

THREE GREAT CONFERENCES IN ONE!



Conference Europe 2006
Data Management,
What Works?

META 2
DATA 0
Europe 0
Conference 6

NEW THIS YEAR: DW/BI TRACK

30 October - 2 November 2006, London, UK

The Trials of Data Modelling in UML

Dagna Gaythorpe

<http://www.seshat.com>

DAMA International Conference

London, 2006

Agenda

- Introduction
- On languages and linguistic shells
- Expressing our ideas in UML
- Where do we go from here?
- Bibliography
- Questions

Introduction - who am I?

Independent consultant - data models for utilities, telecommunications, accountancy, travel, loyalty cards and investment banking.

Member of DAMA UK. (On the committee, so please talk to me if you want to know more.)

What is UML

- Unified Modelling Language
 - it has a vocabulary
 - can be use to express concepts
 - cannot be used to express concepts that aren't in the language

Entity Relationship Diagrams

- ERDs are also in a language
- Graphical language
 - limited concepts
 - clear syntax
 - quite a few ‘dialects’
- Can be translated into other languages
 - RDBMS (e.g. PL/SQL)
 - English (and other ‘normal’ languages)

Linguistic Shells



Linguistic Shells

Not the sort of shell that you find on the beach!

We all grow up with a set of concepts and ideas. We know what they are, and how to express them. We have the words for them.

If we travel somewhere new, we may find new concepts and ideas that we don't have a word for. So we often borrow the word from the new place. And we extend our language - and our linguistic shell.

Linguistic Shells

Graphical languages also have shells.

In an ERD, we can describe a relationships and the paths round the data.

In UML, we can describe the interactions between the data and the processes.

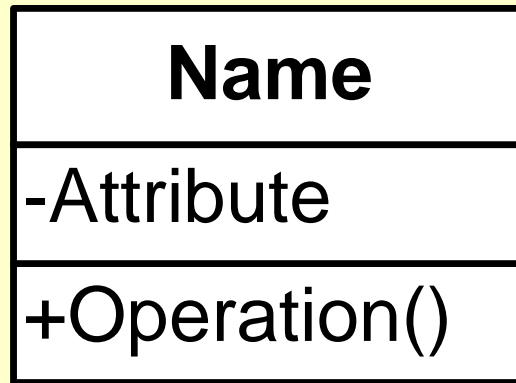
Linguistic Shells

Entity Relationship Diagramming 'doesn't have a word' for Process.

UML 'doesn't have a word' for some of the Relational concepts. Like identifier...

This makes relational modelling in UML difficult, unless we can extend the vocabulary.

UML - Classes



Entities show the name and the attributes.

Classes show the name, attributes and operations.

Basic Concepts

Very simple Employee data:

- Employee No (Mandatory, Primary key)
- Name (Mandatory)
- NI No (Mandatory, Alternate Key)
- Start Date (Mandatory)
- Address (Mandatory)
- Work Location (Mandatory, Foreign Key)
- End Date (Optional)

Basic Concepts

The basic concepts:

- Primary Key
- Alternate Key
- Foreign Key
- Mandatory/Optional

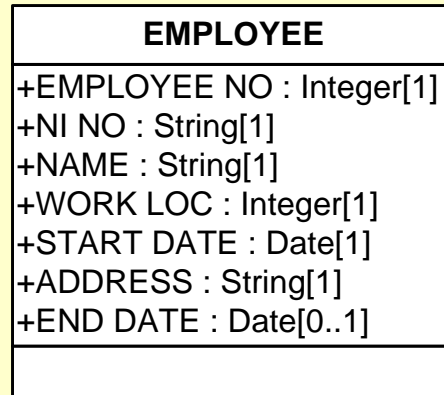
Basic Concepts - ERD

EMPLOYEE

EMPLOYEE NO
A NI NO
NAME
F WORK LOC
START DATE
ADDRESS
o END DATE

ERD notation can handle keys and optionality.

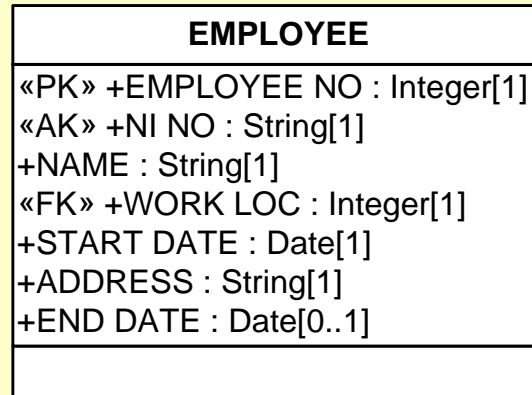
Basic Concepts - UML (1)



UML can handle mandatory and optional (but check the tool!)

It doesn't have a word for 'identifier', so no primary, alternate or foreign keys...

Basic Concepts - UML (2)

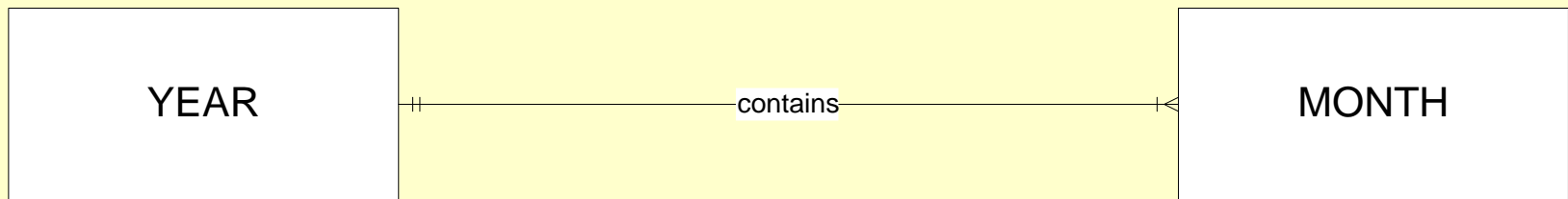


UML does have stereotypes.

Not designed for this, so in order to get what we want, we have to break the language.

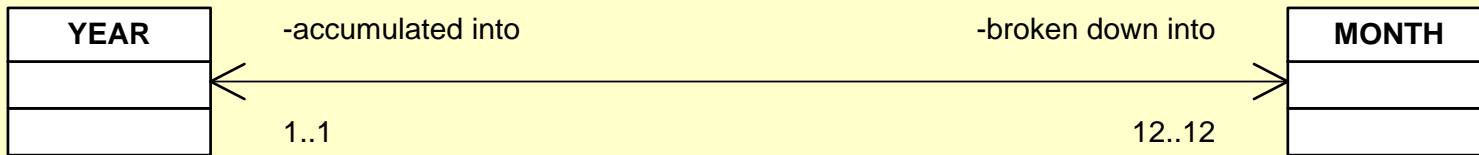
Our own dialect...

Relationship cardinality - ERD



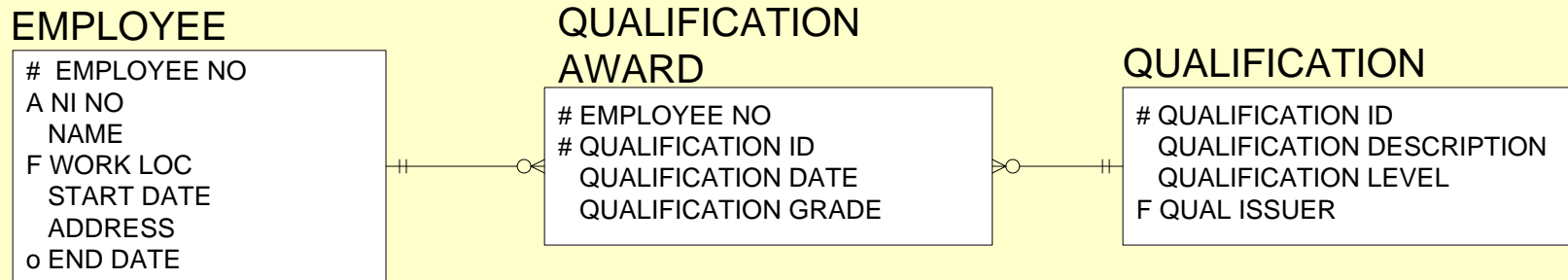
How many months in a year?

Association Multiplicity - UML



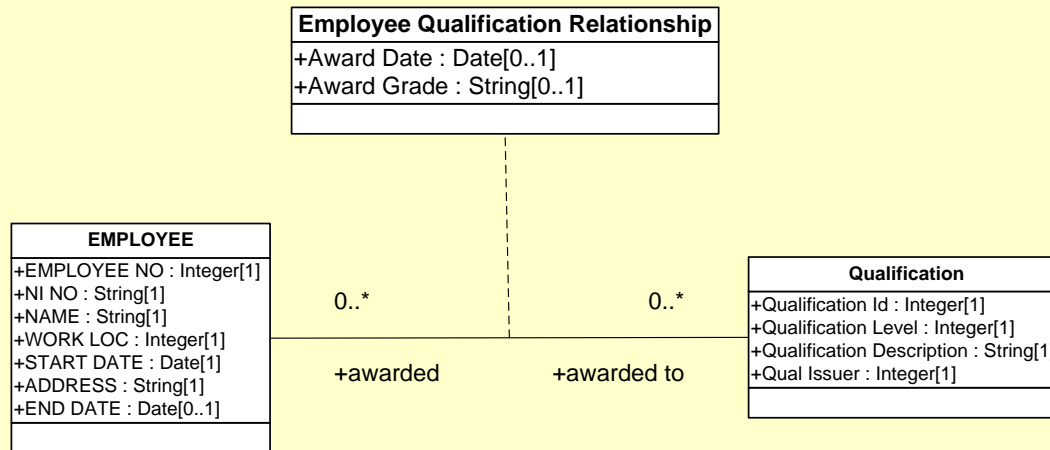
How many months in a year?

Many to Many - ERD



Easy enough in an ERD.

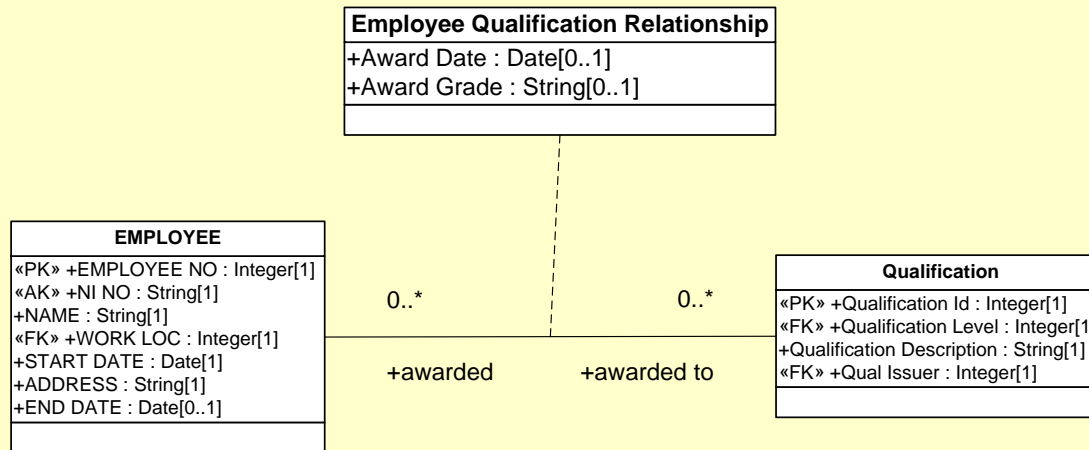
Many to Many - UML (1)



I like this way of describing a relationship. But...

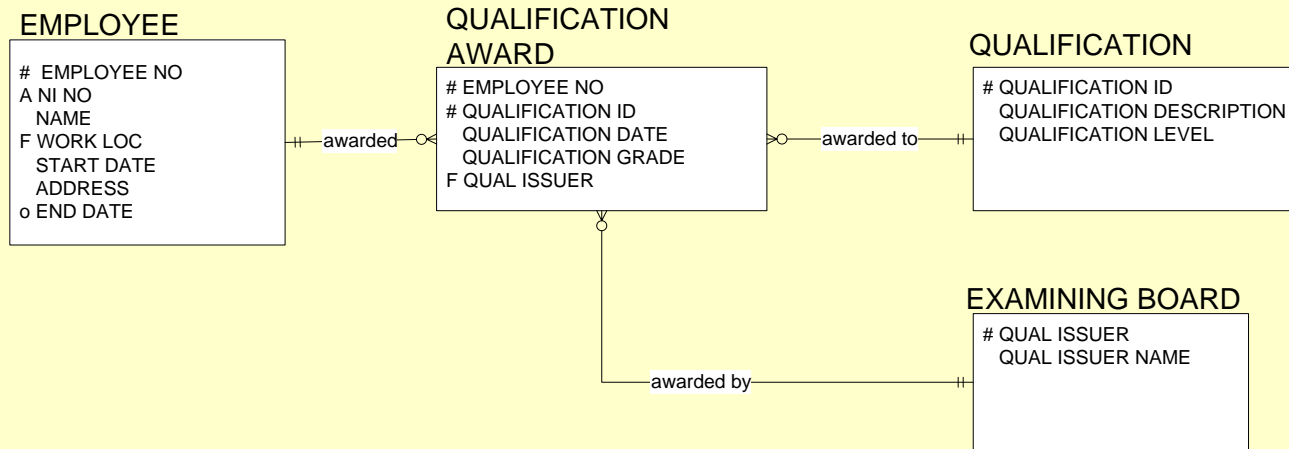
It has an inherent restriction, that there can only be one entry per employee/qualification pair (same as in the ERD version), but what if it was EMPLOYEE and PROJECT, and the association was ROLE? This restricts an employee to one role on a project (and this isn't explicitly obvious, in my opinion).

Many to Many - UML (2)



But it gets messy with the extra stuff we need...

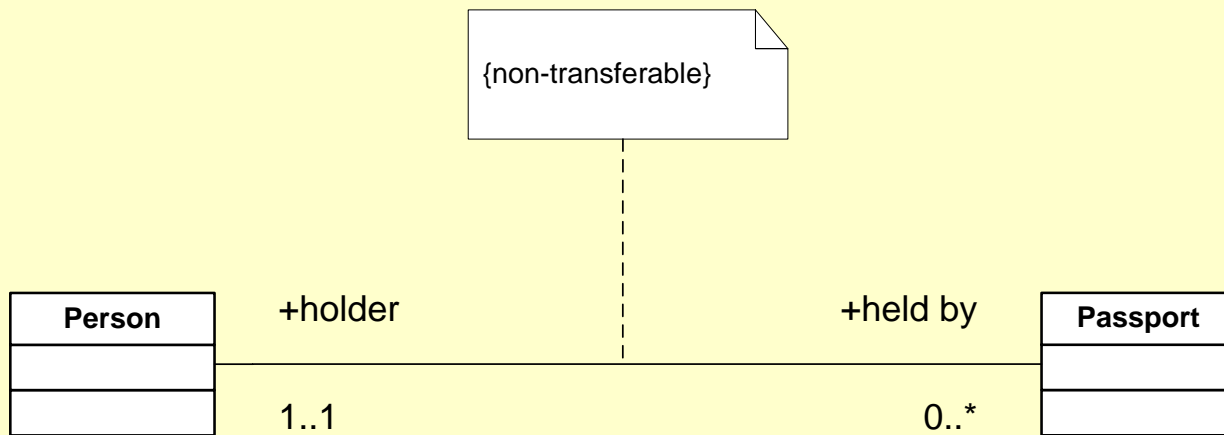
Triples and above



The UML Association won't work here - QUAL AWARD has to be a full class.

Same as when we do it in an ERD...

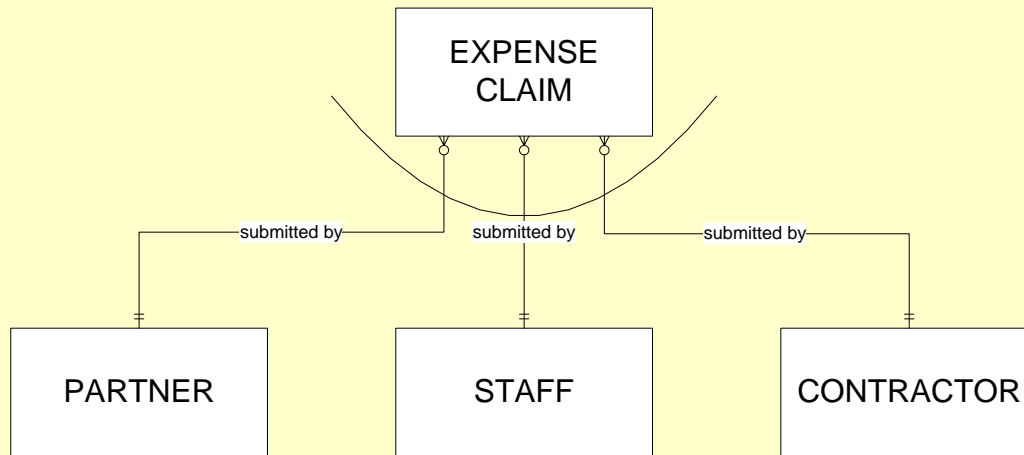
UML has notes...



There is no 'non-transferable' symbol in standard ERDs (but there is a suggestion in Data Modeling Essentials).

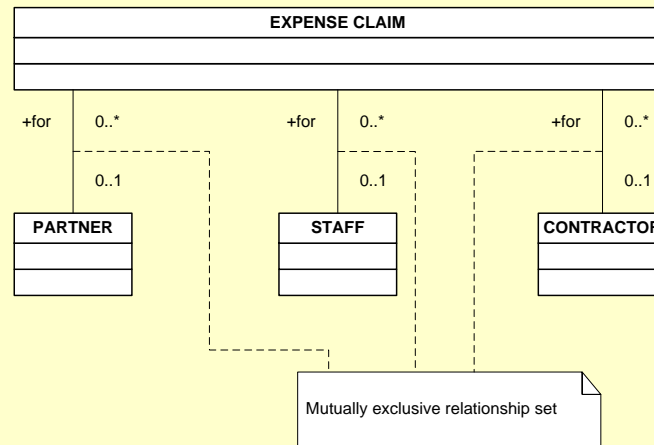
UML has the ability to add a note to a relationship or class, which can add clarity to the model. (But adds more boxes and lines...)

Exclusive relationships - ERD



Easily done in an ERD (although the 'exclusive arc' is not part of the formal language.)

Exclusive relationships - UML



UML only allows pairs of mutually exclusive relationships (using a 2-element OR constraint) not sets of mutually exclusive relationships with more than 2 members, so a UML diagram can't be drawn here. However, it is possible (although not particularly satisfactory) to add a note.

The Future



The Future

The Object Management Group are developing a UML Data Modelling notation.

Due to be released next year (2007).

It will be part of UML 2.0.

We have to hope that it will make this presentation redundant...

Bibliography

Data Modeling Essentials (*Graeme Simsion, Graham Witt, Published by Morgan Kaufman. Additional material available at <http://books.elsevier.com/default.asp?country=United+Kingdom>*)

The Unified Modeling Language Reference Manual (*James Rumbaugh, Ivar Jacobson, Grady Booch. Published by Addison-Wesley.*)

UML Distilled (*Martin Fowler. Published by Addison-Wesley.*)

The Object Management Group - <http://www.omg.org/>

Data Modelling in UML

Questions?

Contact Information

Dagna Gaythorpe

<http://www.seshat.com>

dgaythorpe@seshat.com

THREE GREAT CONFERENCES IN ONE!



Conference Europe 2006
Data Management,
What Works?

META 2
DATA 0
Europe 0
Conference 6

NEW THIS YEAR: DW/BI TRACK

30 October - 2 November 2006, London, UK