Uncertain Databases - Motivation & Foundations

Databases and Information Systems

Fabian Panse

panse@informatik.uni-hamburg.de

University of Hamburg

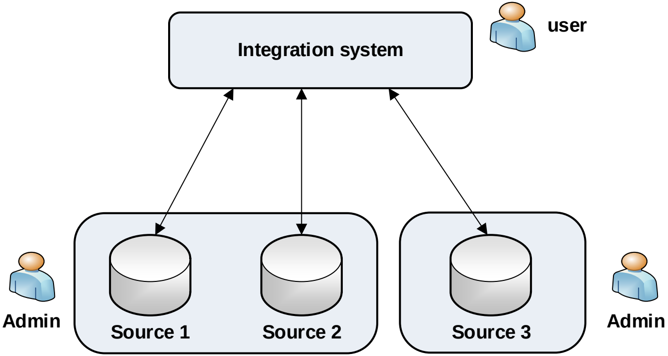
Motivation

Motivation - Why can Uncertain Databases be useful?

- Uncertainty is everywhere in our everyday life
- Uncertainty in data/information processing applications
- Conventional data models ignore such uncertainties
  (in the standard relational data model, the null value is the only concept to deal with uncertainties)

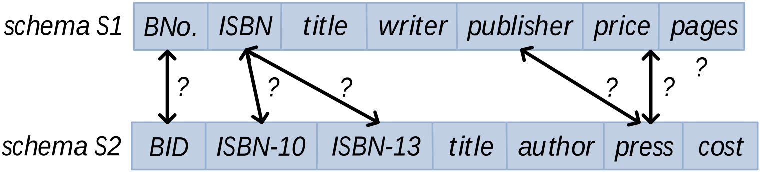
Data Integration

- Large number of autonomous data sources
- Integration system combines the data of the individual sources
- Uncertainty at different points



Data Integration - Uncertainty on Schema Mappings

- **Which schema elements (e.g. attributes) match?**



- Are ‘BID’ and ‘BNo.’ source-specific surrogate keys?
- Which ‘ISBN’ is modeled by source S1?
- ‘press’ obviously match with ‘publisher’, but is more similar to ‘price’
- Has source S2 no attribute which matches with ‘pages’?

Data Integration - Uncertainty on Relevance

- **Are these source tuples relevant for the integration result?**

Intregation System

SELECT *
FROM Books
WHERE year>=2001

Source S1

BID	ISBN-10	ISBN-13	title	author	press	price

All books since 2000

Data Integration - Uncertainty on Duplicates

- **Describe these two tuples the same person?**

PNo.	firstname	lastname	DoB	city
P23	William	Schulz	12.10.1987	HH
P14	Bill	Schultz	10.12.1987	St.Pauli

Data Integration - Uncertainty on Merging Result

- **Duplicate tuples need to be merged to a single tuple. However, which values describe this person best?**

PNo.	*firstname*	*lastname*	*DoB*	*city*
P23	*William*	*Schulz*	*12.10.1987*	*HH*
P14	*Bill*	*Schultz*	*10.12.1987*	*St.Pauli*

Data Integration - Sources of Uncertainty

- **Using different standards and conventions:**

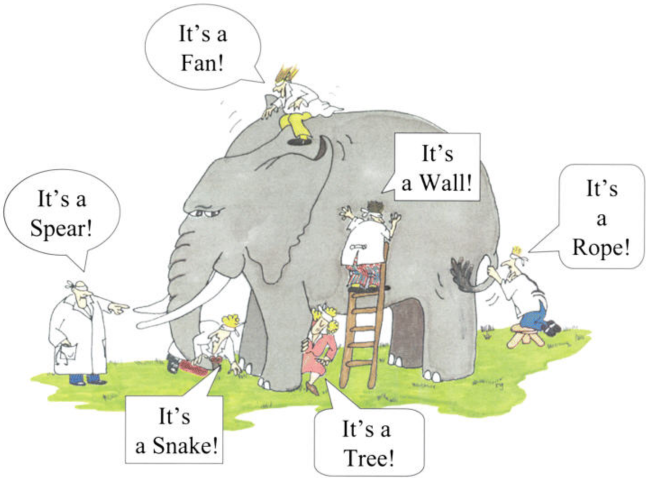
Source: Introduction to Duplicate Detection, Felix Naumann and Melanie Herschel, Morgan&Claypool

Data Integration - Sources of Uncertainty

- **Google reports 593* different spellings of 'Britney Spears'**
  - 'Britney Spears' (488941 searches)
  - 'Brittany Spears' (40134 searches)
  - 'Brittney Spears' (36315 searches)
  - 'Britany Spears' (24342 searches)
  - 'Britny Spears' (7331 searches)
  - ...
- **How many different spellings exist for more complex names?**
  - 'Giannis Antetokounmpo' (Greek basketball player)
  - 'Dharmavarapu Subramanyam' (Indian actor)
  - 'Janice Keihanaikukauakahihuliheekahaunaele' (Hawaiian woman)
  - 'Venkatanarasimharajuvaripeta railway station' (railway station in India)

*Source: http://www.netpaths.net/blog/britney-spears-spelling-variations/

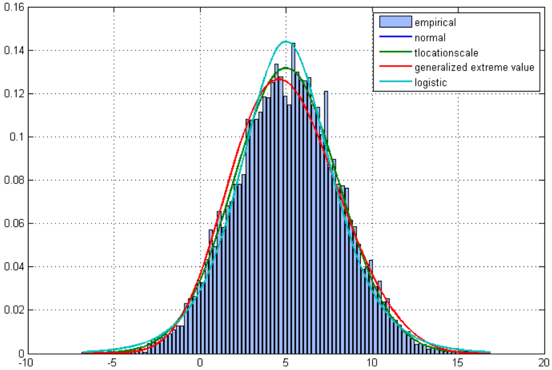
Subjective perceptions



Subjective perceptions

Uncertainty of measurements

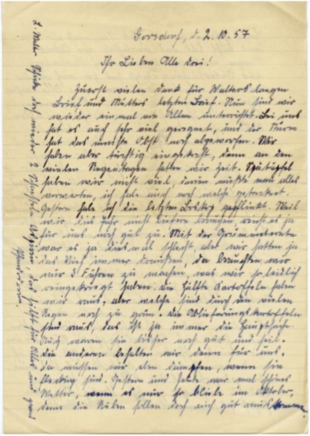
- Imprecise techniques
- Observed value is inconstant
- Confounding factors (e.g. wind)



Optical character recognition

- Handwriting unreadable
- Print faded out
- Medium (e.g. paper) weathered

Optical character recognition



Gersdorf, 2.10.57

Ihr Lieben Alle Drei!

Zuerst vielen Dank für Walters langen Brief und Mutters letzten Brief. Nun sind wir wieder einmal von allem unterrichtet. Bei uns hat es auch sehr viel geregnet und der Sturm has das meiste Obst noch abgeworfen. Wie haben aber tüchtig eingekocht, denn an den vielen Regentagen hatten wir Zeit. Spätäpfel haben wir nicht viel, darum mußte man alles verwerten, ich habe auch noch welche getrocknet. Gestern habe ich die letzten Boskop gepflückt. Weil wir dies Jahr nicht liefern brauchen, reicht es ja für uns noch gut zu. Mit der Grummeternte war es ja diesmal schlecht, aber wir hatten ja das Vieh immer draußen, da brauchten wir nur 3 Fuhren zu machen, was wir so leidlich reingekriegt haben. Die Hälfte Kartoffeln haben wir raus, aber welche sind durch den vielen Regen noch zu grün. Die Ablieferungskartoffeln sind raus, das ist ja immer die Hauptsache. Auch waren sie bisher noch gut unds heil. Die anderen behalten wir dann für uns. Da müssen wir eben dämpfen, wenn sie fleckig sind. Gestern und heute war mal schönes Wetter, wenn es nur so bliebe im Oktober, denn die Rüben sollen doch auch gut rauskommen.

Source: http://familie.berger-odenthal.de/Berger/Oma/

Ambiguity in natural languages

Relevant in:

- Text mining

Ambiguity in natural languages

Relevant in:

- Text mining

SPIEGEL ONLINE SCHULSPIEGEL

Login | Registrierung

Abi - und dann? | Querweltein | Leben U21 | Wissen

Home > SchulSPIEGEL > Polizei > Leibesvisitation: Polizei durchsucht Schüler bis auf die Unterhose

Leibesvisitation bei Schülern: Polizei sucht in Unterhosen Fünf-Euro-Schein

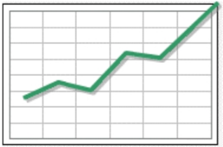
Ambiguity in natural languages

Relevant in:

- Text mining

Nature of Predictions

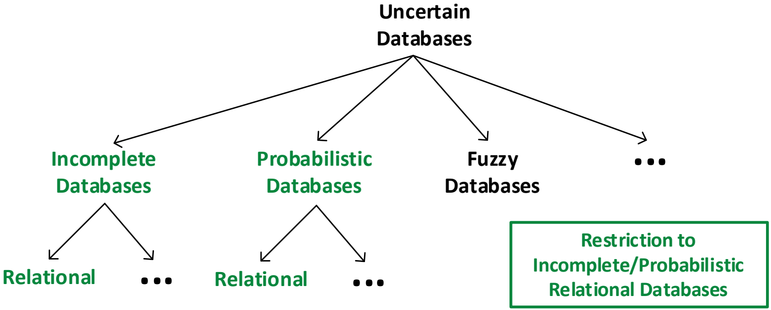
- Network load (street/power/data)
- Available resources
- Price/cost
- Weather



Foundations

Uncertain Databases

- Data models that capture uncertainty on data values
- Several concepts for modeling uncertainty
  (e.g. probability theory, possibility theory, Dempster Shafer theory)
- Several data models as baseline
  (e.g. relational, object-oriented, XML)



Incomplete Databases

Idea:

- Modeling uncertainty by a set of alternative database instances (so-called *possible worlds*)
- All possible worlds are defined on the same schema
- We restrict to finite sets of worlds in this lecture.

Definition: An *incomplete database idb* is a finite set of possible worlds $\mathbf{W} = \{W_1, \ldots, W_k\}$ where each world $W \in \mathbf{W}$ corresponds to a conventional database instance and all these worlds are defined on the same database schema.

Incomplete Databases

Conventional (certain) database:

- Single database instance at a particular time

Person

<u>WK</u>	***name***	***age***
p1	*J.Doe*	*27*
p2	*K.Smith*	*32*
p3	*J.Ho*	*28*
p4	*J.J.Doe*	*31*

Incomplete Databases

Conventional (certain) database:

- Single database instance at a particular time

Incomplete database:

- Several alternative database instances at a particular time

***W*$_1$**

<u>WK</u>	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	41

***W*$_2$**

<u>WK</u>	name	age
p1	J.Doe	72
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	41

***W*$_3$**

<u>WK</u>	name	age
p1	J.Doe	27
p2	K.Smith	32
p4	J.J.Doe	41

Incomplete Databases

Conventional (certain) database:

- Single database instance at a particular time

Incomplete database:

- Several alternative database instances at a particular time

W_1

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Doe	28

W_2

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Doe	29

W_3

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	28

W_4

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	29

W_5

WK	name	age
p1	J.Doe	28
p2	K.Smith	32
p3	J.Doe	28

W_6

WK	name	age
p1	J.Doe	28
p2	K.Smith	32
p3	J.Doe	29

W_7

WK	name	age
p1	J.Doe	28
p2	K.Smith	32
p3	J.Ho	28

W_8

WK	name	age
p1	J.Doe	28
p2	K.Smith	32
p3	J.Ho	29

W_9

WK	name	age
p2	K.Smith	32
p3	J.Doe	28

W_{10}

WK	name	age
p2	K.Smith	32
p3	J.Doe	29

W_{11}

WK	name	age
p2	K.Smith	32
p3	J.Ho	28

W_{12}

WK	name	age
p2	K.Smith	32
p3	J.Ho	29

Incomplete Databases

Semantics:

- Worlds are mutually exclusive:

  "only one of these worlds can be 'true'!"

- Worlds are jointly exhaustive:

  "one of these worlds is assumed to be 'true'!"

Notation:

- Shared schema of possible worlds: *world schema*
- Set of possible worlds: *possible world space*
- Primary key of world schema: *world key* (short WK)

Incomplete Databases

Differences between worlds:

- Different non-key values are assigned to the same key value
- Specific key values are missing (i.e. no tuple with such key values exists)

W_1

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	41

W_2

WK	name	age
p1	J.Doe	72
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	41

W_3

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p4	J.J.Doe	41

Incomplete Databases

Differences between worlds:

- Different non-key values are assigned to the same key value
- Specific key values are missing (i.e. no tuple with such key values exists)

W_1

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	[illegible]
p4	J.J.Doe	[illegible]

W_2

WK	name	age
p1	J.Doe	72
p2	K.Smith	32
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]

W_3

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p4	J.J.Doe	41

different non-key values for p1

Incomplete Databases

Differences between worlds:

- Different non-key values are assigned to the same key value
- Specific key values are missing (i.e. no tuple with such key values exists)

Missing key value p3

W_1

<u>WK</u>	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	41

W_2

<u>WK</u>	na	
p1	J.Doe	72
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	41

	e	age
p1	J.Doe	27
p2	K.Smith	32
p4	J.J.Doe	41

Incomplete Databases

Certain-tuple: Tuple that belongs to every possible world
Maybe-tuple: Tuple missing in some of the possible worlds

Example:

W_1

	WK	name	age
t_1	p1	J.Doe	27
t_2	p2	K.Smith	32
t_3	p3	J.Ho	28
t_4	p4	J.J.Doe	41

W_2

	WK	name	age
t_5	p1	J.Doe	72
t_2	p2	K.Smith	32
t_3	p3	J.Ho	28
t_4	p4	J.J.Doe	41

W_3

	WK	name	age
t_1	p1	J.Doe	27
t_2	p2	K.Smith	32
t_4	p4	J.J.Doe	41

- Certain-tuples: t_2 and t_4
- Maybe-tuples: t_1, t_3, and t_5

Incomplete Databases - Tuple Dependencies

- Two tuples t_r and t_s are exclusive if they do not coexist in any possible world
$$\nexists W \in \mathbf{W}\colon \{t_r, t_s\} \subseteq W$$
- A tuple t_s is positively implicated by a **maybe-tuple** t_r if it belongs to every possible world that contains t_r
$$\forall W \in \mathbf{W}\colon\ t_r \in W \Rightarrow t_s \in W$$
- A tuple t_s is negatively implicated by a **maybe-tuple** t_r if it belongs to every possible world that does not contain t_r
$$\forall W \in \mathbf{W}\colon\ t_r \notin W \Rightarrow t_s \in W$$
- Two tuples t_r and t_s are independent if they are not exclusive and none of them implicates (positively or negatively) the other

Incomplete Databases - Tuple Dependencies

Example:

W_1	t_r	t_s
W_2	t_r	Ø
W_3	Ø	t_s
W_4	Ø	Ø

- t_r and t_s are independent

Incomplete Databases - Tuple Dependencies

Example:

W_1	t_r	t_s
W_2	t_r	Ø
W_3	Ø	t_s
W_4	Ø	Ø

- t_r and t_s are independent
- t_r and t_s are exclusive

Incomplete Databases - Tuple Dependencies

Example:

W_1	t_r	t_s
W_2	t_r	∅
W_3	∅	t_s
W_4	∅	∅

- t_r and t_s are independent
- t_r and t_s are exclusive
- t_r is positively implicated by t_s

Incomplete Databases - Tuple Dependencies

Example:

W_1	t_r	t_s
W_2	t_r	∅
W_3	∅	t_s
W_4	∅	∅

- t_r and t_s are independent
- t_r and t_s are exclusive
- t_r is positively implicated by t_s
- t_r is negatively implicated by t_s and t_s is negatively implicated by t_r

Incomplete Databases - Tuple Dependencies

Example:

W_1	t_r	t_s
W_2	t_r	∅
W_3	∅	t_s
W_4	∅	∅

- t_r and t_s are independent
- t_r and t_s are exclusive
- t_r is positively implicated by t_s
- t_r is negatively implicated by t_s and t_s is negatively implicated by t_r
- t_r is negatively implicated by t_s

Probabilistic Databases

Idea:

- Each possible world is associated with a probability
- We restrict to finite sets of worlds in this lecture.

Definition: A *probabilistic database pdb* is a probability space $(\mathbf{W}, Pr)$ where $\mathbf{W}$ is a possible world space and Pr is a discrete probability distribution over these worlds, i.e. $Pr: \mathbf{W} \rightarrow]0,1]$ is a function so that $\sum_{W \in \mathbf{W}} Pr(W) = 1$.

Probabilistic Databases - Example

W_1, Pr=0.5

	<u>**WK**</u>	***name***	***age***
t_1	**p1**	J.Doe	27
t_2	**p2**	K.Smith	32
t_3	**p3**	J.Ho	28
t_4	**p4**	J.J.Doe	41

W_2, Pr=0.4

	<u>**WK**</u>	***name***	***age***
t_5	**p1**	J.Doe	72
t_2	**p2**	K.Smith	32
t_3	**p3**	J.Ho	28
t_4	**p4**	J.J.Doe	41

W_3, Pr=0.1

	<u>**WK**</u>	***name***	***age***
t_1	**p1**	J.Doe	27
t_2	**p2**	K.Smith	32
t_4	**p4**	J.J.Doe	41

Probabilistic Databases - Marginal Probabilities

Let $pdb = (\mathbf{W}, Pr)$ be a probabilistic database.

- A tuple t belongs to pdb (denoted as $t \in pdb$) if it belongs to any $W \in \mathbf{W}$
- Marginal probability of tuple t is the probability that t exists
- Described by probability mass function p
- Corresponds to the overall probability of all possible worlds that contain t, i.e.

$$p(t) = \sum_{W \in \mathbf{W}, t \in W} Pr(W)$$

- **Certain-Tuple:** tuple t is certain if $p(t) = 1$
- **Maybe-Tuple:** tuple t is maybe if $p(t) \in]0, 1[$

Probabilistic Databases - Marginal Probabilities

Example:

W_1, Pr=0.5

	WK	name	age
t_1	p1	J.Doe	27
t_2	p2	K.Smith	32
t_3	p3	J.Ho	28
t_4	p4	J.J.Doe	41

W_2, Pr=0.4

	WK	name	age
t_5	p1	J.Doe	72
t_2	p2	K.Smith	32
t_3	p3	J.Ho	28
t_4	p4	J.J.Doe	41

W_3, Pr=0.1

	WK	name	age
t_1	p1	J.Doe	27
t_2	p2	K.Smith	32
t_4	p4	J.J.Doe	41

- Marginal probabilities: $p(t_1) = 0.6$, $p(t_3) = 0.9$, $p(t_5) = 0.4$, $p(t_2) = p(t_4) = 1.0$
- Certain-tuples: t_2 and t_4
- Maybe-tuples: t_1, t_3, and t_5

Probabilistic Databases - Tuple Dependencies

Let $pdb = (\mathbf{W}, Pr)$ be a probabilistic database.

- Probability mass function $p_\vee$:

  "At least one of the tuples in set T exists"

  $$p_\vee(T) = \sum_{W \in \mathbf{W}, T \cap W \neq \emptyset} Pr(W)$$

- Probability mass function $p_\wedge$ (joint probability):

  "All of the tuples in set T exist"

  $$p_\wedge(T) = \sum_{W \in \mathbf{W}, T \subseteq W} Pr(W)$$

Probabilistic Databases - Tuple Dependencies

Differences to incomplete databases:

- The number of different kinds of dependencies is infinite
- Independence cannot be concluded from the absence of exclusion and implication

Let $pdb = (\mathbf{W}, Pr)$ be a probabilistic database.

- Two tuples t_r and t_s are independent if

$$p_\wedge(\{t_r, t_s\}) = p(t_r) \times p(t_s)$$

- In this case, it holds that:

$$p_\vee(\{t_r, t_s\}) = 1 - (1 - p(t_r)) \times (1 - p(t_s))$$

Probabilistic Databases - Tuple Dependencies

Independence:

$$p_\wedge(\{t_r, t_s\}) = p(t_r) \times p(t_s)$$

W_1	t_r	t_s	$p(t_r) \times p(t_s)$
W_2	t_r	ø	$p(t_r) \times (1\text{-}p(t_s))$
W_3	ø	t_s	$(1\text{-}p(t_r)) \times p(t_s)$
W_4	ø	ø	$(1\text{-}p(t_r)) \times (1\text{-}p(t_s))$

$$Pr(W_2) = p(t_r) - Pr(W_1) = p(t_r) - p(t_r) \times p(t_s) = p(t_r) \times (1 - p(t_s))$$
$$Pr(W_3) = p(t_s) - Pr(W_1) = p(t_s) - p(t_r) \times p(t_s) = p(t_s) \times (1 - p(t_r))$$
$$Pr(W_4) = 1 - Pr(W_1) - Pr(W_2) - Pr(W_3)$$
$$= 1 - p(t_r) - p(t_s) + p(t_r) \times p(t_s) = (1 - p(t_r)) \times (1 - p(t_s))$$
$$p_\vee(\{t_r, t_s\}) = 1 - Pr(W_4) = 1 - (1 - p(t_r)) \times (1 - p(t_s))$$

Probabilistic Databases - Tuple Dependencies

Let $pdb = (\mathbf{W}, Pr)$ be a probabilistic database.

- Exclusion is defined as in incomplete databases

$\Rightarrow$ Two tuples t_r and t_s are exclusive if

$$\nexists W \in \mathbf{W}\colon \{t_r, t_s\} \subseteq W$$

- In this case, it holds that:

$$p_\wedge(\{t_r, t_s\}) = 0$$

and

$$p_\vee(\{t_r, t_s\}) = p(t_r) + p(t_s)$$

Probabilistic Databases - Tuple Dependencies

Exclusion:

W_1	t_r	t_s	0
W_2	t_r	ø	$p(t_r)$
W_3	ø	t_s	$p(t_s)$
W_4	ø	ø	$1-p(t_r)-p(t_s)$

$$Pr(W_2) = p(t_r)$$
$$Pr(W_3) = p(t_s)$$
$$Pr(W_4) = 1 - Pr(W_2) - Pr(W_3) = 1 - p(t_r) - p(t_s)$$

$$p_\wedge(\{t_r, t_s\}) = Pr(W_1) = 0$$
$$p_\vee(\{t_r, t_s\}) = Pr(W_2) + Pr(W_3) = p(t_r) + p(t_s)$$

Probabilistic Databases - Tuple Dependencies

Let $pdb = (\mathbf{W}, Pr)$ be a probabilistic database.

- Positive implication is defined as in incomplete databases

$\Rightarrow$ A tuple t_s is positively implicated by a **maybe-tuple** t_r if

$$\forall W \in \mathbf{W}\colon\ t_r \in W \Rightarrow t_s \in W$$

- In this case, it holds that:

$$p_\wedge(\{t_r, t_s\}) = p(t_r)$$

and

$$p_\vee(\{t_r, t_s\}) = p(t_s)$$

Probabilistic Databases - Tuple Dependencies

Positive Implication:

W_1	t_r	t_s	$p(t_r)$
W_2	t_r	ø	0
W_3	ø	t_s	$p(t_s)$-$p(t_r)$
W_4	ø	ø	1-$p(t_s)$

$$Pr(W_1) = p(t_r)$$
$$Pr(W_3) = p(t_s) - Pr(W_1) = p(t_s) - p(t_r)$$
$$Pr(W_4) = 1 - p(t_s)$$

$$p_\wedge(\{t_r, t_s\}) = Pr(W_1) = p(t_r)$$
$$p_\vee(\{t_r, t_s\}) = Pr(W_1) + Pr(W_3) = p(t_s)$$

Probabilistic Databases - Tuple Dependencies

Let $pdb = (\mathbf{W}, Pr)$ be a probabilistic database.

- Negative implication is defined as in incomplete databases

$\Rightarrow$ A tuple t_s is negatively implicated by a **maybe-tuple** t_r if

$$\forall W \in \mathbf{W}\colon\ t_r \notin W \Rightarrow t_s \in W$$

- In this case, it holds that:

$$p_{\wedge}(\{t_r, t_s\}) = p(t_r) + p(t_s) - 1$$

and

$$p_{\vee}(\{t_r, t_s\}) = 1$$

Probabilistic Databases - Tuple Dependencies

Negative Implication:

W_1	t_r	t_s	$p(t_s)+p(t_r)-1$
W_2	t_r	ø	$1-p(t_s)$
W_3	ø	t_s	$1-p(t_r)$
W_4	ø	ø	0

$$Pr(W_3) = 1 - p(t_r)$$
$$Pr(W_1) = p(t_s) - Pr(W_3) = p(t_s) - (1 - p(t_r)) = p(t_s) + p(t_r) - 1$$
$$Pr(W_2) = p(t_r) - Pr(W_1) = p(t_r) - (p(t_s) + p(t_r) - 1) = 1 - p(t_s)$$

$$p_\wedge(\{t_r, t_s\}) = Pr(W_1) = p(t_r) + p(t_s) - 1$$
$$p_\vee(\{t_r, t_s\}) = 1 - Pr(W_4) = 1$$

Probabilistic Databases - Tuple Dependencies

Mutual Exclusion:

- The tuples of a set T are *mutually exclusive* if they are pairwise exclusive, i.e.

$$\forall T' \subseteq T, |T'| > 1 \colon \nexists W \in \mathbf{W} \colon T' \subseteq W$$

- In this case, it holds that:

$$p_{\wedge}(T') = 0 \qquad \text{and} \qquad p_{\vee}(T') = \sum_{t \in T'} p(t)$$

for every subset $T' \subseteq T$ with $|T'| > 1$.

Probabilistic Databases - Tuple Dependencies

Mutual Independence:

- The tuples of a set T are *mutually independent*, if
  $$p_\wedge(T') = \prod_{t \in T'} p(t)$$
  for every subset $T' \subseteq T$.

- In this case, it holds that:
  $$p_\vee(T') = 1 - \prod_{t \in T'}(1 - p(t))$$
  for every subset $T' \subseteq T$.

- Mutual independence cannot be concluded from pairwise independence

$$\left.\begin{array}{l} p_\wedge(\{t_r, t_s\}) = p(t_r) \times p(t_s) \\ p_\wedge(\{t_r, t_u\}) = p(t_r) \times p(t_u) \\ p_\wedge(\{t_s, t_u\}) = p(t_s) \times p(t_u) \end{array}\right. \quad \not\Rightarrow \quad p_\wedge(\{t_r, t_s, t_u\}) = p(t_r) \times p(t_s) \times p(t_u)$$

Possible Worlds Semantics

Idea:

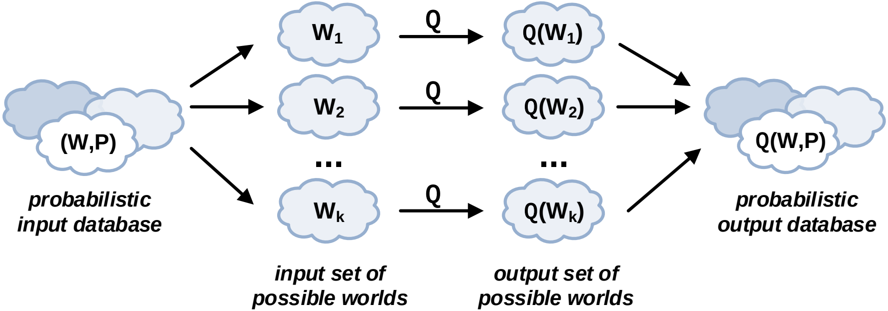
- A query is evaluated in each possible world separately
- Query result is another set of possible worlds
- Identical output worlds are combined (probabilities are summed up)

Definition: Let $pdb = (\mathbf{W}, Pr)$ be a probabilistic database and let Q be a conventional database query, the result of posing Q to *pdb* under the *possible worlds semantics* is the probabilistic database $Q(pdb) = (\mathbf{W}_Q, Pr_Q)$ where

$$\mathbf{W}_Q = \{Q(W) \mid W \in \mathbf{W}\}$$

and $$\forall W \in \mathbf{W}_Q\colon Pr_Q(W) = \sum_{W' \in \mathbf{W}, Q(W')=W} Pr(W')$$

Possible Worlds Semantics -Query principle



Possible Worlds Semantics - Example

Step 1: Evaluate the query in each world separately

Step 2: Combine identical worlds

W_1, Pr=0.5

WK	*name*	*age*
p1	*J.Doe*	*27*
p2	*K.Smith*	*32*
p3	*J.Ho*	*28*
p4	*J.J.Doe*	*31*

W_2, Pr=0.4

WK	*name*	*age*
p1	*J.Doe*	*27*
p2	*K.Smith*	*32*
p3	*J.Ho*	*28*
p4	*J.J.Doe*	*41*

W_3, Pr=0.1

WK	*name*	*age*
p1	*J.Doe*	*27*
p2	*K.Smith*	*32*
p4	*J.J.Doe*	*41*

```
SELECT *
FROM Person
WHERE age<30
```

Possible Worlds Semantics - Example

Step 1: Evaluate the query in each world separately

Step 2: Combine identical worlds

W_1, Pr=0.5

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	31

W_2, Pr=0.4

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p3	J.Ho	28
p4	J.J.Doe	41

W_3, Pr=0.1

WK	name	age
p1	J.Doe	27
p2	K.Smith	32
p4	J.J.Doe	41

```
SELECT *
FROM Person
WHERE age<30
```

Possible Worlds Semantics - Example

Step 1: Evaluate the query in each world separately

Step 2: Combine identical worlds

W'_1, Pr=0.5

<u>WK</u>	name	age
p1	J.Doe	27
p3	J.Ho	28

W'_2, Pr=0.4

<u>WK</u>	name	age
p1	J.Doe	27
p3	J.Ho	28

W'_3, Pr=0.1

<u>WK</u>	name	age
p1	J.Doe	27

```
SELECT *
FROM Person
WHERE age<30
```

Possible Worlds Semantics - Example

Step 1: Evaluate the query in each world separately

Step 2: Combine identical worlds

W'_1, Pr=0.9

WK	name	age
p1	J.Doe	27
p3	J.Ho	28

W'_3, Pr=0.1

WK	name	age
p1	J.Doe	27

```
SELECT *
FROM Person
WHERE age<30
```