

201300180 Data & Information – Tests 1 + 2 (resit) (3 hours)

25 June 2015, 13:45 – 16:45

Please note:

- **You can do only test 1, only test 2, or both tests as you like:**
 - Test 1 consists of questions 1.1, 1.2, 1.3,
 - Test 2 consists of questions 2.1(a+b), 2.2, 2.3(a+b).
- **Please answer questions 1.1, 1.2, 1.3, 2.1, 2.2, and 2.3 each on a separate sheet of paper.**
(Not on the back side of the previous question, the questions will be distributed to different person for grading). **If you do both tests, you need two sets of exam paper.**
- You can give your answers in Dutch or English.
- Reference materials are given in the appendices, you are not allowed to bring any study materials to the test.
- Separate the tests from the appendices; for many questions you want to look at both.

===== Test 1 =====

Question 1.1 (Database queries) (40 points)

Write the following SQL queries for the movie database given in Appendix B (p. 8).

Excerpts from the SQL syntax are given in Appendix C (p. 8).

Use `DISTINCT` only where it is needed to remove duplicates.

- Give the names of directors who directed one or more films jointly with another director.
- Give the names of directors who never directed a film jointly with another director.
- The same film can have a different runtime in different countries.
For films in the database, give the name, the shortest and the longest runtime, as well as the number of countries in which the film has been released.
Order the result from highest to lowest *average* runtime.
Don't include:
 - films with an average runtime less than 60 (minutes),
 - films that were released only in a single country.

Question 1.2 (Web programming) (30 points)

- a) A class that extends the `HttpServlet` class normally implements (overrides) the methods `doGet()`, `doPost()`, `doPut()`, `doDelete()`, etc. What is the typical functionality of these implementations? What do they do? What does this functionality have to do with HTTP?
- b) What are the differences and similarities between Servlets en Java Server Pages (JSPs)? Discuss at least three aspects in which they differ or are similar.
- c) What is the typical role of a Java Bean in the MVC pattern? Why is a Java Bean necessary in the MVC pattern in case JSP and Servlets are properly used (i.e., used for what they are best suited).

Question 1.3 (Quality requirements) (30 points)

Read the case description in Appendix A (p. 5–7).

For this case you are asked to give some quality requirements.

Important: The system that we consider here consists of the central information system of PAZ and the wireless network. Not part of the system are transceivers, pulse-oxymeters, and other measurement equipment. Their quality is assured by the vendors in detailed contracts.

- State three quality characteristics that you consider most important for telemonitoring. See Appendix D (p. 9) for an overview of quality characteristics. Choose from the 31 detailed characteristics, not the 8 top categories.
- For each quality characteristic, explain why you consider it to be a top priority.
- Formulate a (possible) requirement for each of the three characteristics mentioned.

===== Test 2 =====

Question 2.1 (Database Schema) (30 points)**Introduction**

In het Polderland Academisch Ziekenhuis (PAZ) zijn verschillende soorten artsen op verschillende manieren werkzaam. Sommige specialisten zijn in dienst van het ziekenhuis, andere specialisten niet (die vormen een eigen maatschap die diensten levert aan het PAZ). Daarnaast zijn er assistent-artsen (altijd in dienst van het ziekenhuis). Deze zijn afgestudeerd (basis)arts en bekwamen zich nu in een specialisme (bijv. chirurgie, reumatologie, ...). Een assistent-arts wordt opgeleid door één specialist. Tevens is vastgelegd op welk moment iedere assistent-arts met de opleiding begonnen is. Patiënten kunnen behandeld worden door meerdere artsen, maar er is altijd één arts verantwoordelijk voor de behandeling van de patiënt. Dat is uiteraard de specialist. Deze informatie, met nog wat meer gegevens, is vastgelegd in het class diagram in Fig. 1.

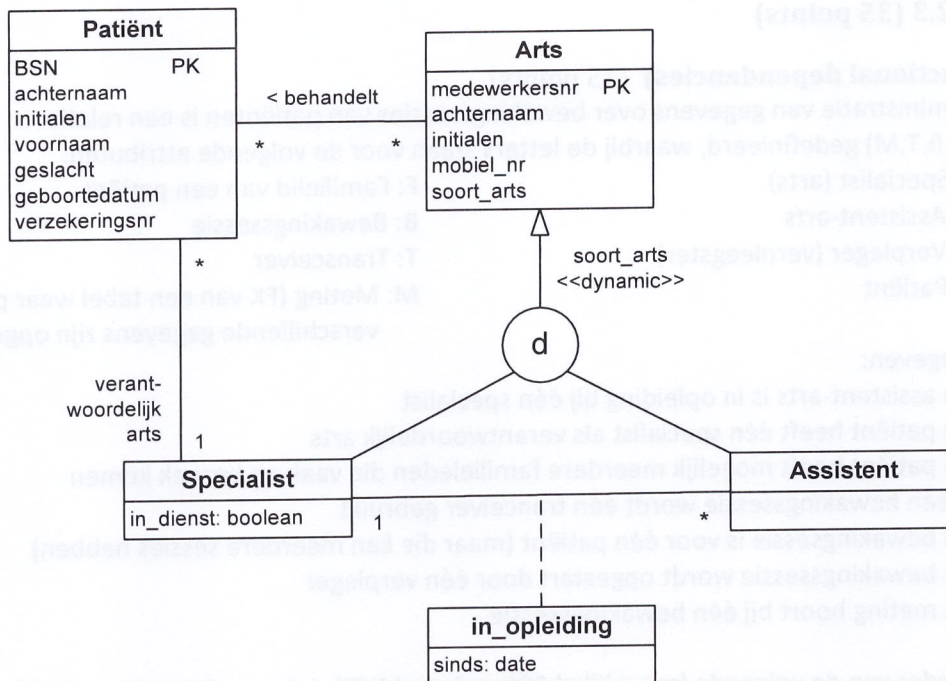


Figure 1: Class diagram for medical doctors and their patients

- a) There are three ways in which a generalization can be represented by database tables:
- separate tables for the superclass and for each of the subclasses
 - separate tables for the subclasses, no table for the superclass
 - only a table for the superclass, no tables for the subclasses.
- Which of (i), (ii), (iii) could – in theory – be applied to the generalization of medical doctors in Figure 1? Why?
- Which of these options is preferable if you take the associations into account? Why?
- b) Define database tables that cover the information in the class diagram above.
(See appendix C (p. 8) for excerpts of the SQL syntax).

Question 2.2 (Class Diagram) (35 points)

A class diagram is asked for the case study in Appendix A (p. 5–7).
Extend Figure 1 to a class diagram for information around telemonitoring of oxygen saturation. This class diagram specifies (part of) the central information system in PAZ (not the information in transceivers, pulse-oxymeters, etc., unless this information is copied into the central information system)

Hint: Please note that attributes in Figure 1 could become attributes of a different class (in case of further generalization).

Question 2.3 (35 points)**2.3a) (Functional dependencies) (15 points)**

Voor de administratie van gegevens over bewakingssessies van patiënten is een relatie $R(S,A,V,P,F,B,T,M)$ gedefinieerd, waarbij de letters staan voor de volgende attributen.

S: Specialist (arts)

A: Assistent-arts

V: Verpleger (verpleegster)

P: Patiënt

F: Familielid van een patiënt

B: Bewakingssessie

T: Transceiver

M: Meting (FK van een tabel waar per meting verschillende gegevens zijn opgeslagen)

Verder is gegeven:

1. Een assistent-arts is in opleiding bij één specialist
2. Een patiënt heeft één specialist als verantwoordelijk arts
3. Een patiënt heeft mogelijk meerdere familieleden die vaak op bezoek komen
4. Bij een bewakingssessie wordt één transceiver gebruikt
5. Een bewakingssessie is voor één patiënt (maar die kan meerdere sessies hebben)
6. Een bewakingssessie wordt opgestart door één verpleger
7. Een meting hoort bij één bewakingssessie

Geef voor ieder van de volgende (mogelijke) FD's a–h en MVD's i–j aan of ze gelden ("ja") of niet ("nee"). Geef ook een motivatie, verwijzend naar de uitspraken 1 t/m 7 hierboven. De motivatie telt voor de helft van het antwoord.

- a) $P \rightarrow S$
- b) $S \rightarrow A$
- c) $M \rightarrow P$
- d) $M \rightarrow S$
- e) $BT \rightarrow S$
- f) $V \rightarrow S$
- g) $T \rightarrow B$
- h) $P \rightarrow F$
- i) $P \twoheadrightarrow F$
- j) $P \twoheadrightarrow BMT$

2.3b) (Normal forms) (20 points)

Consider the relational schema $R(A,B,C,D,E,F)$ with functional dependencies \mathcal{F} , defined by

$$\mathcal{F} = \{ AE \rightarrow F, D \rightarrow BC, F \rightarrow AD \}.$$

- 1) Which functional dependencies violate the BCNF condition? Why? (Hint: compute \mathcal{F}^+ and the candidate keys first)
- 2) Apply the algorithm in appendix E (p. 9) to decompose R into a set of relational schemas that are all in BCNF. For each decomposition step, please give the resulting schemas with their sets of functional dependencies and their candidate keys.
- 3) Which of the functional dependencies in \mathcal{F} were lost in the decomposition?

==== Appendices =====

A. Hospital case description

The case description for questions 1.3 and 2.2 is given in Dutch, as it is based on a case study situated in the Netherlands.

Telemetrie van vitale lichaamsfuncties

Bij sommige patiënten in een ziekenhuis moeten continu bepaalde waarden gemonitord worden. Deze zogenaamde vitale lichaamsfuncties zijn hartslag, bloeddruk, ECG en zuurstofsaturatie van het bloed. Niet bij iedereen zijn alle functies van belang. Bijvoorbeeld bij patiënten die herstellen van een hartinfarct is ECG essentieel, de bewakingsapparatuur kan aan de hand daarvan hartritmestoornissen detecteren en zo nodig onmiddellijk alarm slaan.

Vroeger moesten zulke patiënten vanwege de bewakingsapparatuur de hele dag in bed blijven. Tegenwoordig is dat niet meer nodig, want er zijn draagbare meetsystemen die de gegevens draadloos doorgeven aan een centrale bewakingscomputer. Voor de patiënt maakt dit het verblijf in het ziekenhuis wat dragelijker.

Het Polderland Academisch Ziekenhuis (PAZ) maakt al langer gebruik van telemetrie. Vooral op de afdelingen Cardiologie en Pediatrie (kindergeneeskunde) hebben patiënten hier baat bij. Bij Pediatrie wordt dit gebruikt om kinderen te monitoren die behandeld worden tegen kanker en gemonitord moeten worden op zuurstofsaturatie.

Een recent onderzoek in het PAZ heeft laten zien dat telemetrie in een duidelijke behoefte voorziet, maar dat het gebruiksgemak voor verbetering vatbaar is. Het PAZ zal daarom verder op telemetrie inzetten. Er worden extra *transceivers* (zie hieronder) aangeschaft en er wordt gekeken naar mogelijkheden om het gebruikersgemak te verbeteren.

De transceivers en meetapparatuur wordt geleverd door derden, maar het centrale informatiesysteem, waar de transceivers hun gegevens naartoe sturen, is in beheer bij het ziekenhuis zelf. Nu de mogelijkheden en procedures rond telemonitoring heringericht gaan worden is het van belang ook nog eens goed te kijken naar genoemde informatiesysteem.

Om omvang van de casusbeschrijving en de gegevens in het class diagram beperkt te houden beperken we ons hier tot zuurstofsaturatie.

Rode bloedcellen bevatten hemoglobinemoleculen, die zorgen voor het transport van zuurstof (O_2) en koolstofdioxide (CO_2). Een hemoglobinemolecuul kan 4 moleculen O_2 of CO_2 aan zich binden. Bij verbranding in het lichaam wordt O_2 verbruikt en komt CO_2 vrij. In de longen wordt CO_2 uitgewisseld tegen O_2 . Zuurstofsaturatie is het percentage O_2 -moleculen van het totaal aan O_2 - en CO_2 -moleculen dat is gebonden aan hemoglobine. Wat gemeten wordt is de **perifere zuurstofsaturatie (SpO_2)**, de saturatie in lichaamsdelen als vinger of oorlel. SpO_2 wordt gemeten door een klip (*pulse-oxymeter*) die om een vinger geschoven kan worden. Hierin zit een laser, die aan de hand van kleurverschillen de saturatie kan bepalen (hemoglobine met O_2 heeft een andere kleur dan hemoglobine met CO_2).

Een centrale rol in telemonitoring spelen **transceivers**, draagbare kastjes waar de meetapparatuur aan vastgemaakt kan worden. Een transceiver geeft signalen – de meetgegevens – door aan het centrale informatiesysteem.

Een transceiver communiceert draadloos met het centrale informatiesysteem via een wireless access point. Hij schakelt automatisch over naar het dichtstbijzijnde access point als de drager zich door het ziekenhuis beweegt. Er zijn verschillende modellen transceivers voor deze toepassing op de markt. Het PAZ gebruikt momenteel de modellen M2601A en M2601B van Philips. Daar komt binnenkort waarschijnlijk de TRx4851A bij, dat is in feite de opvolger van de M2601B.

Een bewakingssessie is de tijdsperiode dat een patiënt remote gemonitord wordt. Vóór het opstarten van een bewakingssessie noteert de verpleging in het systeem dat een sessie zal worden opgestart, om welke patiënt en welke transceiver het gaat, en wat de gebruikte grenswaarden zijn voor normale, lage en gevaarlijk lage SpO₂-waarden (over die grenswaarden hieronder meer).

Een sessie begint als een verpleger de patiënt heeft voorzien van een transceiver, de apparatuur heeft aangesloten, de meetwaarden heeft getest, en de transceiver verbinding laat maken met het centrale informatiesysteem.

De transceiver geeft één keer per seconde de meetwaarden door aan het centrale informatiesysteem, totdat een verpleger de sessie beëindigt en de transceiver uitschakelt. Patiënten kunnen de transceiver niet zelf uitschakelen.

Tijdens een bewakingssessie kunnen er verschillende bijzondere situaties optreden.

Een alarmsituatie treedt op als de gemeten waarde bepaalde kritische grenzen overschijdt. Normale SpO₂-waarden zijn tussen 90 en 100 %. Voor SpO₂ zijn drie waarden gedefinieerd:

- normale SpO₂ (groen)
- lage SpO₂ (geel)
- gevaarlijk lage SpO₂ (desaturatie) (rood)

De defaultwaarden voor de kritische grenzen zijn 90 % en 80 %. Met deze instelling is de SpO₂-waarde dus “rood” bij < 80 %, “geel” bij 80–90 % en “groen” bij 90–100 %.

Afhankelijk van de situatie kunnen deze waarden anders worden ingesteld. De verpleger doet dat dan bij het aanmelden van de bewakingssessie in het centrale systeem. Worden de defaults gebruikt, dan hoeft hij niets in te stellen.

Als een kritische waarde wordt overschreden alarmeert het centrale systeem de verpleging door middel van een geluidssignaal bij het monitoringscherm. Aan het signaal is te horen of het een ‘rood’ of een ‘geel’ alarm is. Ook is er een melding op het scherm te zien.

Wat de verpleging precies moet doen bij een alarm hangt van de situatie af. In ieder geval moet iemand naar de patiënt gaan om poolhoogte te nemen en zo nodig hulp te verlenen. Wellicht moet de verpleging ook een arts waarschuwen. Als de waarde snel weer in het groene bereik komt, is er wellicht niets aan de hand.

De locatie van de patiënt is bij benadering bekend. Een transceiver is niet uitgerust met GPS, maar laat bij ieder bericht ook weten via welk wireless access point hij momenteel met het netwerk verbonden is. Aan de hand daarvan is op de monitor de locatie op enige tientallen meters nauwkeurig te zien.

Behalve dat er iets mis kan gaan met de patiënt, kan er ook een storing optreden bij het meten. Veel voorkomende storingen zijn dat de transceiver geen geloofwaardige meetgegevens ontvangt (de clip is losgeschoten van de vinger) of helemaal geen meetgegevens ontvangt (de stekker is losgeraakt).

Is er een storing, dan stuurt de transceiver een signaal naar het informatiesysteem; de verpleging kan op het scherm zien dat er een storing is. Daarnaast geeft de transceiver zelf een geluidssignaal af, zodat de patiënt ook weet dat er een storing is.

Een storing die zo nu en dan voorkomt is dat de patiënt buiten bereik van de wireless access points gekomen is. Merk op: de transceiver kán dan geen signaal naar het informatiesysteem sturen. Als het contact 10 seconden verbroken is, dan wordt dit door beide systemen als storing geïnterpreteerd. De transceiver waarschuwt de patiënt, het centrale informatiesysteem waarschuwt de verpleging.

Vaak kan een patiënt de storing zelf oplossen, bijvoorbeeld door een losgeschoten stekkertje weer vast te zetten, of door terug te keren naar een plek waar de transceiver weer bereik heeft. Als de storing is opgeheven meldt de transceiver dat. Kan de patiënt de storing niet zelf oplossen, dan dient hij terug te komen naar de afdeling. Daar kan de verpleging de storing verhelpen of eventueel de transceiver uitzetten.

Extra information for 2.2 (class diagram)

In aanvulling op de gegevens uit de bovenstaande beschrijving en de gegevens die al opgenomen zijn in het class diagram in Fig. 1 is het volgende van belang.

- Alle SpO₂-metingen die een transceiver doorstuurt worden uiteraard opgeslagen. Bij iedere meting is bekend: tijdstip, locatie (d.w.z. locatie van het access point), en SpO₂-waarde.
- Van een bewakingssessie wordt ook tijdstip van begin en eind opgeslagen, evenals welke kritische waarden zijn ingesteld, welke verpleger de sessie heeft opgestart, en welke storingen er zijn geweest.
- Van een storing wordt opgeslagen: begin- en eindtijdstip evenals aard van de storing. Alarmsignalen worden niet opgeslagen. (Voor het class diagram zijn ze redundant, ze zijn direct af te leiden uit de opgeslagen gegevens).
- Van een transceiver is bekend: het serienummer; de datum waarop deze is aangeschaft; het typenummer (d.w.z. welk model het is) en de daarbij behorende levensduur van de batterij.
- Van verplegers zijn de volgende gegevens in het systeem opgenomen: medewerkersnummer; achternaam; initialen; adres.

B. Movies database tables

Movie	
mid	Integer KEY
name	text
year	numeric(4,0)
plot_outline	text
rating	numeric(2,1)

Acts	
mid	integer
pid	integer
role	text

Genre	
mid	integer
genre	text

Language	
mid	integer
language	text

Person	
pid	integer KEY
name	text

Directs	
mid	integer
pid	integer

Certification	
mid	integer
country	text
certificate	text

Writes	
mid	integer
pid	integer

Runtime	
mid	integer
country	text
runtime	text

C. Excerpts from SQL syntax

(choice is indicated by "|", optional inclusion by "[...]")

Select clause:

```
SELECT [ DISTINCT ] '*' | (aggregate) columns
FROM tables [ WHERE condition ]
[ GROUP BY columns
  [ HAVING condition ] ]
[ ORDER BY columns [ DESC ] ];
```

Create table clause:

```
CREATE TABLE table (
  column [NOT NULL] [UNIQUE] [PRIMARY KEY]
  [ , column ... ]
  [ , PRIMARY KEY (columns) ]
  [ , FOREIGN KEY (columns) REFERENCES table(columns)
  [ , FOREIGN KEY ... ] ]
  [ , CHECK (condition) ] );
```

(Note: you can abbreviate PRIMARY KEY, FOREIGN KEY and REFERENCES to PK, FK, and REF)

Examples of a condition:

```
[ NOT ] EXISTS select clause |
column = value [ (OR | AND) [NOT] column <> value ] |
column IS [NOT] NULL |
column [NOT] IN (value, ...) |
...
```

Aggregate function: COUNT(...) | MIN(...) | MAX (...) | AVG(...) | SUM(...)

D: Quality characteristics (ISO/IEC 25010:2011)

Functional suitability <ul style="list-style-type: none"> – Functional completeness – Functional correctness – Functional appropriateness 	Reliability <ul style="list-style-type: none"> – Maturity – Availability – Fault tolerance – Recoverability
Performance efficiency <ul style="list-style-type: none"> – Time behavior – Resource utilization – Capacity 	Security <ul style="list-style-type: none"> – Confidentiality – Integrity – Non-repudiation – Accountability – Authenticity
Compatibility <ul style="list-style-type: none"> – Co-existence – Interoperability 	Maintainability <ul style="list-style-type: none"> – Modularity – Reusability – Analysability – Modifyability – Testability
Usability <ul style="list-style-type: none"> – Appropriateness recognizability – Learnability – Operability – User error protection – User interface aesthetics – Accessibility 	Portability <ul style="list-style-type: none"> – Adaptability – Installability – Replaceability

E: Lossless BCNF decomposition algorithm

Definition of BCNF:

A relational schema is in BCNF if for every nontrivial functional dependency the left-hand side is a superkey.

Decomposition algorithm:

Let R be a relational schema with a set of functional dependencies \mathcal{F} .

Let $X \rightarrow Y$ be a functional dependency in \mathcal{F} which violates the BCNF constraint.

- Decompose R into
 - $R_1(X^+)$
 - $R_2(Z)$ with $Z = \{X\} \cup \{\text{attributes of } R \text{ not in } X^+\}$.
 - For $i = 1, 2$:
 - determine \mathcal{F}_i for R_i by restricting \mathcal{F}^+ to functional dependencies within R_i
- if R_i is not in BCNF, recursively apply the algorithm

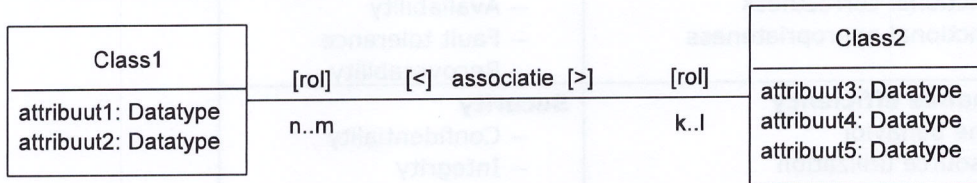
Appendix F: Notations for class diagrams

meta-notation:

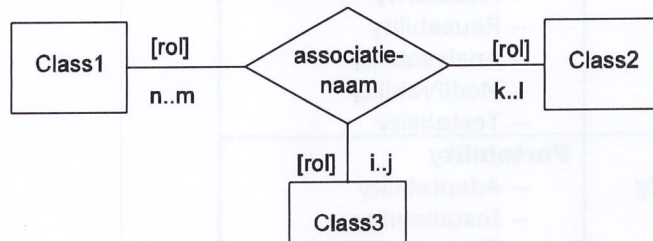
[...] Optional (can be deleted)

.. | .. Choice: one of the given alternatives

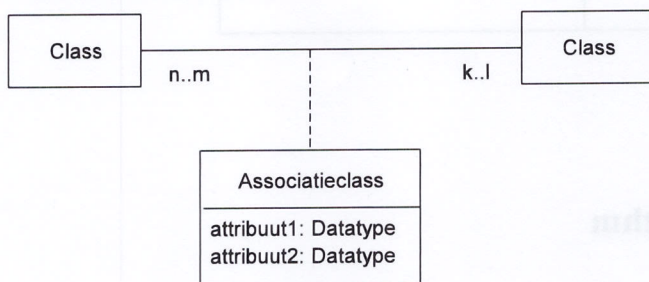
Class and Association



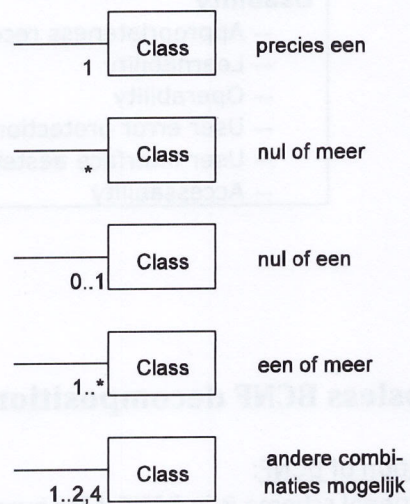
Ternary (or n-ary) association



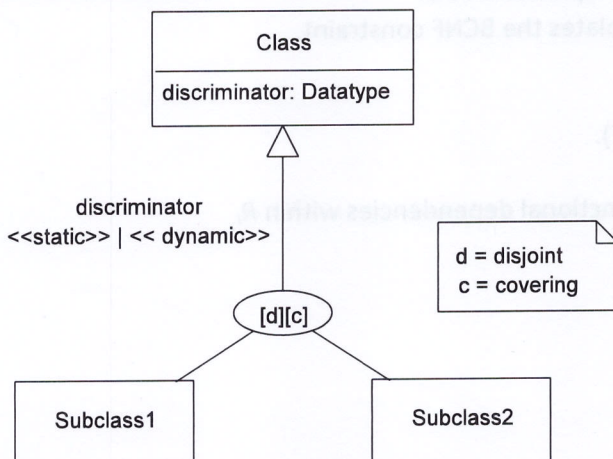
Association class



Multiplicity



Generalization



Composition

