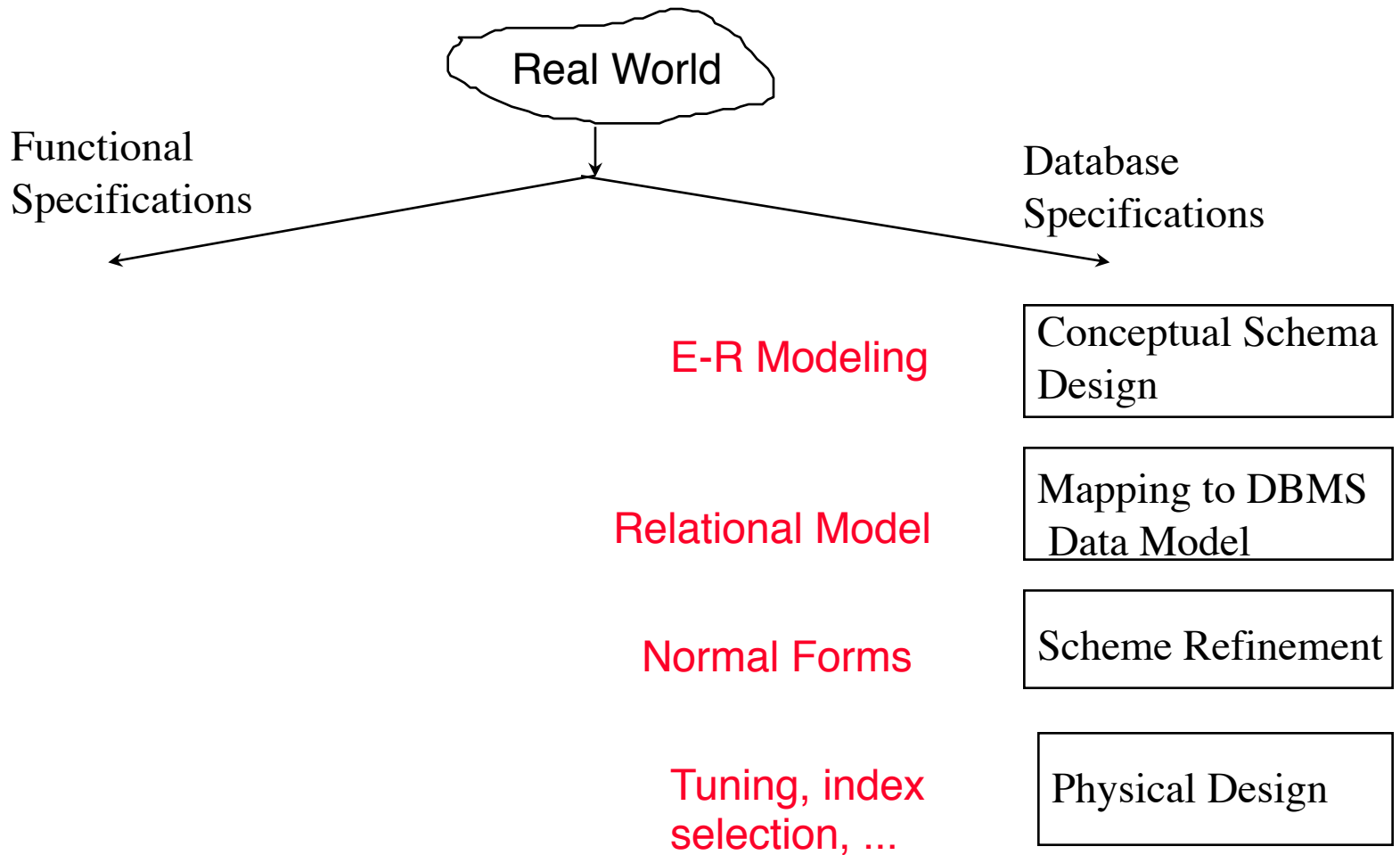


Database Design Process



Example

Vendors

vno	vname	vaddress
v10	Bay	Southgate Mall, Edmonton
v20	Wall Mart	Calgary Trail, Edmonton
v30	Sears	Golden Mall, Calgary

Customers

account	cname	caddress	province	balance	crlimit
a1	J. Doe	101-28 Ave, Edmonton	AB	800	1000
a2	M. Smith	93-105 St, St. Albert	AB	1200	3000
a3	J. Miller	34-75 Ave, Edmonton	AB	500	2500
a4	D. Rafiei	111-22 Ave, Edmonton	AB	0	1200

Transactions

account	vno	amount
a1	v30	68.29
a2	v10	25.86
a1	v10	56.37
a3	v20	12.40
a2	v30	38.12
a3	v10	75.70
a1	v20	68.00
a2	v10	21.60
a1	v30	88.20
a2	v10	25.50



The Same Data (alternative schema)

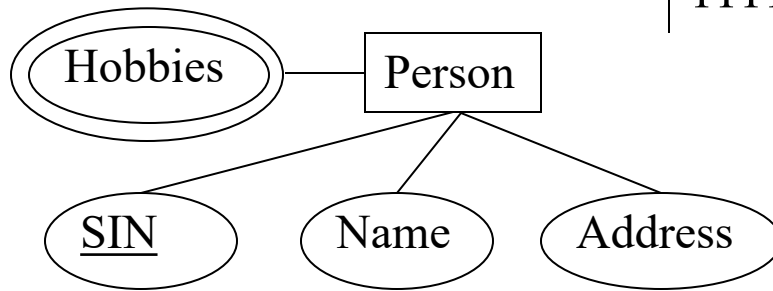
SuperRelation

account	cname	caddress	province	balance	crlimit	vno	vname	vaddress	amount
a1	J. Doe	101-28 Ave, Edmonton	AB	800	1000	v30	Sears	Golden Mall, Calgary	68.29
a2	M. Smith	93-105 St, St. Albert	AB	1200	3000	v10	Bay	Southgate Mall, Edmonton	25.86
a1	J. Doe	101-28 Ave, Edmonton	AB	800	1000	v10	Bay	Southgate Mall, Edmonton	56.37
a3	J. Miller	34-75 Ave, Edmonton	AB	500	2500	v20	Wall Mart	Calgary Trail, Edmonton	12.40
a2	M. Smith	93-105 St, St. Albert	AB	1200	3000	v30	Sears	Golden Mall, Calgary	38.12
a3	J. Miller	34-75 Ave, Edmonton	AB	500	2500	v10	Bay	Southgate Mall, Edmonton	75.70
a1	J. Doe	101-28 Ave, Edmonton	AB	800	1000	v20	Wall Mart	Calgary Trail, Edmonton	68.00
a2	M. Smith	93-105 St, St. Albert	AB	1200	3000	v10	Bay	Southgate Mall, Edmonton	21.60
a1	J. Doe	101-28 Ave, Edmonton	AB	800	1000	v30	Sears	Golden Mall, Calgary	88.20
a2	M. Smith	93-105 St, St. Albert	AB	1200	3000	V10	Bay	Southgate Mall, Edmonton	25.50
a4	D. Rafiei	111-22 Ave, Edmonton	AB	0	1200	NULL	NULL	NULL	NULL

What is wrong with this design?

2nd Example

ER Model



<i>SSN</i>	<i>Name</i>	<i>Address</i>	<i>Hobby</i>
1111	Joe	123 Main	{biking, hiking}

Relational Model

<i>SSN</i>	<i>Name</i>	<i>Address</i>	<i>Hobby</i>
1111	Joe	123 Main	biking
1111	Joe	123 Main	hiking
.....			

Redundancy



3rd Example

- Any redundancy here?

<i>SSN</i>	<i>Name</i>	<i>Town</i>	<i>Zip</i>
1234	Joe	Stony Brook	11790
4321	Mary	Stony Brook	11790
5454	Tom	Stony Brook	11790
.....			

- All addresses in the same town have the same zip code



Guidelines

- Information repetition should be avoided
 - Anomalies: insertion, deletion, modification
- Avoid null values as much as possible
 - Difficulties with interpretation
 - ✓ don't know, don't care, known but unavailable, does not apply
 - Specification of joins
- General Solution: Normalization



Normal Forms

