A Framework for Measuring Information Asymmetry

Yakoub Salhi

CRIL, U. Artois & CNRS
F-62300 Lens, France
salhi@cril.fr

Abstract

Information asymmetry occurs when an imbalance of knowledge exists between two parties, such as a buyer and a seller, a regulator and an operator, and an employer and an employee. It is a key concept in several domains, in particular, in economics. We propose in this work a general logic-based framework for measuring the information asymmetry between two parties. A situation of information asymmetry is represented by a knowledge base and a set of questions. We define the notion of information asymmetry measure through rationality postulates. We further introduce a syntactic concept, called minimal question subset (MQS), to take into consideration the fact that answering some questions allows avoiding others. This concept is used for defining rationality postulates and measures. Finally, we propose a method for computing the MQSes of a given situation of information asymmetry.

Introduction

Information asymmetry refers to situations of interactions between two parties in which one has more relevant pieces of information than another. This concept plays a central role in explaining different important phenomena in several domains, particularly in economics. It was originally introduced in Akerlof's article: *The Market for 'Lemons': Quality Uncertainty and the Market Mechanism* (Akerlof 1970). In particular, Akerlof explained in this article the negative effect of information asymmetry on the quality of goods by considering the example of the used car market. Intuitively, the fact that the buyers and the sellers of cars do not have the same knowledge about the quality provides incentives to favor the cars of less than the average market quality (*adverse selection*); the buyers tend to pay less and the sellers are thereby conduced to reduce the quality of the proposed cars. In order to resolve the problems caused by information asymmetry, the most known approaches are the method of signaling proposed by Spence (Spence 1973; Connelly et al. 2011) and the method of screening proposed by Stiglitz (Stiglitz 1975). For their works on information asymmetry, George Akerlof, Michael Spence and Joseph E.

Stiglitz received the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel in 2001.

Since information asymmetry plays a key role in many economic settings, defining measures in this context has received considerable attention (e.g. (Glosten and Harris 1988; Easley, Hvidkjaer, and O'Hara 2002; Chen, Goldstein, and Jiang 2006; Armstrong, Balakrishnan, and Cohen 2012)). However, to the best of our knowledge, no general framework for defining such measures has been proposed in the literature in the sense that the previous works concern very specific questions in economics and our framework uses an abstract formalism (a formal logic). The main purpose of this work is to introduce a general logic-based framework for measuring information asymmetry. The measures defined through this framework are motivated by different applications. In particular, they can be used for selecting appropriate partner agents: less information asymmetry leads to more appropriate partner agent. They can also be used for better appraising products (e.g. in auctions) by taking into account adverse selection: information asymmetry causes in general erroneous price evaluation. In addition, these measures can be used in combination with methods for resolving problems related to information asymmetry. Indeed, they allow one to detect situations where it is important to involve methods that reduce the negative impact.

In our framework, classical propositional logic is used for representing pieces of knowledge as well as questions. A situation of information asymmetry, called in this paper *IA-instance*, is described through two sets of formulas representing respectively the pieces of knowledge and the questions. More precisely, the first set represents the knowledge state of one of the two involved parties, and the questions are those that this party wants to answer knowing that the other party has the answers. In this context, the notion of information asymmetry measure is defined as a function that associates a non-negative numeric value to every IA-instance: the greater the number returned by a measure, the greater is the amount of information asymmetry. We provide postulates for measuring information asymmetry in order to characterize important rational aspects. In addition, we propose different measures that illustrate different important points. It is worth noting that our approach is similar in a sense to

that used for measuring inconsistency in knowledge bases (e.g. see (Hunter and Konieczny 2010)). Moreover, let us mention that we do not use in this work the word *measure* in the mathematical sense, but only to invoke the informal meaning like in the case of the inconsistency measures.

In order to consider the fact that answering some questions allows avoiding others, we introduce a syntactic concept, called minimal question subset (MQS). Roughly speaking, an MQS is a subset of questions that is minimal w.r.t. the set inclusion and their elements allow obtaining answers for all the considered questions. The MQSes are used for defining rationality postulates and information asymmetry measures. Finally, we introduce a method for generating all the MQSes of an IA-instance using, in particular, the problem of computing X-minimal models (e.g. (Avin and Ben-Eliyahu-Zohary 2001)).

Background

Given a finite set S, we use 2^S and $|S|$ to denote respectively its powerset (the set of its subsets) and its cardinality. Furthermore, we use Prop to denote the set of propositional variables. The propositional formulas of classical propositional logic are built using Prop, the constants $\top$ and $\bot$, denoting respectively true and false, and the unary logical connective $\neg$ and the usual binary connectives $\wedge$, $\vee$ and $\rightarrow$. The set of propositional formulas is denoted Form. We use the letters p, q, r, s to denote the propositional variables, and the Greek letters ϕ, ψ and χ to denote the propositional formulas. Moreover, given a syntactic object o involving propositional variables (a formula, a set of formulas, etc.), we use $Var(o)$ to denote the set of propositional variables occurring in this object.

A *Boolean interpretation* $\mathcal{B}$ of a formula ϕ is defined as a function from a set of propositional variables V, with $Var(\phi) \subseteq V$, to $\{0, 1\}$. It is inductively extended to the propositional formulas as usual. A formula ϕ is *consistent* if there exists a Boolean interpretation $\mathcal{B}$ of ϕ such that $\mathcal{B}(\phi) = 1$, and $\mathcal{B}$ is called a *model* in this case. Given a set of variables S such that $Var(\phi) \subseteq S$, we use $Mod(\phi, S)$ to denote the set of all the models of ϕ defined over S. In addition, ϕ is said to be *valid* if every Boolean interpretation of ϕ is one of its models. For convenience purposes, we also note a Boolean interpretation as a set of expressions of the form $p \mapsto v$ to state that the propositional variable p takes the truth value v.

Given a finite set of formulas Γ, a formula ϕ is said to be a logical consequence of Γ, written $\Gamma \vdash \phi$, if $\bigwedge \Gamma \rightarrow \phi$ is valid, where $\bigwedge \Gamma$ stands for the conjunction of the propositional formulas occurring in Γ with $\bigwedge \emptyset = \top$. In the case where Γ contains a single formula ψ, we write $\psi \vdash \phi$. In addition, if both $\phi \vdash \psi$ and $\psi \vdash \phi$ hold, we write $\phi \equiv \psi$. Furthermore, we use $PA(\Gamma)$ to denote the set of formulas $\Gamma \cup \{\neg\phi : \phi \in \Gamma\}$.

We define an equivalence relation, denoted $\asymp$, over Form as follows: $\phi \asymp \psi$ if and only if $\phi \equiv \psi$ or $\phi \equiv \neg\psi$. Moreover, given a finite set of formulas Γ, we use $Eq^{\asymp}(\Gamma)$ to denote the partition of Γ into equivalence classes w.r.t. the relation $\asymp$.
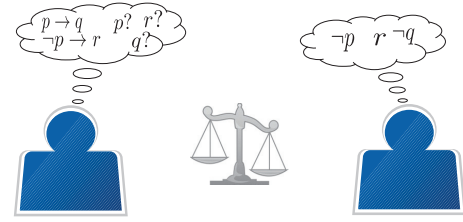

Figure 1: An illustration for information asymmetry

Measuring Information Asymmetry

This section is devoted to the introduction of the notion of information asymmetry measure. Accordingly, we first provide an example describing a case where information asymmetry occurs. Then, using a postulate-based approach, we formally introduce the notion of information asymmetry measure. Some interesting measures are proposed to illustrate different addressed aspects.

Example 1. *In order to illustrate the notion of information asymmetry, we provide a simple example of a voter and three candidates for an election c_1, c_2 and c_3. The voter plans to give his voice according to the answers to three questions:*

- *Is the candidate in favor of the law l?*
- *Is the candidate in favor of the public policy p? (making laws, allocating resources, etc.)*
- *Is the candidate in favor of the public policy q?*

In this context, we assume the voter knows that if a candidate is for the policy p then she/he is also for the law l ($p \rightarrow l$); and if a candidate is against the policy q then she/he is for the law l ($\neg q \rightarrow l$). In addition, the voter has the two following pieces of information: if the candidate c_1 is for q then she/he is also for p ($q \rightarrow p$); the candidate c_2 is in favor of the policy p or the policy q ($p \vee q$). We use K_i to denote the set of pieces of information of the voter about the candidate c_i. Thus, we have $K_1 = \{p \rightarrow l, \neg q \rightarrow l, q \rightarrow p\}$, $K_2 = \{p \rightarrow l, \neg q \rightarrow l, p \vee q\}$ and $K_3 = \{p \rightarrow l, \neg q \rightarrow l\}$. The answers to the three previous questions consist in knowing whether or not each of the following formulas, which corresponds to propositional variables, is true: l, p and q.

Clearly, the knowledge of the voter does not allow having all the desired answers in the case of any candidate. Therefore, there exists a situation of information asymmetry between the voter and each candidate in the sense that the latter has more information than the former. Furthermore, the pieces of information about the candidate c_1 allow the voter to have an answer to the first question since $K_1 \vdash l$, but the pieces of information about the two other candidates do not allow having any answer. Thus, one can reasonably consider that the amount of information asymmetry between the voter and c_1 is lower that those between the voter and the other candidates. It is also possible to consider that the information asymmetry between the voter and c_2 is lower than that between the voter and c_3 since K_3 is included in the set K_2.

The following definition presents the way we describe situations of information asymmetry.

Definition 1 (IA-Instance). *An IA-instance is a structure of the form* $\langle K \mid Q \rangle$ *where* K *(the knowledge base) and* Q *(the questions) are finite subsets of* Form *such that* $K \nvdash \bot$.

We use $\mathsf{IA}_{\mathsf{Form}}$ to denote the set of IA-instances.

Let us consider again Example 1. The situations of information asymmetry between the voter and the candidates corresponds to the three following IA-instances: $\langle \{p \to l, \neg q \to l, q \to p\} \mid \{l, p, q\} \rangle$, $\langle \{p \to l, \neg q \to l, p \wedge q\} \mid \{l, p, q\} \rangle$, and $\langle \{p \to l, \neg q \to l\} \mid \{l, p, q\} \rangle$.

In the following definition, we use $\mathbb{R}_{\infty}^{+}$ to refer to the set of positive real numbers augmented with a greatest element denoted ∞.

Definition 2 (Information Asymmetry Measure). *A function* $\mathcal{A} : \mathsf{IA}_{\mathsf{Form}} \to \mathbb{R}_{\infty}^{+}$ *is an information asymmetry measure if it satisfies the following properties:*

- CompleteInformation: $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$, $\mathcal{A}(I) = 0$ *iff* $\forall \psi \in Q$, *we have either* $K \vdash \psi$ *or* $K \vdash \neg\psi$.
- MoreInformation: $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$, $\forall \phi \in K$ *and* $\forall K' \subset$ Form *with* K' *is finite,* $K \cup K' \nvdash \bot$ *and* $K' \vdash \phi$, $\mathcal{A}(I) \geq \mathcal{A}(\langle (K \setminus \{\phi\}) \cup K' \mid Q \rangle)$.
- Weakening: $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$ *and* $\forall \psi \in Q$, *if* $K \vdash \psi$ *or* $K \vdash \neg\psi$, *then* $\mathcal{A}(I) = \mathcal{A}(\langle K \mid Q \setminus \{\psi\} \rangle)$.
- Equivalence: $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$, $\forall \psi \in Q$ *and* $\forall \chi \in$ Form *with* $\psi \asymp \chi$, $\mathcal{A}(I) = \mathcal{A}(\langle K \mid (Q \setminus \{\psi\}) \cup \{\chi\} \rangle)$.

The rationality postulate CompleteInformation states that an information asymmetry measure must allow distinguishing between the case where the knowledge allows obtaining all the answers and the case where there are missing answers. MoreInformation means that information asymmetry cannot increase if one replaces a piece of knowledge with more informative pieces of knowledge: more information leads to less information asymmetry. The property Weakening states that the amount of information asymmetry depends only on the questions that cannot be answered from the knowledge. This postulate results from the simple fact that information asymmetry is only caused by the ignored answers. Equivalence says that replacing a question with an equivalent formula or a formula equivalent to its negation does not change the amount of information asymmetry. In particular, this property implies that the amount of information asymmetry is free from the syntax of the questions.

Let us note that an information asymmetry measure can be seen as an *incompleteness measure* in the sense that the knowledge part is unable to prove formulas derived from the questions.

As shown throughout this work, the four key rationality postulates that the information asymmetry measures must satisfy bring about interesting measures that take into account several important aspects about the ignored answers.

Similarly to MoreInformation, one can consider an additional property stating that the information asymmetry cannot decrease if one adds new questions:

- MoreQuestions: $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$ *and* $\forall \psi \in$ Form, $\mathcal{A}(I) \leq \mathcal{A}(\langle K \mid Q \cup \{\psi\} \rangle)$.

However, we think that this property is not suitable in the general case, since it can be appropriate in certain cases to consider that adding new questions may allow avoiding many other existing ones. To illustrate this point, let us consider the IA-instance $I = \langle \{r \to (p \wedge q), \neg r \to (\neg p \wedge \neg q)\} \mid \{p, q\} \rangle$. Adding the question r in this case allows clearly avoiding the two questions p and q. Indeed, we may reasonably consider in some cases that we have the same amount of information asymmetry in $\langle \{r \to (p \wedge q), \neg r \to (\neg p \wedge \neg q)\} \mid \{p, q, r\} \rangle$ and $\langle \{r \to (p \wedge q), \neg r \to (\neg p \wedge \neg q)\} \mid \{r\} \rangle$. This can be interpreted by the fact that the agent is aware of a question that allows responding to all the others ("good" questions allow reducing information asymmetry).

The rationality postulate MoreInformation leads to the following proposition.

Proposition 1. *The following properties are satisfied for every information asymmetry measure* $\mathcal{A}$:

- $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$ *and* $\forall \phi \in$ Form *with* $K \vdash \phi$, $\mathcal{A}(I) = \mathcal{A}(\langle K \cup \{\phi\} \mid Q \rangle)$;
- $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$ *and* $\forall \phi \in$ Form *with* $K \cup \{\phi\} \nvdash \bot$, $\mathcal{A}(I) \geq \mathcal{A}(\langle K \cup \{\phi\} \mid Q \rangle)$;
- $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$, $\mathcal{A}(I) = \mathcal{A}(\langle \{\bigwedge_{\phi \in K} \phi\} \mid Q \rangle)$.

The following proposition stems from the postulate Equivalence and the fact that $\psi \asymp \neg\neg\psi$ holds for every formula $\psi \in$ Form.

Proposition 2. $\forall I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$ *and* $\forall \psi \in Q$, $\mathcal{A}(I) = \mathcal{A}(\langle K \mid (Q \setminus \{\psi\}) \cup \{\neg\neg\psi\} \rangle)$.

For instance, one can easily see that the following functions are information asymmetry measures:

$$\mathcal{A}_{bin}(\langle K \mid Q \rangle) = \begin{cases} 0 & \text{if } \forall \psi \in Q, K \vdash \psi \text{ or } K \vdash \neg\psi \\ 1 & \text{otherwise} \end{cases}$$

$$\mathcal{A}_{ma}(\langle K \mid Q \rangle) = |Eq^{\asymp}(\{\psi \in Q \mid K \nvdash \psi \text{ and } K \nvdash \neg\psi\})|$$

where bin and ma refer to respectively *binary* and *missing answers*. It is worth noting that they also satisfy the property MoreQuestions.

Although we do not require the following properties for all information asymmetry measures, they may be appropriate in different contexts:

- VarAdditivity: $\forall I = \langle K \mid Q \rangle, I' = \langle K' \mid Q' \rangle \in \mathsf{IA}_{\mathsf{Form}}$ with $Var(K \cup Q) \cap Var(K' \cup Q') = \emptyset$, $\mathcal{A}(\langle K \cup K' \mid Q \cup Q' \rangle) = \mathcal{A}(I) + \mathcal{A}(I')$.
- Subadditivity: $\forall I = \langle K \mid Q \rangle, I' = \langle K' \mid Q' \rangle \in \mathsf{IA}_{\mathsf{Form}}$ with $K \cap K' = \emptyset$ and $K \cup K' \nvdash \bot$, $\mathcal{A}(\langle K \cup K' \mid Q \cup Q' \rangle) \leq \mathcal{A}(I) + \mathcal{A}(I')$.

The property VarAdditivity states that the amount of information asymmetry of the combination of two IA-instances that have nothing in common (namely they do not share any propositional variable) is the sum of the amounts of information asymmetry at each IA-instance. The property Subadditivity says that the amount of information asymmetry in a combination of two IA-instances that have disjoint knowledge parts is less than or equal to the sum of the amounts of information asymmetry at each IA-instance.

This property is due to the fact that combining pieces of knowledge may allow obtaining additional answers and reducing thereby information asymmetry.

Let us now introduce some interesting information asymmetry measures. Before that, we need the following notational convention. Given an IA-instance $I = \langle K \mid Q \rangle$, we use $\Pi(I)$ to denote the set $\{\{\chi \in PA(Q) : \mathcal{B}(\chi) = 1\} : \mathcal{B} \in Mod(K, Var(K \cup Q))\}$. Intuitively, $\Pi(I)$ corresponds to the set of all possible answer combinations.

Let us consider the two following functions on the IA-instances:

- $\mathcal{A}_\Pi(I) = |\Pi(I)| - 1$;
- $\mathcal{A}_{err}(I) = |Q| - min\{|P \cap P'| : P, P' \in \Pi(I)\}$ with $I = \langle K \mid Q \rangle$.

The first measure $\mathcal{A}_\Pi$ captures the fact that the greater the number of possible answer combinations, the greater is the amount of information asymmetry. Regarding the measure $\mathcal{A}_{err}$, it corresponds to the maximum number of the possible wrong answers by considering an arbitrary possible answer combination.

Proposition 3. *The function $\mathcal{A}_{err}$ is an information asymmetry measure that satisfies* VarAdditivity, Subadditivity *and* MoreQuestions.

Proof. We only consider the case of VarAdditivity, the other cases being simple or similar. Let $I = \langle K \mid Q \rangle, I' = \langle K' \mid Q' \rangle \in \mathsf{IA}_{\mathsf{Form}}$ with $Var(K \cup Q) \cap Var(K' \cup Q') = \emptyset$. Then, let $\mathcal{B}_1, \mathcal{B}_2 \in Mod(K, Var(K \cup Q))$ and $\mathcal{B}'_1, \mathcal{B}'_2 \in Mod(K', Var(K' \cup Q'))$ such that $\mathcal{A}_{err}(I) = |\{\psi \in Q : \mathcal{B}_1(\psi) \neq \mathcal{B}_2(\psi)\}|$ and $\mathcal{A}_{err}(I') = |\{\psi \in Q : \mathcal{B}'_1(\psi) \neq \mathcal{B}'_2(\psi)\}|$. Using the fact $Var(K \cup Q) \cap Var(K' \cup Q') = \emptyset$, we obtain $\mathcal{B}_1 \cup \mathcal{B}'_1, \mathcal{B}_2 \cup \mathcal{B}'_2 \in Mod(K \cup K', Var(K \cup K' \cup Q \cup Q'))$ and, as a consequence, $\mathcal{A}_{err}(\langle K \cup K' \mid Q \cup Q'\rangle) \geq \mathcal{A}_{err}(I) + \mathcal{A}_{err}(I')$ holds. Assume now that $\mathcal{A}_{err}(\langle K \cup K' \mid Q \cup Q'\rangle) > \mathcal{A}_{err}(I) + \mathcal{A}_{err}(I')$. Then, there exists $\mathcal{B}, \mathcal{B}' \in Mod(K \cup K', Var(K \cup K' \cup Q \cup Q'))$ such that $\mathcal{A}_{err}(\langle K \cup K' \mid Q \cup Q'\rangle) = |\{\psi \in Q \cup Q' : \mathcal{B}(\psi) \neq \mathcal{B}'(\psi)\}|$. Thus, using the fact $Var(K \cup Q) \cap Var(K' \cup Q') = \emptyset$, we have $\mathcal{A}_{err}(I) < |\{\psi \in Q : \mathcal{B}(\psi) \neq \mathcal{B}'(\psi)\}|$ or $\mathcal{A}_{err}(I') < |\{\psi \in Q' : \mathcal{B}(\psi) \neq \mathcal{B}'(\psi)\}|$. Therefore, using the definition of $\mathcal{A}_{err}$, we obtain a contradiction, and we thereby deduce that $\mathcal{A}_{err}(\langle K \cup K' \mid Q \cup Q'\rangle) = \mathcal{A}_{err}(I) + \mathcal{A}_{err}(I')$. $\square$

The function $\mathcal{A}_\Pi$ is also an information asymmetry measure, but it does not satisfy the two additivity properties VarAdditivity and Subadditivity.

Proposition 4. *The function $\mathcal{A}_\Pi$ is an information asymmetry measure that satisfies* MoreQuestions.

To be convinced that $\mathcal{A}_\Pi$ does not satisfy the additivity properties, consider the two IA-instances $I = \langle \emptyset \mid \{p, q\}\rangle$ and $I = \langle \emptyset \mid \{p', q'\}\rangle$. Then, we have $\mathcal{A}_\Pi(I) = \mathcal{A}_\Pi(I') = 4$, but we have $\mathcal{A}_\Pi(\langle \emptyset \mid \{p, q, p', q'\}\rangle) = 16$ since every interpretation leads to a possible answer combination.

Example 2. *This example is inspired by Akerlof's article (Akerlof 1970). Let us consider the following statements about a car:*

- *p: it has a high mileage;*
- *q: it was involved in a serious accident;*
- *r: it was repainted;*
- *s: it has worn tires; and*
- *t: tires were replaced.*

Let $I = \langle\{q \rightarrow r, s \rightarrow p, t \rightarrow s\} \mid \{p \vee q, r\}\rangle$ and $I' = \langle\{q \rightarrow r, s \rightarrow p, t \rightarrow s, t \vee q\} \mid \{p \vee q, r\}\rangle$. The Boolean interpretations $\mathcal{B}_1 = \{p' \mapsto 0 \ for \ p' = p, q, r, s, t\}$, $\mathcal{B}_2 = \{p' \mapsto 1 \ for \ p' = p, q, r, s, t\}$, $\mathcal{B}_3 = \{p \mapsto 1, p' \mapsto 0 \ for \ p' = q, r, s, t\}$, $\mathcal{B}_4 = \{r \mapsto 1, p' \mapsto 0 \ for \ p' = q, s, t\}$ are models of $K = \{q \rightarrow r, s \rightarrow p, t \rightarrow s\}$, and we have, as a consequence, $\Pi(I) = \{\{\neg(p \vee q), \neg r\}, \{p \vee q, r\}, \{p \vee q, \neg r\}, \{\neg(p \vee q), r\}\}$. Thus, $\mathcal{A}_\Pi(I) = 3$ and $\mathcal{A}_{err}(I) = 2$ hold. Moreover, knowing that $K \cup \{t \vee q\} \vdash p \vee q$ and $K \cup \{t \vee q\} \nvdash r$, we obtain $\mathcal{A}_\Pi(I') = 2$ and $\mathcal{A}_{err}(I') = 1$.

Relationships between Questions

The information asymmetry measures described previously do not take into consideration, in an explicit way, the fact that answers may allow bringing about other answers, which explains in particular why these measures satisfy the property MoreQuestions. This fact calls for concepts that consider relationships between questions. To this end, we particularly introduce in this section the concept of minimal question subset, which is used for defining interesting properties and information asymmetry measures.

Definition 3 (Minimal Answer Subset)**.** *Let $I = \langle K \mid Q \rangle$ be an IA-instance. A* minimal answer subset *(MAS) S of I is a subset of $PA(Q)$ where the following properties are satisfied:*

1. *$K \cup S \nvdash \bot$;*
2. *$\forall \psi \in Q, K \cup S \vdash \psi$ or $K \cup S \vdash \neg\psi$; and*
3. *$\forall S' \subset S, \exists \psi' \in Q$ s.t. $K \cup S' \nvdash \psi'$ and $K \cup S' \nvdash \neg\psi'$.*

In other words, a minimal answer subset corresponds to a set of possible answers that allows obtaining an answer for every question in the IA-instance and is minimal w.r.t. the set inclusion. We use $\mathsf{MAS}(I)$ to denote the set of the minimal answer subsets of the IA-instance I.

Definition 4 (Minimal Question Subset)**.** *Let $I = \langle K \mid Q \rangle$ be an IA-instance. A* minimal question subset *(MQS) Q' of I is a subset of Q where the following properties are satisfied:*

1. *$\forall S \subseteq PA(Q')$ s.t. $|S| = |Q'|$ and $K \cup S \nvdash \bot$, we have either $K \cup S \vdash \psi$ or $K \cup S \vdash \neg\psi$ for every $\psi \in Q$; and*
2. *$\forall Q'' \subset Q', Q''$ does not satisfy the property (1), i.e., Q' is minimal w.r.t. the set inclusion.*

Similarly to the notion of MAS, an MQS is a subset of questions where every combination of the possible answers of their elements consistent with the knowledge part (1) allows obtaining an answer for every question and (2) it has to be minimal w.r.t the set inclusion. We use $\mathsf{MQS}(I)$ to denote the set of the minimal question subsets of I.

The following proposition relates the notion of minimal answer subset with that of minimal question subset.

Proposition 5. *Given an IA-instance I and $Q' \in \mathsf{MQS}(I)$, $\forall R \subseteq PA(Q')$ s.t. $|R| = |Q'|$ and $K \cup R \nvdash \bot$, there exists $S \in \mathsf{MAS}(I)$ s.t. $S \subseteq R$.*

Definition 5 (Weak Question). *Let $I = \langle K \mid Q \rangle$ an IA-instance. A formula ψ in Q is said to be a* weak question *in I if $\forall Q' \in \mathsf{MQS}(I)$, $\psi \notin Q'$.*

The set of weak questions of an IA-instance I is denoted by $\mathsf{WeakQ}(I)$.

The following proposition comes from the fact that a question that can be answered by using only the knowledge part cannot be useful for answering any other question.

Proposition 6. $\forall I = \langle K \mid Q \rangle \in \mathsf{IA_{Form}}$, $\{\psi \in Q : K \vdash \psi \text{ or } K \vdash \neg\psi\} \subseteq \mathsf{WeakQ}(I)$ *holds.*

Example 3. *Let $I = \langle \{p \to q, \neg p \to r\} \mid \{p, q, r\}\rangle$, $I' = \langle \{(p \wedge q) \leftrightarrow r\} \mid \{p, q, r, \neg p \wedge r\}\rangle$ be two IA-instances. We have $\mathsf{MAS}(I) = \{\{p, r\}, \{\neg r\}, \{\neg p, q\}, \{\neg q\}\}$. For example, $\{\neg q\} \in \mathsf{MAS}(I)$ because we have $\{p \to q, \neg p \to r\} \cup \{\neg q\} \vdash \neg p \wedge r$. Moreover, we obtain $\mathsf{MQS}(I) = \{\{p, q, r\}\}$ and, consequently, $\mathsf{WeakQ}(I) = \emptyset$ holds. Furthermore, we have $\mathsf{MAS}(I') = \{\{p, q\}, \{\neg p, q\}, \{p, \neg q\}, \{\neg p, \neg q\}, \{r\}, \{\neg r, p\}, \{\neg r, q\}\}$. Note that all the possible combinations of p and q are minimal answer subsets. Thus, we obtain $\mathsf{MQS}(I') = \{\{p, q\}\}$ and $\mathsf{WeakQ}(I') = \{r, \neg p \wedge r\}$.*

Let us now introduce a notational convention that is used in the sequel. Given a finite set S and two subsets T and T' of 2^S, we have $T \sqsubseteq T'$ if (i) for all $S' \in T$, there exists $S'' \in T'$ such that $S' \subseteq S''$, and (ii) for all $S'' \in T'$, there exists $S' \in T$ such that $S' \subseteq S''$.

The following additional properties on the information asymmetry measures are based on the notion of minimal question subset:

- WeakQuestion: $\forall I = \langle K \mid Q \rangle \in \mathsf{IA_{Form}}$, $\mathcal{A}(I) = \mathcal{A}(\langle K \mid Q \setminus \mathsf{WeakQ}(I)\rangle)$.

- MQSInclusion: $\forall I, I' \in \mathsf{IA_{Form}}$, if $\mathsf{MQS}(I) \sqsubseteq \mathsf{MQS}(I')$ then $\mathcal{A}(I) \leq \mathcal{A}(I')$.

- EqualAsymmetry: $\forall I, I' \in \mathsf{IA_{Form}}$ with $|\mathsf{MQS}(I)| = |\mathsf{MQS}(I')| = 1$ ($\mathsf{MQS}(I) = \{Q'\}$ and $\mathsf{MQS}(I) = \{Q''\}$), if $|Q'| = |Q''|$ then $\mathcal{A}(I) = \mathcal{A}(I')$.

- MoreAsymmetry: $\forall I, I' \in \mathsf{IA_{Form}}$ with $|\mathsf{MQS}(I)| = |\mathsf{MQS}(I')| = 1$ ($\mathsf{MQS}(I) = \{Q'\}$ and $\mathsf{MQS}(I) = \{Q''\}$), if $|Q'| > |Q''|$ then $\mathcal{A}(I) > \mathcal{A}(I')$.

The property WeakQuestion says that the weak questions do not impact the amount of information asymmetry. The property MQSInclusion states that the information asymmetry cannot decrease if we need in any case to answer more questions to resolve information asymmetry. One can easily see that MQSInclusion is stronger than WeakQuestion. The properties EqualAsymmetry and MoreAsymmetry concern the case where there is a single MQS and say the amount of information asymmetry is strongly swayed by the size of this MQS.

Let us now consider the following simple MQS-based measures:

- $\mathcal{A}_{\mathsf{MQS}}^{min}(I) = min\{|Q'| : Q' \in \mathsf{MQS}(I)\}$

- $\mathcal{A}_{\mathsf{MQS}}^{max}(I) = max\{|Q'| : Q' \in \mathsf{MQS}(I)\}$

- $\mathcal{A}_{\mathsf{MQS}}^{wq}(I) = |Eq^{\succ}(Q \setminus \mathsf{WeakQ}(I))|$

The measure $\mathcal{A}_{\mathsf{MQS}}^{min}$ corresponds to the minimum number of questions that one has to answer in order to obtain answers for all the questions in the instance. In a sense, this measure refers to the size of the best cases, while $\mathcal{A}_{\mathsf{MQS}}^{max}$ corresponds to the size of the worst cases. The measure $\mathcal{A}_{\mathsf{MQS}}$ allows avoiding the questions that can always be answered through other questions.

Proposition 7. *The functions $\mathcal{A}_{\mathsf{MQS}}^{min}$, $\mathcal{A}_{\mathsf{MQS}}^{max}$ and $\mathcal{A}_{\mathsf{MQS}}^{wq}$ are information asymmetry measures that satisfy the following properties:* VarAdditivity, MQSInclusion *(and consequently* WeakQuestion*),* EqualAsymmetry, MoreAsymmetry.

One can see that the property Subadditivity is satisfied by $\mathcal{A}_{\mathsf{MQS}}^{min}$ and $\mathcal{A}_{\mathsf{MQS}}^{max}$, while it is not satisfied by $\mathcal{A}_{\mathsf{MQS}}^{wq}$. For example, let us consider the two IA-instances $I = \langle \{(p \wedge q) \leftrightarrow r\} \mid \{p, q, r\}\rangle$ and $I' = \langle \{p \leftrightarrow q\} \mid \{s\}\rangle$. Similarly to the second IA-instance described in Example 3, $\mathsf{WeakQ}(I) = \{r\}$ holds. Moreover, we clearly have $\mathsf{WeakQ}(I') = \emptyset$; however, $\{r\}$ belongs to the minimal question subsets of $I'' = \langle \{(p \wedge q) \leftrightarrow r, p \leftrightarrow q\} \mid \{p, q, r, s\}\rangle$, and we obtain $\mathsf{WeakQ}(I'') = \emptyset$. Thus, we have $\mathcal{A}_{\mathsf{MQS}}^{wq}(I) = 2$ and $\mathcal{A}_{\mathsf{MQS}}^{wq}(I') = 1$, but $\mathcal{A}_{\mathsf{MQS}}^{wq}(I'') = 4 > \mathcal{A}_{\mathsf{MQS}}^{mq}(I) + \mathcal{A}_{\mathsf{MQS}}^{wq}(I')$.

Furthermore, it is worth noting that the measures $\mathcal{A}_{\mathsf{MQS}}^{min}$, $\mathcal{A}_{\mathsf{MQS}}^{max}$ and $\mathcal{A}_{\mathsf{MQS}}^{wq}$ do not satisfy MoreQuestions. Consider, for instance, the IA-instance $I = \langle \{(p \wedge q_1) \to r_1, \ldots, (p \wedge q_n) \to r_n\} \mid \{q_1, \ldots, q_n, r_1, \ldots, r_n\}\rangle$ where $p, q_1, \ldots, q_n, r_1, \ldots, r_n$ are pairwise distinct propositional variables and $n > 1$. Then, we obtain $\mathsf{MQS}(I) = \{\{q_1, \ldots, q_n, r_1, \ldots, r_n\}\}$ and $\mathsf{WeakQ}(I) = \emptyset$. Thus, $\mathcal{A}_{\mathsf{MQS}}^{min}(I) = \mathcal{A}_{\mathsf{MQS}}^{max}(I) = \mathcal{A}_{\mathsf{MQS}}^{wq}(I) = 2n$. Consider now the IA-instance $I' = \langle \{(p \wedge q_1) \to r_1, \ldots, (p \wedge q_n) \to r_n\} \mid \{p, q_1, \ldots, q_n, r_1, \ldots, r_n\}\rangle$. We have $\mathsf{MQS}(I') = \{\{p, q_1, \ldots, q_n\}\}$ and, as a consequence, $\mathsf{WeakQ}(I') = \{r_1, \ldots, r_n\}$ holds. Therefore, we obtain $\mathcal{A}_{\mathsf{MQS}}^{min}(I) = \mathcal{A}_{\mathsf{MQS}}^{max}(I) = \mathcal{A}_{\mathsf{MQS}}^{wq}(I) = n + 1$.

It is important to point out that the minimal question subsets, along with the minimal answer subsets, can be useful for defining more complex properties and information asymmetry measures than those that are put forward in this section.

Computation Methods

In the present section we tackle the problem of computing the minimal question subsets. We first propose an algorithm for computing the minimal answer subsets. We then describe an approach for computing the MQSes from the MASes, which is based on the X-minimal model problem.

First, let us note that given an IA-instance $I = \langle K \mid Q \rangle$ and $S \subseteq PA(Q)$, the problem of determining if S is a MAS can be solved using at most $|S| + 2$ NP-oracle calls. Indeed, this can be shown by using the following method:

1. if $K \cup S \vdash \bot$ (first call) then S is not a MAS, else let $S' = \{\chi \in PA(Q) : \mathcal{B}(\chi) = 1\}$ ($\mathcal{B}$ a model of $K \cup S$);

2. if $K \cup S \nvdash \bigwedge_{\chi \in S'} \chi$ (second call) then S is not a MAS;

3. if $\exists \chi \in S$, $S \setminus \{\chi\} \vdash \chi$ ($|S|$ calls) then S is not a MAS, else S is a MAS.

Algorithm 1: An approach for computing the MASes w.r.t. a given model of the knowledge part.

Data: An IA-instance $I = \langle K \mid Q \rangle$ and a model $\mathcal{B} \in Mod(K, Var(K \cup Q))$

Result: the set of all the MASes included in $S = \{\chi \in PA(Q) : \mathcal{B}(\chi) = 1\}$

1   $Sol \leftarrow \emptyset$;
2   $S_0 \leftarrow \emptyset$;
3   **for** $\chi \in S$ **do**
4   **if** $K \cup (S \setminus \{\chi\}) \nvdash \chi$ **then**
5       $S_0 \leftarrow S_0 \cup \{\chi\}$;
6   **if** $|S_0| = |S|$ **then**
7     $Sol \leftarrow \{S\}$;
8   **else**
9     $E_1 \leftarrow \{S\}$;
10    $B \leftarrow \emptyset$;
11    $k \leftarrow 2$;
12   **while** $E_{k-1} \neq \emptyset$ **do**
13      $E_k \leftarrow \emptyset$;
14      $R \leftarrow \emptyset$;
15      $f \leftarrow true$;
16   **for** $S' \in E_{k-1}$ **do**
17   **for** $\chi \in S' \setminus S_0$ **do**
18   **if**
             $S' \setminus \{\chi\} \notin E_k$ and $\nexists S'' \in B \cup Sol \; s.t.$
             $S' \setminus \{\chi\} \subseteq S''$ and $K \cup (S' \setminus \{\chi\}) \vdash \chi$
             then
19            $E_k \leftarrow E_k \cup \{S' \setminus \{\chi\}\}$;
20            $f \leftarrow false$;
21   **if** f **then** $R \leftarrow R \cup \{S'\}$;
22      $B \leftarrow B \cup \{S' \setminus \{\chi\} :$
        $S' \in E_{k-1}, \chi \in S' \setminus S_0, S' \setminus \{\chi\} \notin$
        $E_k, \nexists S'' \in B \cup R \; s.t. \; S' \setminus \{\chi\} \subseteq S''\}$;
23      $Sol \leftarrow Sol \cup R$;
24      $k++$;
25 **return** Sol;

It is worth mentioning that the second call allows showing the property *2* in Definition 3, while the last $|S|$ calls allow showing that S is minimal w.r.t. the set inclusion.

Let us now provide a proposition that allows us to reduce the search space by considering only one question per equivalence class w.r.t. $\asymp$.

Proposition 8. *The following properties are satisfied for every IA-instance $I = \langle K \mid Q \rangle$:*

1. $\forall \chi, \chi' \in PA(Q)$ *with* $\chi \neq \chi'$ *and* $\chi \equiv \chi'$, *and* $\forall S \in$ MAS(I), $\chi \notin S$ *or* $\chi' \notin S$;

2. $\forall \chi, \chi' \in PA(Q)$ *with* $\chi \neq \chi'$ *and* $\chi \equiv \chi'$, *and* $\forall S \subseteq PA(Q)$ *with* $\chi, \chi' \notin S$, $S \cup \{\chi\} \in$ MAS(I) *iff* $S \cup \{\chi'\} \in$ MAS(I);

3. $\forall C \in Eq^{\asymp}(Q)$, $\forall \psi, \psi' \in C$ *with* $\psi \neq \psi'$ *and* $\forall Q' \in$ MQS(I), $\psi \notin Q'$ *or* $\psi' \notin Q'$ *holds;*

4. $\forall C \in Eq^{\asymp}(Q)$, $\forall \psi, \psi' \in C$ *with* $\psi \neq \psi'$ *and* $\forall Q' \subseteq Q$ *with* $\psi, \psi' \notin Q'$, $Q' \cup \{\psi\} \in$ MQS(I) *iff* $Q' \cup \{\psi'\} \in$ MQS(I).
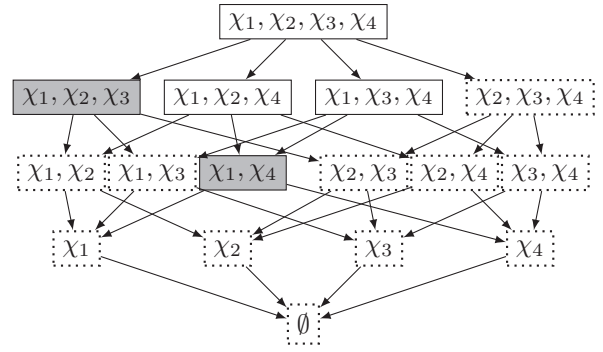

Figure 2: Minimal Answer Subset Lattice

From now on, we consider IA-instances where there are no two questions that are equivalent w.r.t. $\asymp$.

The following propositions provide a simple characterization of the minimal answer subsets through the models of the knowledge part.

Proposition 9. *Let $I = \langle K \mid Q \rangle \in$ IA$_{\mathsf{Form}}$ and $S \subseteq PA(Q)$ s.t. $S \nvdash \bot$. Then, $S \in$ MAS(I) iff $\exists \mathcal{B} \in Mod(K, Var(K \cup Q))$ s.t. (i) $\{\chi \in PA(Q) : S \vdash \chi\} = \{\chi \in PA(Q) : \mathcal{B}(\chi) = 1\}$, and (ii) $\forall \chi \in S, S \setminus \{\chi\} \nvdash \chi$ holds.*

Proposition 10. *Let $I = \langle K \mid Q \rangle \in$ IA$_{\mathsf{Form}}$, $\mathcal{B} \in Mod(K, Var(K \cup Q))$ and $S \subseteq S' = \{\chi \in PA(Q) : \mathcal{B}(\chi) = 1\}$. Then, $S \in$ MAS(I) iff $K \cup S \vdash \bigwedge S'$ and, $\forall \chi \in S, K \cup S \setminus \{\chi\} \nvdash \chi$ holds.*

Proposition 9 shows that all the MASes of an IA-instance can be obtained from the models of its knowledge part, while Proposition 10 shows how to obtain all the MASes associated to a given model.

In Algorithm 1, we describe our method for computing the MASes of a given IA-instance w.r.t. a given model of its knowledge part. This algorithm uses breadth-first search to explore the possible subsets. Indeed, we start with the largest candidate subset, namely S, and the subsets that contain MASes are reduced one element at a time. To know that a subset S' contains at least a MAS, we only need to check the property $K \cup S' \vdash \bigwedge_{\chi \in S \setminus S'} \chi$, which can be performed using a solver for the problem of propositional satisfiability (a SAT solver). In order to reduce the number of explored candidates we use the property that if a subset S' does not contain a MAS or is a MAS, then we can avoid all its subsets. Consider, for instance, the MAS lattice described in Figure 2. Knowing that $\{\chi_1, \chi_2 \chi_3\}$ and $\{\chi_1, \chi_4\}$ are MASes and $\{\chi_2, \chi_3 \chi_4\}$ does not contain any MAS allows avoiding all their subsets. Moreover, to reduce the number of calls to a SAT solver we use the fact that if S' contains a MAS and we have $K \cup (S' \setminus \{\chi\}) \vdash \chi$ then $S' \setminus \{\chi\}$ contains also a MAS for every $\chi \in S'$. This property results from the fact that $K \cup (S' \setminus \{\chi\}) \vdash \bigwedge S$ since $K \cup S' \vdash \bigwedge S$ (S' contains a MAS) and $K \cup (S' \setminus \{\chi\}) \vdash \chi$.

Let us now describe more precisely important parts in Algorithm 1. The first for-loop is used to compute the set S_0 of the formulas that belong to every MAS included in S. It

is worth noting that S_0 is mainly used to reduce the search space (see Line 17 and Line 22). At Line 6, if S_0 contains all the formulas in S, then there is a single MAS which is S; otherwise, we start the approach of breadth-first search described previously. The set E_i contains the candidate subsets that we have to consider at the step i, i.e., those having the size $|S| - (i - 1)$. In the for-loop between Lines 16 and 21, we generate the candidate subsets that we have to explore in the next step (each candidate subset contains at least one MAS). The variable R contains the MASes found at the current step. Further, the variable B contains the greatest subsets that do not contain any MAS and are not included in any MAS.

In order to avoid the MASes found w.r.t. a given model $\mathcal{B}$, we only need to require that the future models satisfy $\bigvee_{\chi \in S} \neg\chi$ with $S = \{\chi \in PA(Q) : \mathcal{B}(\chi) = 1\}$. Indeed, this ensures that the future MASes do not involve the same answer combination.

Let us introduce a notational convention: given an IA-instance $I = \langle K \mid Q \rangle$ and $S \subseteq PA(Q)$, we use $\mathcal{Q}st(S, I)$ to denote the set $\{\psi \in Q : \psi \in S \text{ or } \neg\psi \in S\}$. The set $\mathcal{Q}st(S, I)$ is used to determine the questions related to a given set of answers.

The following proposition states that every MQS corresponds to a set of questions involved in MASes.

Proposition 11. *Let $I = \langle K \mid Q \rangle \in \mathsf{IA}_{\mathsf{Form}}$ and $Q' \subseteq Q$. If $Q' \in \mathsf{MQS}(I)$ then $\exists M \subseteq \mathsf{MAS}(I)$ s.t. $Q' = \bigcup_{S \in M} \mathcal{Q}st(S, I)$.*

In particular, Porposition 11 says that every MQS of an IA-instance I is included in the set $\bigcup_{Q \in \mathsf{MAS}(I)} \mathcal{Q}st(S, I)$.

In order to compute the MQSes from the MASes, we associate to the MASes an instance of the X-minimal model problem such that its solutions corresponds to the MQSes.

Let $\mathcal{B}$ and $\mathcal{B}'$ be two Boolean interpretations over the set of propositional variables V and X a subset of V. Then, $\mathcal{B}$ is said to be smaller than $\mathcal{B}'$ with respect to X, written $\mathcal{B} \preceq_X \mathcal{B}'$, if $\{p \in X : \mathcal{B}(p) = 1\} \subseteq \{p \in X : \mathcal{B}'(p) = 1\}$.

Definition 6 (X-Minimal Model). *Let ϕ be a propositional formula over V, $X \subseteq V$ and $\mathcal{B}$ a model of ϕ. Then, $\mathcal{B}$ is said to be an X-minimal model of ϕ if there is no model $\mathcal{B}'$ s.t. $\mathcal{B}' \preceq_X \mathcal{B}$ and $\mathcal{B} \npreceq_X \mathcal{B}'$.*

Different algorithms have been proposed in the literature for computing minimal models (e.g. (Ben-Eliyahu and Dechter 1996; Avin and Ben-Eliyahu-Zohary 2001; Ben-Eliyahu-Zohary 2005)). Note that the X-minimal model problem has important applications in AI such as in propositional circumscription and in minimal diagnosis (McCarthy 1980; Reiter 1987).

To define our encoding in the X-minimal model problem for an IA-instance $I = \langle K \mid Q \rangle$, we associate to every element in $R \in E = \{\mathcal{Q}st(S, I) : S \in \mathsf{MAS}(I)\}$ a distinct propositional variable x_R. Then, we associate to every $P \in \Pi(I)$ a clause $c_P = \bigvee\{x_R : \exists S \in \mathsf{MAS}(I), S \subseteq PA(R) \cap P\}$. Moreover, we associate to every formula $\psi \in Q$ a distinct propositional y_ψ. We use $\mathcal{E}nc(I)$ to denote the following encoding:

$$\left(\bigwedge_{P \in \Pi(I)} c_p \right) \wedge \bigwedge_{R \in E} \left(x_R \to \bigwedge_{\psi \in R} y_\psi \right)$$

Proposition 12. *Given an IA-instance $I = \langle K \mid Q \rangle$, we have $Q' \in \mathsf{MQS}(I)$ iff there exists a $\{y_\psi : \psi \in Q\}$-minimal model $\mathcal{B}$ of $\mathcal{E}nc(I)$ s.t. $Q' = \{\psi \in Q : \mathcal{B}(y_\psi) = 1\}$.*

Proof. We only consider in this proof the *only if* part, the other being similar.

Part $\Rightarrow$. Let $Q' \in \mathsf{MQS}(I)$. We define the Boolean interpretation $\mathcal{B}$ as follow: $\mathcal{B}(z) = 1$ if $z = y_\psi$ and $\psi \in Q'$, or $z = x_R$ and $R \subseteq Q$; $\mathcal{B}(z) = 1$ otherwise. Clearly, $\mathcal{B}$ satisfies the formula $\bigwedge_{R \in E}(p_r \to \bigwedge_{\psi \in R} y_\psi)$. Moreover, we have for all $T \subseteq PA(Q')$ with $\mathcal{Q}st(T) = Q'$ and $K \cup T \nvdash \bot$, there exists $S \in \mathsf{MAS}(I)$ s.t. $S \subseteq T$. Thus, for all $P \in \Pi(I)$, there exists $S \in \mathsf{MAS}(I)$ s.t. $\mathcal{Q}st(S) \subseteq Q'$ and $S \subseteq P$. Therefore, $\mathcal{B}$ satisfies also $\bigwedge_{P \in \Pi(I)} c_p$, and it is thereby a model of $\mathcal{E}nc(I)$. Furthermore, knowing that Q' is minimal w.r.t. the set inclusion, we obtain that $\mathcal{B}$ is a $\{y_\psi : \psi \in Q\}$-minimal model of $\mathcal{E}nc(I)$. $\square$

Example 4. *We describe in this example our method for computing the set of MQSes from the MASes. Let us consider again the IA-instance $I' = \langle\{(p \wedge q) \leftrightarrow r\} \mid \{p, q, r, \neg p \wedge r\}\rangle$ described in Example 3. We have $\mathsf{MAS}(I') = \{\{p, q\}, \{\neg p, q\}, \{p, \neg q\}, \{\neg p, \neg q\}, \{r\}, \{\neg r, p\}, \{\neg r, q\}\}$ and $\Pi(I') = \{P_1 = \{p, q, r, \neg(\neg p \wedge r)\}, P_2 = \{p, \neg q, \neg r, \neg(\neg p \wedge r)\}, P_3 = \{\neg p, q, \neg r, \neg(\neg p \wedge r)\}, P_4 = \{\neg p, \neg q, \neg r, \neg(\neg p \wedge r)\}\}$. As a consequence, we have:*
$c_{P_1} = x_{\{p,q\}} \vee x_{\{r\}} \vee x_{\{p,r\}} \vee x_{\{q,r\}}$
$c_{P_2} = x_{\{p,q\}} \vee x_{\{p,r\}}$ $c_{P_3} = x_{\{p,q\}} \vee x_{\{q,r\}}$ $c_{P_4} = x_{\{p,q\}}$
Then, we obtain $\mathcal{E}nc(I') = (c_{P_1} \wedge c_{P_2} \wedge c_{P_3} \wedge c_{P_4}) \wedge (x_{\{p,q\}} \to (y_p \wedge y_q)) \wedge (x_r \to y_r) \wedge (x_{\{p,r\}} \to (y_p \wedge y_r)) \wedge (x_{\{q,r\}} \to (y_q \wedge y_r))$. There is a single $\{y_p, y_q, y_r\}$-minimal model $\mathcal{B}$ that assigne 1 to $x_{\{p,q\}}$, y_p and y_q, and 0 to all the other variables. Therefore, $\mathsf{MQS}(I') = \{\{p, q\}\}$ holds.

Conclusion and Perspectives

In this paper, we introduced a framework for measuring information asymmetry, which is based on classical propositonal logic. Indeed, every considered situation of information asymmetry is described by two sets of propositional formulas that represent respectively the pieces of information and the questions. The latter correspond to the possible ignored pieces of information. In this context, we defined the notion of information asymmetry measure through rationality postulates. Moreover, we proposed concepts for considering relationships between questions; in particular, that of minimal question subset (MQS). We also proposed a method for computing all MQSes for a given instance. It is worth noting that we provided different measures to illustrate several addressed aspects.

In our future work, we intend first to improve the proposed framework following two main directions: (1) considering additional properties for information asymmetry measures that capture other interesting aspects; (2) studying

more sophisticated measures than those introduced in this work. We also plan to investigate the computational complexity of computing specific measures.

References

Akerlof, G. A. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* 84(3):488–500.

Armstrong, C. S.; Balakrishnan, K.; and Cohen, D. 2012. Corporate governance and the information environment: Evidence from state antitakeover laws. *Journal of Accounting and Economics* 53(1):185 – 204.

Avin, C., and Ben-Eliyahu-Zohary, R. 2001. Algorithms for Computing X-Minimal Models. In *Logic Programming and Nonmonotonic Reasoning, 6th International Conference, LPNMR 2001, Vienna, Austria, Proceedings*, 322–335.

Ben-Eliyahu, R., and Dechter, R. 1996. On Computing Minimal Models. *Annals of Mathematics and Artificial Intelligence* 18(1):3–27.

Ben-Eliyahu-Zohary, R. 2005. An incremental algorithm for generating all minimal models. *Artificial Intelligence* 169(1):1–22.

Chen, Q.; Goldstein, I.; and Jiang, W. 2006. Price Informativeness and Investment Sensitivity to Stock Price. *The Review of Financial Studies* 20(3):619–650.

Connelly, B. L.; Certo, S. T.; Ireland, R. D.; and Reutzel, C. R. 2011. Signaling Theory: A Review and Assessment. *Journal of Management* 37(1):39–67.

Easley, D.; Hvidkjaer, S.; and O'Hara, M. 2002. Is Information Risk a Determinant of Asset Returns? *The Journal of Finance* 57(5):2185–2221.

Glosten, L. R., and Harris, L. E. 1988. Estimating the components of the bid/ask spread. *Journal of Financial Economics* 21(1):123 – 142.

Hunter, A., and Konieczny, S. 2010. On the measure of conflicts: Shapley Inconsistency Values. *Artificial Intelligence* 174(14):1007–1026.

McCarthy, J. 1980. Circumscription - A Form of Non-Monotonic Reasoning. *Artificial Intelligence* 13(1-2):27–39.

Reiter, R. 1987. A Theory of Diagnosis from First Principles. *Artif. Intell.* 32(1):57–95.

Spence, M. 1973. Job Market Signaling. *The Quarterly Journal of Economics* 87(3):355–374.

Stiglitz, J. E. 1975. The Theory of "Screening," Education, and the Distribution of Income. *The American Economic Review* 65(3):283–300.