(k+1)}{f(k)} = \frac{k+1}{|S|}\ . \tag{5.3}$$

By assigning the weight of 1 to any consistent system, we get $f(0) = 1$ , $f(1) = \dfrac{1}{|S|}$ , and so, $f(k) = \dfrac{k!}{|S|^k}$ . Hence,

$$\textit{for all} \ \ \sigma\ , \ \ 0 < |\sigma| < |S|\ , \quad w(\sigma) = \frac{|S - \sigma|!}{|S|^{\,|S - \sigma|}}\ . \tag{5.4}$$

This *heuristics of weighted mc-subsets* suggests the following definition of evidence, $E(S, F)$ , provided by $S$ for a formula $F$ .

**Definition 5.1.** *Evidence*

$$E(S, F) = \frac{\sum\limits_{\sigma \,\in\, MC(S)} |MOD(\sigma \cup \{F\})| \cdot w(\sigma)}{\sum\limits_{\sigma \,\in\, MC(S)} |MOD(\sigma)| \cdot w(\sigma)}\ . \quad \blacksquare \tag{5.5}$$

If $S$ is consistent, then expression (5.5) becomes identical to (4.1).

**Theorem 5.1.** Definition 5.1 of evidence is in accord with postulates (1)-(5), Section 4.

**Proof.** (1), If $F$ is a logical consequence of all mc-subsets of $S$, then for all $\sigma \in MC(S)$, $MOD(\sigma \cup \{F\}) = MOD(\sigma)$, and $MOD(\sigma \cup \{\neg F\}) = \varnothing$. Hence by Definition 5.1, $E(S, F) = 1$, $E(S, \neg F) = 0$.

(2) If $\phi$ is a tautology, and $\psi$ is an unsatisfiable formula, then $\phi$ and $\neg\psi$ are true in all mc-subsets of $S$. Hence, by (1), $E(S, \phi) = 1$, $E(S, \psi) = 0$.

(3) For all $\phi, \psi$, and all mc-subsets $\sigma \in MC(S)$, the following holds:
$MOD(\sigma \cup \{\phi \wedge \psi\}) \subseteq MOD(\sigma \cup \{\phi\})$, $MOD(\sigma \cup \{\phi \wedge \psi\}) \subseteq MOD(\sigma \cup \{\psi\})$,
$MOD(\sigma \cup \{\phi\}) \subseteq MOD(\sigma \cup \{\phi \vee \psi\})$, $MOD(\sigma \cup \{\psi\}) \subseteq MOD(\sigma \cup \{\phi \vee \psi\})$,
$MOD(\sigma \cup \{\phi \vee \psi\}) = MOD(\sigma \cup \{\phi\}) \cup MOD(\sigma \cup \{\psi\})$,
$|MOD(\sigma \cup \{\phi \vee \psi\})| = |MOD(\sigma \cup \{\phi\})| + |MOD(\sigma \cup \{\psi\})| - |MOD(\sigma \cup \{\phi \wedge \psi\})|$.

Hence, by Definition 5.1,

$$E(S, \phi \wedge \psi) \leq min(\ E(S, \phi), E(S, \psi)\ ),$$
$$E(S, \phi \vee \psi) \geq max(\ E(S, \phi), E(S, \psi)\ ),$$
$$E(S, \phi \vee \psi) = E(S, \phi) + E(S, \psi) - E(S, \phi \wedge \psi).$$

(4) For all $\phi$, $(\phi \vee \neg\phi)$ is a tautology, and $(\phi \wedge \neg\phi)$ is unsatisfiable. Hence, by (2) and (3),

$$E(S, \phi) + E(S, \neg\phi) = E(S, \phi \vee \neg\phi) + E(S, \phi \wedge \neg\phi) = 1 .$$

(5) If $S = \bigcup_{A \in Base(S)} \{A, \neg A\}$, and $|Base(S)| = n$, then $S$ consists of $2n$ formulas, and has $2^n$ mc-subsets of the same cardinality $n$, and the same weight $w = n! \,/\, (2n)^n$. Every atom $A \in Base(S)$ is true in exactly half of the mc-subsets, but false in the rest of them. Hence, by Definition 5.1, $E(S, A) = E(S, \neg A) = 1/2$. ■

Now consider again Example 4.1. System $S$ consists of 6 clauses, and has 4 mc-subsets of the following cardinality:

$$|\sigma_1| = |\sigma_2| = 4, \quad |\sigma_3| = 5, \quad |\sigma_4| = 3,$$

with the following weights:

$$w(\sigma_1) = w(\sigma_2) = \frac{2!}{6^2} = \frac{1}{18}, \quad w(\sigma_3) = \frac{1!}{6} = \frac{1}{6}, \quad w(\sigma_4) = \frac{3!}{6^3} = \frac{1}{36}.$$

Each mc-subset of $S$ has exactly one model, so

$$\sum_{\sigma \in MC(S)} |MOD(\sigma)| \cdot w(\sigma) = \frac{11}{36}.$$

Now the evidences are:

$$E(S, a) = \frac{36}{11}\left(\frac{2}{18} + \frac{1}{6}\right) = \frac{10}{11}, \quad E(S, \neg a) = \frac{1}{11},$$
$$E(S, b) = \frac{36}{11}\left(\frac{2}{18} + \frac{1}{36}\right) = \frac{5}{11}, \quad E(S, \neg b) = \frac{6}{11},$$
$$E(S, c) = \frac{36}{11}\left(\frac{1}{18} + \frac{1}{6}\right) = \frac{8}{11}, \quad E(S, \neg c) = \frac{3}{11}.$$

Since $E(S, a) > E(S, \neg a)$, $E(S, b) < E(S, \neg b)$, and $E(S, c) > E(S\ \neg c)$, this suggests that most likely in reality, $a$ is true, $b$ is false, and $c$ is true. This conclusion conforms to the intuition expressed in Section 4. So, for this particular example, Definition 5.1 is in accord with postulate 6 of evidence (Section 4). The following sections present more general cases demonstrating plausibility of the heuristics of weighted mc-subsets.

## 6. Justifying examples.

In this section we consider a few examples to show that the semantics of weighted mc-subsets conforms to a common sense intuition.

### 6.1. Chain proofs

**Example 6.1.** Consider an inconsistent set of clauses

$$S = \{a,\ a \rightarrow p,\ b,\ b \rightarrow \neg p,\ c,\ c \rightarrow \neg p\}\ ,$$

and let us try to estimate a plausible truth value of $p$. $S$ gives a single proof of $p$, namely, $\{a,\ a \rightarrow p\}$, while there are two independent proofs that $p$ is false: $\{b,\ b \rightarrow \neg p\}$ and $\{c,\ c \rightarrow \neg p\}$. So, it is reasonable to conclude that $p$ is more likely to be false than true. (Cf. Figure 6.1, where arcs present implications, and any unit clause $\alpha$ is displayed as $true \rightarrow \alpha$). ■

$p \leftarrow a \leftarrow true$

$b \leftarrow true$

$\neg p$

$c \leftarrow true$

**Figure 6.1.**

The mc-subsets of the system of Example 6.1, and their models are as follows:

$\sigma_1 = \{a. \ a \rightarrow p, b \rightarrow \neg p, c \rightarrow \neg p\}$ , $\quad m_1 = \{a, \neg b, \neg c, p\}$ ;
$\sigma_2 = \{a, a \rightarrow p, b \rightarrow \neg p, c\}$ , $\quad m_2 = \{a, \neg b, c, p\}$ ;
$\sigma_3 = \{a, a \rightarrow p, b, c \rightarrow \neg p\}$ , $\quad m_3 = \{a, b, \neg c, p\}$ ;
$\sigma_4 = \{a, a \rightarrow p, b, c\}$ , $\quad m_4 = \{a, b, c, p\}$ ;
$\sigma_5 = \{a, b, b \rightarrow \neg p, c, c \rightarrow \neg p\}$ , $\quad m_5 = \{a, b, c, \neg p\}$ ;
$\sigma_6 = \{a \rightarrow p, b, b \rightarrow \neg p, c, c \rightarrow \neg p\}$ , $\quad m_6 = \{ \neg a, b, c, \neg p\}$ .

The weights of the subsets are:

$$|S| = 6, \quad |\sigma_1| = |\sigma_2| = |\sigma_3| = |\sigma_4| = 4, \quad |\sigma_5| = |\sigma_6| = 5 \ .$$
$$w(\sigma_1) = w(\sigma_2) = w(\sigma_3) = w(\sigma_4) = \frac{1}{18} \ , \quad w(\sigma_5) = w(\sigma_6) = \frac{1}{6} \ .$$

Now let us compute the evidence of $p$ that is true in $m_1, m_2, m_3, m_4$ , but false in $m_5, m_6$ :

$$E(S, p) = \frac{4/18}{4/18 + 2/6} = \frac{2}{5} \ , \quad E(S, \neg p) = \frac{2/6}{4/18 + 2/6} = \frac{3}{5} \ .$$

So, $E(S, p) < E(S, \neg p)$ in accord with the abovementioned intuition.

**Definition 6.1.** *Chain proof.* Call a *chain proof of* $l_0$ *of length* $k$ a set of clauses of the form

$$\{ \ l_0 \leftarrow l_1, l_1 \leftarrow l_2, \ldots, l_{k-2} \leftarrow l_{k-1}, l_{k-1} \ \} \ ,$$

where $l_0, \ldots, l_{k-1}$ are literals such that no atomic formula appears in the set more than once. Two chain proofs are *independent* if they have no atomic formula in common except, possibly, the one appearing in $l_0$ . ■

**Example 6.2.** *Multiple chains.* Let us generalize Example 6.1 by considering a system $S$ that provides a single chain proof of $p$ , $Ch_0$ , and $m$ chain proofs of $\neg p$ , $Ch_1, \ldots, Ch_m$ . All these proofs are pairwise independent, and of the same length $k$ . So, if $k > 1$ , then $p$ is more likely to be false than true (cf. Figure 6.2). ■

$$\begin{array}{ll} p \leftarrow a_1 \leftarrow \cdots \leftarrow a_{k-1} \leftarrow true & \mathbf{Ch_0} \\ \quad\quad b_{1,1} \leftarrow \cdots \leftarrow b_{1,k-1} \leftarrow true & \mathbf{Ch_1} \\ \neg p \ \leftarrow \ldots\ldots\ldots\ldots\ldots\ldots & \\ \quad\quad b_{m,1} \leftarrow \cdots \leftarrow b_{m,k-1} \leftarrow true & \mathbf{Ch_m} \end{array}$$

**Figure 6.2.**

The set $CH = \{Ch_1, \ldots, Ch_m\}$ is consistent. Consider a mc-subset $\sigma$ of $S$ such that $CH \subseteq \sigma$ . Since $CH \models \neg p$ , and $Ch_0 \models p$ , $\sigma$ cannot contain the entire chain $Ch_0$ , but one of the $k$ clauses of $Ch_0$ must be excluded to "cut" the chain, and prevent it from implying $p$ . So, there are $k$ different mc-subsets $\sigma \subset S$ , such that $\sigma \models \neg p$ , $|\sigma| = |S| - 1$ , and

$w(\sigma) = \dfrac{1}{|S|}$ . On the other hand, if a mc-subset $\theta$ of $S$ contains $Ch_0$ , and so $\theta \models p$ , then in $\theta$ every one of the chains $Ch_1, \ldots, Ch_m$ must be cut, such that for all $i = 1, \ldots, m$ , mc-subset $\theta$ contains all the clauses of $Ch_i$ except one. Hence, there are $k^m$ mc-subsets $\theta \subset S$ such that $\theta \models p$ , $|\theta| = |S| - m$ , and $w(\theta) = \dfrac{m!}{|S|^m}$ . Every mc-subset of $S$ has a single model (e.g., in the model of a mc-subset containing $CH$ , for all $1 \le i \le m$ and $1 \le j \le k-1$ , $b_{ij}$ is true, while $p$ and $a_n$ are false for all $1 \le n \le k-1$ except one). So,

$$|S| = k(m+1) \ ;$$

$$\sum_{\theta \,\in\, MC(S)} |MOD(\theta \cup \{p\})| \cdot w(\theta) = k^m \cdot \frac{m!}{|S|^m} = \frac{m!}{(m+1)^m} \ ;$$

$$\sum_{\sigma \,\in\, MC(S)} |MOD(\sigma \cup \{\neg p\})| \cdot w(\sigma) = k \cdot \frac{1}{|S|} = \frac{1}{m+1} \ .$$

For $m > 1$,

$$\frac{m!}{(m+1)^m} < \frac{1}{m+1} .$$

Hence, $E(S,\ p) < E(S, \neg p)$, conforming to the common sense intuition expressed in Example 6.2.

**Example 6.3.** *Chains of different length.* Consider a system $S$ consisting of two independent chain proofs: one of $p$ of length $k_1$, and the other of $\neg p$ of length $k_2$ (Figure 6.3). Because $S$ is inconsistent, it appears reasonable to assume that the longer is a chain, the more likely it may contain wrong data. ■

$$p \leftarrow a_1 \leftarrow \cdots \leftarrow a_{k_1 - 1} \leftarrow true$$
$$\neg p \leftarrow b_1 \leftarrow \ldots\ldots\ldots \leftarrow b_{k_2 - 1} \leftarrow true$$

**Figure 6.3.**

There are $k_2$ mc-subsets of $S$ containing the entire proof of $p$ together with the proof of $\neg p$ except one clause, and $k_1$ mc-subsets containing the entire proof of $\neg p$ together with the proof of $p$ without one clause. Each mc-subset has exactly one model, the same size of $k_1 + k_2 - 1$, and the same weight, $w = \frac{1}{k_1 + k_2}$. Hence,

$$E(S,\ p) = \frac{k_2}{k_1 + k_2}\ , \quad E(S, \neg p) = \frac{k_1}{k_1 + k_2}\ .$$

So, if $k_1 < k_2$, then $p$ is more likely to be true than false, and vice versa. This conclusion corresponds to the assumption that in a "polluted" system a shorter chain is more reliable than a longer one.

### 6.2. Rules with exceptions

Formulas constituting a Knowledge system $S$ represent fragments of our knowledge about the world. For instance, the fact that all birds hatch from eggs can be represented by a formula $bird(x) \to hatch(x)$ , which is true for all $x$ , while another formula, $bird(x) \to fly(x)$ , may reflect the knowledge that most birds fly, but is not true for non-flying birds, such as penguins, ostriches, etc. So, $bird(x) \to fly(x)$ is a *defeasible* formula since, being true in general and by default (in the absence of contradicting knowledge), it can be falsified by a specific experience with a non-flying bird. If $S$ contains a defeasible formula $\phi$ , and free variables of $\phi$ are universally quantified, this may lead to a contradiction in $S$ . Therefore sets of defeasible formulas are of certain interest to this study of inconsistency.

A popular proverb says: "There is an exception to every rule". So, a rule is a defeasible statement, and, being accompanied by an exception, may become a source of contradiction. Let a rule claim that objects of a set denoted by *RuleClass* possess a certain *RuleProperty* , while it is known that some of these objects (constituting an *ExceptionalClass* ) do not. Then a rule with exceptions can be presented by a set of clauses in the following way.

(i) Rule: $RuleClass(x) \to RuleProperty(x)$ ;
(ii) Inclusion: $ExceptionalClass(x) \to RuleClass(x)$ ;
(iii) Exception: $ExceptionalClass(x) \to \neg RuleProperty(x)$ .

Let us consider the following example of exception that has recently become a classical one.

**Example 6.4.** $S = \{$
(i) $bird(x) \to fly(x)$ ;
(ii) $penguin(x) \to bird(x)$ ;
(iii) $penguin(x) \to \neg fly(x)$ ;
(iv) $penguin(Quacky)$ $\}$ . ■

Since the free variables of the clauses of Example 6.4 are supposed to be universally quantified, $S$ is inconsistent. Indeed, by clauses (iv), (ii), (i), Quacky can fly, but due to (iv), (iii), she cannot. However, is it possible to learn more about this matter from the system? Let us try to figure it out by applying the semantics of weighted mc-subsets.

$S$ contains 3 mc-subsets:

$\sigma_1 = \{(i), (ii), (iv)\}$, $|\sigma_1| = 3$, $m_1 = \{penguin(Quacky), bird(Quacky), fly(Quacky)\}$ ;
$\sigma_2 = \{(ii), (iii), (iv)\}$, $|\sigma_2| = 3$, $m_2 = \{penguin(Quacky), bird(Quacky), \neg fly(Quacky)\}$ ;
$\sigma_3 = \{(i), (iii), (iv)\}$, $|\sigma_3| = 3$, $m_3 = \{penguin(Quacky), \neg bird(Quacky), \neg fly(Quacky)\}$ .

So, the system of Example 6.4 provides the following evidence

$$E(S,\ fly(Quacky)) = \frac{1}{3}\ , \quad E(S, \neg fly(Quacky)) = \frac{2}{3}\ .$$

Hence, it is most likely that Quacky cannot fly, which resolves the contradiction in $S$ in full accord with the common knowledge about the flying ability of penguins.

The system of Example 6.4 may contain data regarding specific birds other than Quacky, and exceptional classes besides penguins. Since the facts about different birds (and different classes, as well) are pairwise independent, this would increase the number of mc-subsets of $S$

and their models implying $fly(Quacky)$ by the same multiplicative factor as the number of those implying $\neg fly(Quacky)$. So, the evidences would remain the same, suggesting the same belief that poor Quacky cannot fly.

Rules with exceptions have been studied from several points of view. In particular, Kowalski and Sadri [17] have extended Logic programming to include explicit representations of exceptions along with general rules. They have defined a program transformation that accommodates the exceptions and avoids contradiction. The semantics of mc-subsets agrees with the approach presented in [17] on the principle that exceptions prevail over rules (Example 6.4 is considered in [17] leading to the same conclusion). However, the two approaches implement this principle in different ways. In [17] clauses presenting exceptions are designated explicitly, while the approach of mc-subsets does not assume an additional knowledge distinguishing between exceptions and rules in a given logic system. Presence of exceptions manifests itself by making the system inconsistent, and the preference for exceptions stems from the fact that they represent more specific knowledge than rules. As a rather simplistic example showing this difference, consider a system $S = \{p, \neg p\}$ presented in [17], where $p$ is viewed as a rule, and $\neg p$ — as an exception, deriving a conclusion that $p$ is false. For the same system, the semantics of mc-subsets suggests that $p$ is equally likely to be true or false.

### 6.3. Paradoxes

Well known logical systems with an inherent inconsistency are paradoxes. A *paradox* usually contains an apparently true premise $P$ and a statement $Q$ that can be neither true nor false under the premise in the sense that a contradiction is derivable from both $P \wedge Q$ and $P \wedge \neg Q$ .

A popular paradox is the story of *The Island of Knights and Knaves* [30], where the following three *rules of the island* hold:

(i) knights make only true statements,
(ii) knaves make only false statements,
(iii) every native islander is either a knight or a knave.

Suppose that a logician who knows these rules visits the island and meets an inhabitant who makes the following statement:

"You will never know that I am a knight".

Is he a knight or a knave?

The logician thinks: "If the inhabitant is a knave, then indeed I shall never know that he is a knight. So, his statement is true, which contradicts rule (ii). Hence, he cannot be a knave. Then he must be a knight, so, I already know that he is a knight, and therefore his statement is false — again a contradiction to rule (i). Hence, he cannot be a knight either". Thus, if the inhabitant is a native islander, then the story is a paradox indeed, since the inhabitant can be neither a knight nor a knave, and his statement can be neither true nor false. However, if he is not a native, then, being neither a knight nor a knave, he makes a true statement.

Let us now apply the semantics of weighted mc-subsets to the paradox of knights and knaves. The rules of the island can be presented in the following form, where $x$ ranges over persons, and $y$ — over their statements (all free variables are universally quantified).

(i) $Knight(x) \wedge Statement(x, y) \rightarrow True(y)$ ,
(ii) $Knave(x) \wedge Statement(x, y) \rightarrow \neg True(y)$ ,
(iii) $Native(x) \rightarrow ( Knight(x) \wedge \neg Knave(x) \vee \neg Knight(x) \wedge Knave(x) )$ .

The following are the thoughts of the logician ( *'inh'* denotes the inhabitant, and *'st'* — his statement).

(iv) $Statement(inh, st)$ ,
(v) $Native(inh)$ ,
(vi) $Knave(inh) \rightarrow True(st)$ ,
(vii) $Knight(inh) \rightarrow \neg True(st)$ .

Let us denote by $KK$ the set of clauses (i)-(vii). It has 6 mc-subsets shown in Table 6.1. All the mc-subsets of $KK$ have the same cardinality 6 (out of $|S| = 7$ ), and hence the same weight 1/7 . So, by Definition 5.1,

$$E( KK, \neg Knight(inh) ) = \frac{3}{4} , \qquad E( KK, \neg Knave(inh) ) = \frac{3}{4} .$$

Thus, by the semantics of weighted mc-subsets, it is most likely that the inhabitant met by the logician is neither a knight nor a knave, and so is not a native islander. This is indeed a plausible conclusion that helps out of the paradoxical situation.

**Table 6.1.**

| mc-subsets | $\|MOD(\sigma)\|$ | $\|MOD(\sigma \cup \{ \neg Knight(inh)\})\|$ | $\|MOD(\sigma \cup \{ \neg Knave(inh)\})\|$ |
|---|---|---|---|
| $\sigma_1 = S - \{i\}$ | 1 | 0 | 1 |
| $\sigma_2 = S - \{ii\}$ | 1 | 1 | 0 |
| $\sigma_3 = S - \{iii\}$ | 2 | 2 | 2 |
| $\sigma_4 = S - \{v\}$ | 2 | 2 | 2 |
| $\sigma_5 = S - \{vi\}$ | 1 | 0 | 1 |
| $\sigma_6 = S - \{vii\}$ | 1 | 1 | 0 |
| Total | 8 | 6 | 6 |

## 7. Resilience w.r.t. contradiction

As it has been shown in the previous section, in an inconsistent system a shorter chain proof provides a stronger evidence than that of a longer one. However, it is reasonable that under certain conditions a short chain proof can be overruled by a number of longer proofs.

**Example 7.1.** Consider a system $S$ consisting of a single shortest possible proof of $p$, namely, $\{true \rightarrow p\}$, and $m$ longer proofs of $\neg p$. To simplify the analysis, let all the proofs of $\neg p$ have the same length $k$, so $|S| = km + 1$. $S$ contains a single mc-subset that proves $\neg p$. It consists of all the proofs of $\neg p$, but does not contain $\{true \rightarrow p\}$. On the other hand, there are $k^m$ mc-subsets of $S$ proving $p$, such that each one contains $\{true \rightarrow p\}$, and also every proof of $\neg p$ cut by excluding exactly one of its clauses. The mc-subset proving $\neg p$ contains $|S| - 1$ clauses, while every mc-subset proving $p$ consists of $|S| - m$ ones. Every mc-subset of $S$ has exactly one model, so

$$E(S,\ p) = \frac{k^m m! \,/\, (km+1)^m}{1 \,/\, (km+1) + k^m m! \,/\, (km+1)^m}\ , \quad E(S, \neg p) = \frac{1 \,/\, (km+1)}{1 \,/\, (km+1) + k^m m! \,/\, (km+1)^m}\ .$$

The set of long proofs of $\neg p$ overrules the short proof of $p$ if

$$E(S,\ p) < E(S, \neg p)\ ,$$

that is

$$\frac{k^m m!}{(km+1)^m} < \frac{1}{km+1}\ . \tag{7.1}$$

Table 7.1 gives for different $k$ the smallest values of $m$ satisfying condition (7.1). It shows that if proofs of a formula $\phi$ are numerous enough they can overrule a smaller set of much shorter proofs of $\neg\phi$.

**Table 7.1.**

| $k$ | 2 | 3 | 5 | 10 | 20 | 50 | 100 |
|---|---|---|---|---|---|---|---|
| $m$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

This observation suggests that a notion of *resilience* can be associated with any semantics of logic systems. Consider a consistent system $S$, and a formula $\phi$, such that $S \models \phi$.

Suppose that in the course of the system evolution a formula $\psi$ has been added to $S$ such that $S \cup \{\psi\} \models \neg\phi$. Then $S' = S \cup \{\psi\}$ becomes inconsistent. Now we apply to $S'$ a semantics, $SEM$, that determines a certain belief regarding the truth value of $\phi$ in $S'$. Is it the same truth value as it has previously been in $S$? If the belief remains the same, that is, due to $SEM$, $\phi$ is more likely to be true than false in $S'$, then we say that under $SEM$ the belief of $\phi$ exhibits *resilience* with regard to $\psi$.

For a formula $\phi$ true in $S$, let $R_\phi$ denote the set of all formulas $\psi$ such that $S \cup \{\psi\} \models \neg\phi$, but under the given semantics addition of $\psi$ to $S$ does not change the belief of truth value of $\phi$ in $S$. The larger is $R_\phi$ for formulas $\phi$ of $S$, the higher is the resilience of $S$ w.r.t. contradiction under the given semantics. Due to classical logic semantics, $R_\phi$ is empty for all consistent $S$ and $\phi$, since any conclusion can be derived from an inconsistent system. $R_\phi$ is almost empty also under many recent approaches to reasoning with inconsistency [2, 16, 34] that face difficulties in recovering values of ambiguous formulas. On the contrary, under the semantics of weighted mc-subsets $R_\phi$ can be considerably large if $S$ provides a strong evidence of truth of $\phi$, e.g., by numerous proofs of $\phi$ (cf. the examples of Sections 6, 7).

Resilience w.r.t. contradiction is a feature of high importance to Knowledge Systems, since it allows preserving of the meaning of a system under quite trying real circumstances where wrong data penetrate into the system by mistake or even by sabotage. Regarding the latter case, it turns out that under the semantics of mc-subsets it is not at all easy to damage deliberately a set of beliefs that are justified strongly enough, since this semantics reveals and evaluates the degree of the justification. As shown in Example 7.1, even an explicit addition of $\neg\phi$ does not matter if the system provides a strong evidence of truth of $\phi$.

## 8. Special cases

In this section we consider two special cases of inconsistent systems: homogeneous sets of formulas, and sets of ground clauses.

### 8.1. Homogeneous sets

**Definition 8.1.** A set $S$ of formulas is *homogeneous* if all its mc-subsets have the same cardinality. ■

For instance, the set $KK$, describing in Section 6.3 the paradox of Knights and Knaves, is homogeneous (cf. Table 6.1), since all its mc-subsets have the same cardinality 6. (By the way, any consistent set of formulas is trivially homogeneous). Since all mc-subsets of an homogeneous set have the same weight, computing of evidence in a homogeneous set $S$ can be simplified as follows (cf. Definition 5.1):

$$E(S, F) = \frac{\sum_{\sigma \in MC(S)} |MOD(\sigma \cup \{F\})|}{\sum_{\sigma \in MC(S)} |MOD(\sigma)|} . \tag{8.1}$$

Any two different mc-subsets $\sigma_i, \sigma_j$ of any system $S$ have no models in common, since otherwise $\sigma_i \cup \sigma_j$ would be consistent contradicting the assumption that $\sigma_i$ and $\sigma_j$ are maximally consistent subsets. This implies the following lemma.

**Lemma 8.1.** Sets of models of all mc-subsets of any system $S$ are pairwise disjoined,

$$MOD(\sigma_i) \cap MOD(\sigma_j) = \varnothing \quad \text{for all} \quad \sigma_i, \sigma_j \in MC(S) ,$$

and hence, in an homogeneous system

$$E(S, F) = \frac{\left| \bigcup_{\sigma \in MC(S)} MOD(\sigma \cup \{F\}) \right|}{\left| \bigcup_{\sigma \in MC(S)} MOD(\sigma) \right|} . \quad \blacksquare$$

Now, given an homogeneous system $S$, consider a disjunction $S_\vee$ of all its mc-subsets, where every subset is a conjunction of all its formulas:

$$S_\vee = \bigvee_{\sigma \in MC(S)} \sigma , \quad \sigma = \bigwedge_{\phi \in \sigma} \phi .$$

$S_\vee$ is consistent, and $MOD(S_\vee) = \bigcup_{\sigma \in MC(S)} MOD(\sigma)$, so (cf. Lemma 8.1),

$$E(S_\vee, F) = \frac{|MOD(S_\vee \cup \{F\})|}{|MOD(S_\vee)|} = \frac{\left| \bigcup_{\sigma \in MC(S)} MOD(\sigma \cup \{F\}) \right|}{\left| \bigcup_{\sigma \in MC(S)} MOD(\sigma) \right|} = E(S, F) .$$

Thus, for any formula $F$, $S_\vee$ provides the same evidence as $S$, and so, the treatment of an inconsistent homogeneous system, $S$, can be reduced to that of a consistent one, $S_\vee$. By this the complexity of computing evidence can be reduced significantly, since in consistent systems evidence can be estimated in many cases without actual counting of models (as has been shown in [19]).

### 8.2. Sets of ground clauses

One of the generally used standard forms of First Order logic systems (Knowledge systems, in particular) is a skolemized CNF which is a set (a logical conjunction) of clauses, where each clause is a set (a logical disjunction) of literals, and all free variables are universally quantified. With no loss of generality, such a system can be viewed as a set of all ground instances of all its clauses. The following propositions describe important properties of sets of ground clauses, useful for classifying literals (cf. Definition 2.2), and computing their evidence.

**Theorem 8.2.** (a) A consistent subset $\sigma \subseteq S$ is an mc-subset of $S$ iff $\sigma \models \neg L$ for all literals $L$ occurring in $S - \sigma$ ; (b) If a literal $L$ is a non-ambiguous formula in $S$, then every mc-subset of $S$ includes all clauses of $S$ containing $L$ or $\neg L$.

**Proof.** (a) $\rightarrow$ Consider a mc-subset $\sigma$ of $S$, and let $L$ be a literal occurring in a clause $C$ of $S - \sigma$, but $\sigma \not\models \neg L$. So, there must be a model $\mu$ of $\sigma$ asserting $L$, and hence satisfying $C$. Therefore $\sigma \cup \{C\}$ is consistent contradicting the assumption that $\sigma$ is a maximally consistent subset. So, $\sigma \models \neg L$ for all $L$ of $S - \sigma$.

(a) $\leftarrow$ Let $\sigma$ be a consistent subset of $S$ such that $\sigma \models \neg L$ for all literals $L$ occurring in $S - \sigma$. Then every model of $\sigma$ falsifies every clause of $S - \sigma$, so no one of the latter can be added to $\sigma$ without violating its consistency. Hence, $\sigma$ is a maximally consistent subset of $S$.

(b) Consider an mc-subset $\sigma$ of $S$, and a non-ambiguous literal $L$. Suppose that $L$ occurs in a clause $C$ of $S - \sigma$. If $\sigma \not\models \neg L$, then there is a model of $\sigma$ asserting $L$, so,

$\sigma \cup \{C\}$ is consistent contradicting the maximality of $\sigma$. If $\sigma \models \neg L$, then we arrive at a contradiction in the following way. Let $\theta$ denote a subset of $\sigma$ consisting of all clauses of $\sigma$ containing $\neg L$. Then $\sigma - \theta$ does not prove $\neg L$, so, $(\sigma - \theta) \cup \{C\}$ is consistent. Since $L$ is non-ambiguous, and $\sigma \models \neg L$, there is no consistent subset of $S$ proving $L$. Hence, $(\sigma - \theta) \cup \{C\}$ does not prove $L$, and so $(\sigma - \theta) \cup \{C\} \cup \theta$ is consistent. But $(\sigma - \theta) \cup \{C\} \cup \theta = \sigma \cup \{C\}$, again contradicting the maximality of $\sigma$. Thus, $\sigma$ must include all clauses of $S$ containing $L$ or $\neg L$ as well, since $\neg L$ is also a non-ambiguous literal. ■

The converse of Theorem 8.2(b) does not hold, as there can be an ambiguous literal $L$ such that every mc-subset of $S$ includes all clauses of $S$ containing $L$ or $\neg L$, as in the following example.

**Example 8.2.** $S = \{ p, \neg p, p \vee \neg q, \neg p \vee q \}$.
There are two mc-subsets,

$$\sigma_1 = \{ p, p \vee \neg q, \neg p \vee q \}, \quad \sigma_2 = \{ \neg p, p \vee \neg q, \neg p \vee q \},$$

and both include all occurrences of $q$ which is an ambiguous literal, since $\sigma_1 \models q$, $\sigma_2 \models \neg q$. ■

More properties of sets of ground clauses are implied by Theorem 8.2, as follows.

**Corollary 8.2.** If $\sigma$ is a mc-subset of $S$, then
(a) $S - \sigma$ is consistent,
(b) All literals occurring in $S - \sigma$ are ambiguous in $S$,
(c) If a literal $L$ occurs in $S - \sigma$, then $\neg L$ does not (i.e., all literals are pure in $S - \sigma$). ■

Example 4.1 may serve as an illustration of Theorem 8.2 and Corollary 8.2.

## 8.3. Infinite sets of ground clauses.

If $S$ is a logic system with an infinite domain $D$ of its variables, then $Base(S)$ becomes infinite, and a mc-subset of $S$ may have infinitely many models. Let $d$ be a finite subset of $D$, and $E(S_d, F)$ denote the evidence of $F$ in a *version* $S_d$ of $S$ such that the domain of its variables is restricted to $d$. Then the evidence of $F$ in $S$ can be computed as a limit

$$E(S, F) = \lim_{d \to D} E(S_d, F), \tag{8.2}$$

if this limit exists.

The following example shows an application of this approach to a system presented as an infinite set of ground clauses.

**Example 8.3.** Consider a system

$$S = \{P(x) \vee Q(x), \ \neg P(x) \vee R(x), \ \neg P(0), \ \neg Q(0)\},$$

where $x$ is universally quantified and ranges over all natural numbers.

First, let us assess intuitively the evidence provided by $S$ for different ground atoms of the system. For all $k > 0$, $Q(k)$ and $R(k)$ appear more likely to be true than false, since both $Q(k)$ and $R(k)$ occur in $S$ purely unnegated, so, $\neg Q(k)$ or $\neg R(k)$ cannot be proved. On the other hand, there are two rather similar ways to prove either $P(k)$ or $\neg P(k)$, so, $P(k)$ and $\neg P(k)$ seem equally likely to be true. Regarding $x = 0$ the situation is different, since $\neg P(0)$

tilts the balance in favour of $P(0)$ being false, and $Q(0)$ also appears more likely to be false than true, since $\neg Q(0)$ is stated explicitly, while proof of $Q(0)$ depends on a certain value of $P(0)$. Still $R(0)$ is purely unnegated in $S$, and so, is more likely to be true than false.

Now, let us compute evidence due to expression (8.2). Let $d(n)$ be a finite set of $n$ natural numbers including 0 (as 0 occurs explicitly in $S$) such that $d(n) = \{0, i_1, i_2, \ldots, i_{n-1}\}$. Then $S_{d(n)}$ has the following mc-subsets:

$$\sigma_1 = \{P(x) \vee Q(x) \textit{ for all } x \in d(n), \quad \neg P(x) \vee R(x) \textit{ for all } x \in d(n), \quad \neg P(0)\};$$

$$\sigma_2 = \{P(x) \vee Q(x) \textit{ for all } x \in d(n), \quad \neg P(x) \vee R(x) \textit{ for all } x \in d(n), \quad \neg Q(0)\};$$

$$\sigma_3 = \{P(x) \vee Q(x) \textit{ for all } x \in d(n) - \{0\}, \quad \neg P(x) \vee R(x) \textit{ for all } x \in d(n), \quad \neg P(0), \neg Q(0)\}.$$

To compute the number of models of $\sigma_1$, note that for $x = 0$, two truth assignments of ground atoms satisfy $\sigma_1$, namely, $\{\neg P(0), Q(0), R(0)\}$, and $\{\neg P(0), Q(0), \neg R(0)\}$. For all $k > 0$, four truth assignments satisfy $\sigma_1$:
$\{P(k), Q(k), R(k)\}, \{P(k), \neg Q(k), R(k)\}, \{\neg P(k), Q(k), R(k)\}, \{\neg P(k), Q(k), \neg R(k)\}.$

Hence, $|MOD(\sigma_1)| = 2 \cdot 4^{n-1}$.

By a similar argument,

$$|MOD(\sigma_2)| = 4^{n-1} \ ; \quad |MOD(\sigma_3)| = 2 \cdot 4^{n-1} \ .$$

Table 8.1 presents number of models of sets $\sigma_j \cup \{A\}$ for all ground atoms $A$ of $S$, and $j = 1, 2, 3$.

**Table 8.1.** Values of $|MOD(\sigma_j \cup \{A\})|$.

| A | $x = 0$ | | | $x \in (\ d(n) - \{0\}\ )$ | | |
|---|---|---|---|---|---|---|
| | $P(0)$ | $Q(0)$ | $R(0)$ | $P(x)$ | $Q(x)$ | $R(x)$ |
| $\sigma_1$ | 0 | $2 \cdot 4^{n-1}$ | $4^{n-1}$ | $4^{n-1}$ | $6 \cdot 4^{n-2}$ | $6 \cdot 4^{n-2}$ |
| $\sigma_2$ | $4^{n-1}$ | 0 | $4^{n-1}$ | $2 \cdot 4^{n-2}$ | $3 \cdot 4^{n-2}$ | $3 \cdot 4^{n-2}$ |
| $\sigma_3$ | 0 | 0 | $4^{n-1}$ | $4^{n-1}$ | $6 \cdot 4^{n-2}$ | $6 \cdot 4^{n-2}$ |

$S_{d(n)}$ is a homogeneous system (cf. Section 8.1) such that all its mc-subsets have the same weight. So, by expression (8.1),

$$E(S_{d(n)}\, , A) = \frac{\sum_{j=1}^{3} |MOD(\sigma_j \cup \{A\})|}{\sum_{j=1}^{3} |MOD(\sigma_j)|} \ .$$

Then by (8.2),

$$E(S, A) = \lim_{n \to \infty} E(S_{d(n)}, A) .$$

Table 8.2 gives for all ground atoms $A$ of $S$ the values of $E(S, A)$ that are in full agreement with the abovementioned intuition.

**Table 8.2.** Evidence of ground atoms.

| A | $x = 0$ | | | $x \in ( d(n) - \{0\} )$ | | |
|---|---|---|---|---|---|---|
| | $P(0)$ | $Q(0)$ | $R(0)$ | $P(x)$ | $Q(x)$ | $R(x)$ |
| $E(S, A)$ | 1/5 | 2/5 | 3/5 | 1/2 | 3/4 | 3/4 |

In the following section we consider another special case of formulas.

## 9. Doubtless formulas

A basic assumption underlying the semantics of mc-subsets is that any consistent subset of a system $S$ describing a world $W$ may conform to a possible state of the latter. If a formula $\phi \in S$ belongs not to all mc-subsets of $S$, this means that due to $S$ there is some doubt about the truth of $\phi$ in an arbitrary state of $W$. However, there are formulas that are known to be true in every state of the world just by the virtue of their content, and regardless of their representation in a knowledge system. These formulas may express an absolute truth of the world, such as "All living creatures are mortal", or describe some local (in space or time) integrity constrains, e.g., "Every freshman takes a course in programming". Let us call such formulas *doubtless*, and denote by $DLF(S)$ the set of all doubtless formulas of $S$. If $S$ contains some wrong data, there may exist a syntactically consistent subset $\theta$ of $S$ falsifying a doubtless formula $\phi \in DLF(S)$. This means that $\theta$ cannot represent any possible state of the world, and so, should not be considered for determining a semantics of $S$, in particular, for computing evidence. Hence, in the presence of doubtless formulas the definition of mc-subset (cf. Definition 2.1) must be modified as follows.

**Definition 9.1.** *Mc-subset in the presence of doubtless formulas.* A subset $\sigma$ of $S$ is a *maximally consistent subset* if $\sigma$ is consistent and contains all doubtless formulas of $S$, $DLF(S) \subseteq \sigma$, but for all formulas $\phi \in (S - \sigma)$, $\sigma \cup \{\phi\}$ is inconsistent. ■

**Example 9.1.** Consider again the system of Example 4.1,

$$S = \{ a, b, \neg b, a \vee c, \neg a \vee c, \neg b \vee \neg c \} ,$$

and suppose that $DLF(S) = \{ b \}$. This leaves only 3 mc-subsets out of 4 mentioned in Example 4.1:

$\sigma_1 = \{ a, b, a \vee c, \neg a \vee c \}$, $\quad m_1 = \{ a, b, c \}$;
$\sigma_2 = \{ a, b, a \vee c, \neg b \vee \neg c \}$, $\quad m_2 = \{ a, b, \neg c \}$;
$\sigma_4 = \{ b, \neg a \vee c, \neg b \vee \neg c \}$, $\quad m_4 = \{ \neg a, b, \neg c \}$. ■

Let us now compute evidences of the atomic propositions of $S$:

$$|\sigma_1| = |\sigma_2| = 4, \quad |\sigma_4| = 3,$$
$$w(\sigma_1) = w(\sigma_2) = \frac{2!}{6^2} = \frac{1}{18}, \quad w(\sigma_4) = \frac{3!}{6^3} = \frac{1}{36} .$$

Each mc-subset of $S$ has exactly one model, hence

$$\sum_{\sigma \in MC(S)} |MOD(\sigma)| \cdot w(\sigma) = \frac{5}{36} .$$

So, the evidences are:

$$E(S, a) = \frac{36}{5} \cdot \frac{2}{18} = \frac{4}{5} , \quad E(S, \neg a) = \frac{1}{5} ,$$
$$E(S, b) = \frac{36}{5} \cdot \frac{5}{36} = 1 , \quad E(S, \neg b) = 0 ,$$
$$E(S, c) = \frac{36}{5} \cdot \frac{1}{18} = \frac{2}{5} , \quad E(S, \neg c) = \frac{3}{5} .$$

A comparison of this result to that presented at the end of Section 5 shows that introduction of a doubtless formula into the system of Example 4.1 has changed the evidences, and even reversed the belief of $c$ , since now $E(S, c) < E(S, \neg c)$ , so, $c$ is likely to be false.

## 10. A range of semantics

As the evidence of a formula $F$ given by a system $S$ presents a measure of credibility of a belief in truth of $F$, reasoning by evidence provides bases for a wide spectrum of semantics. At one end there is a *basic qualitative semantics of likelihood*, $\Sigma L_1$, as follows.

$$\Sigma L_1 : \textit{Believe that } F \textit{ is } \begin{cases} \textit{true, if } E(S, F) = 1 , \\ \textit{likely to be true, if } 0.5 < E(S, F) < 1 , \\ \textit{equally likely to be true or false, if } E(S, F) = 0.5 , \\ \textit{likely to be false, if } 0 < E(S, F) < 0.5 , \\ \textit{false, if } E(S, F) = 0 . \end{cases}$$

The other extreme of the spectrum is presented by a purely *quantitative semantics*, $\Sigma L_\infty$, under which every formula $F$ is assigned its evidence $E(S, F)$.

$$\Sigma L_\infty : \textit{Believe that } F \textit{ is true with a credibility of } E(S, F) .$$

Inside the range there is room for a family of semantics, $\Sigma L_k$, with different granularity of likelihood ( $k$ is the number of intermediate levels of likelihood), that can be determined by a user, and adapted to specific formulas (becoming $\Sigma L_k(user, formulas)$ ). For instance,

$$\Sigma L_3(John, all) : \textit{Believe that } F \textit{ is } \begin{cases} \textit{true, if } E(S, F) = 1 , \\ \textit{most likely to be true, if } 0.9 \le E(S, F) < 1 , \\ \textit{likely to be true, if } 0.6 \le E(S, F) < 0.9 , \\ \textit{slightly more likely to be true than false, if } 0.5 < E(S, F) < 0.6 , \\ \textit{equally likely to be true or false, if } E(S, F) = 0.5 , \\ \dots\dots\dots \\ \textit{false, if } E(S, F) = 0 . \end{cases}$$

To illustrate an application of this approach, let us consider again Example 4.1 analyzed at the end of section 5. A query regarding the truth value of $a$ receives different (but clearly not contradicting) answers depending on the adopted semantics. Namely,

$\Sigma L_1$ : *a is more likely to be true than false*;

$\Sigma L_3(John, all)$ : *a is most likely to be true*;

$$\Sigma L_\infty : E(S, a) = \frac{10}{11} .$$

A query with free variables, $q = F(x)$, is considered in reasoning by evidence as a request for all values of $x$ such that $F(x)$ is believed to be true, accompanied by the corresponding degree of likelihood (a query may specify a required degree of likelihood). This is, actually, a generalization of the traditional interpretation of $q = F(x)$ as a request for all $x$ such that $S \models F(x)$. For instance, consider the system of Example 6.4 augmented with a clause $bird(Flity)$ such that $E(S, fly(Flity)) = 1$. Given queries $q_1 = \neg fly(x)$ and $q_2 = fly(x)$, the answer to the first query under the $\Sigma L_1$ semantics is *Quacky is likely not to fly*, while $q_2$

gets *Flity can fly*. Note that the latter belief is in accord with Reiter's Default Logic [27] provided that rules are considered as defaults.

The following procedure computes the evidence of a formula $F$, and its likelihood in a given system $S$.

**Procedure Comp_Evidence** (Given $S, F$, returns $E(S, F)$ and the likelihood of $F$);

1. Compute a set $MC(S)$ of all mc-subsets of $S$;
2. For all mc-subsets $\sigma \in MC(S)$ do

   begin
   2.1. Compute $|MOD(\sigma)|$;
   2.2. Compute $|MOD(\sigma \cup \{F\})|$;
   2.3. Compute $w(\sigma)$ by expression (5.4)
   end;
3. Compute $E(S, F)$ by expression (5.5);
4. Determine the likelihood of $F$ due to the adopted semantics. ■

Determining mc-subsets (step 1) and their models (steps 2.1, 2.2) is a hard computational task. Indeed, the satisfiability problem even for propositional formulas is NP-complete [7], the problem of computing the number of models is known to be #P-complete in general [33], and the number of mc-subsets of a highly polluted system $S$ can be exponential in the size of $S$ (cf. Theorem 3.1(ii) ). However, recent research of the average complexity of these problems [11-13, 15, 18, 20, 22, 26] indicates that the semantics of mc-subsets may turn out to be quite tractable on the average.

Procedure Comp_Evidence directly implements the model-theoretic definition of evidence (Sections 4, 5). However, an efficient proof procedure would be most helpful to reasoning by evidence. Presently a few observations provide some hints on constructing of such a procedure.

**Observation 10.1.** If $S$ provides $k$ chain proofs of $F$ (cf. Section 6.1) and $m$ chain proofs of $\neg F$ such that all these proofs are pairwise independent and have the same length, and $k > m$, then $E(S, F) > E(S, \neg F)$.

**Observation 10.2.** Let $S$ provide a set of chain proofs of $F$, $PR(F)$, and a set of chain proofs of $\neg F$, $PR(\neg F)$, such that $|PR(F)| > |PR(\neg F)|$, and the proofs are pairwise independent, but may have different length (cf. Example 6.3). If there exists a bijection $f$ between $PR(\neg F)$ and a subset $PR'(F)$ of $PR(F)$ such that no proof $pr(\neg F) \in PR(\neg F)$ is shorter than its counterpart $f(pr(\neg F)) \in PR'(F)$, but there is a proof $pr'(F) \in PR'(F)$ that is shorter than $f(pr'(F)) \in PR(\neg F)$, then $E(S, F) > E(S, \neg F)$.

**Observation 10.3.** Let $\vdash$ denote a provability relation. We say that a formula $\theta$ consistent with $S$ is a *supplement of S w.r.t. F*, if $S \not\vdash F$, but $S \cup \{\theta\} \vdash F$. A formula $\phi$ *cancels* a model $m$ of $S$ if $m$ is not a model of $S \cup \{\phi\}$. Any supplement $\theta$ of $S$ w.r.t. $F$ cancels all models of $S$ in which $F$ is false. Then the following Theorem holds for ground literals.

**Theorem 10.1.** Consider a consistent subsystem $\sigma$, a ground atom $A$, and a ground literal $L \in \{A, \neg A\}$, such that $\sigma \not\vdash L$ and $\sigma \not\vdash \neg L$. If there exists a formula $\theta$ not involving

$A$ that is a supplement of $\sigma$ w.r.t. $L$ , but there exists no formula $\theta'$ not involving $A$ that is a supplement of $\sigma$ w.r.t. $\neg L$ , then $E(\sigma, L) > E(\sigma, \neg L)$ .

**Proof.** First, we show that if there exists no formula $\theta'$ not involving $A$ that is a supplement of $\sigma$ w.r.t. $\neg L$ , then $|MOD(\sigma \cup \{\neg L\})| \leq |MOD(\sigma \cup \{L\})|$ . Denote by $res(m, L)$ a *residue* of a model $m$ produced by deleting $L$ from $m$ . Consider a formula $\psi = \neg \bigvee_{m \in MOD(\sigma \cup \{L\})} res(m, L)$ . Clearly, $\psi$ cancels all models of $MOD(\sigma \cup \{L\})$ . But $\psi$ must cancel also all models of $MOD(\sigma \cup \{\neg L\})$ , since otherwise $\psi$ (that does not involve $A$ ) would be a supplement of $\sigma$ w.r.t. $\neg L$ , contradicting the premise. The latter implies two conclusions:

$(i)$ $\{ res(m, \neg L) \mid m \in MOD(\sigma \cup \{\neg L\})\} \subseteq \{ res(m, L) \mid m \in MOD(\sigma \cup \{L\})\}$ ,

$(ii)$ $|MOD(\sigma \cup \{\neg L\})| \leq |MOD(\sigma \cup \{L\})|$ .

Now we can prove that $|MOD(\sigma \cup \{\neg L\})| < |MOD(\sigma \cup \{L\})|$ , and hence, by expression (4.3), $E(\sigma , L) > E(\sigma , \neg L)$ . Indeed, if $\{ res(m, \neg L) \mid m \in MOD(\sigma \cup \{\neg L\})\} = \{ res(m, L) \mid m \in MOD(\sigma \cup \{L\})\}$ , and so, $|MOD(\sigma \cup \{\neg L\})| = |MOD(\sigma \cup \{L\})|$ , then there exists no supplement of $\sigma$ w.r.t. $L$ (contradicting the premise), since any formula (not involving $A$ ) cancelling all models of $MOD(\sigma \cup \{\neg L\})$ would also cancel all models of $MOD(\sigma \cup \{L\})$ . ■

The following Corollary is implied by Theorem 10.1.

**Corollary 10.1.** For any system $S$ , if a literal $L$ is pure in a prenex CNF of $S$ , then $E(S, L) > E(S, \neg L)$ .

## 11. Summary

The purpose of a Knowledge System $S$ is to represent (a part of) the world $W$ faithfully, so the meaning of the system must conform to the world in such a way that each model of $S$ corresponds to a possible state of $W$ . Although $S$ is intended for an adequate description of $W$ , the very nature of designing and maintaining of a Knowledge System may affect the degree of its adequacy. In particular, the system may contain contradictory data implying, for instance, that a statement $\alpha$ is true and at the same time false (for a different reason). And there are numerous possible causes of contradiction in Knowledge Systems.

The classical logic, unfortunately, can be of little use for reasoning with inconsistency. The problem is that for any formula $F$ , a conjunction $F \wedge \neg F$ is unsatisfiable in the classical logic, and so contradictory data make a system $S$ trivially meaningless, since any formula $\phi$ becomes its logical consequence. The need of special tools for reasoning with inconsistent information led to developing of paraconsistent logics [4, 8, 9, 14, 16, 25, 28, 29, 34] employing extended sets of truth values. Use of a multiple-valued logic provides an inconsistent system $S$ with an epistemological meaning in the sense that $S$ can reflect beliefs of a reasoner (human or machine) who may happen to hold conflicting beliefs. However, the system remains ontologically inconsistent as far as it is agreed that in the reality, in every possible state of the world every statement is either true or false exclusively.

Thus, given a logic system presenting some knowledge about the real world, one tries to figure out what is the truth value of a given statement in the reality. In the case that the system contains uncertain or even contradictory information, the goal is to decide whether the statement is more likely to be true than false in the world. This paper presents a general approach to resolving contradictions in First Order logic systems. Plausibility of the proposed solution is illustrated by examples appealing to the common sense intuition.

If a Knowledge System $S$ turns out to be inconsistent containing contradictory data, we may view its present inconsistency as a result of information pollution with some wrong data. However, we may reasonably assume that (say, due to a proper maintenance) most of the system content still reflects the world truthfully, and therefore it would be a great loss to allow a small contradiction to depreciate or even destroy a large amount of correct knowledge. So, despite the pollution, $S$ must contain a meaningful subset, and we want the latter to be maximally consistent in order not to lose any correct data. On the other hand, since the world is supposed to be consistent, only consistent subsets of the system have a chance to conform to it. Thus, it appears reasonable to assume that the semantics of a logic system is determined by that of its maximally consistent subsets, *mc-subsets* . This principle, adopted by many researchers in the field, is trivially correct for consistent systems, but is not easy to implement in a general case, since different mc-subsets of an inconsistent system may contain different and even conflicting data.

As a Knowledge System allows assessment of the truth of various statements, it must contain *semantic information* , which we wish to utilize as fully as possible for making plausible decisions. Section 3 gives an estimation of the *quantity of semantic information* , $I(S)$ , contained in a logic system $S$ .

Given the information about $W$ contained in $S$ , we may derive certain conclusions regarding the truth of a formula $F$ in $W$ . In this sense we say that $S$ provides a certain *evidence of F* , $E(S, F)$ , that presents a measure of credibility of a belief in truth of $F$ . Section 4 states common sense postulates governing the evidence. It reveals a close relationship between the evidence, the quantity of semantic information of the system, and the set of models of its mc-subsets. To force the evidence to comply with the postulates, Section 5 introduces the *semantics of weighted mc-subsets* , and defines the value of $E(S, F)$ .

How can it be checked that the evidence and the semantics of weighted mc-subsets, besides satisfying certain common sense postulates, provide bases for reasoning in inconsistent systems? To show that this semantics indeed enables reconciling of contradictions and recovering of plausible truth values of ambiguous formulas, it is applied to a series of justifying examples including chain proofs, rules with exceptions, and paradoxes (Section 6). It turns out (Section 7) that the semantics of weighted mc-subsets possesses an important feature of resilience with respect to contradiction.

Section 8 studies special cases of inconsistent systems, such as homogeneous sets and sets of ground clauses. Section 9 modifies the definition of mc-subsets in the presence of doubtless formulas. Section 10 presents a range of semantics based on reasoning by evidence.

**Acknowledgment.**

Many thanks to the anonymous referees for most inspiring and benevolent comments.

This research has been partly supported by the Israel Ministry of Science and Technology (grant 4469-1-92) and by the S. A. Schonbrunn Research Endowment Fund (grant 162692).

## References

1. J. Aczel and Z. Daroczy, *On Measures of Information and their Characterizations.* Academic Press, New York, 1975.
2. C. Baral, S. Kraus, J. Minker, and V. S. Subrahmanian, Combining knowledge bases consisting of first order theories. *Lecture Notes in Artificial Intelligence,* v. 542, Z. Ras and M. Zemankova, Eds., Springer-Verlag, 1991, 92-101.
3. N. D. Belnap, A useful four-valued logic. *Modern Uses of Multiple-Valued Logic* (G. Epstein and J. M. Dunn, eds.), Reidel, 1977, 5-37.
4. H. A. Blair and V. S. Subrahmanian, Paraconsistent logic programming. *Theoretical Computer Science,* 1989, 135-154.
5. L. Brillouin, *Science and Information.* Academic Press, New York, 1956.
6. W. A. Carnielli, L. Farinas del Cerro, and M. Lima Marques, Contextual negations and reasoning with contradictions. *Proc. 12th IJCAI,* Sydney, Australia, 1991, 532-537.
7. S. Cook, The complexity of theorem-proving procedures. *Proc. 3rd ACM Symp. on Theory of Computing,* 1971, 151-158.
8. N. C. A. da Costa, On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic,* v. 15, 1974, 497-510.
9. N. C. A. da Costa, L. J. Henschen, J. J. Lu, and V. S. Subrahmanian, Automatic theorem proving in paraconsistent logics: theory and implementation. *10th International Conference on Automatic Deduction, Lecture Notes in AI,* v. 449, M. E. Stickel (Ed.), Springer Verlag, 1990, 72-86.
10. R. Fagin, J. Y. Halpern, and M. Y. Vardi, A nonstandard approach to the logical omniscience problem. *Proc. of the 1990 Conference on Theoretical Aspects of Reasoning about Knowledge,* 1990, 41-54.
11. G. Gallo, and G. Urbani, Algorithms for testing the satisfiability of propositional formulae. *J. Logic Programming,* v. 7, 1989, 45-61.
12. A. Goldberg, Average case complexity of the satisfiability problem. *Proc. 4th Workshop on Automated Deduction,* Austin, TX, 1979, 1-6.
13. A. Goldberg, P. Purdom, and C. Brown, Average time analyses of simplified Davis-Putnam procedures. *Information Processing Letters,* v. 15, 1982, 72-75.
14. J. Grant, Classifications for inconsistent theories. *Notre Dame Journal of Formal Logic,* v. 19, 1978, 435-444.
15. K. Iwama, CNF satisfiability test by counting and polynomial average time. *SIAM. J. Comput.,* v. 18, 1989, 385-391.

16. M. Kifer and E. L. Lozinskii, RI: A logic for reasoning with inconsistency. *Proc. 4th IEEE Symp. on Logic in Computer Science,* Asilomar, California, 1989, 253-262.

17 R. A. Kowalski and F. Sadri, Logic programs with exceptions. *Proc. 7th International Conf. on Logic Programming,* ed. by D. H. D. Warren and P. Szeredi, Jerusalem, 1990, 598-613.

18. N. Linial and N. Nisan, Approximate inclusion-exclusion. *Combinatorica,* v. 10, no. 4, 1990, 349-365.

19. E. Lozinskii, Answering atomic queries in indefinite deductive databases. *Intl. J. of Intelligent Systems,* v. 4, no. 4, 1989, 403-430.

20. E. Lozinskii, Counting propositional models. *Information Processing Letters,* v. 41, 1992, 327-332.

21. E. Lozinskii, Information and evidence in logic systems. To appear in *Journal of Experimental and Theoretical Artificial Intelligence.*

22. M. Luby and B. Velickovic, On deterministic approximation of DNF. *Proceedings of STOC'91,* New Orleans, May 1991, 430-438.

23. T. Pequeno and A. Buchsbaum, The logic of epistemic inconsistency. *Proc 2nd Intl. Conf. on Knowledge Representation and Reasoning,* Cambridge, Mass., 1991, 453-460.

24. D. Perlis, Truth and meaning. *Artificial Intelligence,* v. 39, 1989, 245-250.

25. G. Priest, Reasoning about truth. *Artificial Intelligence,* v. 39, 1989, 231-244.

26. P. Purdom and C. Brown, The pure literal rule and polynomial average time. *SIAM J. Comput.,* v. 14, no. 4, 1985, 943-953.

27 R. Reiter, A logic for default reasoning. *Artificial Intelligence,* v. 13, 1980, 81-132.

28. N. Rescher and R. Brandom, *The Logic of Inconsistency.* Oxford, 1980.

29. J. Slaney, The implications of paraconsistency. *Proc. 12th IJCAI,* Sydney, Australia, 1991, 1052-1056.

30. R. Smullyan, *Forever Undecided.* Oxford University Press, 1989.

31. L. A. Stein, Skeptical inheritance: Computing the intersection of credulous extensions. *Proc. 11th IJCAI,* 1989, 1153-1158.

32. D. S. Touretzky, J. F. Horty, and R. H. Thomason, A clash of intuition: The current state of nonmonotonic multiple inheritance systems. *Proc. 10th IJCAI,* 1987, 476-482.

33. L. G. Valiant, The complexity of computing the permanent. *Theoretical Computer Science,* v. 8, 1979, 189-201.

34. G. Wagner, Ex contradictione nihil sequitur. *Proc. 12th IJCAI,* Sydney, Australia, 1991, 538-542.