Confidentiality-Preserving Data Publishing for Credulous Users by Extended Abduction

Lena Wiese
joint work with Katsumi Inoue (NII)
and Chiaki Sakama (Wakayama University)

Inoue Laboratory
National Institute of Informatics
Tokyo

funded by DAAD Deutscher Akademischer Austausch Dienst German Academic Exchange Service

INAP, September 28–30, 2011

Outline

1. Introduction
   - Confidentiality-Preservation
   - Related Work
2. Our Contribution
3. Background
4. Our Approach
5. Conclusion

Confidentiality-Preservation

- Major security goal:
  - confidentiality of data
  - also called privacy, secrecy
- Methods:
  - access control (denial, refusal)
  - k-anonymity (grouping, generalization)
  - inference control (perturbation, noise addition, cover stories, lying, weakening)
  - data fragmentation (breaking sensitive associations)
  - ...

Related Work

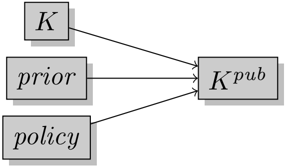
- Already Bonatti et al (1995) introduce incorrect or refused database answers to achieve confidentiality
- Other logic-based mechanisms to ensure data confidentiality:
  - Cuenca Grau et al (2008), Stouppa et al (2009), Toland et al (2010), Biskup (2010), Wiese (2010)
  - all these works do not consider extended disjunctive logic programs (EDPs) with “negation as failure” not and disjunctions in rule heads
- Sakama (2010) surveys several types of dishonesties in multi-agent communication with the help of EDPs

Outline

1. Introduction
2. Our Contribution
   - Application
   - Transformations
3. Background
4. Our Approach
5. Conclusion

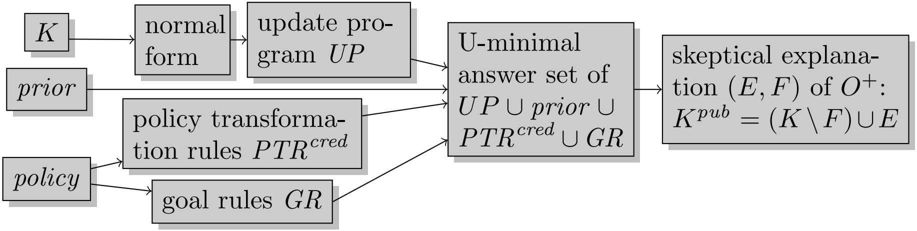
Application

- publish an EDP knowledge base
- user queries knowledge base with credulous reasoning
- preserve confidentiality of elements of a confidentiality policy
- consider invariable background ("a priori") knowledge of such a user
- Aim: compute a secure "view" of the knowledge base such that no confidential information can be inferred by a user based on his knowledge



Transformations

- Use extended abduction:
  - compute skeptical explanation (E, F) for new positive observation O^+
- Can be solved with answer set programming:
  - compute U-minimal answer sets of update programs



Outline

Extended Disjunctive Logic Programs

- literal L: first-order atom or atom preceded by classical negation "$\neg$"
- NAF-literal: $notL$
- literals L_i, disjunction ";", conjunction ",", negation as failure "not", and material implication "$\leftarrow$"
- knowledge base K is an *extended disjunctive logic program* (EDP)
  - set of formulas called *rules* of the form $(n \geq m \geq l \geq 0)$:

$$R = \underbrace{L_1; \ldots; L_l}_{head(R)} \leftarrow \underbrace{L_{l+1}, \ldots, L_m, notL_{m+1}, \ldots, notL_n}_{body(R)}$$

- no function symbols
  - each rule with variables represents a finite set of ground rules
  - elements of Herbrand universe of K substituted in for variables

Extended Disjunctive Logic Programs

Example (medical knowledge base)

$\mathit{Ill}(x, y)$: patient x is ill with disease y
$\mathit{Treat}(x, y)$: x is treated with medicine y
Assume: if one treatment (Medi1) is recorded and another one (Medi2) is not recorded, patient is ill with Aids or Flu

$$
\begin{aligned}
K \quad = \quad \{ & \mathit{Ill}(x, \mathrm{Aids}); \mathit{Ill}(x, \mathrm{Flu}) \leftarrow \mathit{Treat}(x, \mathrm{Medi1}), \mathit{not}\, \mathit{Treat}(x, \mathrm{Medi2}) \ , \\
& \mathit{Ill}(\mathrm{Mary}, \mathrm{Aids}) \ , \\
& \mathit{Treat}(\mathrm{Pete}, \mathrm{Medi1}) \}
\end{aligned}
$$

Answer Set Semantics (Gelfond/Lifschitz 1991)

- answer set S of NAF-free K: subset-minimal set of ground literals satisfying every rule from ground instantiation of K
- if contradiction (inconsistency): all literals $S = \mathscr{L}_K$
- S satisfies ground literal L: $L \in S$
- S satisfies conjunction: satisfies every conjunct
- S satisfies disjunction: satisfies at least one disjunct
- S satisfies ground rule: if body literals in S ($\{L_{l+1}, \ldots, L_m\} \subseteq S$) then at least one head literal L_i is in S $(1 \leq i \leq l)$
- for NAF-literals: use NAF-free reduct K^S

Example

K has two consistent answer sets:
$S_1 = \{\mathit{Ill}(\text{Mary}, \text{Aids}), \mathit{Treat}(\text{Pete}, \text{Medi1}), \mathit{Ill}(\text{Pete}, \text{Aids})\}$
$S_2 = \{\mathit{Ill}(\text{Mary}, \text{Aids}), \mathit{Treat}(\text{Pete}, \text{Medi1}), \mathit{Ill}(\text{Pete}, \text{Flu})\}$

Abduction

- Traditional abduction finds (positive) explanation E for (positive) observation O: $K \cup E \models O$
  - every answer set of K and explanation E together satisfy observation O
- Explanation restricted by specifying a designated set $\mathcal{A}$ of *abducibles*
  - syntactical restrictions on the explanation E: $E \subseteq \mathcal{A} \setminus K$
- Inoue/Sakama, 1995 and 2003 extend this with "negative observations", "negative explanations" F and "anti-explanations"
  - syntactical restrictions for negative explanation $F \subseteq K \cap \mathcal{A}$
- If $\mathcal{A}$ contains a formula with variables, it is meant as a shorthand for all ground instantiations of the formula

Extended Abduction (Inoue/Sakama, 1995 and 2003)

Find (anti-)explanations regarding EDP K
(only *skeptical* (anti-)explanations are needed here):

- given a *positive* observation O, find a pair (E, F) where E is a positive explanation and F is a negative explanation such that
  1. **[skeptical explanation]** O is satisfied in *every* answer set of $(K \setminus F) \cup E$; that is, $(K \setminus F) \cup E \models O$
  2. **[consistency]** $(K \setminus F) \cup E$ is consistent
  3. **[abducibility]** $E \subseteq \mathcal{A} \setminus K$ and $F \subseteq \mathcal{A} \cap K$
- given a *negative* observation O, find a pair (E, F) where E is a positive anti-explanation and F is a negative anti-explanation such that
  1. **[skeptical anti-explanation]** there is *no* answer set of $(K \setminus F) \cup E$ in which O is satisfied
  2. **[consistency]** $(K \setminus F) \cup E$ is consistent
  3. **[abducibility]** $E \subseteq \mathcal{A} \setminus K$ and $F \subseteq \mathcal{A} \cap K$

Normal form of EDPs

For example, rename rules in abducibles $\mathcal{A}$

Example

We transform the example knowledge base K into its normal form based on a set of abducibles that is identical to K: that is $\mathcal{A} = K$
We transform $\langle K, \mathcal{A}\rangle$ into its normal form $\langle K^n, \mathcal{A}^n\rangle$ as follows where we write $n(R)$ for the naming atom of the only rule in $\mathcal{A}$:

$$\begin{aligned} K^n &= \{\mathit{Ill}(\mathrm{Mary}, \mathrm{Aids}), \mathit{Treat}(\mathrm{Pete}, \mathrm{Medi1}), \quad n(R), \\ &\mathit{Ill}(x, \mathrm{Aids}); \mathit{Ill}(x, \mathrm{Flu}) \leftarrow \mathit{Treat}(x, \mathrm{Medi1}), \mathit{not}\, \mathit{Treat}(x, \mathrm{Medi2}), n(R)\} \\ \mathcal{A}^n &= \{\mathit{Ill}(\mathrm{Mary}, \mathrm{Aids}), \mathit{Treat}(\mathrm{Pete}, \mathrm{Medi1}), \quad n(R)\ \} \end{aligned}$$

Update Programs

- Minimal (anti-)explanations can be computed with *update programs* (UPs) (Sakama et al, 2003)
- Update rules
  1. **[Abducible rules]** The rules for abducible literals state that an abducible is either true in K or not. For each $L \in \mathcal{A}$, a new atom $\bar{L}$ is introduced that has the same variables as L
     $abd(L) := \{L \leftarrow not\bar{L}\ ,\ \bar{L} \leftarrow notL\}$
  2. **[Insertion rules]** Abducible literals not contained in K might be inserted into K and hence might occur in the set E of the explanation (E, F). For each $L \in \mathcal{A} \setminus K$, a new atom $+L$ is introduced
     $+L \leftarrow L.$
  3. **[Deletion rules]** Abducible literals contained in K might be deleted from K and hence might occur in the set F of the explanation (E, F). For each $L \in \mathcal{A} \cap K$, a new atom $-L$ is introduced
     $-L \leftarrow notL.$

Update Programs

The **update program** is then defined by replacing abducible literals in K with the update rules; that is, $UP = (K \setminus \mathcal{A}) \cup UR$.

Example

From $\langle K^n, \mathcal{A}^n \rangle$ we obtain $UP =$

$$
\begin{aligned}
\{ \quad & abd(\mathit{Ill}(\text{Mary}, \text{Aids})), \quad abd(\mathit{Treat}(\text{Pete}, \text{Medi1})), \quad abd(n(R)), \\
& -\mathit{Ill}(\text{Mary}, \text{Aids}) \leftarrow not\, \mathit{Ill}(\text{Mary}, \text{Aids}), \\
& -\mathit{Treat}(\text{Pete}, \text{Medi1}) \leftarrow not\, \mathit{Treat}(\text{Pete}, \text{Medi1}), \\
& -n(R) \leftarrow not\, n(R), \\
& \mathit{Ill}(x, \text{Aids}); \mathit{Ill}(x, \text{Flu}) \leftarrow \mathit{Treat}(x, \text{Medi1}), not\, \mathit{Treat}(x, \text{Medi2}), n(R) \}
\end{aligned}
$$

Update minimality

- The set of atoms $+L$ is the set $\mathcal{UA}^+$ of positive update atoms
- The set of atoms $-L$ is the set $\mathcal{UA}^-$ of negative update atoms
- The set of **update atoms** is $\mathcal{UA} = \mathcal{UA}^+ \cup \mathcal{UA}^-$
- From all answer sets of an update program UP we can identify those that are **update minimal** (U-minimal)
  - they contain less update atoms than others

Definition (Update minimality)

S is U-minimal iff there is no answer set T such that $T \cap \mathcal{UA} \subset S \cap \mathcal{UA}$

Outline

Credulous Query Response Semantics

- Credulous query response semantics: a ground formula Q is $true$ in K, if Q is satisfied in *some* answer set of K
- Non-ground Q: set of satisfied ground instantiations

Definition (Credulous query response semantics)

Let U be the Herbrand universe of knowledge base K. For $Q(X)$ with a vector X of free variables, the *credulous query responses* of $Q(X)$ in K are

$$cred(K, Q(X)) = \{Q(A) \mid A \text{ is a vector of elements } a \in U \text{ and there is an answer set of } K \text{ that satisfies } Q(A)\}$$

In particular, for a ground formula Q,

$$cred(K, Q) = \begin{cases} \{Q\} & \text{if } K \text{ has an answer set that satisfies } Q \\ \emptyset & \text{otherwise} \end{cases}$$

Credulous Query Response Semantics

Example (medical knowledge base)

$$\begin{aligned} K \quad = \quad & \{\mathit{Ill}(x, \text{Aids}); \mathit{Ill}(x, \text{Flu}) \leftarrow \mathit{Treat}(x, \text{Medi1}), \mathit{not}\, \mathit{Treat}(x, \text{Medi2}) \ , \\ & \mathit{Ill}(\text{Mary}, \text{Aids}) \ , \\ & \mathit{Treat}(\text{Pete}, \text{Medi1})\} \end{aligned}$$

Ask for all diseases of Pete: $Q(y) = \mathit{Ill}(\text{Pete}, y)$

$$\mathit{cred}(K, Q(y)) \quad = \{\mathit{Ill}(\text{Pete}, \text{Flu}), \mathit{Ill}(\text{Pete}, \text{Aids})\}$$

A priori knowledge

- Set of rules as *invariant* a priori knowledge $prior$
- Additional facts that the user assumes to hold in K, or some rules that the user can apply to data in K to deduce new information.

Example

A user querying K^{pub} might know that a person suffering from flu is not able to work. Hence $prior = \{\neg \mathit{AbleToWork}(x) \leftarrow \mathit{Ill}(x, \mathsf{Flu})\}$.

- We assume that $K \cup prior$ is consistent.

Confidentiality Policy

- Set $policy$ of conjunctions of (NAF-)literals
- Avoid that published knowledge base contains confidential information
- Prevent user from deducing confidential information with the help of his a priori knowledge ("inference problem")

Example

If we wish to declare the disease aids as confidential for any patient x we can do this with

$$policy = \{\mathit{Ill}(x, \text{Aids})\}$$

If we wish to also declare a lack of work ability as confidential, we can add this to the confidentiality policy:

$$policy' = \{\mathit{Ill}(x, \text{Aids})\ ,\ \neg \mathit{AbleToWork}(x)\}$$

Confidentiality-Preservation for Credulous Users

Definition (Confidentiality-preservation for credulous user)

A knowledge base K^{pub} *preserves confidentiality* of a given confidentiality policy under the credulous query response semantics and with respect to a given a priori knowledge $prior$, if for every conjunction $C(X)$ in the policy, the credulous query responses of $C(X)$ in $K^{pub} \cup prior$ are empty:

$$cred(K^{pub} \cup prior, C(X)) = \emptyset.$$

- Subset-minimal change: K^{pub} differs from K only subset-minimally

Definition (Subset-minimal change)

A confidentiality-preserving knowledge base K^{pub} *subset-minimally changes* K (or is *minimal*, for short) if there is no confidentiality-preserving $K^{pub'}$ such that
$((K \setminus K^{pub'}) \cup (K^{pub'} \setminus K)) \subset ((K \setminus K^{pub}) \cup (K^{pub} \setminus K))$.

Confidentiality-Preservation for Credulous Users

Example

For the example K and $policy$ and no a priori knowledge, the fact *Ill*(Mary, Aids) has to be deleted.

But also *Ill*(Pete, Aids) can be deduced credulously, because it is satisfied by answer set S_1.

In order to avoid this, we have three options: delete *Treat*(Pete, Medi1), delete the non-literal rule in K or insert *Treat*(Pete, Medi2).

The same solutions are found for K, $policy'$ and $prior$: they block the credulous deduction of ¬*AbleToWork*(Pete).

Policy transformation

- Elements $policy$ will be treated as negative observations O_i^-
- Transform policy elements to set of rules containing a new positive observation O^+

$$
\begin{aligned}
PTR^{cred} \quad &:= \quad \{O_i^- \leftarrow C_i \mid C_i \in policy\} \\
&\cup \quad \{O^+ \leftarrow not\, O_1^-, \ldots, not\, O_k^-\}
\end{aligned}
$$

Example

The set of policy transformation rules for $policy'$ is

$$
\begin{aligned}
PTR^{cred} \quad = \quad \{&O_1^- \leftarrow \mathit{Ill}(x, \mathsf{Aids}) \ , \ O_2^- \leftarrow \neg \mathit{AbleToWork}(x) \ , \\
&O^+ \leftarrow not\, O_1^-, not\, O_2^-\}
\end{aligned}
$$

Lastly, we consider a **goal rule** GR that enforces the single positive observation O^+: $GR = \{\leftarrow not\, O^+\}$.

Confidentiality with deletions

- We thus obtain a new program P as

$$P = UP \cup prior \cup PTR^{cred} \cup GR$$

- Compute a U-minimal answer set S
- Negative explanation F is obtained from the negative update atoms contained in S: $F = \{L \mid -L \in S\}$
- Check whether

$$(K \setminus F) \cup prior \cup PTR^{cred} \cup \{\leftarrow O^+\} \text{ is inconsistent.} \qquad (1)$$

  - Check for inconsistency with the negation of the positive observation O^+ (which makes F a *skeptical* explanation of O^+)
- Only answer sets of P that are U-minimal among those respecting this inconsistency property (1)

Confidentiality with deletions

Example

We combine the update program UP of K with $prior$ and the policy transformation rules and goal rule. This leads to the following two U-minimal answer sets with only deletions which satisfy the inconsistency property (1):

$$\begin{aligned} S_1 &= \{-\mathit{Ill}(\text{Mary}, \text{Aids}), -\mathit{Treat}(\text{Pete}, \text{Medi1}), n(R), \\ &\quad \overline{\mathit{Ill}(\text{Mary}, \text{Aids})}, \overline{\mathit{Treat}(\text{Pete}, \text{Medi1})}, O^+\} \\ S_2 &= \{-\mathit{Ill}(\text{Mary}, \text{Aids}), \mathit{Treat}(\text{Pete}, \text{Medi1}), -n(R), \\ &\quad \overline{\mathit{Ill}(\text{Mary}, \text{Aids})}, \overline{n(R)}, O^+\} \end{aligned}$$

These answer sets correspond to the previous minimal solutions where $\mathit{Ill}(\text{Mary}, \text{Aids})$ must be deleted together with either $\mathit{Treat}(\text{Pete}, \text{Medi1})$ or the rule named R.

Confidentiality with deletions

Theorem (Correctness for deletions)

A knowledge base $K^{pub} = K \setminus F$ *preserves confidentiality and changes* K *subset-minimally iff* F *is obtained by an answer set of the program* P *that is U-minimal among those satisfying the inconsistency property (1).*

Proof.

(Sketch) Because we chose K to be the set of abducibles $\mathcal{A}$, only negative update atoms from $\mathcal{UA}^-$ occur in UP – no insertions with update atoms from $\mathcal{UA}^+$ will be possible. We obtain an anti-explanation (E, F) where E is empty. We have thus $K^{pub} \cup prior \cup PTR^{cred} \models O^+$ but for every O_i^- there is no answer set in which O_i^- is satisfied. This holds iff for every policy element C_i there is no answer set of $K^{pub} \cup prior$ that satisfies any instantiation of C_i; thus $cred(K^{pub} \cup prior, C_i) = \emptyset$. Subset-minimal change carries over from U-minimality of answer sets. □

Deletions and Insertions

- Allow insertions of literals into K for confidentiality-preservation
- Different set of abducibles $\mathcal{A}$
  - starting from the new negative observations O_i^- used in the policy transformation rules, we trace back all rules in $K \cup prior \cup PTR^{cred}$
  - construct a dependency graph and collect the literals that the negative observations depend on

$$
\begin{aligned}
P_0 \quad = \{L \mid \quad & L \in body(R) \text{ or } not L \in body(R) \\
& \text{where } R \in PTR^{cred} \text{ and } O_i^- \in head(R)\}
\end{aligned}
$$

- Iterate and collect all the literals that the P_0 literals depend on:

$$
\begin{aligned}
P_{j+1} \quad = \{L \mid \quad & L \in body(R) \text{ or } not L \in body(R) \\
& \text{where } R \in K \cup prior \cup PTR^{cred} \\
& \text{and } head(R) \cap P_j \neq \emptyset\}
\end{aligned}
$$

and combine all such literals in a set $\mathcal{P} = \bigcup_{j=0}^{\infty} P_j$.

Deletions and Insertions

As we also want to have the option to delete rules from K (not only the literals in $\mathcal{P}$), we define the set of abducibles as the set $\mathcal{P}$ plus all those rules in K whose head depends on literals in $\mathcal{P}$:

$$\mathcal{A} = \mathcal{P} \cup \{R \mid R \in K \text{ and } head(R) \cap \mathcal{P} \neq \emptyset\}$$

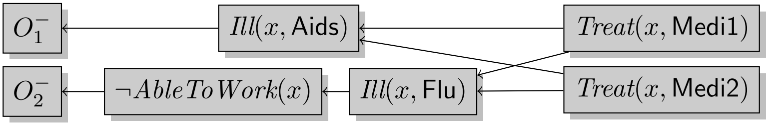
Deletions and Insertions

Example

For the example $K \cup prior \cup PTR^{cred}$, we note that the new negative observation O_1^- directly depends on the literal *Ill*(x, Aids) and the new negative observation O_2^- directly depends on the literal $\neg$*AbleToWork*(x); this is the first set of literals $P_0 = \{\mathit{Ill}(x, \text{Aids}), \neg \mathit{AbleToWork}(x)\}$.
By tracing back the dependencies in the graph, we obtain

$$\begin{aligned} \mathcal{P} \quad = \quad & \{\mathit{Ill}(x, \text{Aids}), \neg \mathit{AbleToWork}(x), \mathit{Ill}(x, \text{Flu}), \\ & \mathit{Treat}(x, \text{Medi1}), \mathit{Treat}(x, \text{Medi2})\} \end{aligned}$$

Lastly, add the rule R of K to $\mathcal{A}$ because literals in its head are in $\mathcal{P}$.



Deletions and Insertions

- obtain the normal form and then the update program UP for K and the new set of abducibles $\mathcal{A}$
- find an answer set of program P where additionally the positive explanation E is obtained as $E = \{L \mid +L \in S\}$ and S is U-minimal among those satisfying

$$(K \setminus F) \cup E \cup prior \cup PTR^{cred} \cup \{\leftarrow O^{+}\} \text{ is inconsistent} \qquad (2)$$

Deletions and Insertions

Example

New set of abducibles leads to additional insertion rules. Among others, the insertion rule for the new abducible $\mathit{Treat}(\mathrm{Pete}, \mathrm{Medi2})$ is

$$+\mathit{Treat}(\mathrm{Pete}, \mathrm{Medi2}) \leftarrow \mathit{Treat}(\mathrm{Pete}, \mathrm{Medi2})$$

With this new rule included in UP, we also obtain the solution where the fact $\mathit{Treat}(\mathrm{Pete}, \mathrm{Medi2})$ is inserted into K (together with deletion of $\mathit{Ill}(\mathrm{Mary}, \mathrm{Aids})$) to protect the two confidential facts $\mathit{Ill}(\mathrm{Pete}, \mathrm{Aids})$ and $\neg \mathit{AbleToWork}(\mathrm{Pete})$.

Theorem (Correctness for deletions & literal insertions)

A knowledge base $K^{pub} = (K \setminus F) \cup E$ preserves confidentiality and changes K subset-minimally iff (E, F) is obtained by an answer set of program P that is U-minimal among those satisfying inconsistency property (2).

Proof.

(Sketch) In UP, positive update atoms from $\mathcal{UA}^+$ occur for literals on which the negative observations depend. For subset-minimal change, only these literals are relevant for insertions; inserting other literals will lead to non-minimal change. By the properties of minimal skeptical (anti-)explanations that correspond to U-minimal answer sets of an update program, we obtain a confidentiality-preserving K^{pub} with minimal change. □

Outline

Contributions

In sum, this paper makes the following contributions:

- it formalizes confidentiality-preserving data publishing for a user who retrieves data under a credulous query response semantics.
- it devises a procedure to securely publish a logic program (with an expressiveness up to extended disjunctive logic programs) respecting a subset-minimal change semantics.
- it shows that confidentiality-preservation for credulous users corresponds to finding a skeptical anti-explanation and can be solved by extended abduction.

Open Questions

- Work out approach for skeptical users
- Work out complexity analysis
- Insertions other than literals
- In online query answering setting, use existential answers to protect secrets:

Example

If we want to hide the fact *Ill*(Mary, Aids) then return the answer
$\exists x$ *Ill*(x, Aids)