Types of Mapping

- Approaches to classifying mappings
  - Structural
  - Semantic
- Examples of mappings
- What does this tell us?

Approaches to classifying mappings

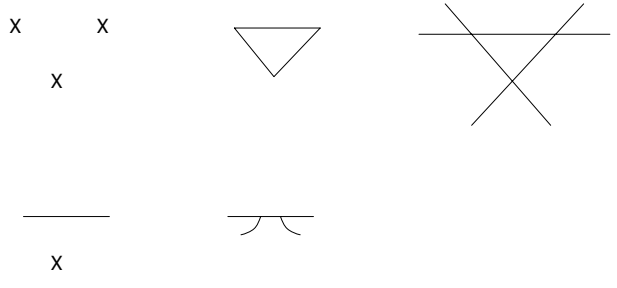
- Still a research problem
  - No complete understanding of 'hardness' of the problem
  - No complete understanding of descriptive power required for mapping languages for different tasks
  - Two viewpoints are to consider structures that can be mapped between, or semantic differences between structures
    - Neither gives a full picture of the problem
- Many very simple mappings are impossible
  - E.g., width * depth = area
  - What types of mapping are not possible?

Structural mappings

- What sorts of data exist?
  - Objects: entity, row, structure, etc
  - Attributes
- In what numbers do they come?
  - Zero: nothing or deleted
  - One
  - Constant
  - Multiple
- Take the cross product of these axes

van Horssen, J.J., Behage, B. and Mooij, M. (1994) Conversion, ESPRIT 7280 - ATLAS, Confidential report.

Representations of Triangles



Structural mappings

	Object→Object	Attr→Attr	Object→Attr	Attr→Object
1→0	*	*		
0→1	*	*		
1→1	*	*	*	*
1→C	*	*	*	*
C→1	*	*	*	*
C→B	*	*	*	*
1→N	*			*
N→1	*		*	
C→N	*			*
N→C	*		*	
N→M	*			

C, B are constant and N, M are variable

Object→Object examples

1→0	Fees object not mapped for academic record display
0→1	Building object created in CAD when mapped from VRML (just geometry)
1→1	Course object in nDeva maps to a course object in Cecil
1→C	Column object in CAD mapped to 6 surface objects in VRML viewer
C→1	Book and publisher object mapped to a single BookInfo object
C→B	Product, manufacturer, supplier mapped to catalog entry and sale point
1→N	Polygon object mapped to a number of line objects
N→1	All course taken objects mapped to a student GPA object
C→N	Shopping basket and customer in broker mapped to many orders
N→C	A day's sale objects mapped to a purchase order and a credit notice
N→M	N points mapped to N(N-1) lines for a connected graph

Attr→Attr examples

1→0	
0→1	
1→1	
1→C	
C→1	
C→B	
1→N	N/A
N→1	N/A
C→N	N/A
N→C	N/A
N→M	N/A

Object→Attr examples

1→0	N/A
0→1	N/A
1→1	Name object mapped to name attribute
1→C	Address object mapped to 4 address fields
C→1	2 point objects mapped to length
C→B	4 point objects mapped to perimeter and area
1→N	N/A
N→1	A month's credit card items mapped to a total expenditure field
C→N	N/A
N→C	A server's request log mapped to total requests and requests per hour
N→M	N/A

Attr→Object examples

1→0	N/A
0→1	N/A
1→1	
1→C	
C→1	
C→B	
1→N	
N→1	N/A
C→N	
N→C	N/A
N→M	N/A

Semantic mappings

- Examine the types of conflict that can occur
  - Categories of conflicts
  - Naming of objects and attributes
  - Structural differences
  - Data level differences

Batini, C., Lenzerini, M. and Navathe, S.B. (1986) A Comparative Analysis of Methodologies for Database Schema Integration, ACM Computing Surveys, 18(4), December, pp. 323-364.

Kim, W. and Seo, J. (1991) Classifying Schematic and Data Heterogeneity in Multidatabase Systems, IEEE Computer, 24(12), December, pp. 12-18.

Conflict categories

- Identical: everything is the same
- Equivalent
  - Where different but equivalent modeling constructs have been applied but the perceptions are still the same and are coherent
    - Behavioral: if the same set of answers to any given query can be obtained from all representations
    - Mapping: instances can be put on a 1-1 correspondence
    - Transformational: if a representation can be obtained by applying a set of atomic transformations that by definition preserve equivalence
- Compatible
  - Not identical or equivalent, but modeling constructs, designer perception and integrity constraints are not contradictory
- Incompatible
  - Contradictory because of the incoherence of the specification

Naming and data conflicts

- Naming conflicts
  - Homonyms: same name for different concepts (e.g., door height)
  - Synonyms: same concept described by different names
    - E.g., dwang and nog
- Data conflicts
  - Mismatched representations
    - Different units
    - Different precisions
  - Wrong data
    - Incorrect data entry
    - Obsolete data

Structural conflicts

- Type: same concept represented by different modeling constructs
  - Missing attributes
  - Missing but implicit attributes
  - Default value conflicts
  - Data type conflicts
  - Integrity constraint conflicts
- Dependency: group of concepts are related with different dependencies (e.g., 1-1 versus n-m)
- Key: different keys assigned to the same concept
- Behavioral: different insertion or deletion policies associated with the same class of object

What does this tell us?

- Generic categories of mappings we may face
- Generic types of conflicts that may occur
- Nothings about the hardness of particular mappings
  - E.g., how hard is it to map XML-based data, or EDI messages, or RDBMS, or…
- What else could we consider?