ALGEBRAIC MODELS OF QUESTION ANSWERING SYSTEMS

Jānis Čīrulis

Department of Computer Science
University of Latvia
jc@lanet.lv

Theory Days at Jõulumäe
October 3–5, 2008
OVERVIEW

1. Background: information systems
   Some simple models
   Dependencies and compatibility in information systems
2. Functional-dependency frames
   Definitions
   Examples
   Knowledge spaces
3. Question answering systems
   Information functions of a frame
   QA-systems and their formal contexts
4. Simulation
   How a QA-system simulates another one
   Essentially incomplete QA-systems, and representable QA-systems
5. The previous work
1. Background

Some simple models

- Information systems (Z. Pawlak)
  (another names: attribute systems, knowledge representation systems)
  
  \( Ob \) is a set of objects,

  \( At \) is a set of attributes of objects,

  \( Val := (Val_a \mid a \in At) \) is a family of sets; each \( Val_a \) is the set of values for the attribute \( a \),

  \( F := (f_o \mid o \in Ob) \text{ on } At \) is a family of descriptions (functions on \( At \), where \( f_o \) assigns to every \( a \) an element of \( Val_a \)).

(b) Incomplete: if the descriptions may be not complete, i.e., if each \( f_o(a) \) is a nonempty subset of \( Val_a \).

\( f_o(o) \) could also be a fuzzy subset of \( Val_a \), or a probability distribution on \( Val_a \), e.c.
Another interpretation: a (simple) *question answering system*:

- Elements of $Ob$ — *states*,
- elements of $At$ — *questions*,
- elements of $Val_a$ — *possible answers* to a question $a$,
- descriptions — *information functions*. 
- Formal contexts (R.Wille & B.Ganter)

A formal context is a triple \( FC := (G, M, I) \), where

- \( G \) is a set of objects (Gegenstände),
- \( M \) is a set of possible properties (Merkmale),
- \( I \) is a binary incidence relation in \( G \times M \).

\( g I m \) means “the object \( g \) has the property \( m \)”.

\( FC \) induces a Galois connection between \( \mathcal{P}(G) \) and \( \mathcal{P}(M) \):

\[
A \subseteq G \mapsto A^* := \{ m \in M | g I m \text{ for all } g \in A \},
\]

\[
B \subseteq M \mapsto B^* := \{ g \in G | g I m \text{ for all } m \in B \}.
\]

A pair \((A, B)\) is a concept of the context if \( A^* = A \) and \( B^* = B \). Concepts are ordered by

\( (A, B) \leq (A', B') \equiv A \subseteq A' \) (and \( B' \subseteq B \))

and form a lattice under this ordering (the concept lattice of the context).
• Information systems as formal contexts

\((Ob, At, Val, F)\) – an information system.

A \textit{descriptor} is a pair \((a, v)\) with \(a \in At\) and \(v \in Val_a\).

The \textit{information space of} \(S\) is the set \(K\) of all descriptors. The \textit{formal context of} \(S\) is the triple \((Ob, K, \models)\), where

\[ o \models (a, v) \equiv v \in f_o(a) \}\}

\(Information\;system \leftrightarrow \text{formal context of this type.}\)
• The inner logic of an information system

\[ IS := (Ob, At, Val, F) \] – an information system.

A pair \((a, V)\) with \(a \in At\) and \(V \subseteq Val_a\) is interpreted as a proposition “the value of \(a\) belongs to \(V\)”.

\(P :=\) the set of all propositions.

The inner logic of \(IS\) is the triple (in fact, a formal context) \((Ob, P, \models)\), where

\[ o \models (a, V) := f_o(a) \subseteq V \] (“the proposition \((a, V)\) is true of \(o\)”).

Information system \(\leftrightarrow\) formal context of this type.
• Many-valued contexts (R. Wille & B. Ganter)

A **many-valued context** is a quadruple \((G, M, W, I)\), where

- \(G\) is a set of objects,
- \(M\) is a set of attributes,
- \(W\) is a set of values (Werte),
- \(I\) is a ternary incidence relation in \(G \times M \times W\) such that
  \((g, m, w) \in I\) and \((g, m, w') \in I\) implies \(w = w'\).

\((g, m, w) \in I\) means "the attribute \(m\) has a value \(w\) for the object \(g\)."

Many-valued context \(\leftrightarrow\) complete information system.
• Chu spaces (W.Pratt)

$K$ – a set of values (or an alphabet).

A *Chu space over $K$* is a triple $(X, r, A)$, where

- $X$ is a set of points
- $A$ is a set of states
- $r$ is a function of type $X \times A \rightarrow K$.

States $\leftrightarrow$ objects,
Points $\leftrightarrow$ attributes,
Chu space $\leftrightarrow$ complete information system
with a common value set $K$ for all attributes.
Dependencies and compatibility in information systems

Let $IS := (Ob, At, Val, F)$ be a complete information system.

Drawbacks?

- Attributes in $IS$ are formally independent, for descriptions in $F$ may be quite arbitrary.

Let $A, B \subseteq At$ (complex attributes).

There is an inclusion dependency between $A$ and $B$ iff $A \subseteq B$.

A functionally depends on $B$ if, for every object, the value of every attribute in $A$ turns out to be uniquely determined by values of attributes in $B$:

$$a \leftarrow B \equiv \text{for all } o_1, o_2 \in Ob,$$

$$f_{o_1}(a) = f_{o_2}(a) \text{ whenever } f_{o_1}(b) = f_{o_2}(b) \text{ for every } b \in B.$$

$$A \leftarrow B \equiv a \leftarrow B \text{ for all } a \in A.$$

Dependencies a posteriori.
A complex attribute $A$ should have a set of complex values:

$$Val_A := \prod (A_a \mid a \in A)$$

i.e., $Val_A$ is the set of all functions $\varphi$ on $A$ such that $\varphi(a) \in Val_a$ for all $a \in A$.

Descriptions should be extended to complex attributes:

$$f_o^+(A) \in Val_A, \quad f_o^+(A)(a) := f_o(a).$$

**Proposition.** If $A \leftarrow B$, then there is a function $d^B_A : Val_B \rightarrow Val_A$ which realises this dependency:

- for every object $o$, $f_o^+(A) = d^B_A(f_o^+(B))$.

This function is unique only if the sets $Val_a$ do not contain “unnecessary” elements:

- for every $a \in At$, $Val_a = \{f_o(a) \mid o \in Ob\}$.

In particular, if $A \subseteq B$, then

- $d^B_A(\varphi) = \varphi|A$,
- the function $d^B_A$ is surjective and
  - is actually the projection of the set $Val_B$ onto $A$. 

10
• It may happen that not all attributes permit simultaneous determination of their values.

Suppose that given is a symmetric and irreflexive rejection relation on $At$.

A complex attribute $A \subseteq At$ is said to be coherent if no attributes from $A$ reject each other.
Two coherent complex attributes $A$ and $B$ are said to be compatible if $A$ and $B$ have a common coherent superset. This is the case if and only if the union $A \cup B$ is coherent.

**Proposition.** The union of any set of parts of a coherent complex attribute is coherent, i.e., the coherent complex attributes form a bounded complete poset under set inclusion.
• Summing up: an extension of an information system.

\[ IS := (Ob, At, Val, F) \] – a complete information system.

Put

\[ At^+ \] – a set of coherent complex attributes together with inclusion and dependence relations,

\[ Val^+ \] – the family of all complex value sets \( Val_A \) together with the family of all dependency functions \( d^B_A \),

\[ F^+ \] – the set of all extended descriptions \( f^+_o \).

Then \( IS^+ := (Ob, At^+, Val^+, F^+) \) is an information system, called an extension of \( S \).

**Problem 0.** Characterise abstractly the class of structures isomorphic to such extensions.

This was the motivation for constructions in the next section.
“Question-answer” interpretation again:

<table>
<thead>
<tr>
<th>Notion:</th>
<th>Notation:</th>
</tr>
</thead>
<tbody>
<tr>
<td>states instead of objects</td>
<td>S instead of $Ob$,</td>
</tr>
<tr>
<td>questions instead of coherent complex attributes</td>
<td>$Q$ instead of $At^+$,</td>
</tr>
<tr>
<td>answers instead of complex values</td>
<td>$A$ instead of $Val^+$,</td>
</tr>
<tr>
<td>information functions instead of descriptions</td>
<td>$F$</td>
</tr>
</tbody>
</table>
2. Functional-dependency frames

**Definition.** An *fd-structure* is a pair $FD := (Q, A)$, where

- $Q$ is a preordered set $(Q, \leftarrow)$ (the *scheme* of the frame),
- $A$ is a system (a *model* of the scheme) consisting of
  - a family of sets $(A_q \mid q \in Q)$, and
  - a family of mappings $d^q_p : A_q \to A_p$ with $p, q \in Q, p \leftarrow q$
    such that
    $$d^q_p(a) = a, \quad d^q_p d^r_q(c) = d^r_p(c).$$

A subset $Q'$ of $Q$ is *compatible* if it is bounded from above: i.e., if there is $r \in Q$ such that $p \leftarrow r$ for every $p \in Q'$.

The set $Q$ is *finitely bounded complete* if every [finite] (possibly, empty) compatible subset of $Q$ has a l.u.b.

Roughly, a l.u.b. of a compatible subset $Q' \subseteq Q$ represents the “complex question” $Q'$ as a one element of $Q$. 
Equivalently, \( Q \) is bounded complete iff every nonempty subset of \( Q \) (in particular, \( Q \) itself) has a g.l.b.

\( Q \) is finitely bounded complete iff

- \( p \triangleright q \) always implies that \( \{p, q\} \) has a l.u.b., and
- \( Q \) has a g.l.b.

We shall assume that all finitely bounded complete preordered sets considered below satisfy also condition

- any finite nonempty subset of \( Q \) has a g.l.b.
Definition. An fd-structure \((Q, A)\) is said to be an \textit{fd-frame} if

- its scheme \(Q\) is finitely bounded complete, and
- the model \(A\) of the scheme satisfies the condition
  - if \(r\) is a l.u.b of a finite subset \(Q'\) of \(Q\), and
  - if \(a, b \in A_r\),
    - then \(d^r_p(a) = d^r_p(b)\) for all \(p \in Q\) implies \(a = b\).

The later condition garanties that an answer to the question \(r\) is completely detemined by the answers to its “components” from \(Q'\).

Dependencies \textit{apriori}
**Definition.** We say that an fd-frame is a *frame with inclusions* if its scheme $Q$ is equipped with an order relation $\subseteq$ (*inclusion*, or *part_of* relation) whose interaction with $\leftarrow$ is subject to the following axioms:

- if $p \subseteq q$, then $p \leftarrow q$,
- if $Q'$ is a finite compatible subset of $Q$ in which
  - $p \leftarrow q$ only if $p \subseteq q$,
  - then $Q'$ has the l.u.b. with respect to $\subseteq$,
- every mapping $d_p^q$ is surjective whenever $p \subseteq q$.

In particular, such a poset $(Q, \subseteq)$ is a finitely bounded complete also with respect to $\subseteq$:

- if $p \not\subseteq q$, then $p \cup q$ exists in $Q'$,
- any two elements $p, q$ of $Q'$ have the meet $p \cap q$ in $Q'$,
- there is the $\subseteq$-least element $\emptyset$ in $Q$. 
**Proposition.** The Armstrong axioms for functional dependencies

- if $p \subseteq q$, then $p \leftarrow q$,
- if $p \leftarrow r, q \leftarrow r, p \bowtie q$, then $p \cup q \leftarrow r$,
- if $p \leftarrow q, q \leftarrow r$, then $p \leftarrow r$

hold in $(Q,\leftarrow,\subseteq)$. 
Examples

E1: Simple frames

Suppose that
Q is a flat domain
(i.e., a poset \((Q, \leq, 0)\) with the least element 0, in which
every chain is of length \(\leq 2\)),

\(A\) is a family \((A_p | p \in Q)\) of non-empty sets
\(A_0\) is a singleton.,

Put
\(p \leftarrow q :\equiv p \subseteq q :\equiv p \leq q,\)

\(d^q_p := \begin{cases} 
\text{the identity function on } A_q \text{ if } p = q, \\
\text{the single function } A_q \rightarrow A_0 \text{ if } p = 0. 
\end{cases}\)

In this way, \(Q\) is converted into a trivial scheme with inclusions, and \(A\), in its model.
So, \((Q, A)\) is an inclusion frame.
We call such frames simple.
Suppose that $IS := ((Ob, At, Val, F)$ is an information system.

$(At, Val)$ is a simple frame “without bottom”; there is a one-to-one correspondence between pairs of kind $(At, Val)$ and simple frames.

In any extension of $IS$, the pair $(At^+, Val^+)$ is a frame with inclusions.
Frames in relational databases

Initially $At, Val$ — as in an information system.

$At^+$ is the set of finite subsets of $At$, (relational types)

$Val_A$ is the set of all rows of type $A$,

a subset of $Val_A$ is a relation with attributes in $A$,

$\subseteq$ is the ordinary set inclusion,

$\leftarrow$ together is given as a kind of constraints,

with functions $d_B^A$ (sintactically and semantically),

$[d_B^A \text{ with } A \subseteq B]$ is the projection of $Val_B$ onto $A$].
E4: Frames from automata

Consider an automaton \((X, Y, Z, \lambda, \mu)\), where

\[ X \] is the input alphabet,
\[ Y \] is the output alphabet,
\[ Z \] is the set of states,
\[ \lambda : Z \times X \to Z \] is the transition function,
\[ \mu : Z \times X \to Y \] is the output function.

Set
\[ Q := X^*, \]
\[ p \subseteq q \] :\( \equiv \) \( p \) is a prefix of \( q \),
\[ p \leftarrow q \] :\( \equiv p \subseteq q \),
\[ A_p := Y^{|p|}, \] where \(|p|\) is the length of \( p \),
\[ d^p_p : A_q \to A_p \] := the function which takes every word from \( A_q \)
into its prefix of length \(|p|\).

Then \((Q, A)\) is an inclusion frame and two questions are compatible if and only if they are comparable.
Let $FD := (Q, V)$ be a frame.

**Definition.** An *information piece* in this frame is any pair $(q, a)$ with $q \in Q, a \in A_q$.

The set of all information pieces is called the *knowledge space* of $FD$.

An information piece $(p, a)$ is said to be

- **entailed by** $(q, b)$ (in symbols, $(p, a) \succeq (q, b)$) if
  
  $p \leftarrow q$ and $a = d_q^p(b)$,
  
  [“Whenever $q$ has the answer $b$, $p$ has the answer $a”]$

- a **restriction of** $(q, b)$ (in symbols, $(p, a) \subseteq (q, b)$) if
  
  $p \subseteq q$ and $a = d_q^p(b)$.

An *abstract knowledge space* is a triple $(K, \succeq, \subseteq)$ that is isomorphic to the knowledge space of some frame with inclusions.

**Theorem 1.** Up to isomorphisms, there is one-to-one correspondence between frames with inclusions and abstract knowledge spaces.

[Axiomatic description of abstract knowledge spaces.]
3. Question answering systems

FD := (Q, A) – an fd-structure with inclusions.

**Definition.** An *information function* in FD is a function \( f \) on \( Q \) that assigns a nonempty subset of \( A_p \) to every \( p \in Q \) so that

- if \( p \leftarrow q \), then \( f(p) = \{d^p_q(b) : b \in f(q)\} = d^q_p(f(q)) \) (i.e., if \( p \) depends on \( q \), then \( f(p) \) contains just the answers that can be calculated out from those in \( f(q) \)),
- if \( r \) is a l.u.b of a finite subset \( Q' \subseteq Q \) with respect to \( \leftarrow \), then \( f(r) := \{c \in A_r | d^r_p(c) \in f(p) \text{ for all } p \in Q'\} \) (i.e., \( f(r) \) contains just the answers that are combined from those belonging to the “components” of \( r \)).

An information function \( f \) is

- *complete* if every \( f(p) \) is a singleton,
- *proper* if there is no other description \( f' \) with \( f'(p) \subseteq f(p) \) for all \( p \in Q \),
- *trivial*, if \( f(p) = A_p \).

Every complete description is proper; the converse may not hold true.
Example: A frame which has only the trivial information function.

Let
\[ Q := \{p_1, p_2, q_1, q_2\}, \]
and
for \( i = 1, 2, \) \( A_{p_i} := \{a_{i1}, a_{i2}\}, A_{q_i} := \{b_{i1}, b_{i2}\}. \)

Assume that
\( p_1, p_2 \subseteq q_1, q_2, \) and \( \leftarrow \) coincides with \( \subseteq. \)

Set
\[
\begin{align*}
  d_{p_i}^{q_1}(b_{1j}) &= a_{ij}, & d_{p_1}^{q_2}(b_{2j}) &= a_{1j}, \\
  d_{p_2}^{q_2}(b_{21}) &= a_{22}, & d_{p_2}^{q_2}(b_{22}) &= a_{21}.
\end{align*}
\]

The trivial information function is proper in this frame, but, clearly, not complete.
Definition. A *question answering system* is a quadruple \((Q, A, S, F)\), where

- \((Q, A)\) is an fd-frame with inclusions,
- \(S\) is a non-empty set,
- \(F := \{f_s \mid s \in S\}\) is a family of information functions.

Elements of
- of \(Q\) are called *questions*, those of \(A_q\), *answers to the question* \(q\).

A QA-system is
- *complete* if all of its information functions are complete,
- *simple* if its frame is simple.

Simple QA-system \(\leftrightarrow\) information system.
Definition. The formal context of a QA-system $QA := (Q, A, S, F)$ is the triple $(S, K, \vdash)$, where

- $K$ is the knowledge space of $QA$, and
- $\vdash$ is a binary relation in $S \times K$ such that
  
  $s \vdash (p, a) \equiv a \in f_s(p)$.

A QA-context is a formal context of some QA-system.

A formal context $(S, K, \vdash)$, where $K$ is an abstract knowledge space, is said to be an abstract QA-context if it is isomorphic to some QA-context. Something like Kripke structures, with $S$ the possible word space, and $K$ the algebra of propositions.

Theorem 2. Up to isomorphisms, there is one-to-one correspondence between QA-systems and abstract QA-contexts.

Problem 1. Characterise the class of abstract QA-contexts.
The concept lattice of the formal context of a QA-system is said to be the concept lattice of this system.

Problem 2. Characterise the class of concept lattices of QA-systems. Have concept lattices of complete QA-systems any distinctive property?

Problem 3. Two QA-systems may be considered as equivalent if their concept lattices are isomorphic. Characterise the class of QA-systems which are equivalent to a complete QA-system.
4. Simulation

In this section we assume that all QA-systems are faithful in the sense that, for every \( q \), \( A_q = \bigcup (f_s(q) \mid s \in S) \).

Let \( QA := (Q, A, S, F) \) \( QA' := (Q', A', S', F') \) be two QA-systems. By a macrostate of a system we understand any non-empty set of its states. (Macrostates are interpreted as vaguely specified states.)

A simulation of a \( QA \) into \( QA' \) is, informally, a triple of “devices” \( (\alpha, \beta, \gamma) \), where

- \( \alpha \) translates every question from \( Q \) into a question in \( Q' \),
- \( \beta \) realizes back translation: for every question \( q \in Q \), it associates a non-empty subset of \( A_q \) with each possible answer \( a' \in A'_{\alpha(q)} \) to the translated question \( \alpha(q) \),
- \( \gamma \) interprets every state of \( QA \) as a macrostate of \( QA' \): if a question was put to \( QA \) in some state, translated back are all answers to the translated question in \( QA' \) obtained in any state from the respective macrostate.
**Definition.** A *simulation* of a $QA$ into $QA'$ is a triple $(\alpha, \beta, \gamma)$, where

- $\alpha$ is a mapping $Q \to Q'$,
- $\beta$ is a family of mappings $\beta_q: A'_{\alpha(q)} \to (\mathcal{P}(A_q) \setminus \emptyset)$ with $q \in Q$, and
- $\gamma$ is a mapping $S \to (\mathcal{P}(S') \setminus \emptyset)$

subject to the following conditions:

- $\beta$ preserves l.u.b.-s of finite compatible subsets of $Q$, (then $\alpha$ is isotone with respect to both $\leftarrow$ and $\subseteq$),
- for all $p, q \in Q$ with $p \leftarrow q$, and every $a' \in A'_{\alpha(q)}$,
  $$\bigcup \{f_p^q(a) \mid a \in \beta_q(a')\} = \bigcup \{\beta_p(b') \mid b' \in f'^{\alpha(q)}_{\alpha(p)}(a')\},$$
- for all $p \in Q$, and every $s \in S$,
  $$f_s(p) = \bigcup \{\beta_p(a') \mid a' \in f'^{s'}_s(\alpha(p))\} \text{ for some } s' \in \gamma(s).$$
Let $QA := (Q, A, S, F)$ be a simple complete QA-system. Extensions of $QA$ are constructed like those of an information system.

Let $QA^+$ stand for the standard extension $(Q^+, At^+, S, F^+)$ in which $At^+$ contains all subsets of $Q$. Such an extension is unique and completely determined by $QA$. Recall that extensions of complete QA-systems are complete.

**Definition.** A QA-system is said to be
- *essentially incomplete* if it can be simulated by no complete QA-system,
- *representable* if it can be simulated by a simple complete QA-system.

**Proposition.** Not every QA-system is representable.

**Problem 4.** Are there essentially incomplete QA-systems? If yes, characterise abstractly those that are not.

**Problem 5.** Characterise abstractly the representable QA-systems.
5. The previous work

[1] and [2] are early papers, where several ideas of the later ones already appeared. Among other things, in [3] discussed are certain formal contexts (without referring to this concept) similar to those appearing here in connection with knowledge spaces. In [4-6] I used another term "knowledge representation system" rather than "question answering system" going back to [2]. In [4], KR-systems without inclusions were treated in terms of category theory. The approach of [5,6] is not so general, and we use there the language of general algebra (as in this presentation). However, some improvements to the model of [4] can be found at the beginning of [6]. Neither in [5] nor [6] inclusion dependencies are explicitly recognized.


