

Veritabanı Yönetim Sistemleri

(Veritabanı Tasarımı)

Normalizasyon



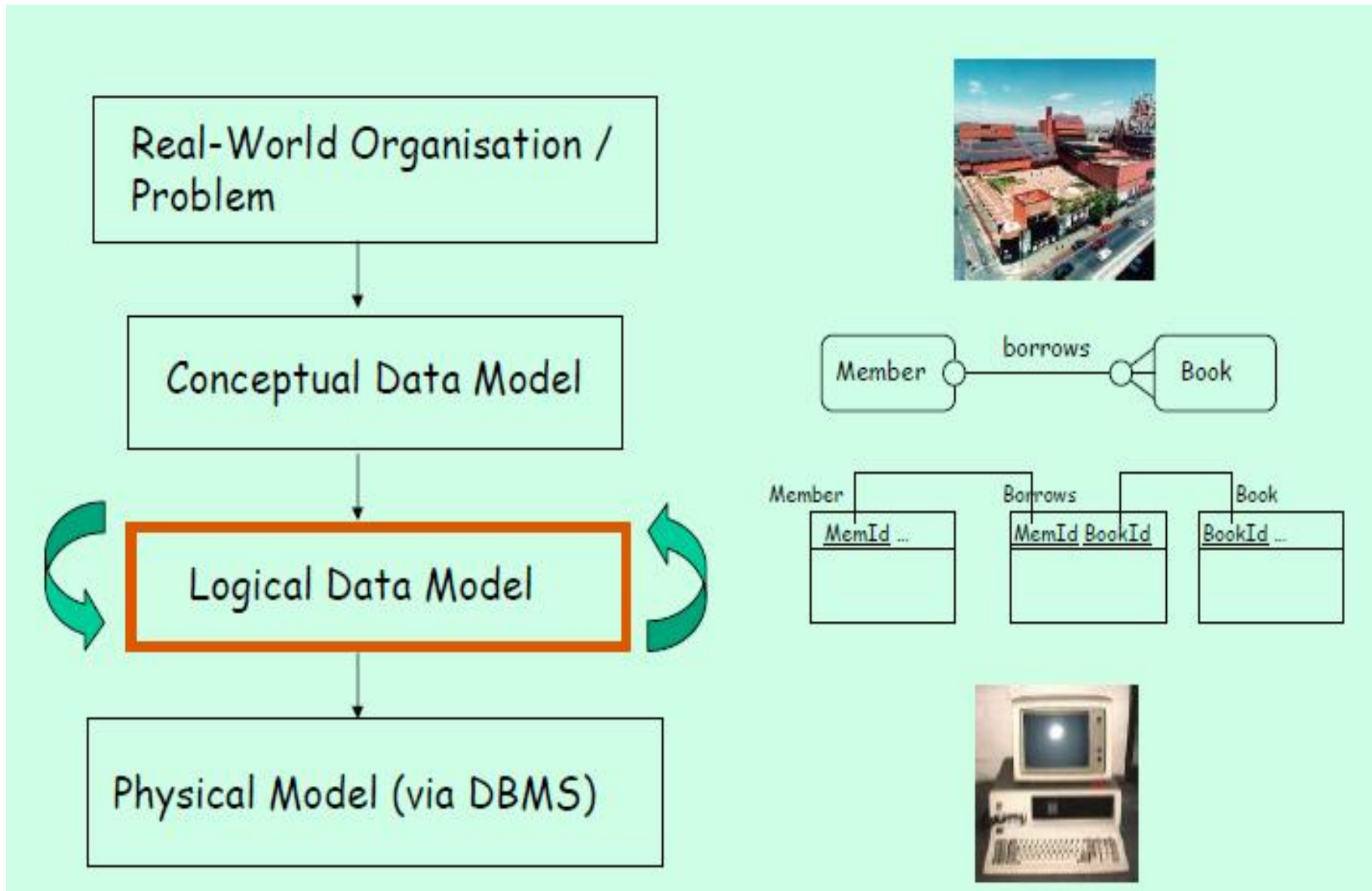
Konular

- ✓ Normalizasyon
- ✓ Birinci Normal Form(1NF)
- ✓ Normalizasyon Neden Yararlıdır?
- ✓ İşlevsel Bağımlılık
- ✓ İkinci Normal Form(2NF)
- ✓ Üçüncü Normal Form(3NF)
- ✓ 1., 2. and 3. Normal Formlar arasındaki İlişki
- ✓ Kaynaklar

Normalizasyon

- ✓ İyi bir veritabanı tasarımı, düzenli ve sağlam tablo tasarımlarını gerektirir.
- ✓ Normalizasyon; veri fazlalıklarını en aza indirerek veri düzensizliklerinin (data anomaly) önüne geçebilmek için tablo yapılarını değerlendirme ve düzeltme işlemi olarak tanımlanabilir
- ✓ Normalizasyon işlemi normal form adı verilen seri işlemlerden meydana gelir. 1NF, 2NF, 3NF, 4NF
- ✓ 2NF 1NF den, 3NF 2NF den, ve 4NF 3NF den daha iyidir.
- ✓ Her tasarım için en yüksek NF daha iyi sonuç verir denemez. Yüksek başarıma ihtiyaç duyulan bazı durumlarda NF nin düşürülmesi (denormalizasyon) gerekebilir. (Yüksek hız fakat veri fazlalığı)

Normalizasyon



Normalizasyon Neden Yararlıdır

Veri Tekrarının Zararları

Aynı bilginin defalarca yazılması:

1. Kaynak ve zaman israfı
2. Veri tutarsızlıklarına (düzensizliklerine) neden olabilir.

An anomaly is a situation where inconsistent data is introduced into a table, or data is lost unintentionally

Normalizasyon Neden Yararlıdır

Veri Tekrarının Zararları

<u>StaffId</u>	Name	Staff Address	Branch	Branch Address
S1	Fred Bloggs	23 Acacia Gardens	B1	42 Victoria Rd
S2	Bill Sykes	17 Mafeking Terrace	B1	42 Victoria Rd
S3	George Shaw	42 Privet Drive	B1	42 Victoria Rd
S4	John Doe	5 Mornington Crescent	B2	112 King St
S5	Tommy Atkins	10 Rillington Place	B2	112 King St

Veri tekrarı tablo içerisine başka tablonun gömüldüğü anlamına gelir...

Normalizasyon Neden Yararlıdır

Veri Girişindeki Tutarsızlık

<u>StaffId</u>	Name	Staff Address	Branch	Branch Address
S1	Fred Bloggs	23 Acacia Gardens	B1	42 Victoria Rd
S2	Bill Sykes	17 Mafeking Terrace	B1	42 Victoria Rd
S3	George Shaw	42 Privet Drive	B1	42 Victoria Rd
S4	John Doe	5 Mornington Crescent	B2	112 King St
S5	Tommy Atkins	10 Rillington Place	B2	112 King St
S6	Ed Grundy	Keeper's Cottage	B2	121 King St

Potential for inconsistent data (branch addresses)

			B3	26 Garthdee Rd
--	--	--	----	----------------

Can't add a new branch until it has some staff

Normalizasyon Neden Yararlıdır

Veri Güncellenmesindeki Tutarsızlık

S1	Fred Bloggs	23 Acacia Gardens	B1	20 Union Street
S2	Bill Sykes	17 Mafeking Terrace	B1	20 Union Street
S3	George Shaw	42 Privet Drive	B1	20 Union Street
S4	John Doe	5 Mornington Crescent	B2	112 King St
S5	Tommy Atkins	10 Rillington Place	B2	112 King St
S6	Tim Penman	21b Baker Street	B1	42 Victoria Rd

Bölüm adresi değiştiğinde bir çok yerde değişiklikler yapmak gerekir. Değişiklik unutulabilir ya da yanlış olabilir.

Normalizasyon Neden Yararlıdır

Veri Silmede Tutarsızlık

S1	Fred Bloggs	23 Acacia Gardens	B1	42 Victoria Rd
S2	Bill Sykes	17 Mafeking Terrace	B1	42 Victoria Rd
S3	Georgae Shaw	42 Privet Drive	B1	42 Victoria Rd
S6	Tim Penman	21b Baker Street	B1	42 Victoria Rd

B2 de çalışan personel kalmadığı zaman B2 adresi kaybedilmiş olur.

Normalizasyon Neden Yararlıdır

İyileştirilmiş Tasarım

StaffId Name Staff Address Branch# Branch Address

S1	Fred Bloggs	23 Acacia Gardens	B1	42 Victoria Rd
S2	Bill Sykes	17 Mafeking Terrace	B1	42 Victoria Rd
S3	George Shaw	42 Privet Drive	B1	42 Victoria Rd
S4	John Doe	5 Mornington Crescent	B2	112 King St
S5	Tommy Atkins	10 Rillington Place	B2	112 King St

StaffId Name Staff Address Branch#*

S1	Fred Bloggs	23 Acacia Gardens	B1
S2	Bill Sykes	17 Mafeking Terrace	B1
S3	George Shaw	42 Privet Drive	B1

Branch# Address

B1	20 Union street
B2	112 King St
B3	26 Garthdee Rd

S6	Ed Grundy	Keeper's Cottage	B1
----	-----------	------------------	----

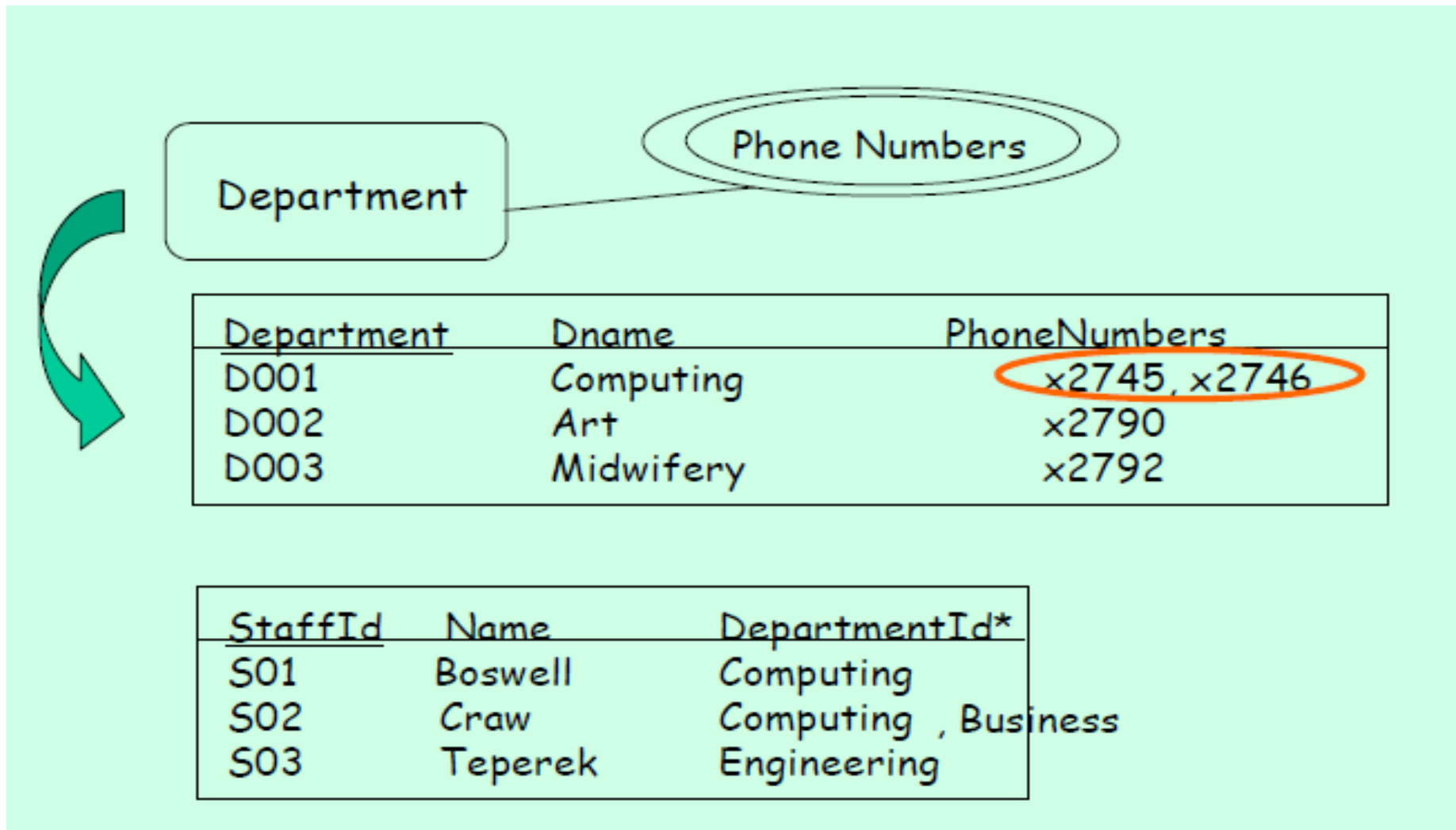
Birinci Normal Form(1NF)

Tabloları 1. Normal forma dönüştürmek için:

- ✓ Tüm alanlar tek değerli olmalıdır.
- ✓ Tüm alanlar birincil alan tarafından belirlenebilmelidir. (Alanlar, birincil anahtara fonksiyonel bağımlı olmalıdır)

Birinci Normal Form(1NF)

Tablolar 1.Normal Formda Değillerdir



Birinci Normal Form(1NF)

Department Tablosunu 1.Normal Forma Dönüştürme

<u>ID</u>	Name	Department
S1	Boswell	Computing, Art
S2	Craw	Computing, Business

Result of normalisation

<u>ID</u>	Name
S1	Boswell
S2	Craw

<u>ID*</u>	<u>Department</u>
S1	Computing
S1	Art
S2	Computing
S2	Business

Birinci Normal Form(1NF)

Tekrarlanan Gruplar

<u>OrderNo</u>	SuppNo	PartNo	Descr	UnitPrice	Quantity	TotalCost
1	S1	P1	Screw	6p	3	18p
2	S1	P2	Nut	7p	4	28p
3	S1	P3	Bolt	10p	1	10p
4	S2	P1	Screw	6p	3	46p
5	S2	P3	Bolt	7p	4	28p

Yukarıdaki tablo 1. Normal Formdadır.

Tabloda birincil anahtar var

Her alanda tek değer vardır.

Buna rağmen veri tekrarı vardır.... Nedeni?

İkinci Normal Form(2NF)

Purchase_Item Tablosundaki Tutarsızlık

Purchase_Item

<u>OrderNo</u>	<u>PartNo</u>	Descr	Quantity	UnitPrice	ItemPrice
1	P1	Screw	3	6p	18p
1	P2	Nut	4	7p	28p
1	P3	Bolt	1	10p	10p
2	P1	Screw	3	6p	18p
2	P3	Bolt	4	10p	40p

İkinci Normal Form(2NF)

Fonksiyonel Bağımlılık

- ✓ Kayıtların ayırt edilebilmesi, tablolar arasında ilişki kurulabilmesi sağlar.
- ✓ Anahtarlar belirlenirken fonksiyonel bağımlılık kavramının bilinmesi gerekir.

ogrenciNo alanı kullanılarak öğrenci adı belirlenebilir. (Tersi doğru değildir...)

Bu durumda:

ogrenciAdi alanı, ogrenciNo alanına fonksiyonel bağımlıdır.

ogrenciNo alanı, ogrenciAdi alanını belirler.

ogrenciNo → ogrenciAdi şeklinde gösterilir.

Ogrenciler										
ogrenciNo TC Kimlik No	adi	soyadi	sifre md5 formatında saklanıyor...	telefonNo	eposta	babaAdi	adres	dogumTarihi	il	ilce
00000000001	Ayşe	Demirr	a7f4e18520f1a28fb9b1edb53f9fd6b6		ad@a.com	Hasan	Bilinmiyor	0000-00-00	34	409
00000000003	Hasan	Çelik	hasancelik		hc@a.com	Hasan	Bilinmiyor	NULL	01	001
00000000004	Tamer	Yorulmaz	e1e6205a7c630320a8f854df101905fb		ty@a.com	Yılmaz	Bilinmiyor	1975-05-01	01	008
00000000008	Ayşe	Eren	e78c265a4f809993ccb24c6ea5c308dc		aer@a.com	Mustafa	Konya	1994-06-07	42	560
00000000009	Ayşe	Yılmaz	9693bb4495eae586d84e2001f1d665ac		ay@a.com	Ahmet	Kocaeli	1999-05-01	41	533
00000000021	Ayten	Gül	035e15c85c630a56ebfd9d44f7796da1	1234567892	Girilmemiş	Girilmemiş	Gebze	1993-09-01	00	940

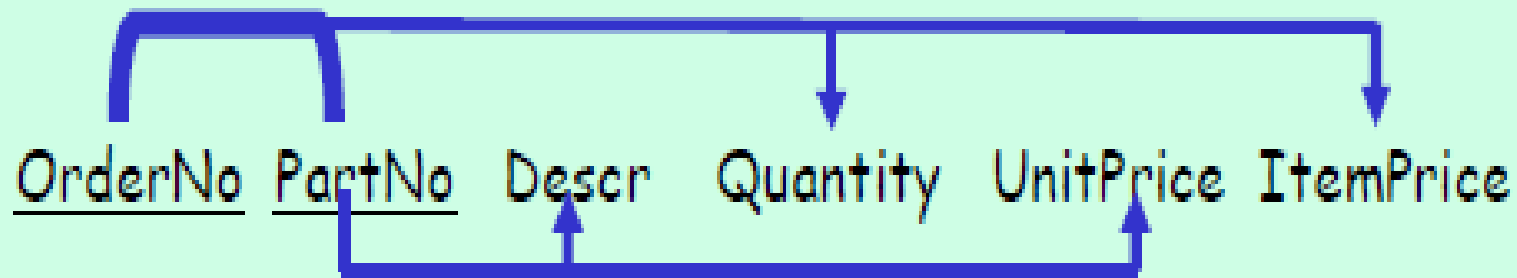
İkinci Normal Form(2NF)

Tam Fonksiyonel Bağımlılık

- Attributes may depend on a *set* of other attributes
 - StudentId, ModuleName \rightarrow ExamMark
 - OrderNo, PartNo \rightarrow Quantity
- D is *fully functionally dependent* on A, B, C if
$$A, B, C \rightarrow D \text{ but } A, B \not\rightarrow D \quad B, C \not\rightarrow D \dots$$
i.e. all the attributes on the LHS are needed to determine the RHS

A table is in 2nd normal form, if all non-primary attributes are fully functionally dependent on the primary key

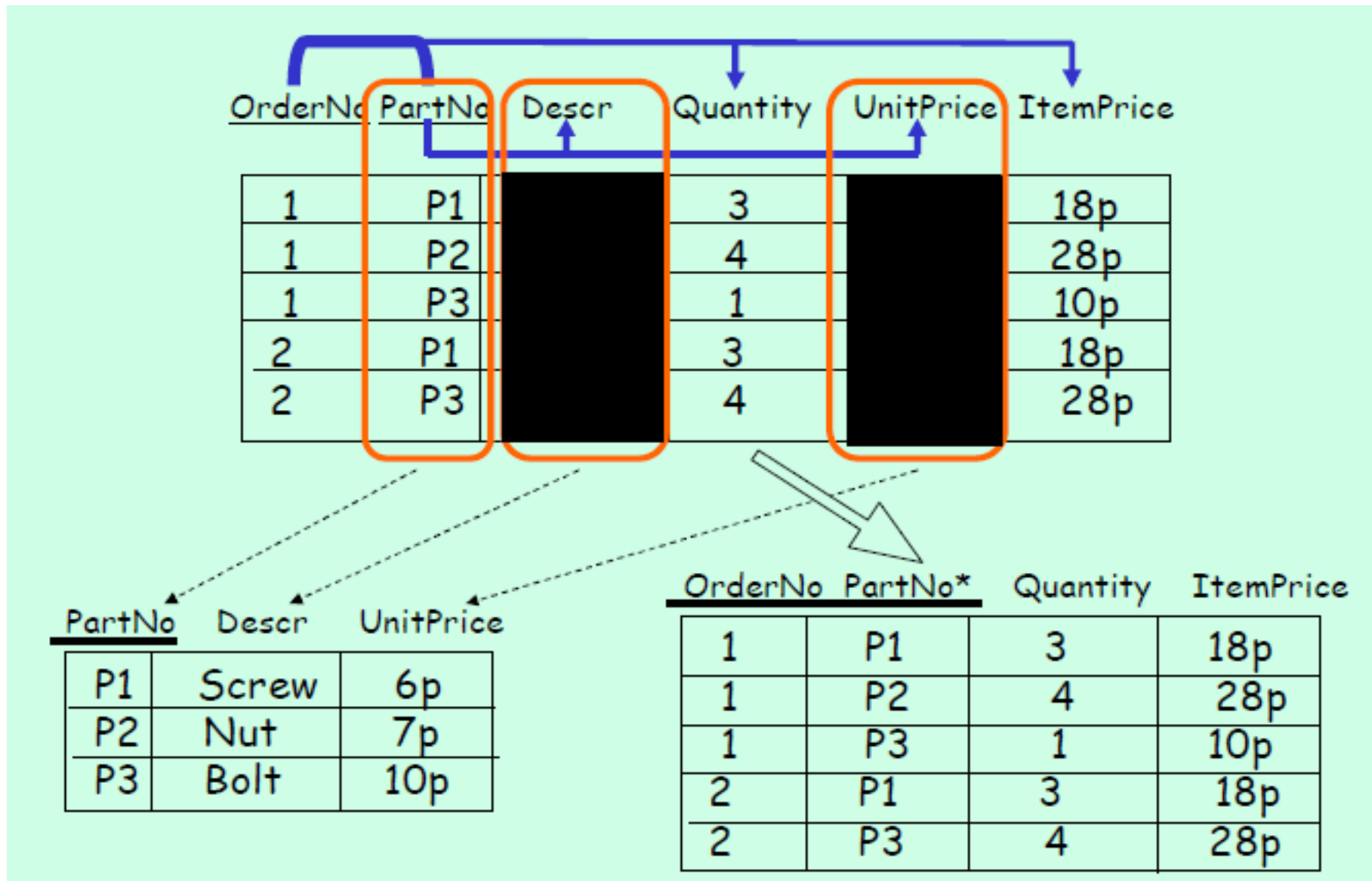
Tam Fonksiyonel Bağımlılık



1	P1	Screw	3	6p	18p
1	P2	Nut	4	7p	28p
1	P3	Bolt	1	10p	10p
2	P1	Screw	3	6p	18p
2	P3	Bolt	4	10p	40p

- Quantity is fully functionally dependent on OrderNo and PartNo
- Description is NOT FFD on OrderNo and PartNo, because it is entirely determined by PartNo
 - E.g., Part P3 is a Bolt, regardless of which Order it's in

Kısmi Bağımlılıkların Giderilmesi



İkinci Normal Form(2NF)

Second normal form

A table is in *second normal form* if

- it is in first normal form

- and there are no partial dependencies

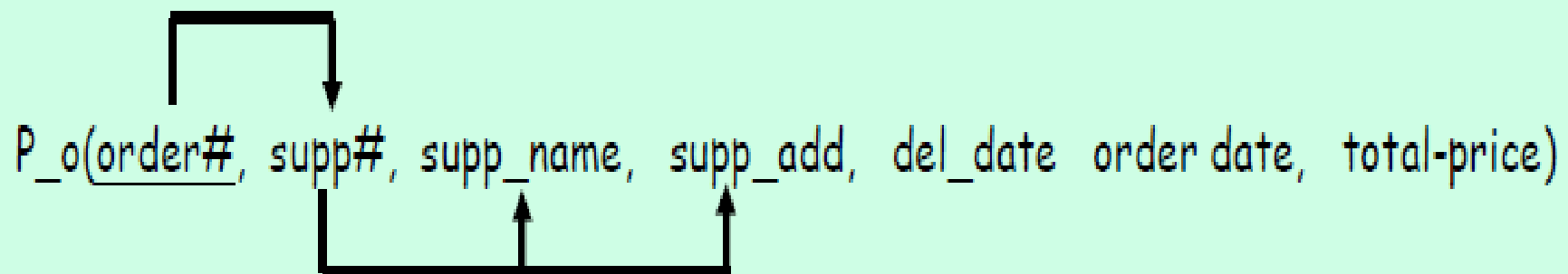
- (more precisely, if every non-key attribute is fully functionally dependent on the primary key)

Note, a (1NF) table is *automatically* in second normal form if its primary key is atomic (has just one attribute)

Üçüncü Normal Form(3NF)

order# supp# supp_name supp_add del_date order date total-price

01	S1	Safeways	King St	Feb 1st	Jan 28th	£50
02	S1	Safeways	King St	Feb 12th	Feb 4th	£85
03	S1	Safeways	King St	Feb 14th	Feb 12th	£30
05	S2	Sommerfield	George St	Feb 5th	Feb 8th	£20
06	S2	Summerfield	George St	Feb 7th	Feb 9th	£100



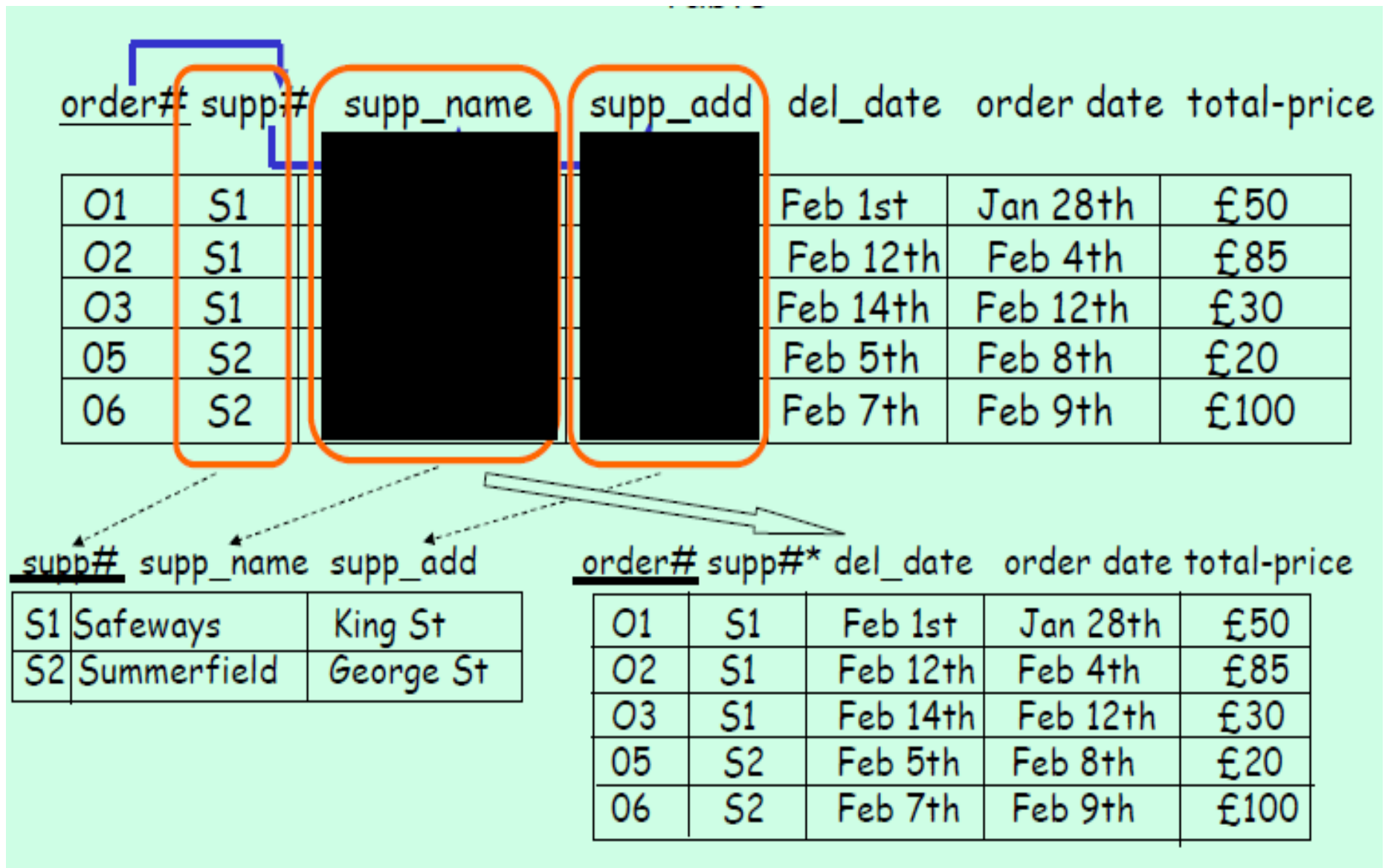
Üçüncü Normal Form(3NF)

Geçişken Bağımlılık

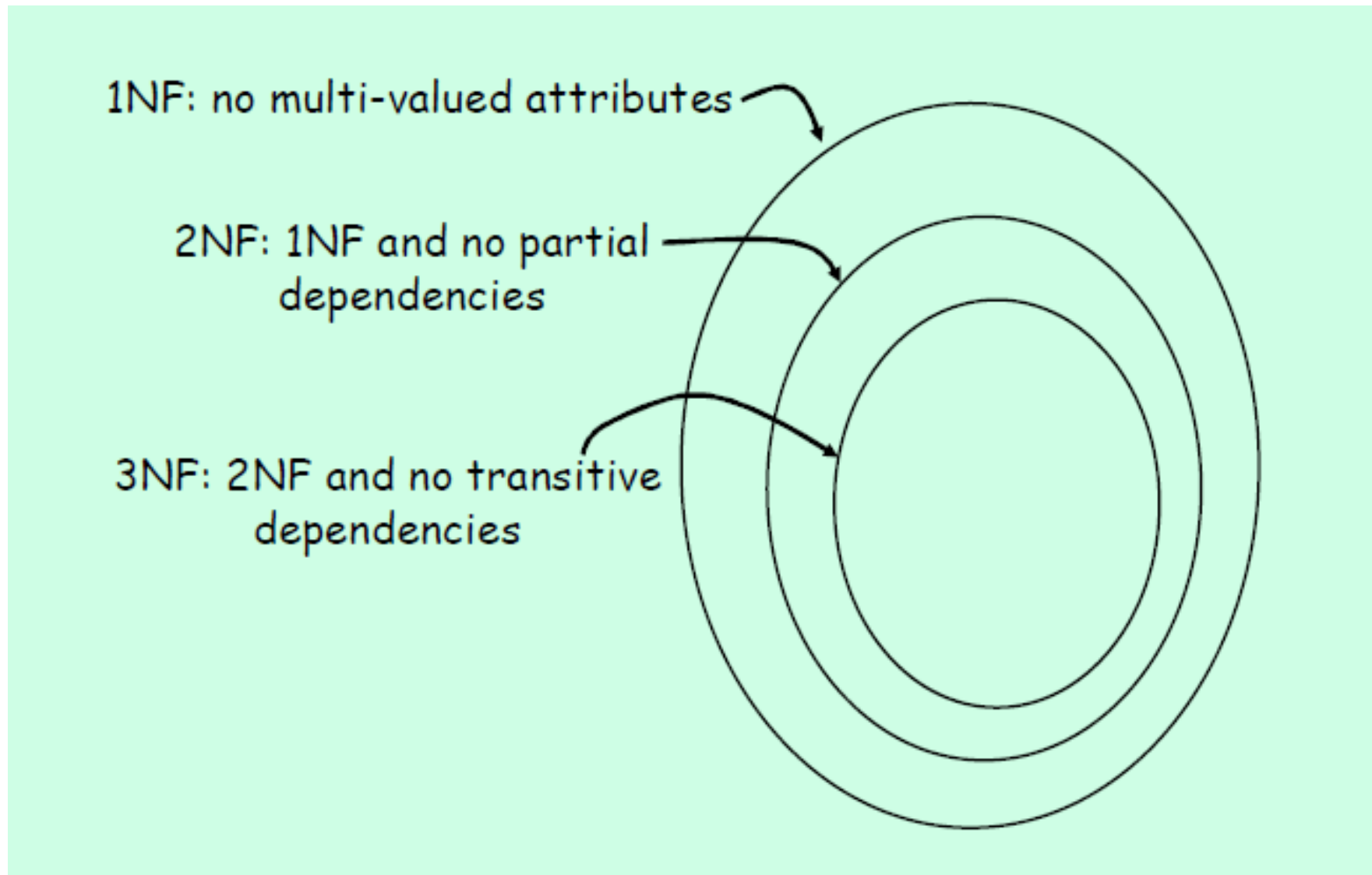
- If $A \rightarrow B$ and $B \rightarrow C$, then we can write
 - $A \rightarrow B \rightarrow C$
 - $\text{OrderNo} \rightarrow \text{SupplierNo} \rightarrow \text{SupplierName}$
- We say
 - "*C is transitively dependent on A*"
 - "*A determines C via B*"
- So for the supplier table we say
 - "Supplier name is transitively dependent on OrderNo"
 - "OrderNo determines SupplierName via SupplierNo"

*A table is in 3rd normal form
if it is in 2nd normal form
and there are no transitive dependencies*

Purchase_Order Tablosundan Geçişken Bağımlılık



1., 2. ve 3. Normal Formlar arasındaki İlişki



Tablolar 3. Normal Formda mıdır?

1

<u>StudentID</u>	<u>ModuleID</u>	CWGrade	ExamGrade
u010123	CMM003	6	5
u010999	CMM003	5	4
u010123	CMM007	5	3
u010999	CMM007	2	5

2

<u>StudentID</u>	<u>ModuleID</u>	CWGrade	CWResult
u010123	CMM003	6	Pass
u010999	CMM003	5	Pass
u010123	CMM007	5	Pass
u010999	CMM007	2	Fail

3

<u>StudentID</u>	<u>ModuleID</u>	CWResult	CWGrade
u010123	CMM003	Pass	6
u010999	CMM003	Pass	5
u010123	CMM007	Pass	5
u010999	CMM007	Fail	2

Normalize Edilmemiş Bir Tablo

<u>StudentID</u>	<u>Module</u>	CWGrade	ExamGrade
u010123	CMM003	6	5
u010999	CMM003	5	4
u010123	CMM007	5	3
u010999	CMM007	4	5

$\text{StudentID, Module} \rightarrow \text{CWGrade, ExamGrade}$

Equivalently, CWGrade and ExamGrade are single-valued

Therefore, the table is in 1NF

CWGrade and ExamGrade are fully functionally dependent on StudentID and Module

Equivalently, there are no partial dependencies

Therefore, the table is in 2NF

There are no transitive dependencies

Therefore, the table is in 3NF

Kaynaklar

- ✓ Dr. Brayant, <http://www.comp.rgu.ac.uk/staff/chb/teach.html>