

# English Sentence Structures and EER Modelling

Sebastian Link



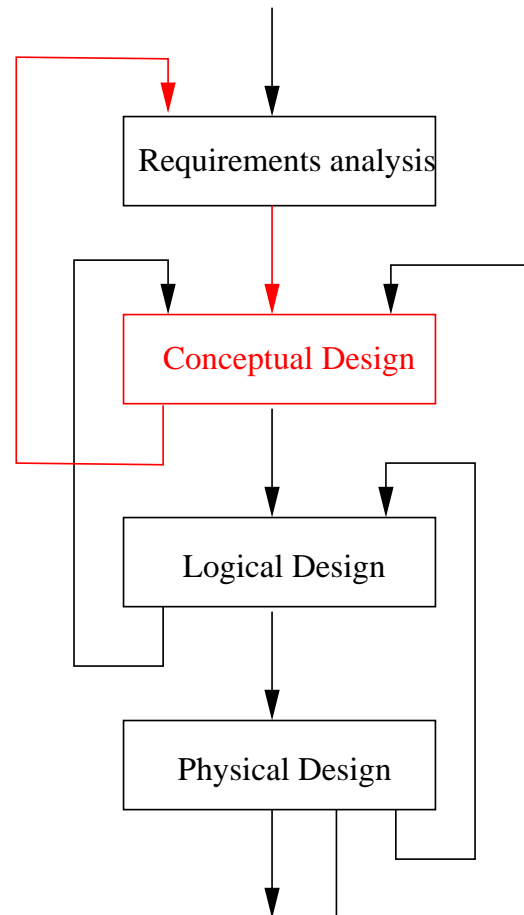
*School of Information Management,  
Victoria University of Wellington  
New Zealand*



This research is supported by the Marsden Fund Council from Government funding, administered by the Royal Society of New Zealand.

## Conceptual Modelling in the Design Phase

- ▶ ER can provide well-defined and natural features
- ▶ ER can provide safe features leading to good database design
- ▶ IDNF



- ▶ conceptual models must provide means for communication between designer and user
- ▶ NL constructs must find counterparts within model
- ▶ need guidelines for conversion
- ▶ provide justification of modelling features

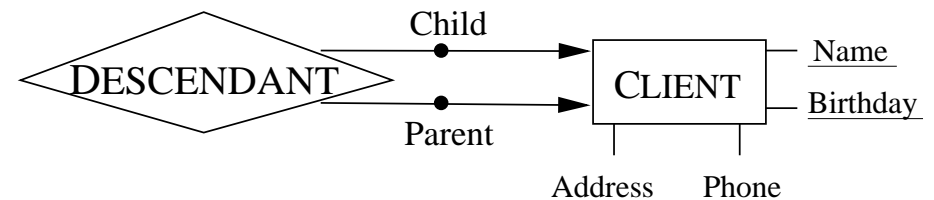
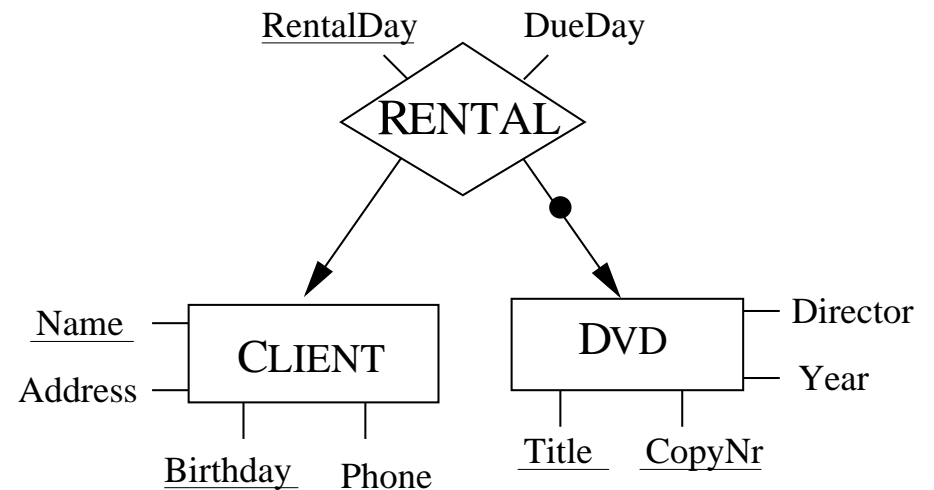
## Output of Requirements Analysis

*The database stores information about a university. Each person should have a name, and an address that person lives at. Students, general staff, lecturers and professors are all persons. Every student has a student ID and a majoring subject. For the general staff we keep track of their position. Lecturers have a department where they work and a certain teaching area. Professors are also associated with a department, and have several areas of research expertise. Graduate students are students with a degree and study a specific topic. They are supervised by either lecturers or professors within a semester. Courses have a course number and a title, and are taught by a number of lecturers and professors within a certain semester. The course co-ordinators teach courses on the basis of a list of recommended textbooks each of which comes with a title and an ISBN.*

- ▶ Objective: find natural counterparts to basic constructs such as
  - ↳ nouns, verbs, adjectives, and adverbs
  - ↳ gerund, clauses, sentences, text
- ▶ challenge: sentences depend on each other

## ER Features: Entity and Relationship Types

- ▶  $E = (attr(E), id(E))$
- ▶  $R = (comp(R), attr(R), id(R))$
- ▶ *order* of object type  $O$  is
  - 0, if  $O$  entity type,
  - $k$ , if  $k - 1$  is max order of any component of  $O$



# EER Features: Specialisation and Generalisation

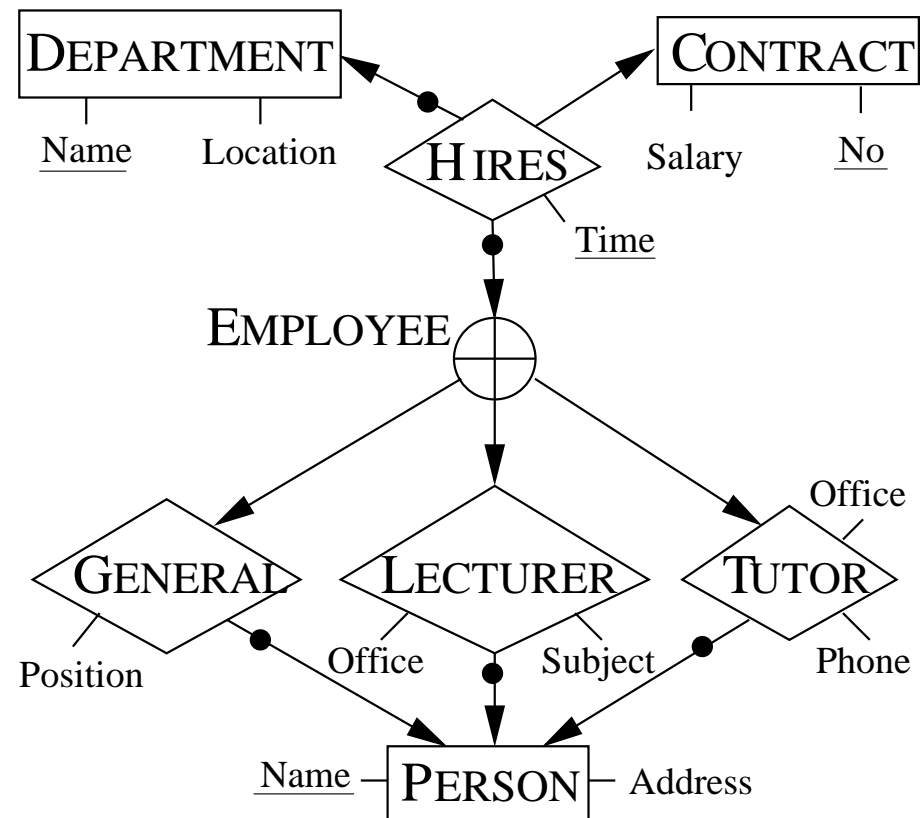
- specialisation:

$$S = (\{C\}, attr(R), \{C\})$$

- generalisation:

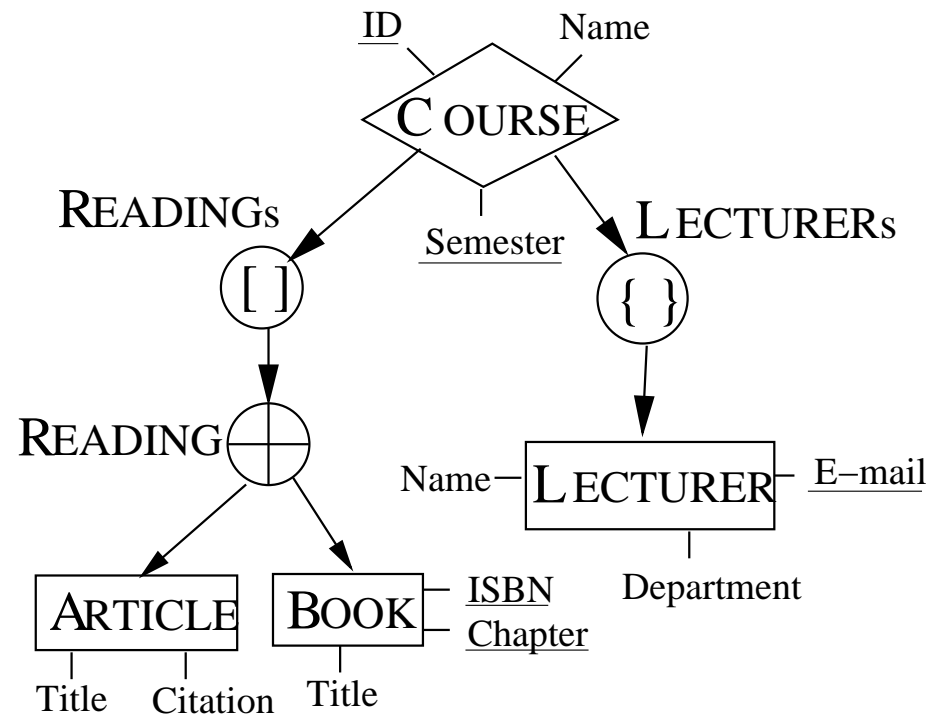
$$C = O_1 \oplus \dots \oplus O_k$$

with  $comp(C) = \{O_1, \dots, O_k\}$



## EER Features: Collection Types

- ▶ collections:
  - lists (duplicates, order),
  - sets (no duplicates, no order),
  - bags (duplicates, no order),
  - rankings (no duplicates, order)
- ▶ *list-, set-, bag-, ranking-type*  $U$  with  $comp(U) = \{C\}$ 
  - ↪ list type  $U[C]$
  - ↪ set type  $U\{C\}$
  - ↪ bag type  $U\langle C \rangle$
  - ↪ ranking type  $U[C]$



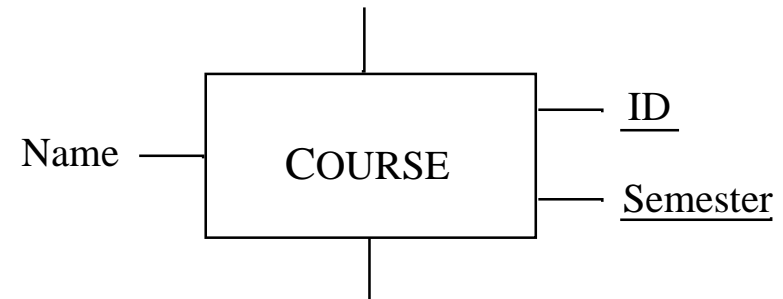
## EER Features: Nested Attributes

▶ flat attributes  $\mathcal{A}$ , labels  $\mathcal{L}$  and null attribute  $\lambda$

▶  $\mathcal{N}$  over  $\mathcal{A}$  and  $\mathcal{L}$ :

- $\mathcal{A} \subseteq \mathcal{N}$ ,  $\lambda \in \mathcal{N}$
- $N_1, \dots, N_k \in \mathcal{N}$ ,  $L \in \mathcal{L}$ :  
 $L(N_1, \dots, N_k) \in \mathcal{N}$ ,  
 $L(N_1 \oplus \dots \oplus N_k) \in \mathcal{N}$
- $N \in \mathcal{N}$ ,  $L \in \mathcal{L}$ :  
 $L[N], L\{N\}, L\langle N \rangle, L[N] \in \mathcal{N}$

LECTURERS<sub>s</sub> { LECTURER (Name,E-Mail,Department)}



READINGS<sub>s</sub> [ R EADING ( A RTICLE (Title,Citation)  $\oplus$   
 B OOK (ISBN,Chapter,Title)) ]

## EER Features vs. English Sentence Structures

English sentence concept	EER feature
transitive verb	relationship type
common noun	component of relationship type
adjective	attribute of component
adverb	attribute of relationship type
numerical expression	attribute of object type
preposition	role name of component
gerund	relationship type that is component of another relationship type
clause	relationship type with components
complex sentence	relationship type of order higher than 1
alternative phrase	cluster type
plural	collection type/nested attribute
“is a” sentence	specialisation

## Comparison to Chen's original correspondences

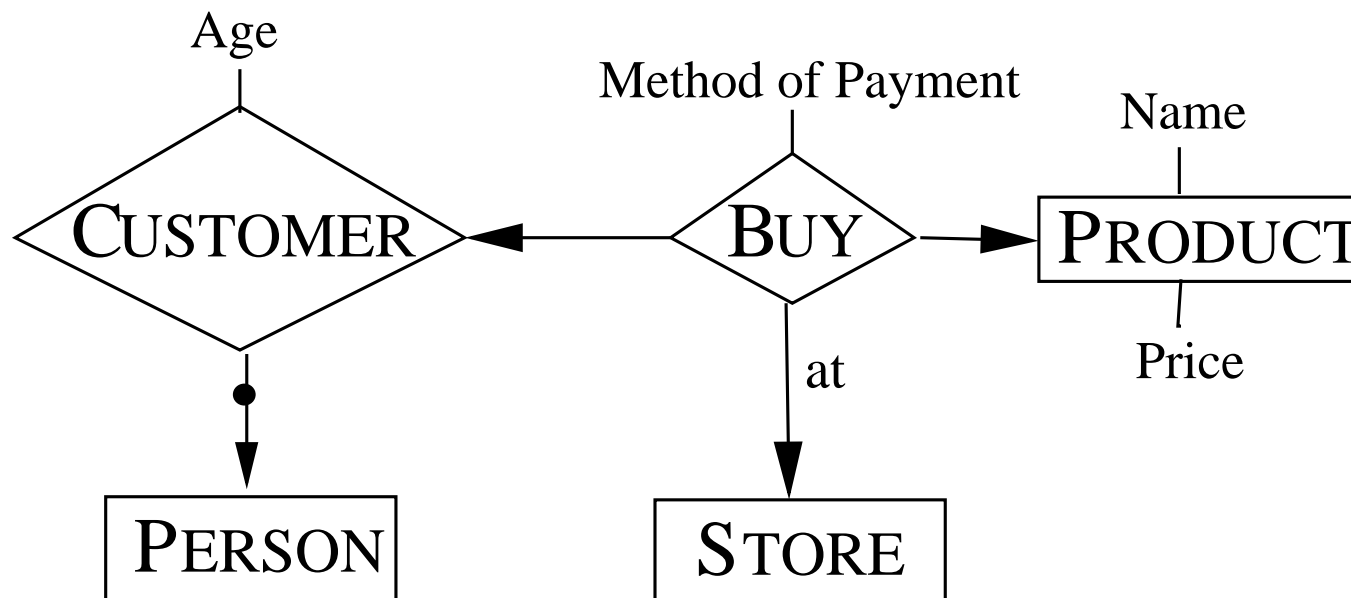
- ▶ Peter P.-S. Chen: *English Sentence Structure and ER Diagrams*, Inf. Sci. 29(2-3): 127-149, 1983

English sentence concept	ER feature
transitive verb	relationship type
common noun	entity type
adjective	attribute of entity type
adverb	attribute of relationship type
numerical expression	attribute of entity or relationship type
gerund	relationship-converted entity type
clause	high-level entity type abstracted from group of interconnected low-level entity and relationship types
complex sentence	one or more entity types connected by relationship type in which each entity type can be decomposed recursively into low-level entity types interconnected by relationship types

- ▶ EER reflects (English) sentence structures more soundly and naturally
- ▶ higher-order object types reflect dependence between sentences
- ▶ this provides justification for introduction of new ER features

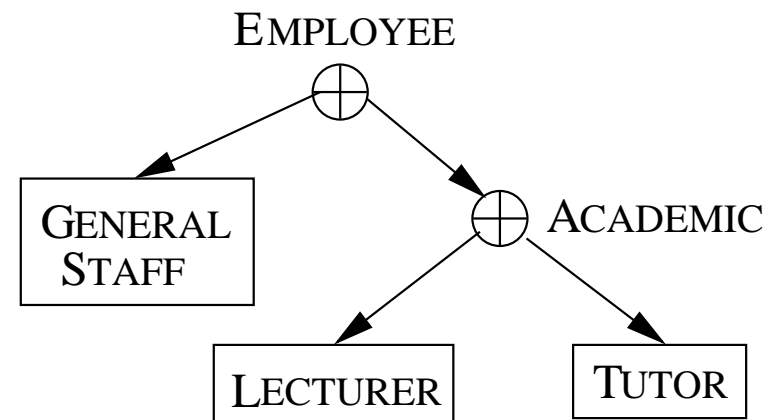
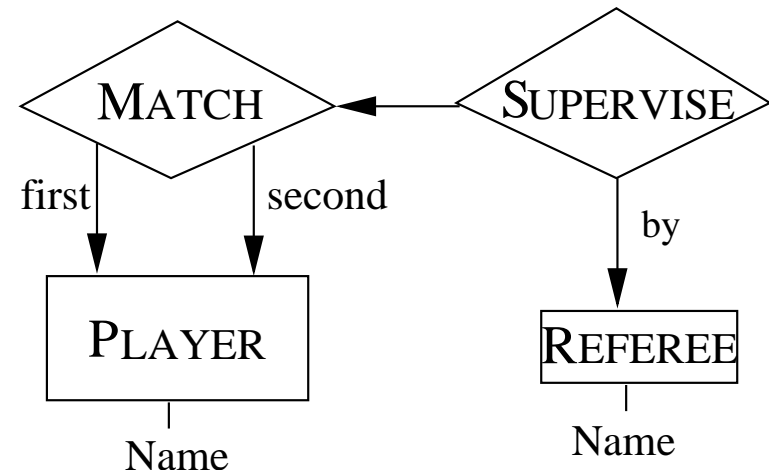
## Examples: Nouns, Verbs, Preps, Adjectives, Adverbs

- ▶ *A customer is a person who buys products at a store. A 25-year old customer buys a 200 dollar watch paying with her credit card.*



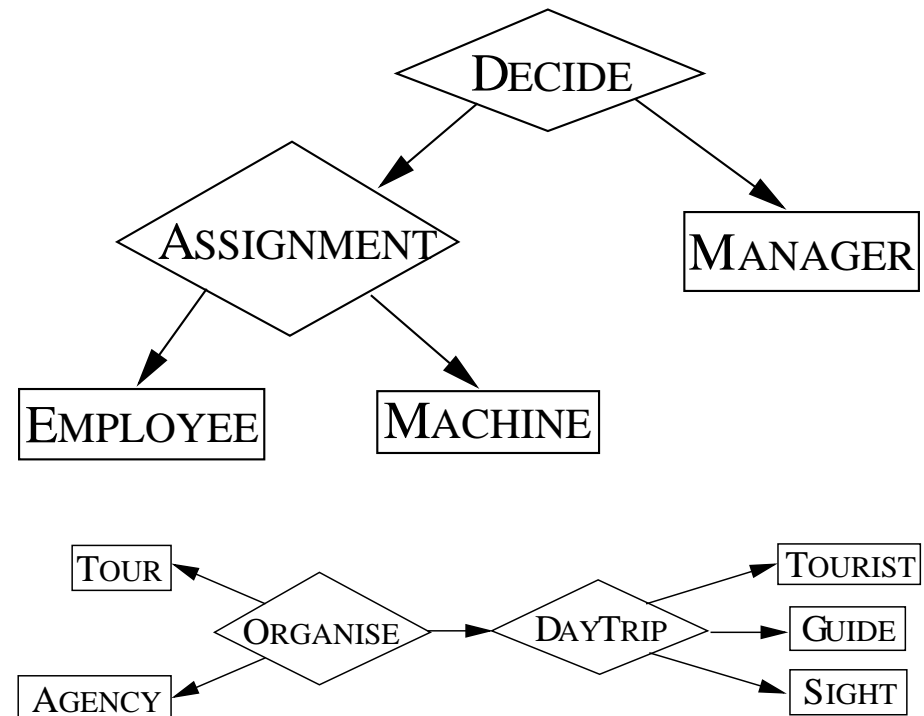
## Examples: Gerund and Alternatives

- ▶ *Nadal plays Federer, and playing is supervised by Lars Graf.*
- ▶ *Academics are either lecturer or tutors, and employees are either academics or general staff*



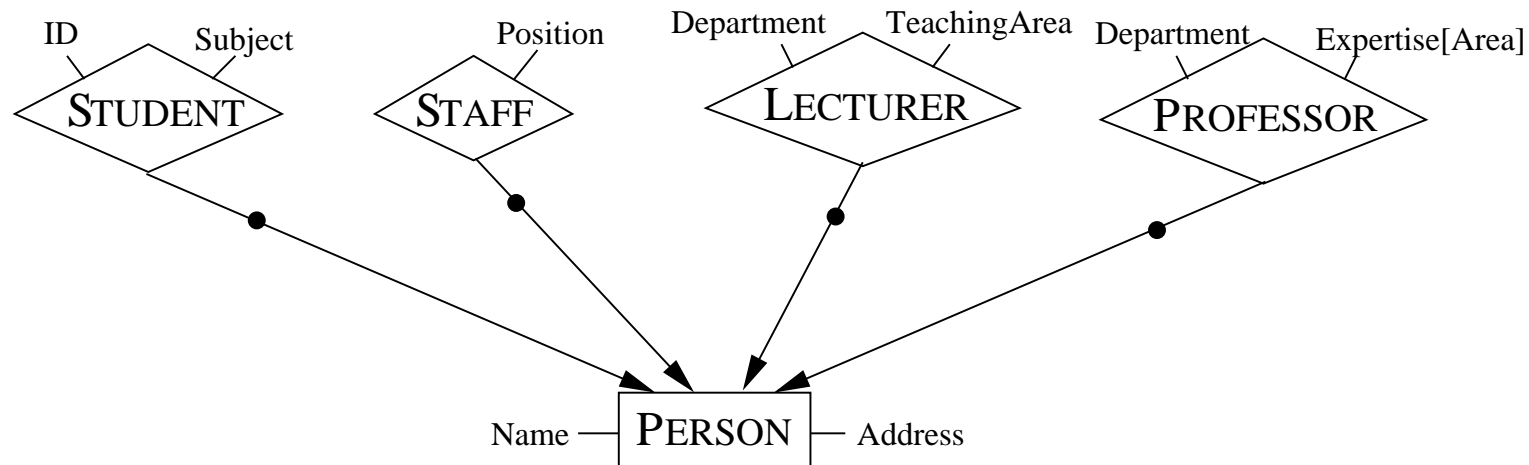
## Examples: Clauses and Sentences

- ▶ *Managers decide which machine is assigned to which employee.*
  
- ▶ *A tour is organised by a travel agency into day trips on which tourists visit various sights and are led by a tour guide.*



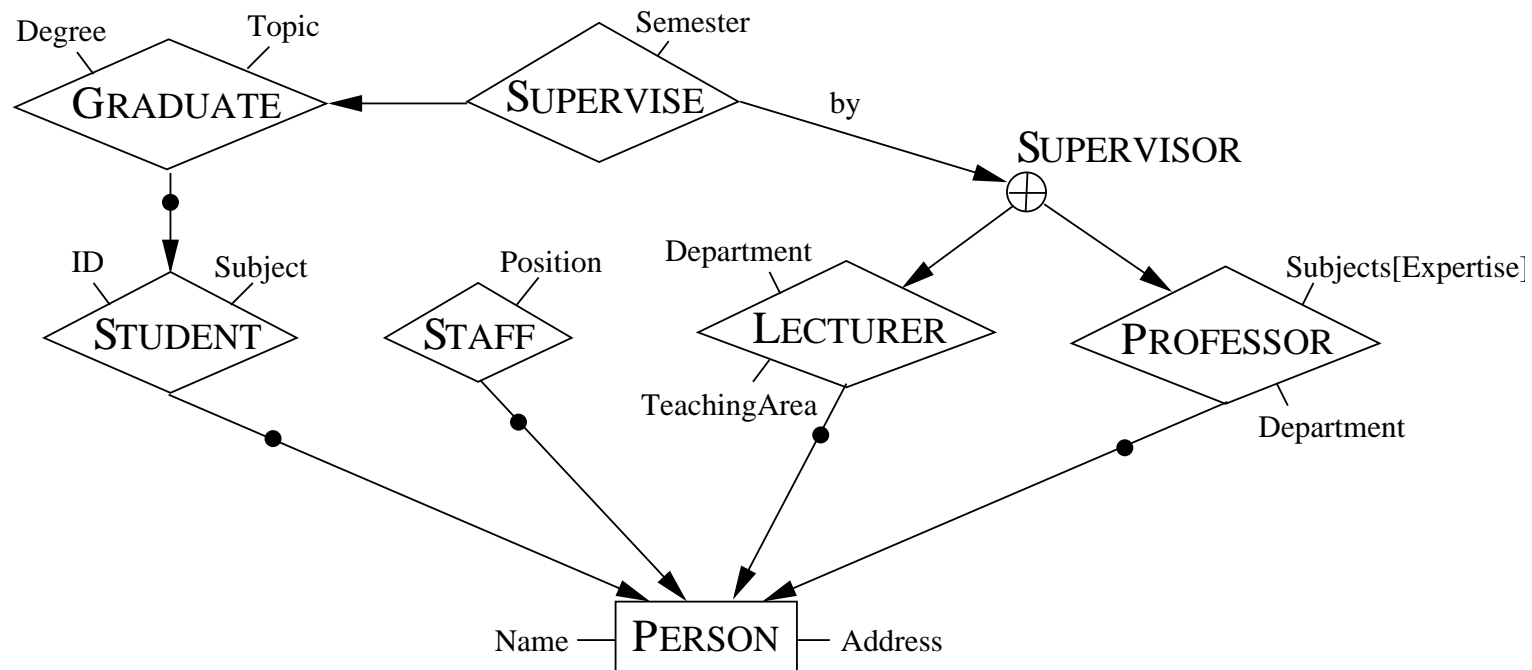
## The complex Example - Step 1

*Each person has a name. Each person has an address that person lives at. A student is a person. Each general staff is a person. A lecturer is a person. A professor is a person. Every student has a student ID and a majoring subject. Professors have several areas of research expertise.*



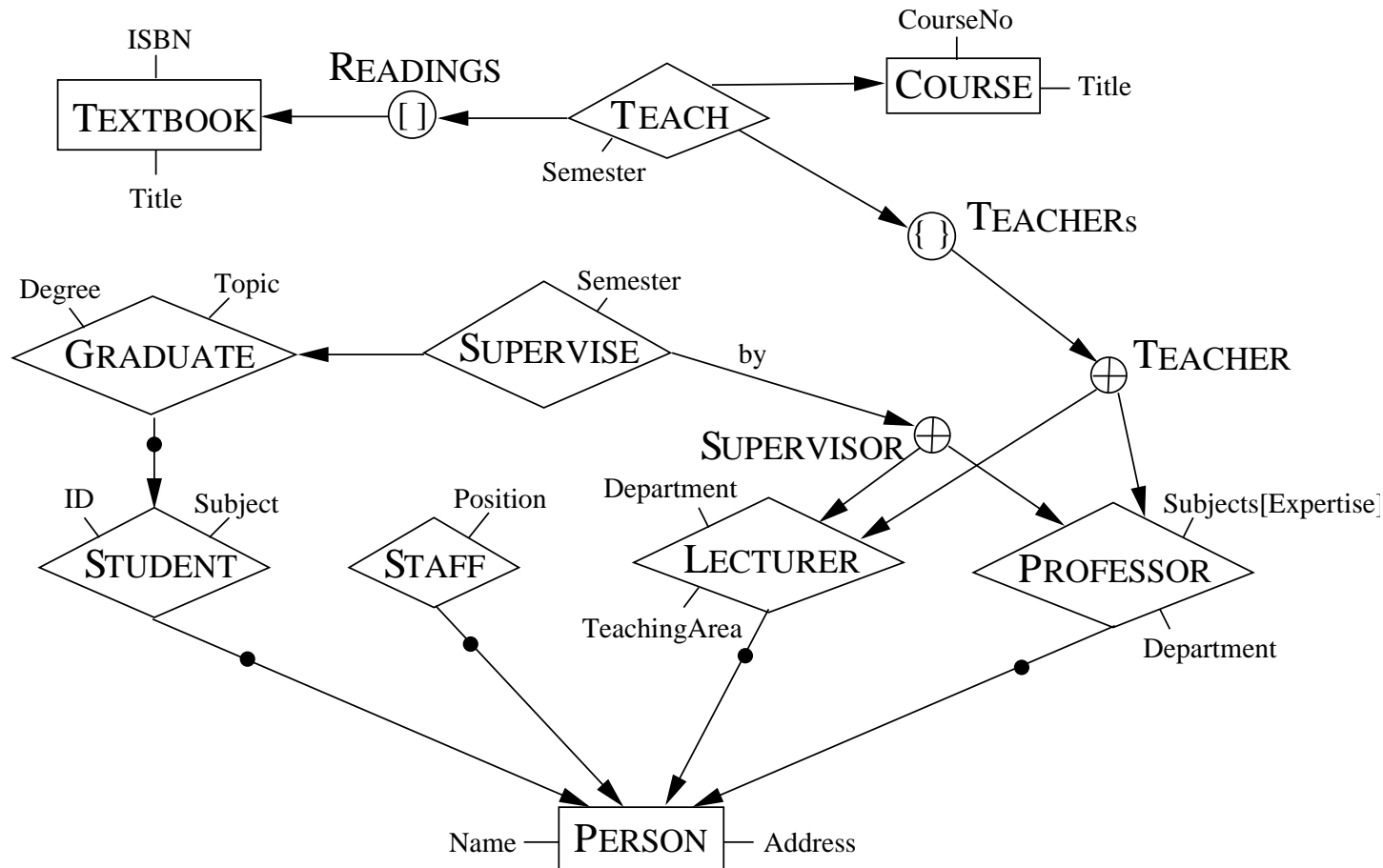
## The complex Example - Step 2

*A graduate student is a student. A graduate student has a degree and studies a specific topic. A supervisor is either a lecturer or a professor. Graduate students are supervised by a supervisor within a semester.*



## The complex Example - Step 3

*Courses have a course number and a title. Courses are taught by a number of teachers within a certain semester. A teacher is either a lecturer or a professor. A textbook has a title and an ISBN. Teachers teach courses on the basis of a list of recommended textbooks.*



## Conclusion

- ▶ ER model does not just provide safe constructs that result in good database design, but also features that enable good communication between designer and user
- ▶ essential to best *approximate* requirements
- ▶ additional EER features justified in the sense that modelling becomes more natural
- ▶ provides also a justification why the EER features exist
- ▶ higher-order object types reminiscent of nested sentence structure in natural language text

# Future Work?



## Questions?

