CSCI317 Database Performance Tuning

Transformations of Generalization Hierarchies

Dr Janusz R. Getta

School of Computing and Information Technology -
University of Wollongong

Transformations of Generalization Hierarchies

Outline

Methods

Superset method

Subset method

Association method

Hybrid method

Methods

Superset method

- Move all attributes to the top level class

Subset method

- Move all attributes to the leaf level classes

Association method

- Replace generalization with association

Hybrid method

- A mixture of superset, subset, and association methods

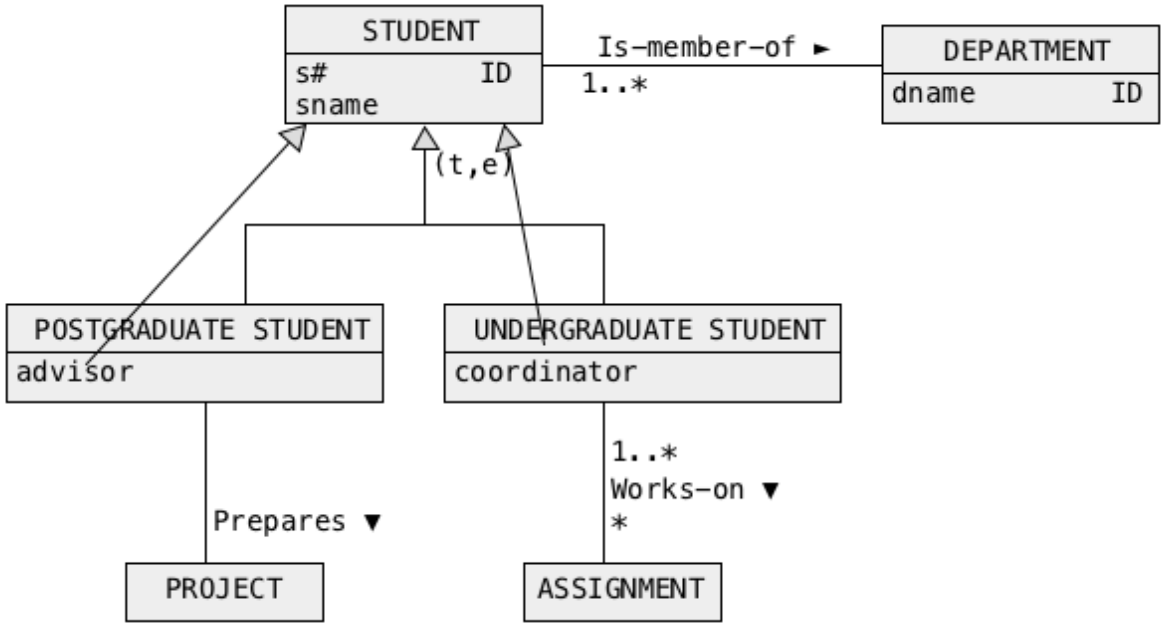
Transformations of Generalization Hierarchies

Outline

- Methods
- Superset method
- Subset method
- Association method
- Hybrid method

Superset method

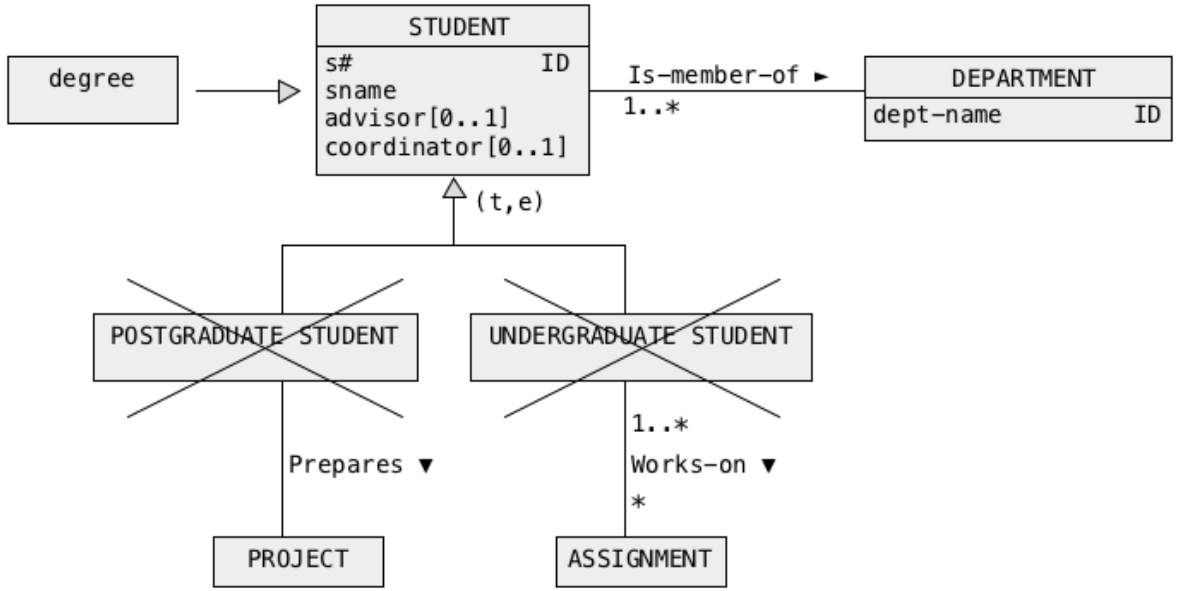
Move the attributes from subclasses to a superclass



The attributes `advisor` from a class `POSTGRADUATE STUDENT` and `coordinator` from a class `UNDERGRADUATE STUDENT` are moved to a superclass `STUDENT`

Superset method

Eliminate subclasses

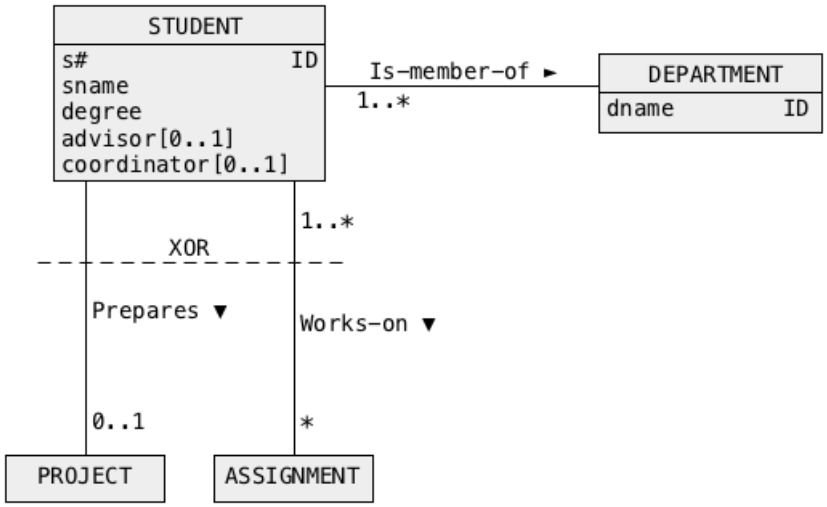


The classes `POSTGRADUATE STUDENT` and `UNDERGRADUATE STUDENT` are removed

An attribute `degree` is added to a class `STUDENT`

Superset method

Final design



The attributes `advisor` and `coordinator` are optional in a class `STUDENT` becasue only some students have a coordinator while the others have an advisor

`XOR` above a dashed line means that each student is linked either to a coordinator or to an advisor and not to both of them

Superset method

Performance related observations

Superset method is beneficial when the majority of queries access the attributes from many different levels of generalization hierarchy and when aggregations over different subclasses are performed

Superset method is opposite to horizontal and vertical decompositions

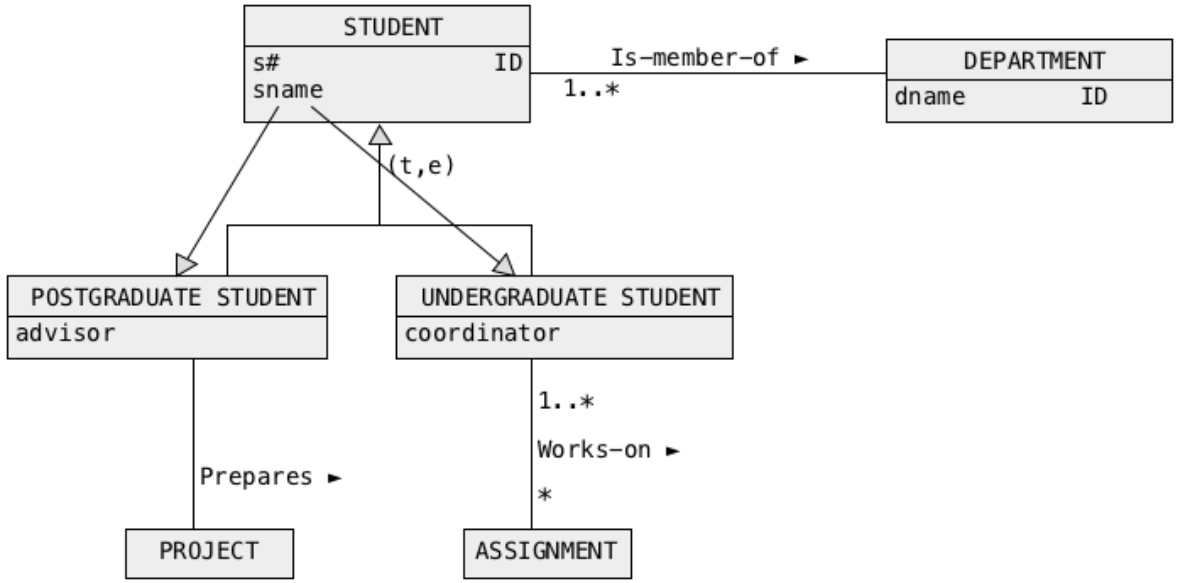
Superset method increases the total number of objects to be traversed and a size of each object

Transformations of Generalization Hierarchies

Outline

Subset method

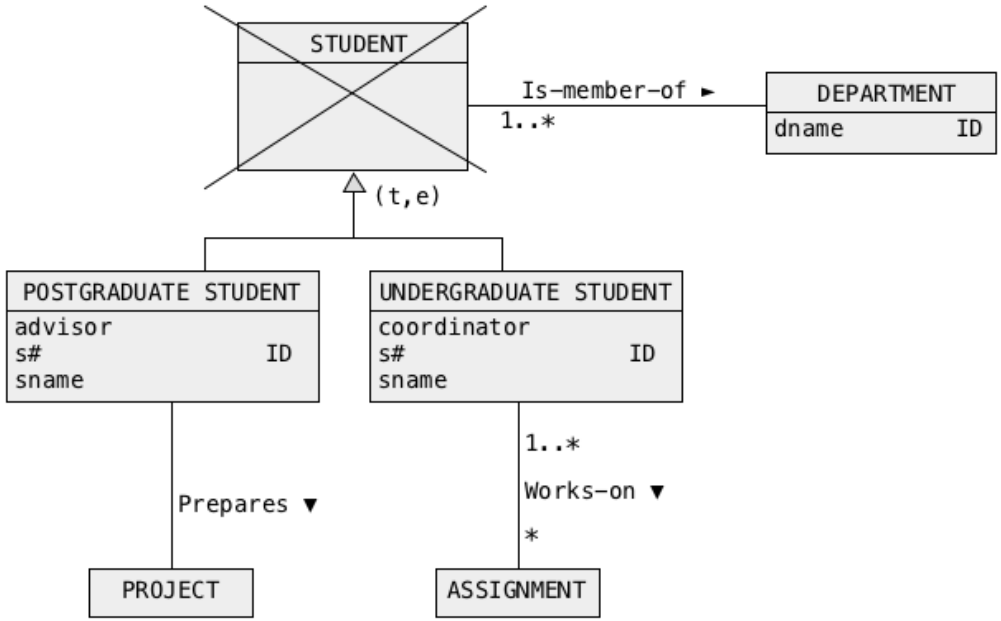
Move the attributes down to the subclasses



The attributes `s#` and `sname` are moved from a class `STUDENT` to both classes `POSTGRADUATE STUDENT` and `UNDERGRADUATE STUDENT`

Subset method

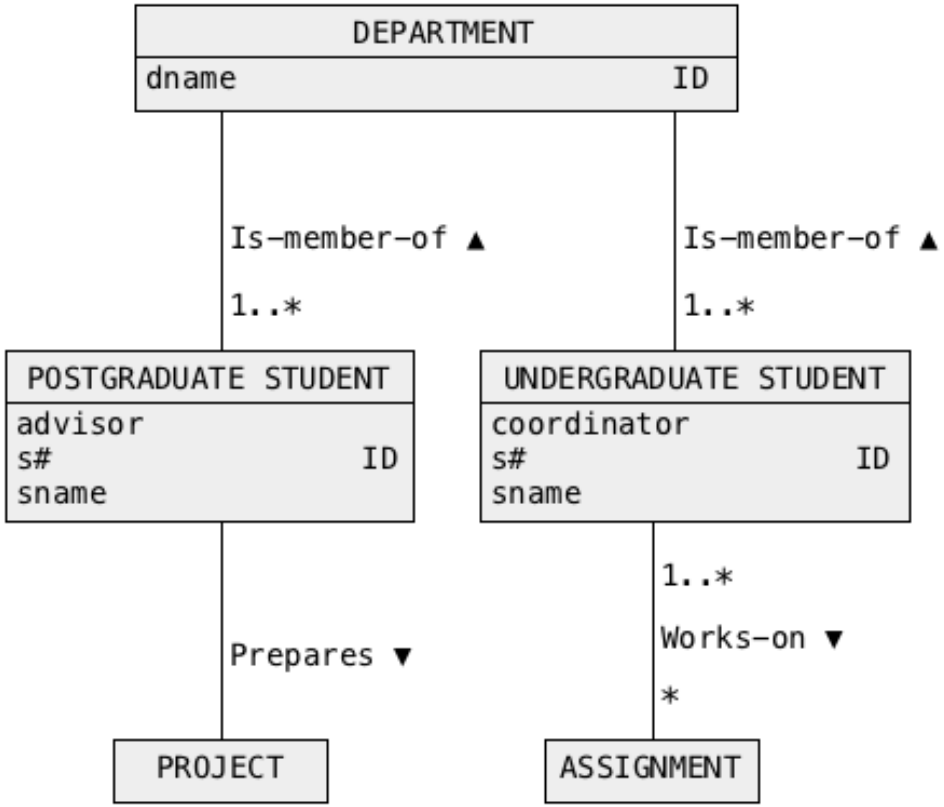
Eliminate the superclass



A class `STUDENT` is deleted and an association `Is-member-of` is replicated to connect a class `DEPARTMENT` with both classes `POSTGRADUATE STUDENT` and `UNDERGRADUATE STUDENT`

Subset method

Final design



Subset method

Performance related observations

Subset method is beneficial when the majority of queries access individually the lowest level classes in a generalization hierarchy

Subset method is equivalent to vertical decomposition of classes (horizontal decompositions of relational tables)

Subset method decreases the total number of objects to be traversed and it increases a size of each object in a class

Subset method increases a number of joins when the majority of queries accesses the attributes from many different classes

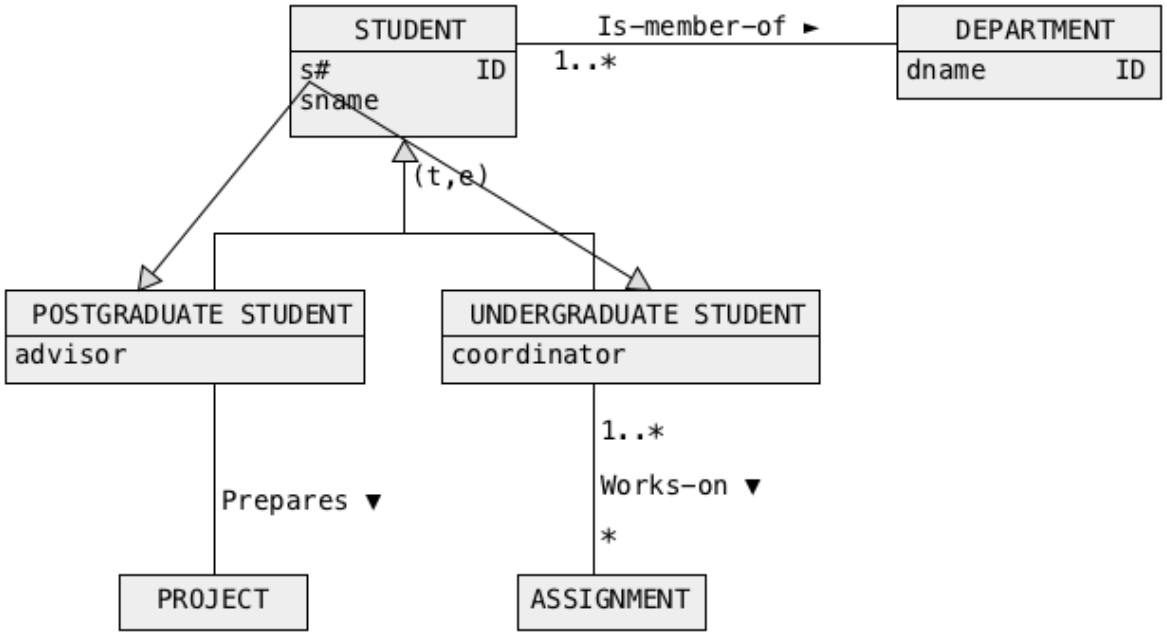
Transformations of Generalization Hierarchies

Outline

- Methods
- Superset class method
- Subset class method
- Association method
- Hybrid method

Association method

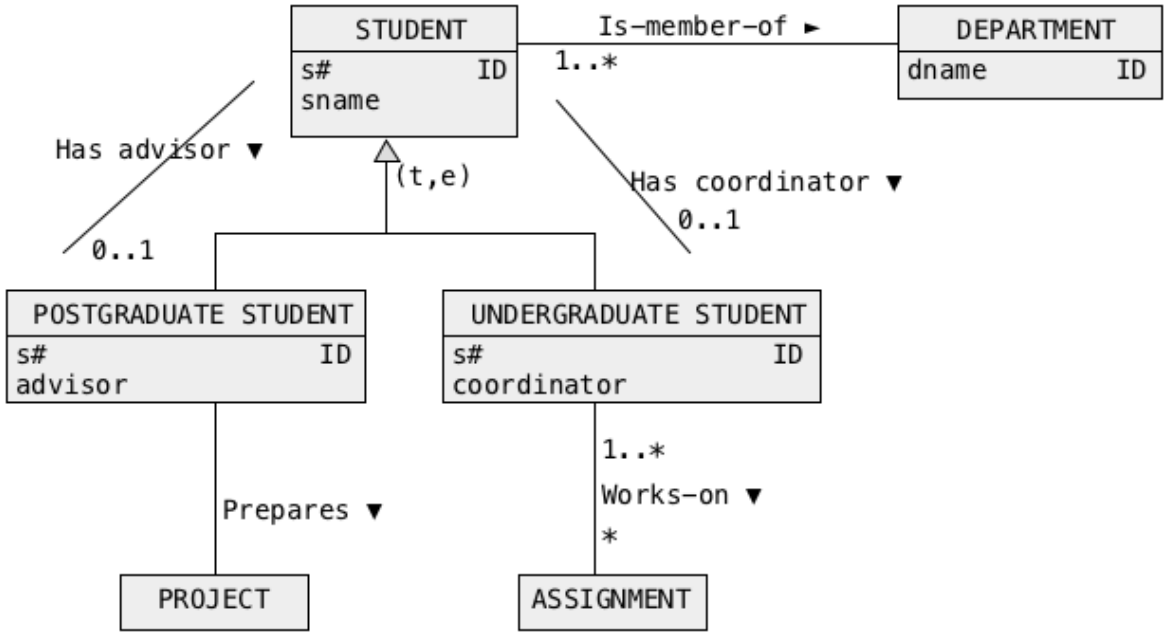
Copy an identifer to the subclasses



An attribute `s#` is copied to both subclases `POSTGRADUATE STUDENT` and `UNDERGRADUATE STUDENT`

Association method

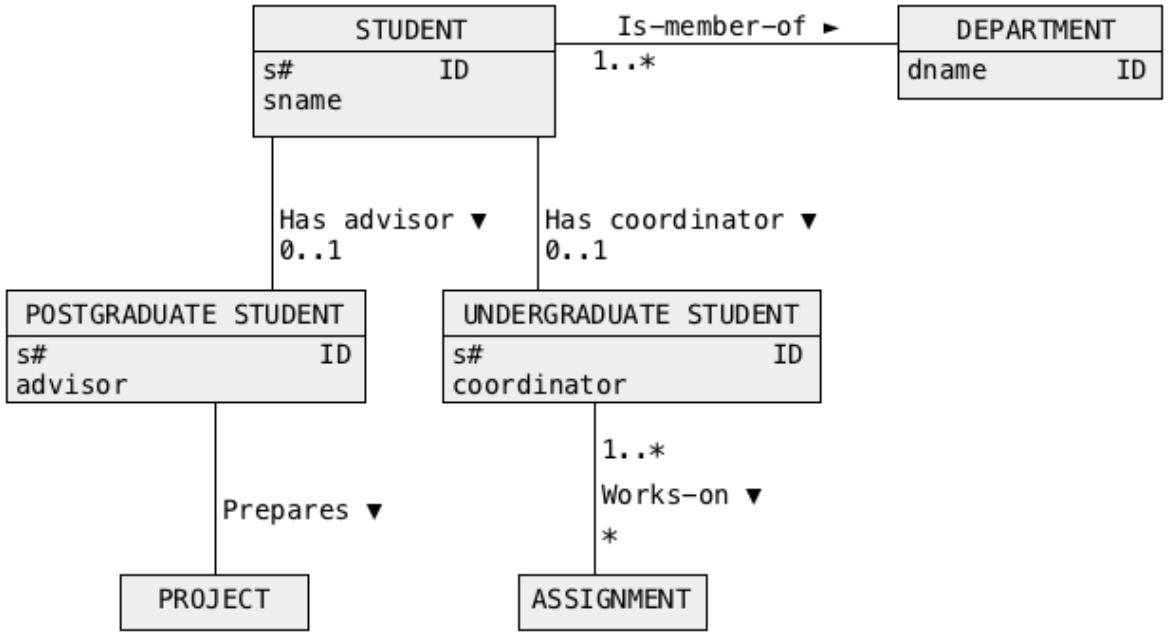
Replace generalization with one-to-one associations



Semantics of the classes `POSTGRADUATE STUDENT` and `UNDERGRADUATE STUDENT` is changed and a generalization is replaced with one-to-one associations

Association method

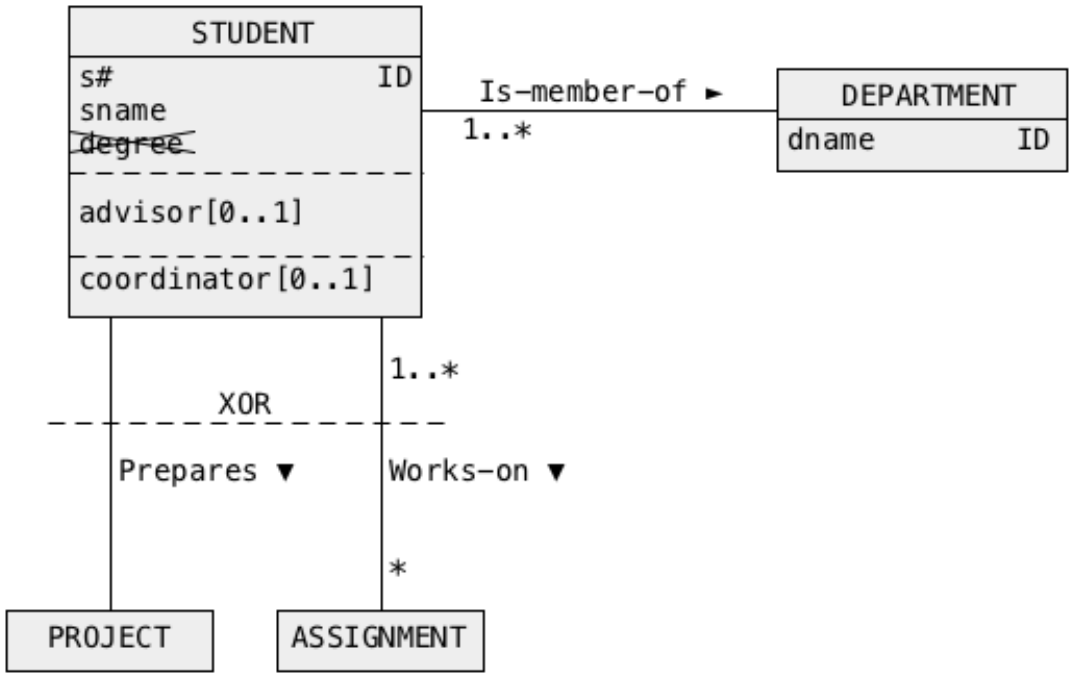
Final design



A class `POSTGRADUATE STUDENT` represents now a part of a description of postgraduate students; a class `UNDEGRADUATE STUDENT` represents now a part of a description of undergraduate students

Association method

Another way is to use superset method, apply a horizontal decomposition, and remove categorization attribute



After application of superset method a class `STUDENT` is decomposed horizontally and an attribute `degree` is removed.

Association method

Performance related observations

Association method is beneficial when the majority of queries access individually the classes in a generalization hierarchy

Association method is equivalent to horizontal decomposition of classes (vertical decompositions of relational tables)

Associtions method decreases a size of each object in a class

Associations method increases a number of joins when the majority of queries accesses the attributes from many different classes

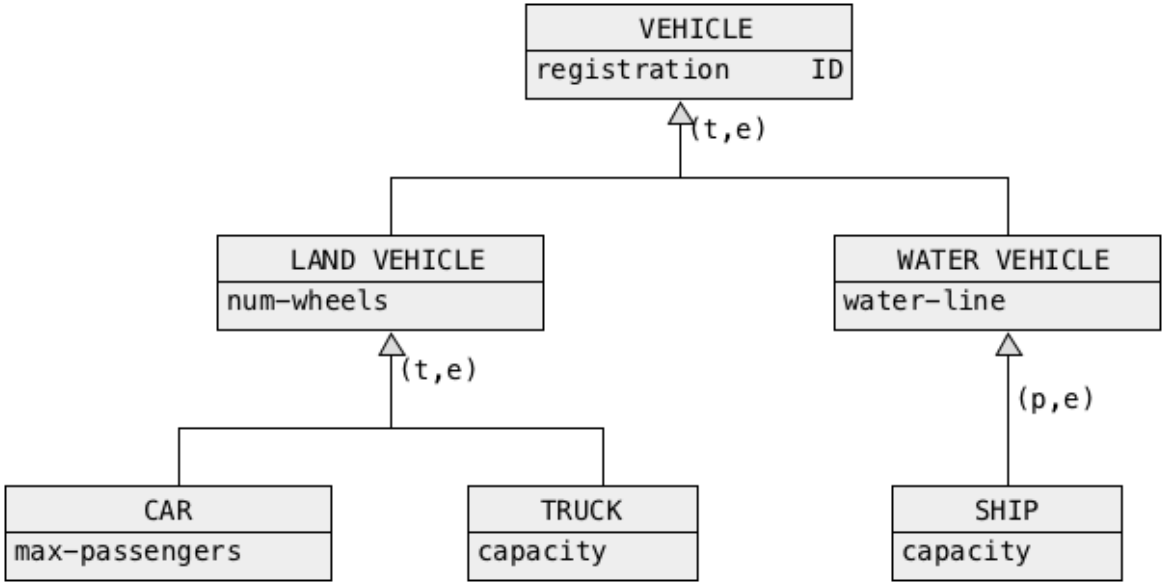
Transformations of Generalization Hierarchies

Outline

- Methods
- Superset class method
- Subset class method
- Association method
- Hybrid method

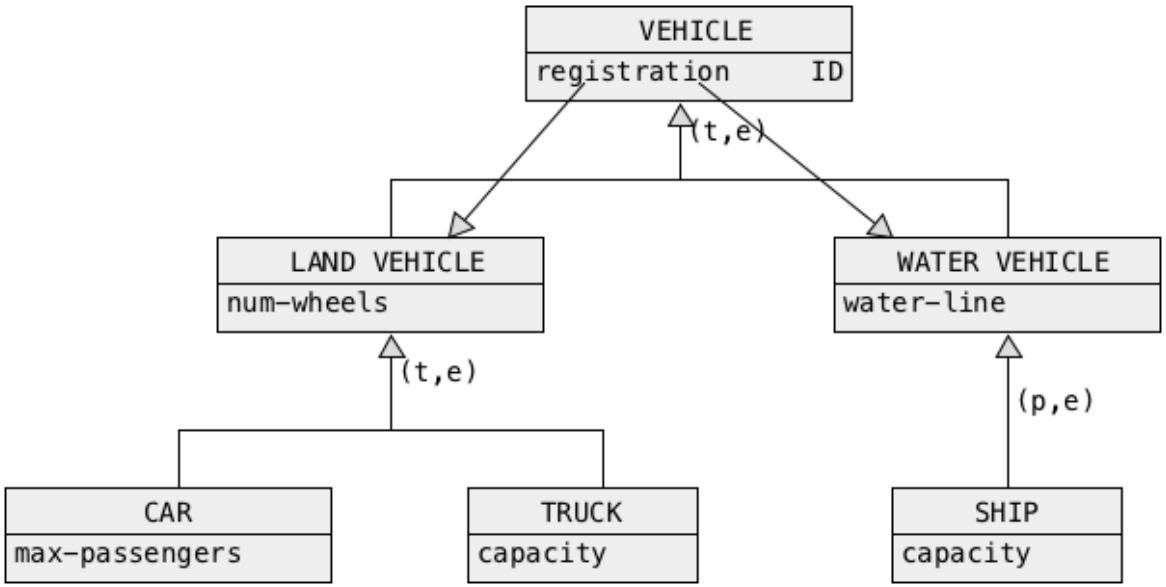
Hybrid method

Original generalization hierarchy



Hybrid method

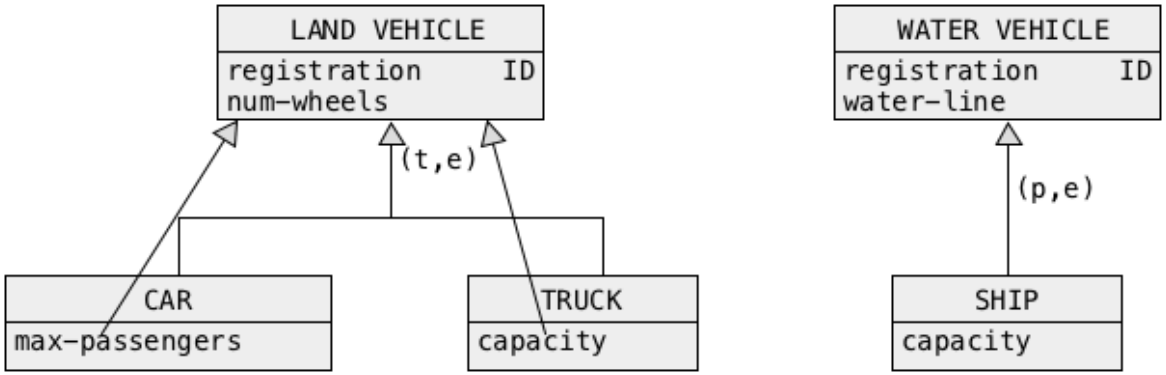
Application of subset method



An attribute `registration` is moved to both subclasses `LAND VEHICLE` and `WATER VEHICLE`

Hybrid method

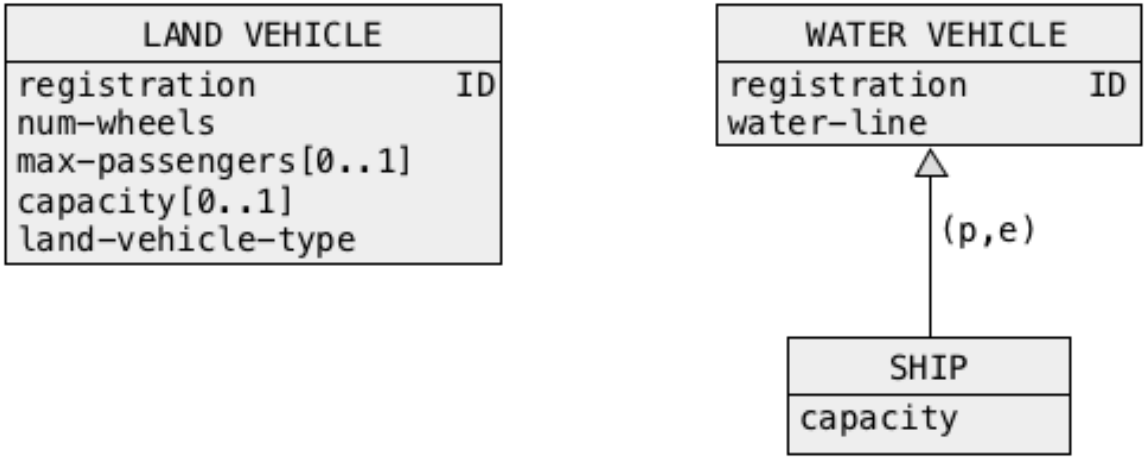
Application of superset method



The attributes `max-passengers` and `capacity` are moved to a superclass `LAND VEHICLE`

Hybrid method

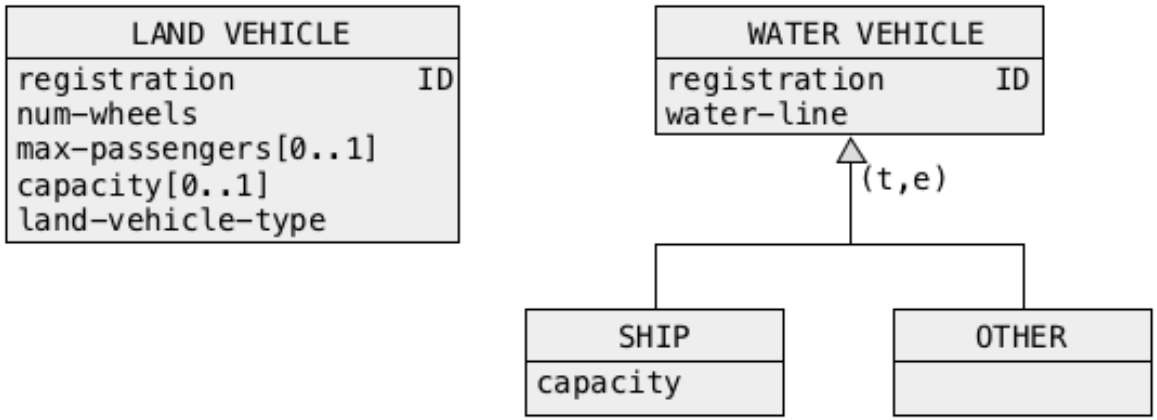
Transformation of (p,e) generalization into (t,e) generalization



A `p-e` generalization `SHIP ISA WATER VEHICLE` is replaced with a `t-e` generalization `SHIP ISA WATER VEHICLE` and `OTHER (WATER VEHICLE) ISA WATER VEHICLE`

Hybrid method

Application of association method

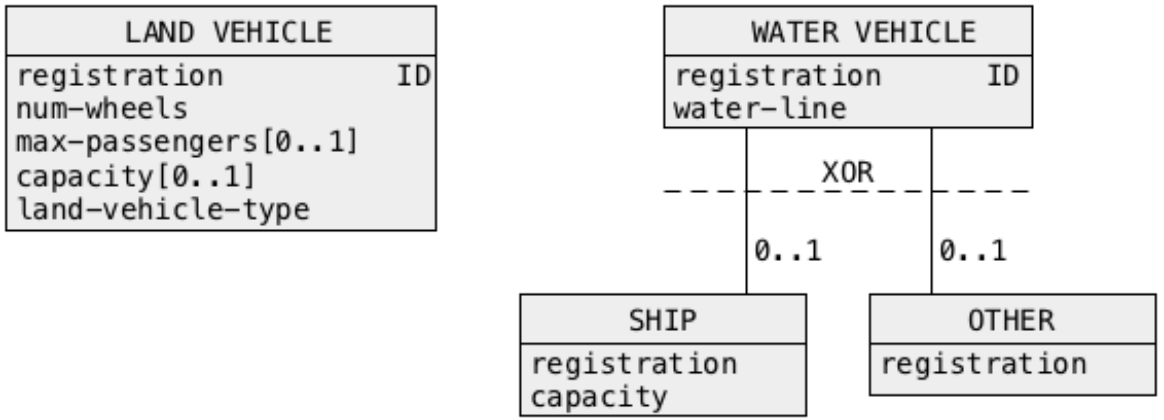


Association method is applied to `t-e` generalization

Semantics of classes `SHIP` and `OTHER` is changed such that a class `SHIP` represents a part of a description of ship and a class `OTHER` represents a part of a description of other water vehicle

Hybrid method

Final design



A description of water vehicle consists of either a description of ship or a description of other water vehicle