



Minimal Disclosure in Hierarchical Hippocratic Databases with Delegation

Nicola Zannone

Dep. of Information and Communication Technology

University of Trento

joint work with Fabio Massacci and John Mylopoulos



Motivation

- Privacy and data protection is becoming essential in IS
 - Customers viewpoint
 - Protect their personal sensible information
 - Enterprises viewpoint
 - Losing market share
- Many countries have promulgated privacy legislation
 - The US Privacy Act of 1974
 - The EU Directives on Privacy of 1995
- Privacy principles
 - Purpose specification
 - Consent
 - Minimal collection
 - Minimal disclosure
 - Limited retention

Running Example

- **Mississippi: an on-line bookseller**
 - **Purchase:** delivery, credit assessment, and notification
 - **Delivery:** direct delivery or by post
 - **Notification:** by email or by fax
 - **Information:** name, shipping address, and credit card info
- **Worldwide Express (WWE_x): a delivery company**
 - **Direct delivery:** door-to-door delivery
 - **Information:** name, shipping address
- **Local Delivery Companies (LDCs):**
 - **Door-to-door delivery**
 - **Information:** name, shipping address
- **Credit Card Company (CCC):**
 - **Credit assessment:** credit rating, credit resolution
 - **Information:** name, credit card info, transaction
- **Credit Rating Company (CRC):**
 - **Credit rating**
 - **Information:** credit card info, transaction



Summary

- Hippocratic Databases
- Hierarchy of Purposes
- Minimum Cost Algorithms
- Minimal Authorization Table
- Conclusion and future work

Hippocratic Databases

- Privacy-aware technology
- Use purpose as a central concept
 - special attribute occurring in every table of the database
 - associated with each piece of data
- Together with purpose, collect
 - external-recipients: the actors to whom data items can be disclosed
 - retention-period: the period during which data items should be maintained in the database
 - authorized-users: the users entitled to access data items
- Metadata Schema
 - Privacy Policy Table
 - Privacy Authorization Table



Privacy Policy Table

- Contains the privacy policies of the enterprise
- Stores
 - purpose
 - external-recipients
 - retention-period

Privacy Policy Table

purpose	table	attribute	external-recipients	retention
purchase	customer	name	<i>empty</i>	1 month
purchase	customer	address	<i>empty</i>	1 month
purchase	customer	email	<i>empty</i>	1 month
purchase	customer	fax-number	<i>empty</i>	1 month
purchase	customer	credit-card-info	<i>empty</i>	1 month
purchase	order	transaction	<i>empty</i>	1 month
purchase	order	status	<i>empty</i>	1 month
delivery	customer	name	<i>empty</i>	1 month
delivery	customer	address	<i>empty</i>	1 month
direct delivery	customer	name	{ delivery-company }	1 month
direct delivery	customer	address	{ delivery-company }	1 month
delivery by post	customer	name	{ post-office }	1 month
delivery by post	customer	address	{ post-office }	1 month
credit-assessment	customer	name	{ credit-card-company }	1 month
credit-assessment	customer	credit-card-info	{ credit-card-company }	1 month
credit-assessment	order	transaction	{ credit-card-company }	1 month
notification	customer	name	<i>empty</i>	1 month
notification	customer	email	<i>empty</i>	1 month
notification	customer	fax-number	<i>empty</i>	1 month
notification	order	status	<i>empty</i>	1 month
notification by email	customer	name	<i>empty</i>	1 month
notification by email	customer	email	<i>empty</i>	1 month
notification by email	order	status	<i>empty</i>	1 month
notification by fax	customer	name	<i>empty</i>	1 month
notification by fax	customer	fax-number	<i>empty</i>	1 month
notification by fax	order	status	<i>empty</i>	1 month



Privacy Authorization Table

- Contains the access controls policies that implement privacy policies
- Represents the effective disclosure of information
- Created from Privacy Policy Table by instantiating each external recipient with the corresponding users.
- Stores
 - purpose
 - authorized-users

Privacy Authorization Table

purpose	table	attribute	authorized-users
purchase	customer	name	{ Mississippi }
purchase	customer	address	{ Mississippi }
purchase	customer	email	{ Mississippi }
purchase	customer	fax-number	{ Mississippi }
purchase	customer	credit-card-info	{ Mississippi }
purchase	order	transaction	{ Mississippi }
purchase	order	status	{ Mississippi }
delivery	customer	name	{ Mississippi }
delivery	customer	address	{ Mississippi }
direct delivery	customer	name	{ Mississippi, WWEx }
direct delivery	customer	address	{ Mississippi, WWEx }
delivery by post	customer	name	{ Mississippi, Post Office }
delivery by post	customer	address	{ Mississippi, Post Office }
credit-assessment	customer	name	{ Mississippi, CCC }
credit-assessment	customer	credit-card-info	{ Mississippi, CCC }
credit-assessment	order	transaction	{ Mississippi, CCC }
notification	customer	name	{ Mississippi }
notification	customer	email	{ Mississippi }
notification	customer	fax-number	{ Mississippi }
notification	order	status	{ Mississippi }
notification by email	customer	name	{ Mississippi }
notification by email	customer	email	{ Mississippi }
notification by email	order	status	{ Mississippi }
notification by fax	customer	name	{ Mississippi }
notification by fax	customer	fax-number	{ Mississippi }
notification by fax	order	status	{ Mississippi }

Beyond Hippocratic Databases

- Complex strategies
 - Enterprises may provide their services in different ways
 - Each different method could require different data
- Dynamic coalitions & Delegation of Information
 - Business process may be not executed by a single enterprise
 - We can have a host of partners
 - Enterprises may outsource some information to partners
 - Different partners can offer the same service

Beyond Hippocratic Databases (II)

- Agrawal's proposal
 - Split a purpose into multiple purposes and store them in the database
 - Opt-in and opt-out data items for a certain purpose
- This solution cannot be used to reason about the fulfillment of the root purpose
 - The system may collect a set of information that is not sufficient to fulfill the root purpose



Hierarchy of Purposes

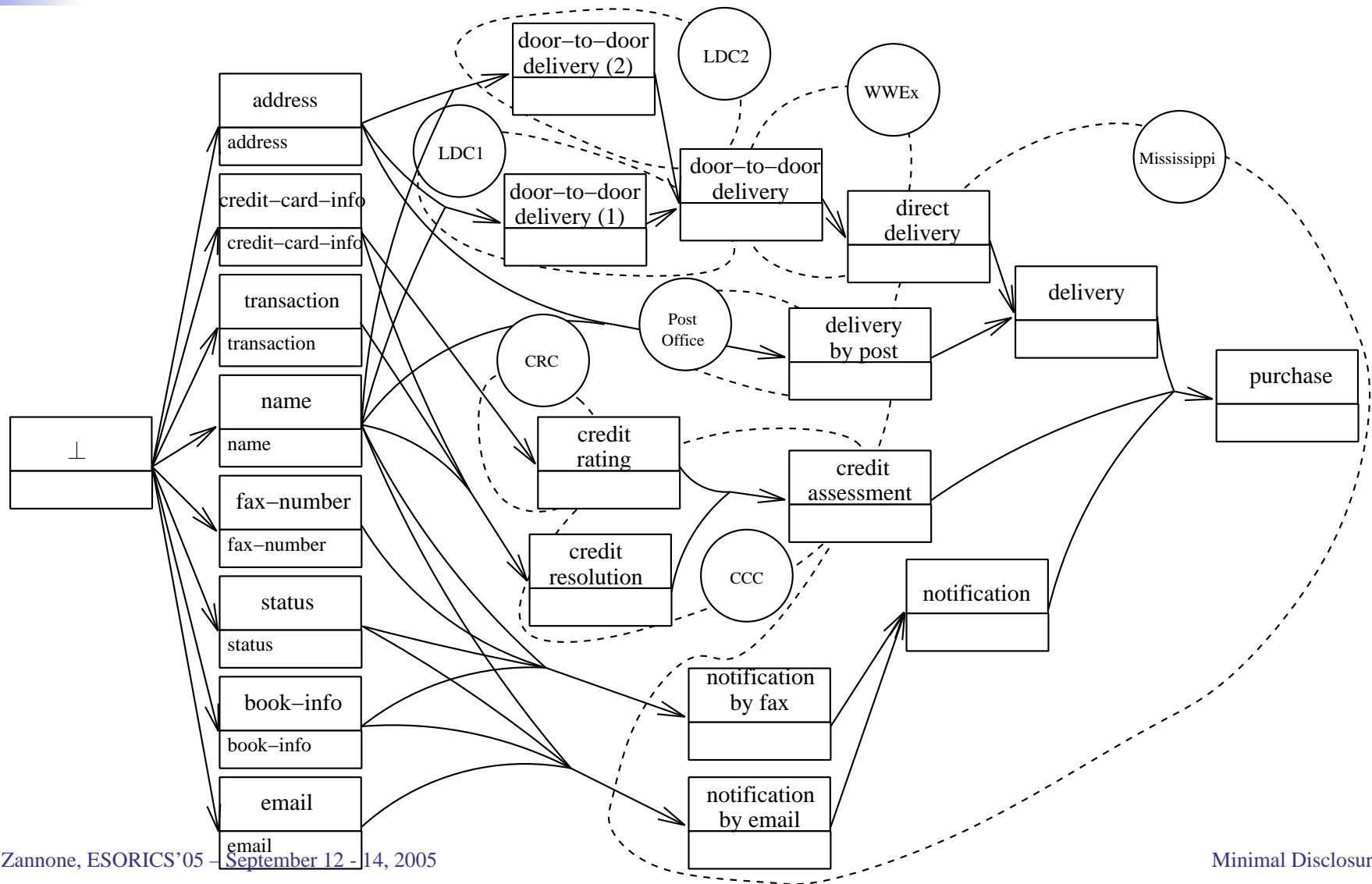
- Goal analysis

- Decomposing purposes through an AND/OR refinement
- If a purpose is AND-decomposed, then all of its sub-purposes must be fulfilled to fulfill it
- If a purpose is OR-decomposed, then at least one of its sub-purposes must be fulfilled to fulfill it

From Hippocratic DBs to Purpose DAGs

- Individual Partner's Privacy Policy Table
 - Purposes are analyzed through a goal refinement process
 - Build a purpose hierarchy (or purpose DAG)
- Privacy policy of the entire business process
 - Merge purpose DAGs associated with each partner
 - Delegation arcs link nodes across PPTs
- Purposes node are linked to the data items
 - Data item nodes are linked to a source node
- Privacy penalty is associated with arcs
 - Arcs joining source node and data item nodes
 - Delegation arcs

Purpose DAG



Minimum Cost Algorithms

- Customers want to know which is process that more protect their privacy
 - A path represents a possible process to fulfill a service
 - Find the minimum cost path from the source to the root
- Different cost functions can be used to measure the same path
 - The cost of a path is the sum of the weights of its arcs
 - Minimum cost set of data items
 - The cost of one edge is counted as many times as it is traversed
 - Effective use of information
 - More a datum is used, more it might be compromised

Selection of Privacy Preferences

■ Off-line Requirements Capture

- initialization
- delete arcs
- add arcs
- increase weights
- decrease weights

■ On-Line Privacy Assessment

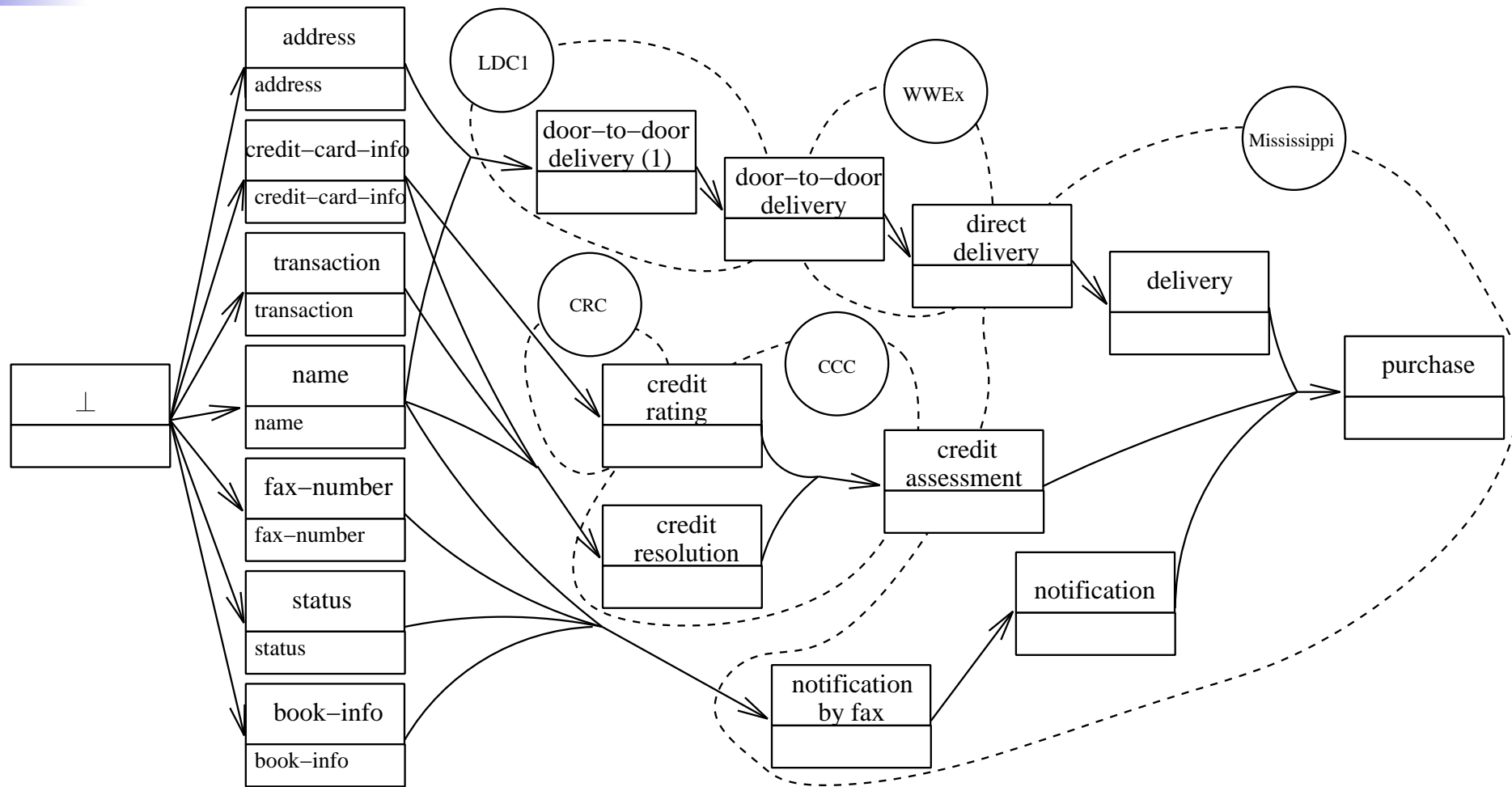
- delete arcs
- increase weights
- decrease weights
- ~~add arcs~~
 - Customers cannot impose a new method for delivering a service
 - Customers cannot add a partner to a business process
 - Solve problems such as system integration and commercial agreement
- minimum cost path cannot be computed by the enterprise
 - Each customer may associate a different privacy penalty with the same data item

Initialization

- Find the minimum cost path of a new business process
- Default preferences
 - Statistical evidence over the customer choices
 - Marketing strategies

Data Item	Cost	Delegation	Cost
name	1	CCC	2
address	5	CRC	4
email	4	WWEx	2
fax-number	2	LDC ₁	2
credit-card-info	10	LDC ₂	3
transaction	5	Post Office	5
book-info	2		
status	3		

Minimum Cost Path



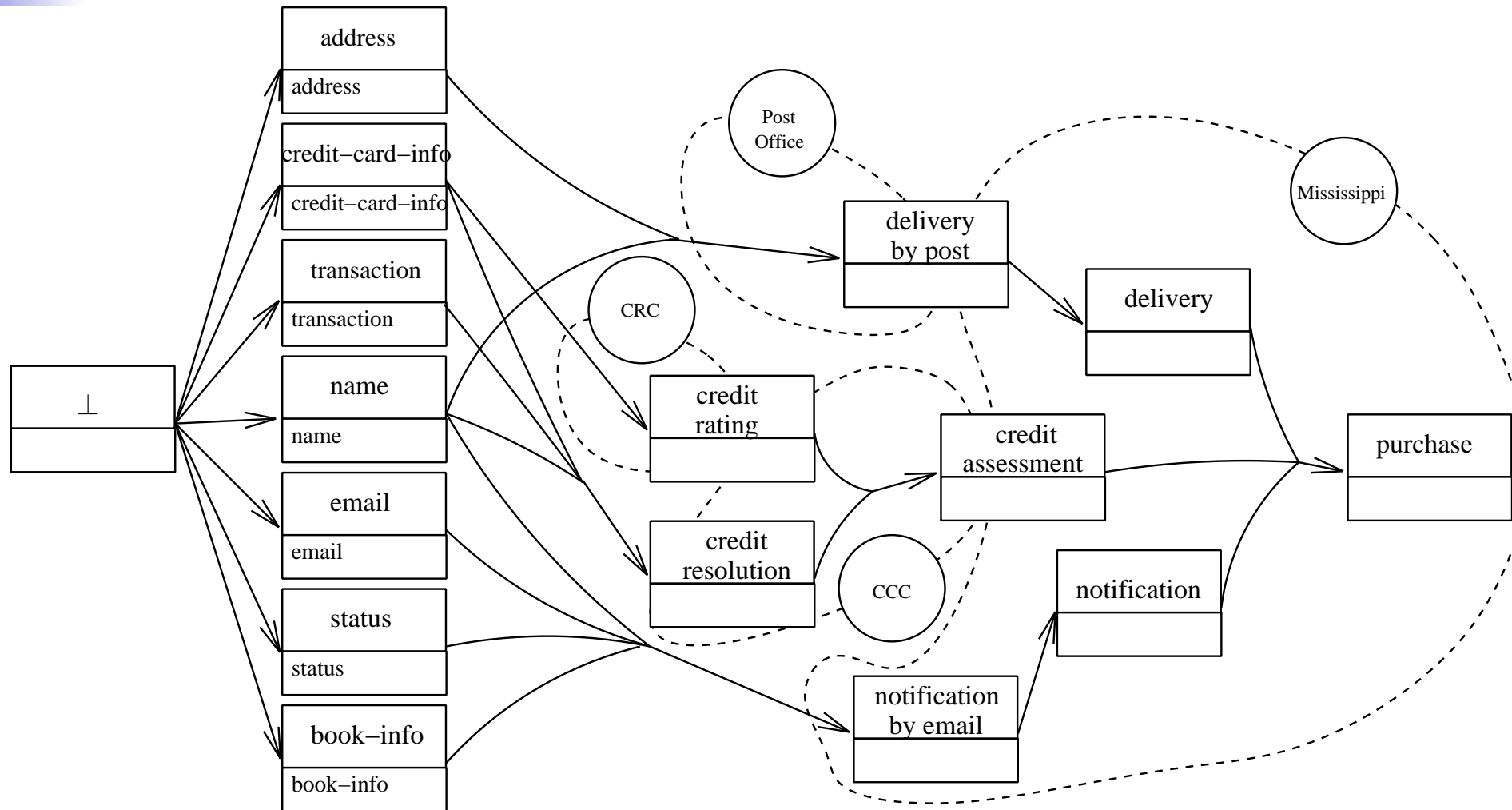
On-the-fly Update

- Privacy assessment phase requires that data structures are maintained and that operations are performed on-line
 - Avoid to recompute the entire path
 - Reuse the valid part of the old solution as much as possible
- The problem of dynamically updating the purpose DAG can be divided in two distinct classes
 - Adding new arcs or decreasing the privacy penalty of an existing arc
 - Deleting an existing arc or increasing the privacy penalty of an existing arc

Example

Data Item	Cost	Delegation	Cost
name	1	CCC	2
address	5	CRC	4
email	4	WWEx	∞
fax-number	20	LDC ₁	2
credit-card-info	10	LDC ₂	3
transaction	5	Post Office	5
book-info	2		
status	3		

Minimum Cost Path



Minimal Authorization Table

- The minimum cost path is used to build the minimal authorization table
- Minimal authorization table describes the access control of the entire business process
 - Minimum set of authorizations needed to fulfill the service
- Authorization table created only if a path exists
 - Information disclosed only if the purpose can be fulfilled
- Authorizations are the minimum cost set
 - Individual customer preferences

Minimal Authorization Table

purpose	table	attribute	authorized-users
purchase	customer	name	{ Mississippi }
purchase	customer	address	{ Mississippi }
purchase	customer	fax-number	{ Mississippi }
purchase	customer	credit-card-info	{ Mississippi }
purchase	order	transaction	{ Mississippi }
purchase	order	status	{ Mississippi }
delivery	customer	name	{ Mississippi }
delivery	customer	address	{ Mississippi }
direct delivery	customer	name	{ WWEx }
direct delivery	customer	address	{ WWEx }
door-to-door delivery	customer	name	{ LDC ₁ }
door-to-door delivery	customer	address	{ LDC ₁ }
credit-assessment	customer	name	{ CCC }
credit-assessment	customer	credit-card-info	{ CCC }
credit-assessment	order	transaction	{ CCC }
credit scoring	customer	credit-card-info	{ CRC }
credit resolution	customer	name	{ CCC }
credit resolution	customer	credit-card-info	{ CCC }
credit resolution	order	transaction	{ CCC }
notification	customer	name	{ Mississippi }
notification	customer	fax-number	{ Mississippi }
notification	order	status	{ Mississippi }
notification by fax	customer	name	{ Mississippi }
notification by fax	customer	fax-number	{ Mississippi }
notification by fax	order	status	{ Mississippi }



Conclusions

- Improve Hippocratic approach for dynamic business processes
 - New Hippocratic Database model with Delegation
 - Algorithms for computing minimum set of authorizations during design
 - Algorithms for computing minimum set of authorizations by clients
- Future works
 - Actor hierarchy
 - e.g. company-division-department-individual worker
 - Ensure complete and correct answers to queries
 - Build global certificates

FD-graph

- Two types of nodes
 - Single nodes
 - Compound nodes
- Two type of arcs
 - OR-edge
 - AND-edge
- A decomposition arc is represented by a compound node with a leaving OR-edge and one or more incoming AND-edges

MinimumCost

Algorithm MinimumCost

begin

make- PQ -empty;

insert arc $\langle \perp, \perp \rangle$ into PQ

mark source node \perp as visited

while PQ -nonempty **do begin**

extract from the queue PQ the node t with minimum priority

initialize node t

for each OR-edge $\langle t, x \rangle$ leaving node t **do** ScanMC(t, x);

for each AND-edge $\langle t, z \rangle$ leaving node t **do begin**

decrement($TODO[z]$);

if node z is marked visited **then begin**

compute disclosure penalty of z

compute list of data items needed to fulfil z

for each OR-edge $\langle z, x \rangle$ leaving node z **do** ScanMC(z, x);

end

end

end

end

ScanMC

```
Procedure ScanMC( $t$ : node;  $x$ : simple-node);  
begin  
  update disclosure penalty and list of data items associated with  $x$   
  if node  $x$  has not been previously visited  
    then begin  
      mark node  $x$  as visited  
      insert arc  $\langle t, x \rangle$  into PQ  
    end  
  else if arc  $\langle t, x \rangle$  improves minimal path  
    then priority associated with arc  $\langle t, x \rangle$  is decreased  
end
```

WeightIncrease

```
Procedure WeightIncrease( $\langle X, y \rangle$ : decomposition arc,  $\omega$ : weight);
begin
  if  $|X| = 1$ 
  then  $x :=$  the single element of  $X$ ;
  else  $x :=$  Compound( $X$ );
  if arc  $\langle x, y \rangle$  belongs to minimal path then begin
    update disclosure penalty associated with  $y$ 
    for each OR-edge  $\langle s, y \rangle$  incoming to node  $y$  do ScanIWI( $s, y$ );
    while PQ-nonempty do begin
      extract from the queue PQ the node  $t$  with minimum priority
      initialize node  $t$ 
      for each OR-edge  $\langle t, x \rangle$  leaving node  $t$  do
        if arc  $\langle t, x \rangle$  belongs to minimal path then
          for each OR-edge  $\langle s, x \rangle$  incoming to node  $x$  do ScanIWI( $s, x$ );
      for each AND-edge  $\langle t, z \rangle$  leaving node  $t$  do begin
        compute disclosure penalty of  $z$ 
        compute list of data items needed to fulfill  $z$ 
        if arc  $\langle t, z \rangle$  improves minimal path then
          update disclosure penalty and list of data items associated with  $z$ 
          for each OR-edge  $\langle z, x \rangle$  leaving node  $z$  do
            if arc  $\langle z, x \rangle$  belongs to minimal path
              for each OR-edge  $\langle s, x \rangle$  incoming to node  $x$  do ScanIWI( $s, x$ );
      end
    end
  end
end
```

ScanlWI

```
Procedure ScanlWI( $t$ : node;  $x$  : simple-node);  
begin  
  update disclosure penalty and list of data items associated with  $x$   
  if arc  $\langle t, x \rangle$  improves minimal path  
    then if  $\langle t, x \rangle \notin PQ$   
      then insert arc  $\langle t, x \rangle$  into PQ  
    else priority associated with arc  $\langle t, x \rangle$  is decreased  
end
```