Malaysian Journal of Mathematical Sciences 15(S) December: 91–104 (2021)

Malaysian Journal of Mathematical Sciences

Journal homepage: https://einspem.upm.edu.my/journal

The Impossibility of Keeping Secrets

Studer, T.*

Institute of Computer Science, University of Bern, Switzerland

E-mail: thomas.studer@inf.unibe.ch
**Corresponding author*

Received: 29 October 2020
Accepted: 5 July 2021

Abstract

Controlled query evaluation (CQE) is an approach to guarantee data privacy for database and knowledge base systems. CQE-systems feature a censor function that may distort the answer to a query in order to hide sensitive information. We introduce a high-level formalization of controlled query evaluation in the language of modal logic. We then use this to define several desirable properties of CQE-systems. Finally we establish two impossibility theorems, which show that certain combinations of these properties cannot be obtained.

Keywords: no-go theorem; data privac; controlled query evaluation; modal logic.

1 Introduction

Controlled query evaluation (CQE) refers to a data privacy mechanism where the database (or knowledge base) is equipped with a censor function. This censor checks for each query whether the answer to the query would reveal sensitive information to a user. If this is the case, then the censor will distort the answer. Essentially, there are two possibilities how an answer may be distorted:

i) The CQE-system may refuse to answer the query, see [17], or

ii) The CQE-system may give an incorrect answer, i.e. it lies, see [8].

Biskup and Bonatti [6] study censors that combine lying and refusing. The censor based approach has the advantage that the task of maintaining privacy is separated from the task of keeping the data. This gives more flexibility than an integrated approach (like hiding rows in a database) and guarantees than no information is leaked through otherwise unidentified inference channels. Controlled query evaluation has been applied to a variety of data models and control mechansims. The case of incomplete databases is treated in [7] and censors for description logic knowledge bases are dealt with in [20].

Impossibiltiy theorems, also called no-go theorems, are well-known from theoretical physics where they describe particular situations that are not physically possible. Often the term is used for results in quantum mechanics like Bell's theorem [4], the Kochen–Specker theorem [14], or, for a more recent example, the Frauchiger–Renner paradox [11]. [15] provide a modal logic analysis of the latter paradox. Arrow's theorem, see [2], in social choice theory also is an impossibility theorem stating that no voting system can be designed that meets certain given fairness conditions. [16] present a version of independence logic in which Arrow's theorem is derivable.

In the present paper we develop a highly abstract model for dynamic query evaluation systems like CQE using the language of modal logic. We formulate several desirable properties of CQE-systems in our framework and establish two impossibility theorems saying that certain combinations of those properties are impossible. The main contribution of this paper is the presentation of the abstract logical framework as well as the high-level formulation of the no-go theorems. Note that some particular instances of our results have already been know: our first no-go theorem is a generalization of Th. 50 in [20] and our second no-go theorem is a generalization of results in [5].

There are many different notions of privacy available in the literature. For our results, we rely on *provable privacy*, which has been defined in [18] and is applied in the context of description logics in [19]. Note that provable privacy is a rather weak notion of data privacy. Using such a weak definition of privacy makes our impossibility theorems actually stronger since they state that under certain conditions not even this weak form of privacy can be achieved.

Clearly our work is also connected to the issues of lying and deception. Logics dealing with notions of lying are introduced and studied in [1, 10]. In his PhD thesis [12], Icard examines logics for deception and strategic omission.

2 Logical Preliminaries

Let X be a set. The notation $\mathcal{P}(X)$ is used to describe the power set of X. For sets Γ and Δ we use Γ, Δ for $\Gamma \cup \Delta$. Moreover, in such a context we write A for the singleton set $\{A\}$. Hence Γ, A stands for $\Gamma \cup \{A\}$.

Definition 2.1. *A* logic L *is given by*

i) A set of formulas Fml_L*,*

ii) A consequence relation $\vdash_\mathsf{L}$ *for* L *that is a relation between sets of formulas and formulas, i.e.*

$$\vdash_\mathsf{L} \subseteq \mathcal{P}(\mathsf{Fml}_\mathsf{L}) \times \mathsf{Fml}_\mathsf{L}$$

satisfying for all $A, C \in \mathsf{Fml}_\mathsf{L}$ *and* $\Gamma, \Delta \in \mathcal{P}(\mathsf{Fml}_\mathsf{L})$*:*

(a) *Reflexivity:* $\{A\} \vdash_\mathsf{L} A$

(b) *Weakening:* $\Gamma \vdash_\mathsf{L} A \implies \Gamma, \Delta \vdash_\mathsf{L} A$

(c) *Transitivity:* $\Gamma \vdash_\mathsf{L} C$ *and* $\Delta, C \vdash_\mathsf{L} A \implies \Gamma, \Delta \vdash_\mathsf{L} A$.

Transitivity is sometimes called *cut*. The previous definition gives us single conclusion consequence relations, which is sufficient for the purpose of this paper. For other notions of consequence relations see, e.g., [3] and [13].

As usual, we write $\vdash_\mathsf{L} A$ for $\emptyset \vdash_\mathsf{L} A$. A formula A is called a *theorem of* L if $\vdash_\mathsf{L} A$.

We do not specify the logic L any further. The only thing we need is a consequence relation as given above. For instance, L may be classical propositional logic with $\vdash_\mathsf{L}$ being the usual derivation relation (see Section 4) or L may be a description logic with $\vdash_\mathsf{L}$ being its semantic consequence relation, see [20].

Definition 2.2.

i) A logic L *is called* consistent *if there exists a formula* $A \in \mathsf{Fml}_\mathsf{L}$ *such that* $\not\vdash_\mathsf{L} A$.

ii) A set Γ *of* Fml_L*-formulas is called* L-consistent *if there exists a formula* $A \in \mathsf{Fml}_\mathsf{L}$ *such that* $\Gamma \not\vdash_\mathsf{L} A$.

We need a simple modal logic M over L.

Definition 2.3. *The set of formulas* Fml_M *is given inductively by:*

i) If A *is a formula of* Fml_L*, then* $\Box A$ *is a formula of* Fml_M

ii) $\bot$ *is a formula of* Fml_M

iii) If A *and* B *are formulas of* Fml_M*, so is* $A \to B$*, too.*

The remaining classical connectives $\top$, $\wedge$, $\vee$, and $\neg$ is defined as usual. Note that M is not a fully-fledged modal logic. For instance, it does not include nested modalities.

We give semantics to Fml_M-formulas as follows.

Definition 2.4. *An* M*-model* $\mathcal{M}$ *is a set of sets of* Fml_L*-formulas, that is*

$$\mathcal{M} \subseteq \mathcal{P}(\mathsf{Fml}_\mathsf{L}).$$

Definition 2.5. *Let* $\mathcal{M}$ *be an* M*-model. Truth of an* Fml_M*-formula in* $\mathcal{M}$ *is inductively defined by:*

i) $\mathcal{M} \Vdash \Box A$ *iff* $w \vdash_\mathsf{L} A$ *for all* $w \in \mathcal{M}$

ii) $\mathcal{M} \nVdash \bot$

iii) $\mathcal{M} \Vdash A \to B$ *iff* $\mathcal{M} \nVdash A$ *or* $\mathcal{M} \Vdash B$.

We use the following standard definition.

Definition 2.6. *Let* Γ *be a set of* Fml_M*-formulas.*

i) We write $\mathcal{M} \Vdash \Gamma$ *iff* $\mathcal{M} \Vdash A$ *for each* $A \in \Gamma$

ii) Γ *is called* satisfiable *iff there exists an* M*-model* $\mathcal{M}$ *with* $\mathcal{M} \Vdash \Gamma$

iii) Γ entails *a formula* A*, in symbols* $\Gamma \models A$*, iff for each model* $\mathcal{M}$ *we have that*

$$\mathcal{M} \Vdash \Gamma \quad \textit{implies} \quad \mathcal{M} \Vdash A.$$

3 Privacy

Definition 3.1. *A* privacy configuration *is a triple* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *that consists of:*

i) The knowledge base $\mathsf{KB} \subseteq \mathsf{Fml}_\mathsf{L}$*, which is only accessible via the censor*

ii) The set of a priori knowledge $\mathsf{AK} \subseteq \mathsf{Fml}_\mathsf{M}$*, which formalizes general background knowledge known to the attacker and the censor*

iii) The set of secrets $\mathsf{Sec} \subseteq \mathsf{Fml}_\mathsf{L}$*, which should be protected by the censor.*

A privacy configuration $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *satisfies the following conditions:*

i) KB *is* L*-consistent (consistency)*

ii) $\{\mathsf{KB}\} \Vdash \mathsf{AK}$ *(truthful start)*

iii) $\mathsf{AK} \not\models \Box s$ *for each* $s \in \mathsf{Sec}$ *(hidden secrets).*

Note that in the above definition, KB and Sec are sets of Fml_L-formulas while AK is a set of Fml_M-formulas. Thus AK may not only contain domain knowledge but also knowledge about the structure of KB. This is further explained in Section 4.

A *query* to a knowledge base KB is simply a formula of Fml_L.

Given a logic L, a query q can be evaluated over a knowledge base KB. There are two possible answers: t (true) and u (unknown).

Definition 3.2. *The evaluation function* eval *is defined by:*

$$\mathsf{eval}(\mathsf{KB}, q) := \begin{cases} t & \textit{if} \quad \mathsf{KB} \vdash_{\mathsf{L}} q, \\ u & \textit{otherwise}. \end{cases}$$

If the language of the logic L includes negation, then one may also consider an evaluation function that can return the value f (false), i.e. one defines $\mathsf{eval}(\mathsf{KB}, q) := f$ if $\mathsf{KB} \vdash_{\mathsf{L}} \neg q$. However, in the general setting of this paper, we cannot include this case.

A censor has to hide the secrets. In order to achieve this, it can not only answer t and u to a query but also r (refuse to answer). We denote the set of possible answers of a censor by

$$\mathbb{A} := \{t, u, r\}.$$

Let X be a set. Then X^ω denotes the set of infinite sequences of elements of X.

Definition 3.3. *A* censor *is a mapping that assigns an answering function*

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})} : \mathsf{Fml}_{\mathsf{L}}^{\omega} \longrightarrow \mathbb{A}^{\omega},$$

to each privacy configuration $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$. *By abuse of notation, we also call the answering function* $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}$ *a* censor. *A sequence* $q \in \mathsf{Fml}_{\mathsf{L}}^{\omega}$ *is called* query sequence.

Usually, the privacy configuration will be clear from the context. In that case we simply use Cens instead of $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}$.

Given a sequence s, we use s_i to denote its i-th element. That is for a query sequence $q \in \mathsf{Fml}_{\mathsf{L}}^{\omega}$, we use q_i to denote the i-th query and $\mathsf{Cens}(q)_i$ to denote the i-th answer of the censor.

Example 3.1. *Let* $A, B, C \in \mathsf{Fml}_{\mathsf{L}}$. *We define a privacy configuration with* $\mathsf{KB} = \{A, C\}$, $\mathsf{AK} = \emptyset$, *and* $\mathsf{Sec} = \{C\}$. *A censor* Cens *yields an answering function* $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}$, *which applied to a query sequence* $q = (A, B, C, \ldots)$ *yields a sequence of answers, e.g.,*

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q) = t, u, r, \ldots.$$

In this case, $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}$ *gives true answers since* $\mathsf{eval}(\mathsf{KB}, A) = t$ *and* $\mathsf{eval}(\mathsf{KB}, B) = u$ *and it protects the secret be refusing to answer the query* C.

Another option for the answering function would be to answer the third query with u, *i.e., it would lie (instead of refuse to answer) in order to protect the secret.*

A further option would be to always refuse the answer, i.e.

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q) = r, r, r, \ldots.$$

This, of course, would be a trivial (and useless) answering function that would, however, preserve all secrets.

In this paper, we will consider continuous censors only, which are given as follows.

Definition 3.4. *A censor* Cens *is* continuous *iff for each privacy configuration* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *and for all query sequences* $q, q' \in \mathsf{Fml}_{\mathsf{L}}^{\omega}$ *and all* $n \in \omega$ *we have that*

$$q|_n = q'|_n \quad \Longrightarrow \quad \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_n = \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q')|_n,$$

where for an infinite sequence $s = (s_1, s_2, \ldots)$, *we use* $s|_n$ *to denote the initial segment of* s *of length* n, *i.e.* $s|_n = (s_1, \ldots, s_n)$.

Continuity means that the answer of a censor to a query does not depend on future queries, see also Lemma 3.1.

A censor is called truthful if it does not lie.

Definition 3.5. *A censor* Cens *is called* truthful *iff for each privacy configuration* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$*, all query sequences* $q = (q_1, q_2, \ldots)$*, and all sequences*

$$(a_1, a_2, \ldots) = \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q),$$

we have that for all $i \in \omega$*,*

$$a_i = \mathsf{eval}(\mathsf{KB}, q_i) \quad \textit{or} \quad a_i = r.$$

Hence a truthful censor may refuse to answer a query in order to protect a secret but it will not give an incorrect answer.

In the modal logic M over L, we can express what knowledge one can gain from the answers of a censor to a query. This is called the content of the answer.

Definition 3.6. *Given an answer* $a \in \mathbb{A}$ *to a query* $q \in \mathsf{Fml}_\mathsf{L}$*, we define its* content *as follows:*

$$\begin{aligned} \mathsf{cont}(q, t) &:= \Box q, \\ \mathsf{cont}(q, u) &:= \neg\Box q, \\ \mathsf{cont}(q, r) &:= \top. \end{aligned}$$

Assume that we are given a privacy configuration $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *and a censor* Cens*. We define the content of the answers of the censor to a query sequence* $q \in \mathsf{Fml}_\mathsf{L}^\omega$ *up to* $n \in \omega$ *by*

$$\mathsf{cont}(\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), n) := \bigcup_{1 \le i \le n} \{\mathsf{cont}(q_i, a_i)\} \cup \mathsf{AK},$$

where $a = \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)$*. Note that here we have also included the a priori knowledge.*

The following is a trivial observation showing the role of continuity.

Lemma 3.1. *Let* Cens *be a continuous censor. The content function is monotone in the second argument: for* $m \le n$ *we have*

$$\mathsf{cont}(\mathsf{Cens}(q), m) \subseteq \mathsf{cont}(\mathsf{Cens}(q), n).$$

We call a censor credible if it does not return contradicting answers.

Definition 3.7. *A censor* Cens *is called* credible *iff for each privacy configuration* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *and for every query sequence* q *and every* $n \in \omega$*, the set* $\mathsf{cont}(\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), n)$ *is satisfiable.*

Definition 3.8. *The* full content *of a knowledge base* KB *is given by*

$$\mathsf{full}(\mathsf{KB}) := \bigcup_{A \in \mathsf{Fml}_\mathsf{L}} \mathsf{cont}(A, \mathsf{eval}(\mathsf{KB}, A)).$$

Lemma 3.2. *For any knowledge base* KB*, we have that*

$$\{\mathsf{KB}\} \Vdash \mathsf{full}(\mathsf{KB}).$$

Proof. Let A be an $\mathsf{Fml_L}$-formula. We dinstinguish:

i) $\mathsf{KB} \vdash_{\mathsf{L}} A$. Then $\Box A \in \mathsf{full}(\mathsf{KB})$ and further $\{\mathsf{KB}\} \Vdash \Box A$

ii) $\mathsf{KB} \not\vdash_{\mathsf{L}} A$. Then $\neg\Box A \in \mathsf{full}(\mathsf{KB})$ and further $\{\mathsf{KB}\} \Vdash \neg\Box A$.

□

Lemma 3.3. *We let* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *be a privacy configuration. Further we let* $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}$ *be a truthful censor. For every query sequence q and $n \in \omega$, we have that*

$$\mathsf{cont}(\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), n) \subseteq \mathsf{full}(\mathsf{KB}) \cup \{\top\} \cup \mathsf{AK}.$$

Proof. By induction on n. The base case $n = 0$ is trivial since

$$\mathsf{cont}(\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), 0) = \mathsf{AK}.$$

Induction step. Since Cens is truthful, we have

$$a_{n+1} \in \{r, \mathsf{eval}(\mathsf{KB}, q_{n+1})\}.$$

We distinguish:

i) $a_{n+1} = r$. Then $\mathsf{cont}(\mathsf{Cens}(q), n+1) = \mathsf{cont}(\mathsf{Cens}(q), n) \cup \{\top\}$ and the claim follows immediately from the induction hypothesis.

ii) $a_{n+1} = \mathsf{eval}(\mathsf{KB}, q_{n+1})$. Then,

$$\begin{aligned}\mathsf{cont}(\mathsf{Cens}(q), n+1) = \mathsf{cont}(\mathsf{Cens}(q), n) \cup \\ \mathsf{cont}(q_{n+1}, \mathsf{eval}(\mathsf{KB}, q_{n+1})).\end{aligned}$$

The claim follows from the induction hypothesis and

$$\mathsf{cont}(q_{n+1}, \mathsf{eval}(\mathsf{KB}, q_{n+1})) \in \mathsf{full}(\mathsf{KB}),$$

which holds by Definition 3.8.

□

The following corollary is a generalization of Cor. 30 in [20].

Corollary 3.1. *Every truthful censor is credible.*

Proof. Let $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ be a privacy configuration and Cens be a truthful censor for it. By Definition 3.1, we have $\{\mathsf{KB}\} \Vdash \mathsf{AK}$. Thus by the two previous lemmas, we find that for each $n \in \omega$,

$$\{\mathsf{KB}\} \Vdash \mathsf{full}(\mathsf{KB}) \cup \{\top\} \cup \mathsf{AK},$$

and

$$\mathsf{full}(\mathsf{KB}) \cup \{\top\} \cup \mathsf{AK} \supseteq \mathsf{cont}(\mathsf{Cens}(q), n),$$

that means $\mathsf{cont}(\mathsf{Cens}(q), n)$ is satisfiable for each $n \in \omega$ and thus Cens is credible. □

There are several properties that a 'good' censor should fulfil. We call a censor effective if it protects all secrets.

Definition 3.9. *A censor* Cens *is called* effective *iff for each privacy configuration* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *and for every query sequence* $q \in \mathsf{Fml}_{\mathsf{L}}^{\omega}$ *and every* $n \in \omega$*, we have*

$$\mathsf{cont}(\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), n) \not\models \Box s \quad \textit{for each } s \in \mathsf{Sec}.$$

A 'good' censor should only distort an answer to a query when it is absolutely necessary, i.e. when giving the correct answer would leak a secret. We call such a censor minimally invasive.

Definition 3.10. *Let* Cens *be an effective and credible censor. This censor is called* minimally invasive *iff for each privacy configuration* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *and for each query sequence* $q \in \mathsf{Fml}_{\mathsf{L}}^{\omega}$*, we have that whenever*

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)_i \neq \mathsf{eval}(\mathsf{KB}, q_i),$$

replacing

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)_i \quad \textit{with} \quad \mathsf{eval}(\mathsf{KB}, q_i),$$

would lead to a violation of effectiveness or credibility, that is for any censor $\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}$ *such that*

$$\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_{i-1} = \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_{i-1},$$

and

$$\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)_i = \mathsf{eval}(\mathsf{KB}, q_i),$$

we have that for some n

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), n) \models \Box s \quad \textit{for some } s \in \mathsf{Sec},$$

or

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), n) \textit{ is not satisfiable.}$$

It is a trivial observation that a truthful, effective and minimally invasive censor has to answer the same query always in the same way.

Lemma 3.4. *Let* Cens *be a truthful, effective and minimally invasive censor. Further let* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *be a privacy configuration and* q *be a query sequence with* $q_i = q_j$ *for some* i, j*. Then*

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)_i = \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)_j.$$

Consider a truthful, effective, continuous and minimally invasive censor and a given query sequence. If the censor lies to answer some query, then giving the correct answer would immediately reveal a secret.

Lemma 3.5. *Let* Cens *be a truthful, effective, continuous and minimally invasive censor. Further let* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *be a privacy configuration and* q *be a query sequence. Let* i *be the least natural number such that*

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)_i \neq \mathsf{eval}(\mathsf{KB}, q_i).$$

Let $\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}$ *be such that*

$$\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_{i-1} = \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_{i-1},$$

and

$$\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)_i = \mathsf{eval}(\mathsf{KB}, q_i).$$

Then, it holds that

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), i) \models \Box s \quad \textit{for some } s \in \mathsf{Sec}.$$

Proof. Consider the query sequence q' given by $q'_j := q_j$ for $j < i$ and $q'_j := q_i$ for $j \geq i$, i.e. q' has the form $(q_1, q_2, \ldots, q_{i-1}, q_i, q_i, q_i, \ldots)$. In particular, we have $q|_i = q'|_i$. Thus by continuity of the censor we find

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_i = \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q')|_i.$$

Thus, $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q')_i \neq \mathsf{eval}(\mathsf{KB}, q_i)$. By the definition of minimally invasive we find that for some n

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q'), n) \models \Box s \quad \text{for some } s \in \mathsf{Sec}, \tag{1}$$

or

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q'), n) \text{ is not satisfiable.} \tag{2}$$

Since the censor is truthful and by Corollary 3.1, we find that (2) is not possible. Thus (1) holds for some n.

By the definition of q' and the previous lemma we find

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q'), n) = \mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q'), i),$$

if $i \leq n$. Thus, in case $i \leq n$, (1) implies

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), i) \models \Box s \quad \text{for some } s \in \mathsf{Sec}. \tag{3}$$

In case $i > n$, we find by Lemma 3.1 that

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), n) \subseteq \mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q), i).$$

Thus again (1) implies (3), which finishes the proof. □

Next, we define the notion of a repudiating censor, which garantees that there is always a knowledge base in which no secret holds and which, given as input to the answering function, produces the same results as the actual knowledge base. Hence this definition provides a version of plausible deniability for all secrets.

Definition 3.11. *A censor* Cens *is called* repudiating *iff for each privacy configuration* $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ *and each query sequence* q*, there are knowledge bases* KB_i $(i \in \omega)$ *such that*

i) $(\mathsf{KB}_i, \mathsf{AK}, \mathsf{Sec})$ *is a privacy configuration for each* $i \in \omega$

ii) $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_n = \mathsf{Cens}_{(\mathsf{KB}_n,\mathsf{AK},\mathsf{Sec})}|_n$*, for each* $n \in \omega$

iii) $\mathsf{KB}_i \not\models_{\mathsf{L}} s$ *for each* $s \in \mathsf{Sec}$ *and each* $i \in \omega$.

Now, we can establish our first no-go theorem, which is a generalization of Th. 50 in [20].

Theorem 3.1 (First No-Go Theorem)**.** *A continuous and truthful censor satisfies at most two of the properties effectiveness, minimal invasion, and repudiation.*

Proof. Let the censor Cens be continuous, truthful, effective, and minimally invasive. We show that Cens cannot be repudiating. We let S be an $\mathsf{Fml}_{\mathsf{L}}$-formula and consider the privacy configuration $(\mathsf{KB}, \mathsf{AK}, \mathsf{Sec})$ given by

$$\mathsf{KB} := \{S\} \qquad \mathsf{AK} := \emptyset \qquad \mathsf{Sec} := \{S\},$$

and the query sequence $q := (S, S, \ldots)$. We set

$$a := \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q).$$

Obviously, we have $a = (r, r, \ldots)$ since otherwise Cens would either be lying (i.e. not be truthful) or revealing a secret (i.e. not be effective).

Now, asssume that Cens is repudiating. Then there exists a knowledge base KB_1 such that

i) $(\mathsf{KB}_1, \mathsf{AK}, \mathsf{Sec})$ is a privacy configuration

ii) $\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_1 = \mathsf{Cens}_{(\mathsf{KB}_1,\mathsf{AK},\mathsf{Sec})}(q)|_1$

iii) $\mathsf{KB}_1 \not\vdash_{\mathsf{L}} S$.

Let $(a'_1) := \mathsf{Cens}_{(\mathsf{KB}_1,\mathsf{AK},\mathsf{Sec})}(q)|_1$. Because of $\mathsf{KB}_1 \not\vdash_{\mathsf{L}} S$ and Cens being truthful, we find that $a'_1 = u$ or $a'_1 = r$.

Suppose towards a contradiction that

$$a'_1 = r. \tag{4}$$

Now, let Cens′ be a censor as in Lemma 3.5, i.e. such that

$$\mathsf{Cens}'_{(\mathsf{KB}_1,\mathsf{AK},\mathsf{Sec})}(q)_1 = u = \mathsf{eval}(\mathsf{KB}_1, S). \tag{5}$$

By Lemma 3.5 we get

$$\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB}_1,\mathsf{AK},\mathsf{Sec})}(q), 1) \models \Box S. \tag{6}$$

However, by (5) we also have $\mathsf{cont}(\mathsf{Cens}'_{(\mathsf{KB}_1,\mathsf{AK},\mathsf{Sec})}(q), 1) = \{\neg\Box S\}$, which contradicts (6).

Hence, (4) is not possible and thus we have $a'_1 = u$. This, however, contradicts

$$\mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)|_1 = \mathsf{Cens}_{(\mathsf{KB}_1,\mathsf{AK},\mathsf{Sec})}(q)|_1.$$

We conclude that Cens cannot be repudiating. □

4 Non-Refusing Censors

In this section we study censors that do not refuse to answer a query.

Definition 4.1. *A censor is* non-refusing *if it never assigns the answer r to a query.*

Of course, a non-refusing censor has to lie in order to keep the secrets. That means if a censors of this kind shall be effective, then it cannot be truthful.

Even if we consider lying censors, we work with the assumption that

Assumption 1. *An attacker believes every answer of the censor.*

Otherwise, we are in a situation where an attacker cannot believe any answer because the attacker does not know which answers are correct and which are wrong, which means that any answer could be a lie. In that case, querying a knowledge base would not make any sense at all.[1]

Because of the assumption (1), we can use our notions of effectiveness (Definition 3.9) and credibility (Definition 3.7) also in the context of lying censors: an attacker should not believe any secret and the beliefs should be satisfiable.

Theorem 3.1 about truthful censors did not make any assumptions on the underlying logic L. The next theorem about non-refusing censors is less general as it is based on classical logic. We will use $a, b, c, \ldots$ for atomic propositions and $A, B, C, \ldots$ for arbitrary formulas.

Moreover, we assume that the knowledge base KB only contains atomic facts (we say KB is *atomic*). That is if $F \in \mathsf{KB}$, then F is either of the form p or of the form $\neg p$ where p is an atomic proposition. Hence we find that if $\mathsf{KB} \vdash_{\mathsf{L}} a \to b$ for two distinct atomic propositions a and b, then $\mathsf{KB} \vdash_{\mathsf{L}} \neg a$ or $\mathsf{KB} \vdash_{\mathsf{L}} b$. We can formalize this using the set of a priori knowledge by letting

$$\Box(a \to b) \to (\Box\neg a \lor \Box b) \in \mathsf{AK}.$$

Now, we can establish our second no-go theorem, which is a generalization of results by [5].

Theorem 4.1 (Second No-Go Theorem)**.** *Let* L *be based on classical logic. A continuous and non-refusing censor cannot be at the same time effective and minimally invasive.*

Proof. Let the censor Cens be continuous, non-refusing, and minimally invasive. We show that Cens cannot be effective. Let L be classical propositional logic. We consider the knowledge base

$$\mathsf{KB} := \{a, b\},$$

where both a and b shall be kept secret, i.e.

$$\mathsf{Sec} := \{a, b\}.$$

Further we assume that it is a priori knowledge that KB is atomic. Thus, in particular,

$$\Box(c \to a) \to (\Box\neg c \lor \Box a) \in \mathsf{AK},$$
$$\Box(\neg c \to b) \to (\Box c \lor \Box b) \in \mathsf{AK}.$$

We consider the query sequence $q := (c \to a, \neg c \to b, c, \ldots)$ and set $a := \mathsf{Cens}_{(\mathsf{KB},\mathsf{AK},\mathsf{Sec})}(q)$.

We find $\mathsf{Cens}(c \to a) = t$ since Cens is minimally invasive and KB might contain $\neg c$. Further, we find $\mathsf{Cens}(\neg c \to b) = t$ since Cens is minimally invasive and KB might contain c.

[1] This is, of course, not completely true. It is possible to distort knowledge bases in such a way that privacy is preserved but statistical inferences are still informative, see, e.g. [9].

Note that after issuing the first two queries of the sequence q, an attacker knows that a or b must be entailed by KB. But since the attacker does not know which one is the case, no secret is leaked. Formally, we have

$$\mathsf{cont}(\mathsf{Cens}(q), 2) \vdash_{\mathsf{M}} \Box(c \to a), \tag{7}$$

and

$$\mathsf{cont}(\mathsf{Cens}(q), 2) \vdash_{\mathsf{M}} \Box(\neg c \to b). \tag{8}$$

By basic modal logic, (7) and (8) yield

$$\mathsf{cont}(\mathsf{Cens}(q), 2) \vdash_{\mathsf{M}} \Box c \to \Box a, \tag{9}$$

and

$$\mathsf{cont}(\mathsf{Cens}(q), 2) \vdash_{\mathsf{M}} \Box\neg c \to \Box b, \tag{10}$$

respectively. Using the a priori knowledge AK, we obtain from (7) and (8)

$$\mathsf{cont}(\mathsf{Cens}(q), 2) \vdash_{\mathsf{M}} \Box\neg c \lor \Box a, \tag{11}$$

and

$$\mathsf{cont}(\mathsf{Cens}(q), 2) \vdash_{\mathsf{M}} \Box c \lor \Box b. \tag{12}$$

Because of $\Box c \lor \neg\Box c$, we get by (9) and (12) that

$$\mathsf{cont}(\mathsf{Cens}(q), 2) \vdash_{\mathsf{M}} \Box a \lor \Box b.$$

Thus, at this stage, it is known that a secret holds, but an attacker does not know which one and hence privacy is still preserved.

Now comes the third query, which is c. There are two possibilities for a non-refusing censor to choose from:

i) $(a)_3 = u$ (which is true). We find $\mathsf{cont}(\mathsf{Cens}(q), 3) \vdash_{\mathsf{M}} \neg\Box c$. By (12) we get

$$\mathsf{cont}(\mathsf{Cens}(q), 3) \vdash_{\mathsf{M}} \Box b,$$

and a secret is leaked.

ii) $(a)_3 = t$ (which is a lie). We find $\mathsf{cont}(\mathsf{Cens}(q), 3) \vdash_{\mathsf{M}} \Box c$. By (9) we get $\mathsf{cont}(\mathsf{Cens}(q), 3) \vdash_{\mathsf{M}} \Box a$ and a secret is leaked.

In both cases, a secret is leaked. Thus the censor cannot be effective. □

To avoid this problem, a censor must not only protect the single elements of Sec but also their disjunction, see [5]. For the privacy configuration of the previous proof that means Cens must also protect $a \lor b$. Then, already the second query, $\neg c \to b$ would be answered with u because the answer t, as shown above, reveals $a \lor b$.

Note that protecting the disjunction of all secrets is not as simple as it sounds. Consider, for instance, a hospital information system that should protect the disease a patient is diagnosed with. In this case, protecting the disjunction of all secrets means protecting the information that the patient has some disease. This, however, is not feasible as it is general background knowledge that everybody who is a patient in a hospital has some disease. Worse than that, sometimes the disjunction of all secrets may even be a logical tautology, which cannot be protected.

5 Conclusion

In this paper, we have established two impossibility theorems for data privacy using tools from modal logic. We are confident that logical methods will play an important role for finding new impossibility theorems or for better understanding already known ones, see, e.g., the logical analyses carried out by [15] and by [16].

Another line of future research relates to the fact that refusing to answer a query can give away the information that there exists a secret that could be infered from some other answer. Similar phenomena may occur in multi-agent systems when one of the agents refuses to communicate. For example, imagine the situation of an oral exam where the examiner asks a question and the student keeps silent. In this case the examiner learns that the student does not know the answer to the question for otherwise the student would have answered.

It is also possible that refusing an answer can lead to knowing that someone else knows a certain fact. Consider the following scenario. A father enters a room where his daughter is playing and he notices that one of the toys is in pieces. So he asks who has broken the toy. The daughter does not want to betray her brother (who actually broke it) and she also does not want to lie. Therefore, she refuses to answer her father's question. Of course, then the father knows that his daughter knows who broke the toy for otherwise the daughter could have said that she does not know.

We believe that it is worthwhile to study the above situations using general communication protocols that include the possibility of refusing an answer and to investigate the implications of refusing in terms of higher-order knowledge.

Acknowledgement We acknowledge the support the Swiss National Science Foundation grant 200020_184625. We are also grateful the reviewers of an earlier version of this paper. Their comments helped to substantially improve this paper.

Conflicts of Interest The authors declare no conflict of interest.

References

[1] T. Ågotnes, H. van Ditmarsch & Y. Wang (2018). True lies. *Synthese*, *195*(10), 4581–4615.

[2] K. J. Arrow (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, *58*(4), 328–346.

[3] A. Avron (1991). Simple consequence relations. *Information and Computation*, *92*(1), 105–139.

[4] J. S. Bell (1964). On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, *1*, 195–200.

[5] J. Biskup (2000). For unknown secrecies refusal is better than lying. *Data & Knowledge Engineering*, *33*(1), 1–23.

[6] J. Biskup & P. A. Bonatti (2004). Controlled query evaluation for known policies by combining lying and refusal. *Annals of Mathematics and Artificial Intelligence*, *40*(1), 37–62.

[7] J. Biskup & T. Weibert (2008). Keeping secrets in incomplete databases. *International Journal of Information Security*, *7*(3), 199–217.

[8] P. A. Bonatti, S. Kraus & V. S. Subrahmanian (1995). Foundations of secure deductive databases. *IEEE Transactions on Knowledge and Data Engineering*, *7*(3), 406–422.

[9] F. P. Calmon & N. Fawaz (2012). Privacy against statistical inference. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1401–1408. IEEE, Monticello, IL.

[10] H. V. Ditmarsch (2014). Dynamics of lying. *Synthese*, *191*(5), 745–777.

[11] D. Frauchiger & R. Renner (2018). Quantum theory cannot consistently describe the use of itself. *Nature Communications*, *9*, Article number : 3711. https://doi.org.10.3929/ethz-b-000293141.

[12] B. Icard (2019). *Lying, Deception and Strategic Omission : Definition et Evaluation*. PhD thesis, Université Paris Sciences et Lettres, Paris.

[13] R. Iemhoff (2016). Consequence relations and admissible rules. *Journal of Philosophical Logic*, *45*(3), 327–348.

[14] S. Kochen & E. P. Specker (1967). The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, *17*, 59–87.

[15] N. Nurgalieva & L. Rio (2019). Inadequacy of modal logic in quantum settings. In *15th International Conference on Quantum Physics and Logic -QPL 2018*, pp. 267–297. Open Publishing Association, Halifax.

[16] E. Pacuit & F. Yang (2016). Dependence and independence in social choice: Arrow's theorem. In *Dependence Logic*, pp. 235–260. Springer International Publishing, Switzerland.

[17] G. L. Sicherman, W. De Jonge & R. P. Van de Riet (1983). Answering queries without revealing secrets. *ACM Transactions on Database Systems*, *8*(1), 41–59.

[18] K. Stoffel & T. Studer (2005). Provable data privacy. In *Database and Expert Systems Applications*, pp. 324–332. Springer, Berlin, Heidelberg.

[19] P. Stouppa & T. Studer (2009). Data privacy for $\mathcal{ALC}$ knowledge bases. In S. Artemov & A. Nerode (Eds.), *Logical Foundations of Computer Science*, pp. 409–421. Springer, Berlin, Heidelberg.

[20] T. Studer & J. Werner (2014). Censors for boolean description logic. *Transactions on Data Privacy*, *7*, 223–252.